Is Having an Educationally Diverse Social Network Good for Health?

Mark C. Pachucki
University of Massachusetts-Amherst
Department of Sociology
Computational Social Science Institute
Email: mpachucki@umass.edu

Diego F. Leal
University of South Carolina
Department of Sociology
Email: leald@mailbox.sc.edu

For correspondence, please send email to the corresponding author, Mark C. Pachucki (mpachucki@umass.edu) or letters to the Department of Sociology, 200 Hicks Way, Thompson Hall 934, University of Massachusetts-Amherst, Amherst, MA 01003

Acknowledgements: We thank Nicholas A. Christakis, Thomas Keegan, and the Human Nature Lab at Yale University for providing access to the data used in this research, and Liza Nicoll in particular for help facilitating answers to questions regarding the data collection process and server administration. This research was funded by a pilot project grant from the The Roybal Center for the Study of Networks and Well Being at Yale University (funded by National Institute on Aging grant # 2P30AG034420-06).

Is Having an Educationally Diverse Social Network Good for Health?

Abstract: While network research often focuses on social integration as a predictor of health, a less-explored idea is that connections to dissimilar others may benefit well-being. As such, this study investigates whether network diversity is associated with changes in four health outcomes over a 3-year period of time in the United States. Specifically, we focus on how an underexplored measure of network diversity – educational attainment assortativity – is associated with common self-reported outcomes: propensity to exercise, body-mass index, mental health, and physical health. We extend prior research by conducting multilevel analyses using this measure of diversity while adjusting for a range of socio-demographic and network confounders. Data are drawn from a longitudinal probability sample of US adults (n=10,679) in which respondents reported information about themselves and eight possible alters during three yearly surveys (2013-2015). We find, first, that higher educational attainment is associated with more educationally insular networks, while less-educated adults have more educationally diverse networks. Results further suggest that having educationally similar networks is associated with higher BMI among the less-educated. Further exploration of the relationship between ego network diversity, tie strength, and health is warranted.

Keywords: egocentric networks, health, network diversity, assortativity

1. Introduction

Despite a strong tendency for people to affiliate with people who are similar to themselves (Lazarsfeld & Merton, 1954; Marsden, 1988; McPherson, Smith-Lovin, & Cook, 2001; J. A. Smith, McPherson, & Smith-Lovin, 2014), being socially connected to different types of people appears to be an important factor in a range of domains, including community-level economic development (Eagle, Macy, & Claxton, 2010), higher-order cognitive processing (Molesworth, Sheu, Cohen, Gianaros, & Verstynen, 2015), and pro-social communication with others (Alshamsi, Pianesi, Lepri, Pentland, & Rahwan, 2016).

Social relationships – often theorized as a form of social capital (Bourdieu, [1986] 2018; Coleman, 1990; Lin, 2001; Putnam, 2001) – have been shown to be strongly implicated in health behaviors and health status through a range of pathways (Berkman & Krishna, 2014), and much research on social capital and health has been conducted in an egocentric research framework (Perry, Pescosolido, & Borgatti, 2018). An egocentric perspective on one's relationships provides a glimpse of the social contours of the range of social confidants who may provide social support, access to resources, opportunities for social influence, channels for disease spread, and who may help individuals to buffer against stress.

Although a great deal of research has focused on the benefits of social relationships – and specifically, integration and network size – as predictors of lower rates of morbidity and mortality (Holt-Lunstad, Smith, & Layton, 2010), a less-frequently explored idea is that network composition, and more specifically, network diversity – a form of social capital that serves as a relational asset (Lin 1999) – may possibly serve as a social determinant of health. While in many realms of network life, birds of a feather do flock together, and connections to those who are *similar* on some dimensions may provide for increased social support and buffer one from ill health. Yet having connections to people who are *dissimilar* in some way may facilitate access to new information about appropriate health behaviors, provide a range of models for appropriate behavior, and serve as channels for social influence.

In this study, we review prior research on network diversity and health and find that, on balance, network diversity appears to benefit modifiable health behaviors, mental health, physical/cognitive function, and overall mortality. However, several conspicuous gaps merit further investigation. To a great extent, much of this foundational work has evaluated hypotheses related to heterogeneity of network roles (based upon measures of relationship type – or role – diversity). While role diversity has been established as a reliable measure and it has important properties, we argue that encountering other forms of attribute-based diversity in one's everyday experience – for instance, diversity in socioeconomic status – may also be important to health. Second, much of the existing research on network diversity relies upon individually-based measures of diversity that do not explicitly measure ties *between* socially-connected alters. Third, the majority of research in this area tends to be based on cross-sectional designs, and so movement towards understanding mechanisms through which network diversity may be shaping health has been understandably restricted.

