DISENTANGLING CULTURAL HOMOPHILY FROM CONFOUNDING MECHANISMS

A Thesis

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Master of Arts

by

Brandon R. Sepulvado

________________________________________

Omar Lizardo, Director

Graduate Program in Sociology

Notre Dame, Indiana

December 2014
A high level of clustering distinguishes social networks from other types of networks. Homophily, or the tendency for similar individuals to befriend each other, is commonly purported to generate this phenomenon. While much extant research on homophily emphasizes factors that structurally induce similarity among individuals, less work has investigated how cultural tastes result in individuals preferentially selecting homogeneous alters and how best to approach this methodologically. Using novel data, this study employs statistical network analyses to infer how music tastes produce different structures within an emerging social network by influencing the number and similarity of friends that a person has. I find music tastes vary in the extent to which they produce homophily but that, in order to uncover these dynamics, one must account for the number of people with whom an individual interacts. I conclude with a theoretical discussion of the findings and suggest future research directions.
CONTENTS

Figures ................................................................................................................................. iii

Tables ................................................................................................................................. iv

Acknowledgments ................................................................................................................ v

Chapter 1: Introduction ....................................................................................................... 1

Chapter 2: Musical taste and homophily ............................................................................ 5

Chapter 3: Sources of spurious homophily ....................................................................... 9

Chapter 4: Data and methods ............................................................................................ 12

Chapter 5: Results ............................................................................................................... 18

Chapter 6: Discussion and conclusion ............................................................................... 28

References ............................................................................................................................ 34
FIGURES

Figure 1: Goodness of Fit Simulation Results ........................................................................ 27
TABLES

Table 1: Statistical models ................................................................. 19
I would like to thank Omar Lizardo, David Hachen, and David Gibson for the guidance throughout the thesis process. This thesis was not accomplished as a singular task but rather came about through a roundabout path of reading, data analysis, writing, re-reading, re-analyzing, and re-writing. None of it would have been possible without their help.

I also benefitted greatly from financial support. This work was made possible by a grant from the National Science Foundation (CBET-09-41565) and was sponsored by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053.
A high level of clustering distinguishes social networks from other types of networks
(Newman and Park 2003). One mechanism generating this unique characteristic is
assortative mixing (Newman 2002, 2003) or homophily (Lazarsfeld and Merton 1954;
Louch 2000; McPherson, Smith-Lovin, and Cook 2001). The basic idea is that individuals
who are similar on some dimension will tend to form ties, and that this “mixing of a
network by a discrete characteristic will tend to break the network up into separate
communities” (Newman 2003:1). While homophily is an important characteristic of
social networks and their formation because of its contribution to their distinctive
structure, it can be difficult to study for reasons discussed below.

Homophily is generally described along three dimensions. The first concerns the
presence of homophily and whether similar people tend to be connected more than one
would expect by chance. Similarity between individuals due purely to population
composition and random matching has been described as baseline homophily, and
similarity above and beyond what is expected by chance has been termed inbreeding
homophily (McPherson et al. 2001:9). For example, in a recent paper, Smith,
McPherson, and Smith-Lovin (2014) state that “compositional changes often swamp the
effect of other macro-level forces on absolute homophily” and go on to show a large discrepancy between the homophily expectations based solely upon demographic considerations and those uncovered after statistical controls. Baseline homophily is therefore spurious homophily because it results from random processes that generate dyadic similarity due solely to population composition.

The second dimension of homophily is the basis on which it is formed. Discussion about homophily’s basis usually takes the form of a status versus values characterization (McPherson et al. 2001). Status homophily results from externally imposed or “ascribed characteristics”, while value homophily is rooted in “internal states presumed to shape our orientation toward future behavior” (McPherson et al. 2001:419). The former has been documented on the basis of race, sex, education, and age to name just a few (Brashears 2008; Louch 2000; McPherson et al. 2001), and the latter can be seen with regard to musical tastes, religion, and ethnicity (Brashears 2008; Lewis, Gonzalez, and Kaufman 2012; Smith et al. 2014; Steglich, Snijders, and West 2006; Wimmer and Lewis 2010). While homophily with regard to these traits is generally stable, its magnitude is not (Smith et al. 2014). For instance, race, religion, sex, age, and education remained salient dimensions along which people sorted themselves from 1985 to 2004, yet the strength of religious, age, and educational homophily changed. Age and educational homophily strengthened for close confidants, and religious homophily weakened.

The final dimension concerns the mechanisms that generate homophily. Some homophily is induced by external factors, and some homophily results from preferential selection of actors (McPherson et al. 2001). Organizations are one source for structural
inducement of social ties. Organizations serve as conduits directing social behaviors (Feld 1981), and the degree to which organizations take advantage of this fact can actually influence their survival. Some organizations serve as very restrictive foci, “requiring each pair of individuals [members] to devote [much] time and energy to participating in joint activities associated with that focus” (Feld 1981:1025).