In this paper, we aim to expand existing efforts in these areas by turning towards a unique longitudinal and nationally-representative dataset of more than 10,000 Americans surveyed about their personal networks and health between 2013 and 2015. We pose the following questions: (1) *To what extent is the educational diversity in one's personal network associated with having better or worse health?;* and (2) *Whose health benefits the most from having social ties to those with diverse educational attainment?*

2. Background

2.1. Why Should Network Diversity be Good or Bad for One's Health?

It is not a foregone conclusion that any given structural dimension of one's network – whether density, cohesion, diversity, or number of social ties – should necessarily be a benefit to one's health. As Lin (1999) explained in elaborating network dimensions of social capital, there can be a relational advantage in having cohesive networks for maintaining one's resources, while having structurally different alters that bridge across locations in the network may be more important for obtaining new resources. From one perspective, having large and diverse networks increases the number and range of types of individuals that an individual must maintain contact with, which may come at the expense of cognitive burdens, role conflict, and stress (Burt, 2004; Cornwell, 2009; Dunbar, 2018; Goldman & Cornwell, 2015). Indeed, having more ties can be stressful and be associated with poor mental health, especially for women (Kawachi & Berkman, 2001). Yet in terms of health, it is also reasonable to think that being connected with different kinds of people may confer health benefits – as a form of network health externality that emerges above and beyond one's own resources (K. P. Smith & Christakis, 2008; VanderWeele & Christakis, 2019).

Overall, research in this area suggests that network diversity – generally defined as being socially connected with people of different backgrounds – is indeed associated with better health, including better overall self-rated health (Cattell, 2001), less susceptibility to upper respiratory infections (Cohen, Doyle, Skoner, Rabin, & Gwaltney, 1997), lower risk of heart disease (Barefoot, Gronbaek, Jensen, Schnohr, & Prescott, 2005), and indications of improved mental health indicators such as lower depression levels (Erickson, 2003). Although during the past two decades there has been sporadic attention to the topic of network diversity in health in an egocentric framework, this field of inquiry has rapidly expanded since Moore and colleagues (2009) showed in a cross-sectional study that greater network diversity (as part of a greater multi-measure construct of network social capital) was associated with lower risk of excess adiposity. More recently, findings from a longitudinal investigation in the same cohort largely comported with the earlier cross-sectional findings (Wu, Moore, & Dube, 2018).

Cross-sectionally, greater network diversity has been associated with better mental health in the case of homeless California adults and depression (Rice, Kurzban, & Ray, 2012), in lower incidence of post-traumatic stress disorder among US adults (Platt, Keyes, & Koenen, 2014), and greater dispositional optimism in US adults (Andersson, 2012). Yet a study of Canadian adults and depression found that greater geographic diversity of alters was associated with more depressive symptoms (Bassett & Moore, 2013).

Egocentric studies of modifiable health behaviors have shown that greater network diversity is associated with more salubrious levels of physical activity among Canadian adults (Legh-Jones & Moore, 2012) and US-based older adults (Shiovitz-Ezra & Litwin, 2012); and with lower odds of smoking in Canadian adults (Moore, Teixeira, & Stewart, 2014) and among adolescents in multiple countries (Choi & Smith, 2013). Yet interestingly, in a longitudinal study, Child, Stewart, and Moore (2017) found that having a greater range of occupations in one's personal network (greater network extensity) was associated with higher incidence of binge drinking (a poor health behavior).

Last, research among older adults has shown that greater network diversity is associated with an absence of disability (Escobar-Bravo, Puga-Gonzalez, & Martin-Baranera, 2012), and that higher proportions of family-based ties in respondents' networks (i.e., less diversity) is associated with higher levels of disability (Cornwell & Laumann, 2015). Greater network diversity was associated with higher white matter integrity in the brain and neuronal myelination processes among middle-age US adults (Molesworth et al., 2015). Among older Dutch adults, having more diverse networks was associated with less cognitive decline concurrently, and over time (Ellwardt, Van Tilburg, & Aartsen, 2015). Finally, across a multi-national sample of approximately 14,000 older adults in several developing countries, having less diverse and integrated networks (i.e., networks with few friends or community contacts and restricted to family) was associated with earlier mortality (Santini et al., 2015).

2.2. How Network Diversity is Conceptualized and Measured in Studies of Health

A substantial majority of studies of network diversity and health status/behavior have operationalized the concept of *role diversity* (Barefoot et al., 2005; Cornwell & Laumann, 2015; Ellwardt et al., 2015; Escobar-Bravo et al., 2012; Kelly, Patel, Narayan, Prabhakaran, & Cunningham, 2014; Legh-Jones & Moore, 2012; Molesworth et al., 2015; Moore et al., 2014; Rice et al., 2012; Song, Pettis, & Piya, 2017; Viruell-Fuentes, Morenoff, Williams, & House, 2013; Zhang et al., 2012). This concept is often measured using a form of a network position generator (Lin & Dumin, 1986) that seeks to enumerate characteristics of an individual's ties to a set of alters with different social roles (e.g., as family members, church members, friends, neighbors, and so forth). The most common instrument used in this context has been Cohen's Social Network Index (SNI), which evaluates ego's access to 12 different roles; a recent example is Mowbray, Quinn, and Cranford (2014). Usually, network size (i.e., number of alters within each role) and range (difference between the highest- and lowest-status alters in terms of their occupational prestige) are also measured (Molesworth et al., 2015; Moore et al., 2014).

Of course, the high prevalence of Cohen's SNI is a reflection of an individual-level/egocentric analytic bias of this literature, which reflects data-collection norms in the public health domain where ties between socially-connected alters were not explicitly taken into account during the period of the scale development. There are exceptions, however. For example, Choi and Smith (2013) depart from this egocentric tendency by doing a meta-analysis of the role of nodes' network position (as isolates, members, or liaisons) and their association with smoking behaviors in the context of 8 different adolescent friendship networks. It would appear, then, to be a worthwhile endeavor to build upon these efforts by more carefully accounting for the structure of respondents' personal networks and incorporating information on the social connections between ego's alters.

2.3. Mechanisms Linking Educational Diversity to Health

What mechanisms might explain the observed relationships between health and network diversity? It is quite likely that across different health measures, different mechanisms link social ties with health (Thoits, 2011).

Similarity. Given that individuals with similar attributes tend to form ties with one another (McPherson et al., 2001), and also that network diversity has been associated with largely positive health outcomes, for individuals without access to health care, information, or other resources, *similarity by socioeconomic status to trusted confidants, as well as their ties to one another* may contribute to the reproduction of health inequalities via insulating them from health-based opportunities, and by inhibiting forms of social support. In addition, how those trusted others are connected to one another is likely to matter as well. In a word, similarity along key attributes can significantly impede the flow of new resources and information. For instance, (Schaefer, Kornienko, & Fox, 2011) report that marginalized individuals such as depressed teenagers tend to befriend other marginalized and depressed individuals. In that context, mentally unhealthy (healthy) individuals, and their friends, are unlikely (likely) to have the resources to assist their likewise mentally unhealthy (healthy) friends in times of need. Similarity thus acts as a possible mechanism for the reproduction of health disparities between individuals.

Tie strength. To obtain resource benefits from social contacts, one must not just have a relationship with others, but also be able to mobilize those resources – which implies that *tie strength* is also consequential (Granovetter, 1973). Greater closeness between two people is associated with a greater likelihood of similarity between them (Cornwell, 2009), and this may be associated with a less diverse network. There is evidence suggesting that individuals in poor health tend to have weaker friendships (Haas, Schaefer, & Kornienko, 2010), which will further amplify the effects of similarity on the reproduction of health inequalities mentioned above. In short, if healthier individuals are indeed more likely to have both access to resources through their ties to similarly healthier others *and* to be able to actively mobilize such resources through their relatively strong connections, then it is easy to see how tie strength could be considered an important mechanism connecting network diversity and health.

Cohesion. Having a cohesive network may allow greater access to health-related social support if one is trying to maintain one's health (Lin, 1999). There could be a 'differential access' mechanism wherein individuals having different levels of education, access social capital through different network pathways (Moore, Daniel, Gauvin, & Dubé, 2009). For instance, in research on the psychological well-being of Canadian adults, Moore and colleagues found that individuals with less education tend to rely upon friends and family (i.e., strong ties) – who tend to have similar socioeconomic status – for resources, whereas individuals with more education rely on acquaintances (i.e., weak ties). The authors suggested that those who are less educated may have less educationally diverse networks. And indeed, in comparing the 1985 and 2004 General Social Survey samples, J. A. Smith et al. (2014) found higher levels of education homophily between egos and their close confidants at the lower end of the education distribution.

Subjective social status. Last, the *perception of one's position in society* may be a relevant mechanism linking social capital and health outcomes. In examining the relationship between a social capital scale and psychological distress, Song (2011) finds a significant mediating role for subjective social status – i.e., one's perception of status relative to others. This is important because there is evidence suggesting that ego's subjective social status (e.g., perceptions of relative class identification) is in itself dependent on the socioeconomic status of their social contacts, above and beyond ego's socioeconomic status (Hodge & Treiman, 1968). In that context, assortative ego networks with relatively high (low) levels of education might increase (reduce) ego's subjective social status, which in turn might positively (negatively) affect their health.

Given the state of research on network diversity and homophily, we hypothesize that social networks of less-educated individuals are likely to have less educationally diverse (more assortative) personal networks than those of more educated individuals (H1). Testing this hypothesis is both a necessary first step to then evaluating the relationship between network diversity and health, and also novel in that prior studies have focused on education homo-/heterogeneity at the individual/egocentric level (e.g., ego's level of education), rather than network measures of assortativity. Based on prior research grounded in studies of role diversity, we hypothesize that in general, less education assortativity (i.e., having social contacts with a more diverse range of educational attainment) will be associated with better general physical and mental health, more physical activity, and lower BMI (Hypothesis 2).

Last, we also predict (Hypothesis 3a) that it is more likely that the health of individuals in a less-educated tier (who are more likely to have a health resource deficit due to their low-SES status) may benefit from access to more-educated others (greater network diversity). Yet maintaining educational diversity within one's network may also be more burdensome, particularly for low status egos (for whom educational diversity necessarily means keeping ties with higher status alters). Thus, a competing hypothesis (Hypothesis 3b) could also be that network educational diversity is more likely to lead to poorer health and/or less health-promoting behaviors among low-SES individuals.

3. Data & Methods

To test these propositions, we utilize three yearly waves of longitudinal egocentric survey data obtained through an online survey administered by the Gallup organization between 2013 and 2015 as part of its ongoing, longitudinal, probability-based panel of American households. The Gallup Panel (Gallup, 2014) contacts U.S. households at random via random-digit-dialing (RDD) of landline telephones and cellphones or address-based sampling (ABS). This is an online panel, which the polling firm acknowledges only includes individuals with internet access (80% of the US population). The social network instruments used here queried who respondents spent free time with and discussed important matters with, and we describe further details about the enumeration process below. These questions were adapted from the GSS and National Social Life and Health Survey and piloted in a smaller sample (O'Malley, Arbesman, Steiger, Fowler, & Christakis, 2012). The present sample was drawn from an enumeration of 20,373 respondents

(Year 1), 27,829 (Year 2), and 24,087 (Year 3). Nearly half of this sample (n=10,679) provided a response at all three waves, allowing for models to adjust for changes in network composition. There were no covariate-based exclusion criteria for this study.