Organizations that do so promote homophily among their members and specialize their ecological niche in relation to other organizations. Ultimately, requiring a great deal of time and other resources and (at least) indirectly discouraging ties to external members has been shown to result in lower membership loss and turnover (McPherson 1983; Popielarz and McPherson 1995; Ranger-Moore and McPherson 1991). Religiosity can serve as an example of actor selected homophily. American evangelicals see themselves as living in a corrupted world where they need to spread the gospel. While they might make many transient contacts who are not saved, the evangelicals tend to form close ties with only like-minded others (Smith et al. 1998).

Cultural homophily is particularly interesting because it often precludes an easy characterization of the bases and mechanisms driving homophily. For instance, some music genres are commonly linked with specific racial groups. An example of this will be included in the later analyses. Further, cultural homophily might emerge from preferential selection, and the results of which might then drive the construction of a group that encourages further structurally induced homophily. Religious beliefs exemplify this feedback loop (Riley 2005). There are many colleges and universities in the United States that were founded on the basis of integrating religious beliefs into
higher education. In addition to the initial establishment of these colleges, prospective students enroll because they seek an environment structured largely around their ideology. That is, religious students can preferentially select into an environment which structurally induces religious homophily.
CHAPTER 2:
MUSICAL TASTE AND HOMOPHILY

Cultural tastes, including music preferences, provide interactional fodder, which promotes (or dissuades) contact between individuals (DiMaggio 1987). Over the course of many interactions, differences in musical tastes get translated into structural and compositional patterns within social networks. Research regarding musical taste(s) originally examined how preferences relate to each other, but more recent work has begun to investigate how preferences impact social network characteristics.

What was once a distinction between elite and mass cultures [8] evolved into that between cultural omnivores and univores (Peterson and Kern 1996; Peterson 1992) and now into an even greater differentiation among consumers, e.g., exclusive versus inclusive highbrow, inclusive non-highbrows, and exclusive lowbrows (Ollivier 2008). While patterns of cultural consumption have become increasingly diverse over the past three decades (López-Sintas and Katz-Gerro 2005) and particular classifications and terminologies have shifted, the underlying idea is that the music genres to which an individual listens signify certain information about the person and erect boundaries between co-consumers and those who do not listen to the genre(s). What might at first appear as an unimportant or innocent choice--to which song one chooses to listen--
becomes a way through which symbolic boundaries and various inequalities are reproduced.

In the last two decades or so, omnivorousness--listening to and being well acquainted with a diverse range of music genres--has become a mark of privilege. It takes a good deal of resources to familiarize oneself with a lot of music genres and even more to cultivate a critical appreciation (Bryson 1996). Therefore, it should be no surprise that it is the economically privileged omnivores who use their class status to distinguish themselves by appreciating and engaging in many forms of consumption, which can lead to further economic advantages (DiMaggio 1987; Erickson 1996), yet omnivores do not uncritically listen to every music genre. For example, the majority of the U.S. public enjoys country and gospel music, yet the most omnivorous among the citizenry dislike the two genres (Bryson 1996:894). On the other hand, low-status Americans' consumption profiles tend to be much more limited, and they often couple musical taste with racial, ethnic, religious, and geographic boundaries (Bryson 1997).

These boundaries result in different sets of relations between genres. Goldberg (Goldberg 2011) finds three logics pertaining to music consumption: omnivore versus univore, highbrow versus lowbrow, and contemporary versus traditional. In the first group, there are no dislikes and no hierarchical ordering of genre preferences. Individuals with this consumption style are generally open to all genres, though they listen to some more than others. The highbrow/lowbrow cluster is organized largely around “an opposition between canonical, instrumental, and traditional music on the one side and popular contemporary music on the other” (Goldberg 2011:1418). Finally,
the contemporary versus traditional logic revolves around a tension between a cosmopolitan appreciation as opposed to a preference for “local identity that is rooted in white American tradition” (Goldberg 2011:1419). Even among individuals who are open to myriad genres, boundaries get constructed between acceptable and unacceptable diversity (Ollivier 2008). Some see omivorousness as a way to cultivate personal worth, while others appreciate a wide range of genres without a critique of the culturally benighted. Still, others structure their consumption habits around practical routines or needs.

In addition to examining which music genres tend to be related, insofar as the same individuals listen to them, research also has examined the link between music tastes and ego network structure. Lizardo (2006) finds that engaging in highbrow cultural activities (e.g., going to museums, reading books and poetry, and listening to classical music) is associated not only with having a dense personal network of strong ties but also with having few weak ties. Lowbrow (or popular) cultural consumption (e.g., going to the movies or a non-classical concert) is associated with an ego network full of weak ties. Further, the number of cultural preferences an individual holds and the intensity with which the person holds them is related to network structure (Lizardo 2011). An individual who is relatively indifferent with regard to many cultural preferences tends to have a sparse personal network, while people with many strong preferences tend to have dense personal networks.

A few recent studies have investigated the relationship between culture and network composition. For example, one's moral worldview predicts one's ego network
composition better than a spate of sociodemographic traits, such as race, sex, parental education, and income (Vaisey and Lizardo 2010:1608). With regard to musical tastes, Steglich, Snijders, and West (2006) find homophily in adolescent friendships with regard to classical music, and Lewis et al. find that college students who listen to lite or classical rock and classical and jazz music have a disproportionate number of Facebook friends who listen to the same genre(s) (2012).

While there has been an increased focused on how music preferences shape social networks, further work still needs to be done. Lizardo (2006, 2011) convincingly demonstrates the effects of cultural tastes on social network structure, yet this begs the question of what mechanisms drive the relationships he finds. For example, a dense personal network could occur because individuals with a particular trait exhibit homophily or because they tend to have more contacts in general and the network has many people with that trait. The pieces that examine music preferences and homophily (Lewis et al. 2012; Steglich et al. 2006) do control for mechanisms that can potentially obfuscate the connection between cultural tastes and network characteristics. However, they do not present a systematic investigation into the relative effects of confounding mechanisms on the relationship between tastes and network structure. The next section reviews these potential confounders.
CHAPTER 3: SOURCES OF SPURIOUS HOMOPHILY

There is clearly reason to believe a link between cultural consumption and social network structure exists. However, there are several potential confounding relationships that can cause spurious homophily. In this section, I present a brief list of mechanisms that can obfuscate actual homophily within a network.

The first is sociability. Some individuals have more friends than others, and, while this is unsurprising, not taking this into account can have detrimental consequences for conclusions about homophily. As an example, consider Person A and Person B. Person A has twice the number of friends that Person B has. Given the large difference, we would expect Person A to have more friends of all possible characteristics, but do Person A and Person B exhibit the same levels of homophily among their contacts? The only way to know this is to condition on each individual's degree. Previous investigations of ethnic and racial (Wimmer and Lewis 2010) as well as music preference (Lewis et al. 2012; Steglich et al. 2006) have illustrated the importance of taking sociality into account. Investigations employing name generator data, such as that provided by the General Social Survey, do not permit researchers to condition on an individual's degree.
Transitivity is another potentially confounding mechanism in the study of homophily. If people tend to befriend those with whom they have mutual contacts at a higher rate than those with whom they share no contacts, then not conditioning on this phenomenon can falsely inflate the observed homophily. We need to be able to distinguish between friendships that stem from transitivity—even if the people are similar on some trait(s)—and those that result from preferential selection due to some trait(s). Louch (2000) has shown that there is in fact common variation between homophily and transitivity.

Social influence is another possible source of bias. The effects of homophily and social influence can be incredibly difficult to differentiate (Shalizi and Thomas 2011), and each has been shown to have unique influence with regard to music tastes and friendship ties (Lewis et al. 2012; Steglich et al. 2006). These phenomena are so difficult to tease apart because the general effects manifest in the same phenomenon: clustering of similar individuals. Do these clusters exist because similar people searched for each other or because they changed attributes after contact? If one wants to know exactly how much impact homophily has on network formation, then one must find a way to control for social influence.

Unlike the previous confounding mechanisms that can lead to spurious homophily, the next potential source of spuriousness is not something other than homophily. Rather, we might observe spurious homophily on trait p—whether or not homophily actually exists for trait p—if it is correlated with trait q and individuals with trait q exhibit homophily. One common example of this concerns instances of status and
value homophily. As mentioned earlier, homophily is commonly found with regard to race (Brashears 2008; McPherson et al. 2001), yet some of this observed racial homophily is attributable to ethnic homophily, stemming from shared experiences, feelings, and knowledge (Wimmer and Lewis 2010).
CHAPTER 4:
DATA AND METHODS

The data in this paper come from an ongoing study called NetSense (Striegel et al. 2013). The NetSense study equipped roughly 200 incoming first year students at the University of Notre Dame with smartphones and unlimited texting, data, and calling plans. The research team uses the phones to keep track of calls and texts made and received, the duration of calls, the numbers of individuals who call or are called by participants, and the proximity of participants to each other via Bluetooth proximity detection. Additionally, participants completed surveys before they arrived on campus and every four months thereafter. The surveys ask questions about demographic traits, attitudes, cultural tastes, and certain behaviors. Participants also are also asked a narrower subset of questions about their 20 closest contacts.

The network used in the current project was constructed from participants’ calls and texts. Since the effects of homophily and social influence can be difficult to distinguish if measured in a single time period (Shalizi and Thomas 2011), responses to the survey before campus arrival are used to predict patterns of ties formed during the first 18 weeks as indicated by calling and texting. This use of time ordering helps minimize the effect of the confounding social influence mechanism. Ties are undirected,
and they are said to exist when an ego contacts an alter who then responds to the ego. Goodreau (2007:13) calls this mutualization and argues that a “dually nominated tie may be taken as a form of validated relationship”. Mutualization is particularly important for calls and texts among undergraduate students because there are many situations in which someone can call or text another without reciprocation. In addition, we restricted the network to include only Netsense participants because we want to ensure the ties were new, i.e., formed after arrival on campus, and we need the detailed survey information about all individuals in the network in order to examine whether cultural homophily is a predictor of tie formation. The network consists of 186 nodes and 647 edges, resulting in a density of 0.0376.