We chose health outcomes that have been commonly examined in prior studies of network diversity – these include generalized measures of physical and mental health, as well as a common health behavior (exercise/physical activity) and body-mass-index, which is an indicator of cardiometabolic risk. Information on health outcomes include self-rated physical health *("How would you describe your own physical health at this time?")* with ordinal responses being Poor, Fair, Good, and Excellent. This measure was then dichotomized into "Excellent vs. Other" to examine the contrast between the best health possible and other categories. A similar question was asked of self-rated mental health (*"How would you describe your own mental health or emotional wellbeing at this time?"),* with the same response categories and dichotomization. Self-reported height and weight were used to generate a continuous BMI score (height/weight^2), and outliers below 15 and over 60 were coded as missing. A weight-related behavior question was asked of the form *"Please indicate whether or not you have done any of the following to try and improve your health in the past three months – exercised regularly (at least 3 times per week).*"

Socio-demographic covariates include individual characteristics such as continuous age, sex (male/female), race (White/Black/Asian/Native Hawaiian and American Indian/Other, which was re-coded into White/Black/Asian/Other/Multi-racial), ethnicity (Hispanic/non-Hispanic), education (<HS/HS/Some College/College/Post-graduate, coded 1-5), income (ordinal tiers, coded as categorical 0-8), and household assets (ordinal tiers, coded as categorical 0-7). An adapted form of the original MacArthur network subjective social status ladder (Adler, Epel,

Castellazzo, & Ickovics, 2000) asked respondents *"Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you place yourself?"* (coded as continuous 1-10), employment status (employed full-time/part-time but not a full-time student/full-time student/retired/homemaker/not employed, which was re-coded into employed full-time/part-time/other), region (Northeast/Midwest/South/West), and marital status (single/married/separated/divorced/widowed /never married/living with a partner, which was then re-coded into "married or living with a partner" vs. "other").

Egocentric network measures

For the key independent variable, education assortativity, we relied on (a) ego's nomination of up to eight alters, (b) information about whether the alters were connected to one another *("Please select the option that best describes the current connection between [alter x's name prompted by what the respondent wrote in on (a)] and [alter y's name, prompted by what the respondent wrote in on (a)]."*), and (c) ego-reported educational attainment of alters *("As far as you know, what is the highest level of education ([alter's name]) has completed?"*).[1]

[1] To elicit alter names, a first name generator question asked, "*Looking back over the past 12 months, think of up to four adults (ages 16 and over) with whom you spend the most free time. By free time, we mean time spent for your enjoyment after work or school or on the weekend. These adults could be members of your household, friends from work or school or elsewhere, family members or relatives, or others. Please enter the first names (or initials, nicknames) of these adults.*" A next name generator asked, "*From time to time, most people discuss important matters with others. Looking back over the past 12 months, think of up to four adults (ages 16 and over) with whom you most often discussed important matters. These adults could be members of your household, friends from work or school or elsewhere, family members or relatives, or others. Please enter the first names (or initials, nicknames) of these adults.*" Then respondents were instructed, "*Please review the full list of names you just provided. If any people appear twice on the list, click the box next to the name to REMOVE duplicate names so that each person will only appear once on the list below.*" A series of follow-up questions were then asked about characteristics of each of the confirmed alters.

Assortativity Coefficient. Much prior research relies on adapting a position-generator method to measure network diversity, and such research is able to reach outside the stronger ties that name generators typically enumerate. Yet given a wealth of research that family and friends (typically strong ties) are consequential to health (Christakis & Fowler, 2007; Yang et al., 2016), the richness of information on between-alter ties available to us from name generators, we instead opt to use the assortativity coefficient (Newman, 2003) to indicate the extent to which a given ego network is segregated along the exogenous attribute of education as a key predictor. Assortative (disassortative) mixing is a dyadic process where ties tend to emerge between nodes within (outside) the same categorical attribute (Goodreau, Kitts, & Morris, 2009). A low assortativity coefficient is indicative of a diverse/desegregated network, while a high assortativity coefficient indicates the presence of a non-diverse/segregated network (Bojanowski & Corten, 2014).

We measure assortative mixing along categorical node-level attributes of alters. Since we are modeling self-reported relationships between a given ego and their alters, all ties are considered to be symmetric. This means that the personal networks here are always undirected. Following the notation of Newman (2003), and using a binary education attribute (i.e., high school graduate vs. college graduate) as an example, we define the assortative mixing coefficient as:

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i} \quad \text{Eq. 1}$$

Where $\sum_i e_{ii}$ is the sum of the fraction of within-category ties among all pairs (e.g., sum of the fraction of college-to-college ties and of the fraction of high school (HS)-to-HS ties), a_i is the sum of the fraction of within-category ties of nodes type *j* (e.g., HS-to-HS ties) plus the fraction of outside-category ties of nodes type *j* to nodes type *k* (e.g., HS-to-college ties) and b_i is the sum of the fraction of within-category ties of nodes type *j* (e.g., HS-to-HS ties) times the fraction of outside-category ties of nodes type *k* to nodes type *j* (e.g., college-to-HS ties). Since all ties under analysis are undirected, the fraction of outside-category ties is, by definition, the same irrespective of the node type (e.g., fraction of college-to-HS-ties = fraction of HS-to-college ties).