Exponential-family random graph models (ERGMs) offer an approach for the statistical modeling of networks. Unlike many traditional statistical methods in the social sciences that estimate actor or event attributes, ERGMs predict the probability of a tie between two nodes using graph level indices. They assume that the global structure of ties within a network stems from local processes (Lusher and Robins 2013b) and that the presence of a tie between two nodes depends upon other extant ties in the network (Harris 2014; Hunter and Handcock 2006; Lusher and Robins 2013b). ERGMs compare observed networks to the other possible network configurations, conditioned on structural characteristics of the observed graph, using “a distribution of graphs where the observed data are central in that distribution” (Robins and Lusher 2013b:33). In addition to endogenous network effects, ERGMs permit investigation of the impact of actors' attributes on network structure. A particularly relevant benefit of these
statistical models is that we can interpret a model including structural and actor attribute parameters as providing both information on the effect of actor attributes controlling for structural dynamics and the presence and magnitude of structural processes net of effects related to actor characteristics (Robins and Daraganova 2013:93).

As previously noted, ERGMs estimate the probability of a tie being present between a randomly chosen pair of nodes conditional upon the set of variables in the model. The survey contains questions that ask respondents if they listen to 22 different genres of music: big band, bluegrass, blues or R&B, broadway musicals or showtunes, choral or glee club, classic rock or oldies, classical or chamber music, country, dance music, ethnic or national music, folk music, hymns or gospel music, jazz, Latin, easy listening, new age, opera, operetta or musicals, parade or marching band, rap or hip-hop, reggae, rock, or heavy metal. After preliminary analyses, only big band, classic rock/oldies, country, dance, latin, hymns/gospel, new age, and heavy metal genres were included for two main reasons. Multicollinearity was a problem among several genres, and genres that did not achieve significance on any of the various models' parameters were excluded to help with model convergence. Only degree parameters (to be discussed shortly) are included for new age and heavy metal because preliminary analyses indicated that homophily terms for the two did not achieve significance.

There are three general categories of variables used in the proceeding models, each of which is a graph--not dyadic--level statistic. First, there are several instances of a homophily parameter
\[
\sum_{i<j} x_{ij} y_i y_j
\]

where \(x_{ij}\) is a binary variable indicating the presence of a tie between node \(i\) and node \(j\), \(y_i\) is a binary variable indicating whether node \(i\) has trait \(y\), and \(y_j\) is a binary variable indicating whether node \(j\) exhibits trait \(y\) (Robins and Daraganova 2013). These terms insert into the model a count of the number of ties with adjacent nodes that share a value on a specified trait. These variables will be used to test for homophily on the basis of musical genre preference and to control for homophily on a set of sociodemographic variables.

Second, I use a degree term (Robins and Daraganova 2013:94), based on

\[
\sum_{i<j} x_{ij}(y_i + y_j)
\]

where each variable has the same meaning as before. However, unlike the homophily statistics, the degree ones can assume a third value (i.e., 0-2). This term counts the number of stubs in the network attached to a node with a specified attribute. For a given edge \(x_{ij} = 1\) in the network, node \(i\) may have trait \(y\), node \(j\) may have trait \(y\), both may have trait \(y\), or neither may have it. These terms are necessary to accurately estimate homophily on a characteristic because it controls for the degree of nodes with the attributes of interest. If one did not include these activity measures, it would be impossible to conclude that the observed level of preferential selection did not simply occur due to an increased number of total ties. The degree parameters control for spurious homophily resulting from sociality being associated with a given trait.
The last type of variable pertains to the structure of the network, and I utilize two of these terms. The first term

\[ \sum_{i<j} x_{ij} \]

conditions on the number of edges in the graph. It may be interpreted like the constant in a traditional OLS regression, providing the intercept and expected probability when all the other model terms are set to zero (Lusher and Robins 2013a:42). This base probability that two nodes will be connected is simply a function of the graph’s density.

The other term accounts for triadic closure. The parameter that captures closure in the proceeding models assumes a form of dependency in which two ties can be related—even though they do not share any nodes—if the two dyads are connected by two other ties, forming a four cycle (Harris 2014; Koskinen and Daraganova 2013). Snijders and colleagues originally termed this social circuit dependency (2006), and Hunter (2007) and Hunter and Handcock (2006) further revised this formulation to create the geometrically weighted edgewise shared partner (GWESP) distribution. The GWESP statistic (Goodreau, Kitts, and Morris 2009; Hunter and Handcock 2006; Hunter 2007) equals

\[ e^{\alpha} \sum_{i=1}^{n-2} \{1 - (1 - e^{-\alpha})^i\}p_i \]

where \( p_i \) is the number of edges that share \( i \) contacts and \( \alpha \) “controls the geometric rate of decline in the effect of triad closure on tie probability for an increasing number of shared partners” (Goodreau et al. 2009:112–13). The GWESP statistic counts the
number of nodes to which both adjacent nodes on an edge are connected and then
conditions on the distribution of these counts. Controlling for the GWESP distribution is
important for investigations into homophily because it accounts for triadic closure, the
tendency for two nodes to be connected due to their common association with a third
node. Ignoring this triadic process could lead to an upwardly biased estimate for a
homophily coefficient because spurious homophily could also be caused by befriending
friends' friends.

Several control variables are included in the final model. Homophily has been
demonstrated on the basis of gender, race/ethnicity, religion, and socioeconomic status
(Brashears 2008; Louch 2000; McPherson et al. 2001; Smith et al. 2014; Wimmer and
Lewis 2010), and I include homophily and degree parameters to disentangle these
effects from musical taste homophily. I also include a term for residence hall homophily
because the University randomly assigns students to same-sex residence halls. Only
significant control terms are presented in the models.
CHAPTER 5:

RESULTS

Table 1 presents the results of several models. The first model uses only homophily variables to predict the probability of a tie between two randomly chosen vertices in the graph. Model 2 adds degree parameters to the first model, and Model 3 adds the GWESP term, to condition on the level of triadic closure in the network, to Model 1. Model 4 combines Models 2 and 3, and the final model includes the control variables on top of Model 4's contents.
## TABLE 1:

### STATISTICAL MODELS

<table>
<thead>
<tr>
<th>Homophily</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Band</td>
<td>0.61*</td>
<td>0.64*</td>
<td>0.50*</td>
<td>0.61*</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.30)</td>
<td>(0.21)</td>
<td>(0.28)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>Classic Rock/Oldies</td>
<td>-0.21</td>
<td>0.36*</td>
<td>-0.16</td>
<td>0.33</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.16)</td>
<td>(0.08)</td>
<td>(0.17)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Country</td>
<td>0.29**</td>
<td>0.49**</td>
<td>0.21**</td>
<td>0.46**</td>
<td>0.38*</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.16)</td>
<td>(0.08)</td>
<td>(0.16)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Dance</td>
<td>0.18*</td>
<td>0.41*</td>
<td>0.11</td>
<td>0.39*</td>
<td>0.45*</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.16)</td>
<td>(0.07)</td>
<td>(0.16)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Latin</td>
<td>0.94***</td>
<td>1.23***</td>
<td>0.71***</td>
<td>1.09***</td>
<td>0.76**</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.24)</td>
<td>(0.13)</td>
<td>(0.21)</td>
<td>(0.27)</td>
</tr>
</tbody>
</table>

### Degree

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Band</td>
<td>-0.05</td>
<td>-0.07</td>
<td>-0.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.08)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>Classic Rock/Oldies</td>
<td>-0.41***</td>
<td>-0.31**</td>
<td>-0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>Classical</td>
<td>0.13</td>
<td>0.08</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>-0.15</td>
<td>-0.17</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>Dance</td>
<td>-0.26</td>
<td>-0.24</td>
<td>-0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.09)</td>
<td>(0.11)</td>
<td></td>
</tr>
<tr>
<td>Hymns/Gospel</td>
<td>0.23**</td>
<td>0.14*</td>
<td>0.19*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td></td>
</tr>
<tr>
<td>Latin</td>
<td>-0.14</td>
<td>-0.16</td>
<td>-0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.08)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>New Age</td>
<td>0.27***</td>
<td>0.16**</td>
<td>0.13*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>Heavy Metal</td>
<td>-0.18**</td>
<td>-0.11*</td>
<td>-0.12*</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 1 cont.

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female-Degree</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Af. American-Degree</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino-Degree</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian American-Degree</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other-Degree</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atheist-Degree</td>
<td>-0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racial Homophily</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Res. Hall-Homophily</td>
<td>1.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental edu.-Homophily</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atheist-Homophily</td>
<td>1.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Structural Items</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edges</td>
<td>-3.35</td>
<td>-2.79</td>
<td>-4.20</td>
<td>-3.72</td>
<td>-4.63</td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.13)</td>
<td>(0.08)</td>
<td>(0.13)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Triadic Closure</td>
<td>0.97</td>
<td>0.96</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>5479.18</td>
<td>5435.89</td>
<td>5225.24</td>
<td>5196.18</td>
<td>4596.08</td>
</tr>
<tr>
<td>BIC</td>
<td>5525.70</td>
<td>5552.19</td>
<td>5279.51</td>
<td>5320.23</td>
<td>4805.41</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-2733.59</td>
<td>-2702.95</td>
<td>-2605.62</td>
<td>-2582.09</td>
<td>-2271.04</td>
</tr>
</tbody>
</table>

*** p < 0.001, ** p < 0.01, * p < 0.05
Model 1 includes six parameters. The edges parameter is present in every model, and it may be recalled that this parameter conditions on the number of edges present in the network. Basically, this term indicates the probability of a tie for an individual who listens to none of the genres for which homophily parameters are included. Moving to the substantive variables of interest, we can see that each of the five genres for which homophily parameters are included are significant. Listening to big band music during high school increases the probability that an individual in this network will have a tie with another individual who listens to big band. Homophily is also present on the basis of country, dance, and Latin music. However, there are significantly fewer ties between students who listened to classic rock and oldies during high school than would be expected with random matching.