Assortative mixing is often interpreted by some network theorists as a "preference" (Newman, 2003: 208701-1) of nodes to attach to other within-category (e.g., same-race or same-education level) nodes, as well as exposure to opportunities to interact or shared foci of activity (Feld, 1981, 1982). Decades of research suggest that, at the individual level, and assuming permeability of institutional barriers (McPherson and Smith-Lovin 1987), nodes' *preferences* to form connections to similar others – that is, *choice* homophily – is a ubiquitous force behind the formation of ties (Blau and Schwartz, 1984; McPherson, Smith-Lovin, and Cook, 2001; Centola 2015). In this context, the assortativity coefficient is a summary measure designed to compare the pattern of within- versus outside-category ties in a given (ego) network. Thus, it is an important tool to shed light on the prevalence of social processes like choice homophily by measuring assortative mixing.[2]

[2] As suggested by Bojanowski and Corten (2014), the assortativity coefficient can be interpreted as an index that represents the (proportional) weight of the main diagonal of the cross-tabulation of within- and outside-category ties of a given (ego) network. If all ties happen to be within-group ties (i.e., if all ties are located in the main diagonal) that indicates perfect assortative mixing in the ego's personal network. The assortativity coefficient ranges from complete disassortativity (indicating greater diversity, or heterogeneity) to complete assortativity (indicating a lack of diversity, or homogeneity) in the mixing pattern of the observed ties within and across a given exogenous category (e.g., education level) (see technical details in Newman (2003)]). Because the algorithm calculates an infinite value in the case of perfect assortativity (see details below), we assigned a value of 0.5 to "fully assortative" ego networks as just outside the maximum observed value (0.49). Alternative analyses not reported here were conducted that modified this by changing the upper bound to 1.0. Findings were robust to this measurement change. Following equation 1, a_i is the sum of the fraction of *existing* within-category ties of nodes type *j* plus the fraction of *existing* outside-category ties of nodes type *j* to nodes type *k*, and b_i is the sum of the fraction of *existing* within-

Network controls. Covariates were also included to adjust for the size of an ego's personal network, as well as the density of one's personal network given as the proportion of existing ties out of all possible ties of that network size. The correlation between these two covariates was low and did not pose a threat to model estimation. Due to the pervasiveness of (homophily-based) selection and social influence in social networks, some factors may systematically reduce the probability of observing a diverse network. More precisely, high frequency of contact or closeness with contact members will be important in this regard since these two factors are known to increase the likelihood that alters could be similar to each other (Cornwell, 2009). Thus, adjusting for average strength of "close" and "liking" ties among alters (two different indicators of strong ties here) serve as important controls. Closeness was measured using "*How close do you feel to (display alter's name, as appropriate)?*" (1 = "Not close at all", and 10 = "Extremely close/closer than any other person I know"). Liking was measured with, "*How much do you like (display alter's name, as appropriate)?*" (1 = "Do not like at all", and 10 = "Like a lot/Like more than any other person I know.")

3.1. Analytic approach

We employ a multilevel modeling strategy (2-level panel data), wherein egos (level 2) are nested in time (level 1) (Perry et al., 2018). This framework has the advantage of accounting for an ego's dependence with its prior observations. A random coefficient for education

---

category ties of nodes type *j*. In that context, the denominator in equation 1 will necessarily reduce to 0 when there is only one alter or when all ties are within-category ties since in those scenarios the available alter(s) will represent the entirety (i.e., a proportion of 1) of the *existing* (within- and outside-category) ties in the ego network, thus reducing the denominator in eq. 1 to 0 (1 – 1 = 0). Once the denominator of equation 1 becomes 0, the assortativity coefficient becomes undefined. Relatedly, since a_i, b_i, and e_{ii} are all a function of the pattern of *existing* ties, adding isolates (i.e., non-contacts) does not affect the pattern of existing ties at all and, for that matter, the assortativity coefficient. For a technical explanation of this property, called intransitivity to adding isolates, see Bojanowski & Corten (2014).

assortativity, included after testing a null random intercept model, allows change in assortativity to vary across egos. Conceptually, across all model frameworks, we theorize network diversity as part of the experiential social context – and thus a characteristic – of the ego. The appropriate form of each model was determined by the outcome variable specification, namely, multi-level logistic regression for dichotomized exercise, self-rated physical health (hereafter, SRPH), and self-rated mental health (SRMH), and multi-level linear OLS for continuous BMI.

Covariate missingness. At baseline, 10,679 individuals were present at all three panels. All four dependent variables were missing at a low level, approximately <5.0% of respondents. Education assortativity could only be calculated for 85% of egos (n=9,108) for two main reasons. First, to calculate assortativity requires at least 2 alters (see footnote 2); those with less than two alters for whom an assortativity value was not calculable for this reason comprised 14.7% (n=1507) of the baseline sample. Second, if an alter was nominated by ego but was missing an education value, we dropped that observation (<1%).

Analyses relied on complete-case analysis rather than partial imputation of individual-level covariate values because of our skepticism at imputing information on covariates without also imputing network ties as well. Although some recent work has offered promising steps in imputing edge information in sociocentric datasets (Huisman, 2014; J. A. Smith, Moody, & Morgan, 2017; Wang, Butts, Hipp, Jose, & Lakon, 2016), this branch of the field of network science, and especially in egocentric settings, is at present relatively underdeveloped. At baseline, other covariate information was also missing on dependent variables (exercise regularly, 4.7%; BMI, 2.4%; self-rated physical health, 0.8%, self-rated mental health, 1.4%), and socio-demographic covariates (household assets, 13.5%; income, 8.1%; marital status, 6.4%; employment status, 5.4%; subjective social status, 1.8%; ethnicity, 1.7%; race, 0.9%; region of country, 0.2%. Of special note is the unusual completeness on respondent education; only 2 of the 10,679 respondents were missing this information.

Observations without complete covariate information across any two waves were dropped from multilevel models (analyses were missing ~15% of participants from the full sample). Those retained in the multilevel model sample have higher average income, household assets, subjective social status, and are slightly more educated. Because of known difficulties in using population weights in multilevel settings, we do not use these weights, and spend additional time in the discussion section speculating on how the patterns of missing covariate data may bias the observed results. Because these data were deidentified to investigators, a human subjects approval waiver was granted by University of Massachusetts, Amherst. Data management, cleaning, and analyses were conducted using Stata 13 (StataCorp, 2013) and the R programming language (R Core Team, 2018).

4. Results

Figure 1 illustrates that education assortativity is approximately normally distributed (μ=-0.21, SD=0.23) though with a higher frequency of participants at the fully assortative end of the scale. Still, in general, there are very few individuals whose educational attainment networks are fully disassortative (more diverse) or fully assortative (more homogeneous), though there is overall a slight tendency towards assortativity. Bivariate associations describe a largely linear relationship between education assortativity and network size (where having a larger network is associated with greater alter educational homogeneity), and no relationship between assortativity and average alter closeness, average alter liking, or graph density.

[Insert Figure 1 here]

Table 1 further describes the baseline analysis sample, and that on average, respondents report 4 alters (of a possible 8), that networks are relatively dense (μ=0.87), and that they like their nominated alters (μ= 8.65) slightly more than they feel close to them (μ=8.24), though it bears keeping in mind these are likely to be largely strong ties. Roughly twice as many individuals report regular exercise as not, and although only 21% report being in excellent self-rated physical health, twice that amount report excellent self-rated mental health. The sample skews slightly male, with a strong majority who are non-Hispanic white, and the mean age of respondents is in their late 50s. Socioeconomically, the sample skews towards middle-to-upper class. Roughly half are employed full-time, average income is 5.8 (where 5=$75-<100K/year, and 6=$100-<150K), and the modal category of educational attainment is postgraduate (39.8%), household assets average 3.9 (where 3=$100-<$250K, and 4=$250-<500K), and self-perception of subjective social status is 7.3 of a possible 10.

[Insert Table 1 here]

Figure 2 illustrates the variation in education assortativity by dependent variables (SRPH, SRMH, Exercise, BMI), as well as by respondents' educational attainment, household income, wealth, and subjective social status. P-values for significance are reported across levels for a given covariate. Those respondents who regularly exercise, have lower BMI, and report excellent SRPH, and SRMH tend towards more educationally assortative (more educationally homogeneous) networks. Figure 3 illustrates variation in education assortativity by different measures of socioeconomic status. At the extremes of each measure, there is more assortativity at higher SES tiers, and lower assortativity at lower SES tiers, with a somewhat monotonic trend in the middle categories.

[Insert Figure 2 and 3 about here]

Association of network and socio-demographic covariates with education assortativity

Table 2 reports on network and socio-demographic characteristics associated with education assortativity. Models include an OLS regression (Year 1 only), and a multilevel regression (Years 1-3). The question asked here is: to what extent does one's educational attainment predict one's network assortativity on education, net of socioeconomic and structural network characteristics? Estimates in both model specifications reveal that relative to college-aged respondents, having a higher level of education is associated with greater assortativity (less diversity in one's network), while having less education is associated with more educational diversity in one's personal network. We also observe that having more alters and a denser network is associated with greater assortativity (more homogeneity), but there is no association between the two measures of tie strength (alter closeness and alter liking) and assortativity.

Multilevel regression estimates of education assortativity and health

Having documented evidence of a relationship between educational attainment and education assortativity, we next turn to evaluating relationships between educational assortativity and our suite of health indicators (Table 3). Three stepwise models are reported. The best-fitting model series (according to BIC minima) includes main effects of education assortativity and education attainment, and subjective social status, though a more expansive series can be found in the Supplementary Information Appendix. This model also includes four network attributes (number of alters, personal network density, average alter closeness, average alter liking), and a vector of confounders (gender, age, race, ethnicity, region, marital status), and a continuous panel measure.

Across all health outcomes, and contrary to our expectations, education assortativity (more homogeneity) is positively and significantly associated with propensity to regularly exercise (OR=1.28, CI=1.02-1.61, p=0.03). Network characteristics vary in their associations across outcomes; network density is negatively associated with propensity to exercise regularly, for instance (OR=0.55, CI=0.38-0.81, p=0.002). Greater average liking of alters is associated with being in excellent SRPH (OR=1.26, CI=1.08-1.46, p=0.003) and SRMH (OR=1.25, CI=1.11-1.39, p<0.001), while greater average alter closeness is negatively associated with excellent SRPH (OR=0.87, CI=0.76-1.00, p=0.046). Educational attainment is as expected – positively associated with propensity to exercise, negatively associated with BMI, and positively associated with SRPH. Subjective social status is consistently associated in expected directions with all four outcomes.