Model 2 includes the degree parameters because they control for one possible mechanism leading to a spurious connection between homophily and musical taste. If people who listen to a specific genre have a higher degree than those who do not listen to it, then even with random matching we would expect more dyads where both adjacent nodes listen to the genre. Thus, we need both parameters to make an accurate inference about the presence and effect size of homophily. The homophily coefficient for the big band genre changed little: it increased by .03, and the degree parameter was not significant. Each of the remaining homophily coefficients substantially increased. The beta for classic rock and oldies changed direction and is now .36, and the degree parameter for this genre is significant and negative. The homophily coefficient for
country music increased by .20 to .49, and the coefficient for this genre's degree is significant at -.15. Similarly, the homophily coefficients for the dance and Latin genres increased by .23 and .29 respectively, and both genres had significant negative coefficients for the degree parameters, -.26 and -.14 respectively.

The common findings for classic rock/oldies, country, dance, and Latin are (1) that the homophily parameters indicate that the addition of an edge where both adjacent nodes listen to the genre significantly increases the log odds of a tie existing between a randomly chosen node pair from the network and (2) that the addition of a stub in which the adjacent node listens to the genre actually decreases the log odds of a tie being present between two individuals in the social network. These findings suggest that the consumption of these musical tastes is accompanied by a communitarian or subcultural disposition, but these results unfortunately do not permit a conclusion regarding the mechanism of such dispositions. For example, it could be that these tastes require more time and effort to adequately consider oneself a consumer and that doing so requires or is at least aided by having many similar friends, or this could result from external inducement, e.g. the genres are stigmatized by a majority of individuals in the social environment. While this latter explanation might be able to account for the significantly lower degree finding, it should be noted that the former explanation cannot.

The last major finding from Model 2 is that a group of genres has degree effects that are separate from homophily effects. A significant degree parameter indicates that “nodes with the attribute tend to have higher network activity (i.e., more ties) than
nodes without the attribute” (Robins and Daraganova 2013:95). Both individuals who
listen to gospel/hymn music and those who listen to new age music have larger degrees
than individuals who do not listen to the genres. On the other hand, individuals who
listen to heavy metal music have lower degrees than those who do not listen to it. This
begs an explanation that links cultural tastes with the number of ties an individual has.

Model 3 looks at the effect that the GWESP term has on the naive homophily
parameters. It should be recalled that the GWESP term controls for transitivity. First, the
GWESP coefficient (.97) is significant, indicating that two individuals within the network
are more likely to befriend each other if they share a common friend or friends than if
they have no mutual contacts. The effect of all the genres except classic rock/oldies on
the probability of a tie between two people actually decreases from Model 1, and the
coefficient for classic rock/oldies slightly increases, though not substantially.

Conditioning on the GWESP parameter versus sociability or degree produces quite
different results. Whereas in Model 2 all of the homophily statistics increased in
magnitude, the coefficients in Model 3--again with the exception of that for classic
rock/oldies--went down, and the term for homophily with respect to dance music
actually became non-significant.

Model 4 conditions the homophily parameters on both sociability (degree) and
transitivity. While the coefficient for classic rock/oldies is no longer significant,

homophily still seems to be an important mechanism at work in the network, and

ignoring overall sociability (degree) falsely inflates observed homophily to a greater
extent than triadic closure. The remaining significant homophily statistics closely
resemble those in Model 2, yet they are consistently lower than Model 2 predicts, indicating that transitivity-based spurious homophily exists even though it has less influence than degree-based spurious homophily.

In this model, we also begin to see that types of activity-homophily dynamics emerge. First, for big band, country, and Latin music, the homophily statistics are significant and positive, while there is no activity-based effect. It is worth noting that homophily is only observed to have a positive effect on tie formation. Since there is no negative effect, it seems that consumers of various genres do not have fewer co-consumers than one would expect from chance. Second, sociability can influence network formation in isolation from homophily dynamics, but it has both positive and negative effects. People who listen to classical rock/oldies and heavy metal tend to have lower degrees than those who do not, and students who listen to hymns/gospel music and new age music have on average higher degree than those who do not. Finally, we also observe positive homophily statistics in conjunction with negative degree statistics on the dance genre. While I will leave theorizing for the discussion section, these findings, if corroborated with the final model, will necessitate an explanation that can account for (1) positive homophily dynamics, (2) a variable link between musical taste and sociality, and (3) an inverse relationship between homophily and sociability at least insofar as they pertain to cultural consumption.

Model 5 includes the sociodemographic control variables in addition to the contents of Model 4.10 The general results are consistent with those from Model 4—the magnitude of coefficient change ranges from .01 to .08—with the exception of three
parameters. First, the Latin music homophily statistic drops in magnitude a substantial .33 points, but this should not be alarming given that this is the only genre directly aligned with a racial/ethnic category. Second, whereas in Model 4 the homophily parameter for big band music was significant and the degree parameter was not, the opposite is true in Model 5.