The second series adds additional socioeconomic status variables to models, including household asset tiers (categorical), income tiers (categorical), employment status (categorical). Across all models, higher levels of education assortativity (less diversity) remains significantly associated with greater propensity to regularly exercise. The associations between network characteristics and the outcome variables reported above are consistent between the first and the second series of models. In other words, the association between average liking and being in excellent SRPH and SRMH, as well as the association between average alter closeness and excellent SRPH are all statistically significant and remain the same direction.

Turning towards the fully-specified third model series which include an interaction between education assortativity and education tiers, we observe little evidence of moderation for exercise frequency, SRPH, or SRMH. However, there is evidence that the relationship between assortativity with BMI is modified somewhat by ego's education level as predicted by H3a. Relative to college-educated adults, education assortativity among HS-educated adults ($b = 0.69$, $p<0.008$) appears associated with higher BMI (marginal effect is 0.62 kg/h^2 greater BMI). Thus, greater assortativity (less diversity) is associated with higher BMI among low-education individuals. This relationship is visualized in Figure 4 below.[3]

[Insert Figure 4, BMI predictive margins plot, about here]

In addition to moderation of assortativity and health by ego education, we also tested for moderation of this relationship by tie strength.[4] Table 4 reports that although tie strength does not moderate the relationships between education assortativity and exercise frequency, BMI, or mental health, it does appear to moderate the relationship with physical health, and rather strongly. The most straightforward interpretation is that among people with very assortative (educationally similar) networks, having a higher average strength of ties (as measured by ego's liking of alters) significantly increase ego's odds of being in excellent health by a large order of

[3] In the model predicting self-reported physical health, we do find some crossover effects of education assortativity by educational attainment on SRPH wherein among college-educated adults, being in a highly educationally assortative (similar) personal network is associated with marginally better self-reported health, but the mean difference (0.04) on a scale of 0.0-1.0 suggests statistical significance, but little practical significance.

[4] Additionally, we tested for moderation by subjective social status, and observed no association (not presented here, though available from authors).

magnitude. This comports with the position that being in a homophilous network may be linked with better health due to having less burden to maintain relationships.

5. Discussion & Conclusions

Although the field of egocentric network analysis has enjoyed several decades of development from its early roots (Fischer, 1977; Laumann & Pappi, 1973; Lin, 2001; Wellman, 1979), careful attention to measuring network diversity as part of inquiring how social capital shapes health has been a less-developed area. Recent efforts to renew conceptual and accompanying measurement attention to forms of network composition are encouraging (Bojanowski & Corten, 2014). While examination of role diversity and health has been a consistent focus in network studies cutting across population health and social science, there has been very little attention given to attribute-based diversity and health in network settings. A novel aspect of the present study is that our analyses prospectively test relationships between education assortativity and multiple health indicators. In so doing, we hope to spur additional interest in examining attribute-based diversity in studies of health.

The present research finds that, contra to our expectations for Hypothesis 1, social networks of less-educated individuals are less assortative – that is, they are more, not less, educationally diverse than networks of more-educated individuals, whose network members tend to be more similar to their own education levels. This is not consistent with recent findings in the GSS that suggest greater education homophily at the lower end of the SES spectrum than the higher end (J. A. Smith et al., 2014). Yet the present measure of homophily is also quite a different measure than a simple tie homophily measure, in that assortativity takes into account educational homophily between alters and the structure of an ego's network.[5]

We had also hypothesized (Hypothesis 2) that less education assortativity (i.e., having social contacts with a more diverse range of educational attainment) would be associated with better general physical and mental health, more physical activity, and lower BMI. Though we saw support in bivariate models, in multi-level multivariable specifications we see no evidence to support this claim. In fact, we find the opposite in terms of physical activity, where less diversity (greater assortativity) is associated with propensity to exercise regularly. These findings stands in contrast to recent findings of a protective effect on BMI for network diversity in a 5-year longitudinal setting (Wu et al., 2018). This may be attributable to measurement differences of network diversity between these studies, and differences in measurement period (3 years in the present study, vs. 5 years in the prior study). Though we do not wish to overinterpret what are essentially null associations between education assortativity and separate models of BMI, SRPH, and SRMH, a conservative interpretation would be that measuring personal network assortativity reveals a different story than prior research that has found evidence of a diversity benefit to health using a role diversity measure.

For the third hypothesis, we predicted that it would be more likely that the health of individuals in a lower SES tier would benefit from access to higher-SES others. The finding that network education diversity is associated with higher BMI status for low-education egos

[5] As a sensitivity analysis, following Smith, et al. (2014) we constructed a measure of average categorical education distance between ego and their nominated alters in order to compare how our assortativity measure comports with their dyadic measure. We then examined how educational attainment and network characteristics are associated with education distance in a multi-level model that adjusts for the full range of socio-demographic characteristics (available from authors). Briefly, we find that less-educated individuals (<HS) and the most-educated individuals (postgraduate) have the least education homophily (see Supplementary Information Appendix). Though a more thorough comparison and interpretation of measures is beyond the scope of this study, future work may productively interrogate this direction.

provides limited evidence for Hypothesis 3a, consistent with an explanation that relationships with educationally similar alters was deleterious to the weight status of lower-SES participants. There was no evidence to support Hypothesis 3b, that maintaining education diversity in one's network was burdensome in a way that was associated with poorer health.