Is Model 5 a good or even acceptable model? This question can be more difficult to answer for models within the exponential random graph framework that include dependencies between edges as opposed to traditional statistical models or even ERGMs that do not include dependencies. Models that do not include dependencies can use relatively simple goodness of fit measures because they do not violate the observational independence assumption (Harris 2014), but models including dependencies must rely on a more complicated, simulation-based approach (Goodreau et al. 2009; Harris 2014; Hunter, Goodreau, and Handcock 2008). The general idea behind utilizing simulation to assess goodness of fit is that of generative sufficiency. ERGMs attempt to model mechanisms that produce the observed social network's structure, and simulation-based goodness of fit assessments use the statistical model's specifications to create synthetic networks and then compare various structural properties between the simulated and observed networks. If the simulated networks do not generally significantly differ from the observed network on the specified global characteristics, then the model is said to be acceptable. This paper follows the recommendation of Goodreau, Kitts, and Morris (2009) and compares the simulated
and observed networks with regard to the distributions of degree, edge-wise shared partners, and minimum geodesic distance.

Figure 1 presents the graphical results for the simulated goodness of fit tests. In these graphs, the observed statistics are plotted in the bold line, and the 95% confidence intervals for the simulated network statistics are shown with light grey lines connecting the various box plots. The first graph compares the degree statistics, and the two types of networks do not significantly differ at any point. While there is a decent amount of variation in the degree values for the simulated networks, the middle 50% of the simulated networks' degree distributions generally closely align with the observed distribution. The next graph provides the same information with respect to edge-wise shared partners, and this distribution is particularly important because because edge-wise shared partners, the number of mutual contacts for adjacent nodes on a dyad, are influenced at least in part by homophily-based pairing. The figure does show a small number of significant deviations of the simulated network statistics from the observed network statistics. However, these differences are small, and there are few significant differences, as well. Finally, the third graph indicated that Model 5 generates networks with sufficiently similar geodesic distance distributions.
Figure 1: Goodness of Fit Simulation Results
CHAPTER 6:  
DISCUSSION AND CONCLUSION

To recapitulate the results, there are three main types of findings pertaining to the differential effects of homophily and sociability; transitivity was present in every model in which it was included; and there are homophily and degree effects from sociodemographic control variables that cloud the relationship between cultural tastes and homophily. First, for the country genre, only homophily produced an effect on network structure. Second, for another large set of genres (big band, classic rock/oldies, hymns/gospel, new age, and heavy metal), sociability influenced network structure without a homophily effect. Third, for dance and Latin music, both homophily and sociability played a role. Finally, the country, dance, and Latin homophily variables experience a change in magnitude with the inclusion of control variables, and the big band homophily variable even loses significance. The remainder of this paper will attempt to make sense of these findings and suggest directions for future research linking culture and dynamics of social network formation.

While this paper began as ostensibly about homophily vis-a-vis music tastes, the homophily findings are the least piquant, though they are still empirically interesting. Unlike previous studies (Lewis et al. 2012; Steglich et al. 2006), I did not find homophily among individuals who listen to classical, jazz, or classic rock music. Classical and jazz
parameters never achieved significance, which is partly why they are not presented in the five models, but the classic rock homophily parameter maintained significance through the first three models. It was only with the contemporaneous inclusion of terms for both sociability and transitivity that the classic rock homophily parameter became insignificant. On the other hand, this paper does find country, dance, and Latin homophily.

The series of models illustrate complicated relationships between cultural homophily, sociability, transitivity, and homophily on other dimensions. Country homophily is actually suppressed by variation for which Model 2 controls, yet it is upwardly biased without conditioning on transitivity and the control variables. Dance homophily is also downwardly biased without accounting for the confounding mechanisms. Homophily on the basis of Latin music is downwardly biased without conditioning on sociability, while controlling for transitivity without sociability downwardly biases Latin music homophily more than the naive estimator in Model 1. It is also necessary to control for race because not doing so falsely inflates the observed homophily.

There are more sociability effects than homophily effects, but, unlike homophily, sociability (degree) effects did not occur in a consistent direction. Degree parameters for big band, classic rock/oldies, dance, and heavy metal music were negative, though the big band effect only becomes significant in the final model. The effects for hymns/gospel and new age music were positive. These findings are interesting yet
uneasy to interpret because most literature on culture and social networks does not address sociability, instead focusing mostly on homophily.

Unfortunately, there is no clear trend between the percentage of students who listen to a genre and the direction of the degree statistic. Among the three genres for which there is a negative effect, 60% and 63% listen to classic rock/oldies and heavy metal respectively, but only 15% listen to big band music. The two genres which have a positive degree statistic, hymns/gospel and new age music, are consumed by a small minority of the respondents: 16% and 24% respectively. This might suggest a relationship between a positive degree statistic and listening to a minority genre, though there is hardly enough evidence to make this conclusion, but there are two problems with this line of thought. First, this reasoning does not hold for a negative degree effect. While that is not an a priori reason to dismiss a potential explanation, we have no theoretical reason to assume that dynamics for positive and negative sociability effect are distinct. Second, there might be a better explanation that can account for both positive and negative effects without reference to environmental composition of music consumption. I will discuss this shortly.

The final category of results consists of positive homophily and negative degree statistics. One initial attempt at interpretation might suggest that this combination could stem from being a minority, but, over half of the students listen to dance music. I have already shown that there is no consistent relationship between the degree statistic's direction and the percent of the sample who listen to a genre, but it might be that there is at least a compositional effect for the positive effect of homophily.
However, we can see this is also not the case if we look at the other genres for which there is a positive homophily effect, as country listeners constitute 46% of students and Latin listeners comprise only 17% of the students. Thus, the compositional argument does not work for homophily-sociability interaction, as well.

I suggest that a more plausible explanation should be rooted in students' general dispositions toward cultural consumption. This explanation starts with the premise that there exists within a population a social logic that affects the way people interact with each other and their environment (Tarde 1963). This logic is at least partially attributable to the fact that various cultural traits tend to co-occur and that this co-occurrence has obvious roots in the configuration of interests and information that characterize particular niches in the social structure (Converse 1964:211). Further, the content of culture is an important cognitive determinant of social actions (Sperber 1996; Whitehouse 2004), providing a feedback loop that can reinforce extant social boundaries.

There is a continual interaction between cultural and social boundaries. The culture we consume and embody constructs symbolic boundaries that both stem from and reinforce social divisions, and this further creates uneven distributions of cultural resources (Bryson 1996, 1997; Emirbayer and Goodwin 1994; Erickson 1996; López-Sintas and Katz-Gerro 2005; Ollivier, Gauthier, and Truong 2009; Ollivier 2008; Peterson and Kern 1996; Peterson and Simkus 1992; Peterson 1992; Tampubolon 2008). Given the uneven distribution of cultural elements created by social boundaries, social logics governing the associations of cultural traits and individuals tend to emerge. Myriad
pockets in the social structure induce the conjunction of certain people and ideas, and
the habitual experience and enactment of *de facto* cultural configurations can lead to
the perception of obdurate relations between constituent elements of a culture
(Converse 1964; Martin 2011; Smith and Decoster 2008). While it is important to
remember that relationships between symbolic elements are not always logically
necessary or present (Converse 1964; Martin 2000), an associative web of relations does
exist through the individuals embodying and utilizing different patterns of culture.

From childhood, individuals are socialized within certain socio-cultural settings,
and, particularly early in life, individuals absorb “practical schemes of perception,
appreciation, and action” as well as “conscious rules or . . . explicit discourses” (Lizardo
and Skiles 2012:266). Permeating both implicit and explicit modes of reasoning and
memory (see Bargh 2006; Sloman 1996; Smith and Decoster 2008), psychologically
embedded culture facilitates inter alia action by unconsciously emphasizing and de-
emphasizing aspects of the environment, and it can help an individual make sense of
previous action(s) through consciously available meanings, as well (Emirbayer and
Goodwin 1994; Lizardo and Skiles 2012; Vaisey 2008, 2009). These socially grounded
symbolic boundaries thus pervade our conscious and unconscious. An example of this
can be seen in Americans' coupling of moral absolutism and relativism with “traditional
values” and “secular-rational values,” respectively (Baker 2005:88).

These dispositions resulting from one's position in the sociocultural cartography
manifest themselves in terms of affordances (Martin 2011). These affordances provide
implicit evaluations and impulsions to act toward or against certain phenomena. Thus,
we would expect homophily in terms of a music genre to come about because students who listen to the genre acquired a disposition (before matriculating to the university) that provides positive feedback for befriending similar consumers, and we likewise would think that higher sociability levels for those who listen to certain genres stem from a disposition that endows a positive valence on both those genres and having more contacts. A linkage between positive homophily and negative degree ericts vis-a-vis music genres would simply be related to another type of disposition.

This explanation is far from complete, nor is it a conclusion forced wholly from the results. However, it does have more simplicity and cognitive validity than a mélange of explanatory bits that only occasionally points to compositional and relational aspects of the sociocultural environment. Further research should do several things to evaluate the logical validity and soundness of this explanation. One research direction should examine the acquisition of these habitual dispositions earlier in the students' lives and see if there are systematic difference between their childhood and adolescent locations in a variety of fields and their homophily and sociability patterns as young adults. Another task will be to investigate how these levels of homophily and sociability with regard to music consumption change longitudinally. Perhaps, such information will help arbitrate between competing explanations or provide data that suggests an altogether new explanatory scheme. Finally, a more complete analysis would also include parameters in analytical models to condition on social influence vis-a-vis music genres.
REFERENCES


Bargh, John A. 2006. “What have we been priming all these years? on the development, mechanisms, and ecology of nonconscious social behavior.” *European journal of social psychology* 36(2):147–68.