Prior literature has suggested several different mechanisms that might link (ego) network diversity and health outcomes. In that context, we discussed, and tested, moderation through three additional mechanisms: tie strength, differential access to social capital by ego's educational level, and subjective social status. Our results suggest that tie strength does moderate the association between assortativity and health, where higher average strength of ties reduces ego's odds of being in excellent health. We thus strongly believe that fully incorporating tie strength into future analyses by using a weighted assortativity coefficient is a highly desirable next step in the study of ego network diversity and health. Also worthy of note and further exploration is that of the multiple covariates used to indicate socioeconomic status, subjective social status tended to be more predictive of the health outcomes being scrutinized than more objective measures (e.g., household assets, income, educational attainment), a finding that comports with earlier work (Singh-Manoux, Marmot, & Adler, 2005; Tan, Kraus, Carpenter, & Adler, 2018) and is increasingly recognized as an important social determinant of health.

It is also important to note how demographic patterns of structural advantage and disadvantage affect network selection. Namely, not all respondents have the same underlying probability of *access* to diverse networks. On average, people of color ("collective blacks", in the words of Bonilla-Silva (2004)) living in the US are systematically at a disadvantage. We know, for instance, that blacks, Hispanics/Latinos, and Native Americans are at a greater risk of being incarcerated (e.g., Kim, Losen & Hewitt 2010; Hirschfield 2018; Alexander, 2012; Roberts, 2003). Given this fact, these individuals should be, in all likelihood, less able to have access to educationally heterogeneous networks. Therefore, it can be expected that a person of color may have less diversity in their networks when compared to an average white individual whose neighbors, parents, friends, and other close social contacts, are much less likely to be forcefully removed (e.g., incarcerated, deported, killed) from their network.

With these important structural limitations in mind, this is believed to be the first study to investigate educational assortativity (and among the first to investigate a form of socioeconomic status diversity) and its relationship with several different commonly-investigated indicators of health. This form of attribute-based diversity measures a different dimension of lived experience than the role-based measure more often used. Importantly, the investigation of this new dimension revealed results that were not hypothesized based upon prior research based upon measures of role-based relationship type diversity. A particular strength of this study is its large size and national scope; the Gallup network panel represents one of the largest longitudinal egocentric network datasets currently available to investigators. Although other currently available datasets such as the National Social Life Health and Aging Project (Cornwell & Laumann, 2015), and the UC Berkeley Social Networks Study (Offer & Fischer, 2018) offer more depth of focus on health-related traits, both have fewer, smaller panels.

Other possible considerations that may affect interpretation of results are measurement-related. First, network diversity as measured by a ‘name generator’ is less likely to reach the same range as may be found using a position generator (Lin, 1999; Lin & Dumin, 1986). It is worth noting that this study does not include perfectly isolated people (egos with no ties) and those who are highly isolated (egos with one tie only) because of demands in calculating assortativity. Yet given this, we suspect that observing the associations we do even with a relatively blunt instrument such as a name generator to calculate education assortativity suggests that we may be underestimating the strength of a relationship between network diversity and health. Additionally, there may be measurement error in that ego report of alter traits (such as educational attainment) may not be as accurate as if those alters reported upon their traits directly. To this point, Marsden (1990) reviewed research on reports of alter attributes and found ego-proxied alter attributes to be largely similar to alter direct report. Next, although research on personal network composition has been long presumed that the egocentric network reflects only the closest of ties, recent research suggests that especially around major life transitions, individuals reach out to more peripheral social contacts for emotional support (Small, 2018). Thus, although models adjust for network density, incorporating additional information about tie strength beyond simple measures of average alter closeness and alter liking to weighted measures of assortativity could be revealing.

While controls of network size and density are important structural measures, consideration of the consistency of these networks – specifically, identification of which specific alters were dropped and which were newly added – was beyond the scope of this research. Last, those in the sample tended to be of higher socioeconomic status, women, not Hispanic, with more homogeneous types of relationships with others, and larger and more sparse networks. Given this overrepresentation of high-SES individuals, it is possible that our findings underestimate the strength of a mechanism linking low-SES educational assortativity with health. Additionally, though population weights were available, we opted not to include them because of known difficulties in using population weights in multi-level models.

In sum, the pursuit of social diversity in everyday life – of thought, of experience, and in interpersonal relationships – is a noble idea, and arguably fundamental in some sense to the human experience, even mirrored at a biological level in the role of genetic diversity in human evolution. However, research which focuses on how diversity matters in everyday life reminds us that investigating different forms of diversity in one's interpersonal life are critical to obtaining a more comprehensive picture of how social context shapes health.

Conflicts of Interest: The authors have no conflicts to disclose.

References

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. *Health Psychology, 19*(6), 586-592.

Alshamsi, A., Pianesi, F., Lepri, B., Pentland, A., & Rahwan, I. (2016). Network Diversity and Affect Dynamics: The Role of Personality Traits. *Plos One, 11*(4).

Andersson, M. A. (2012). Dispositional Optimism and the Emergence of Social Network Diversity. *Sociological Quarterly, 53*(1), 92-115.

Barefoot, J. C., Gronbaek, M., Jensen, G., Schnohr, P., & Prescott, E. (2005). Social network diversity and risks of ischemic heart disease and total mortality: Findings from the Copenhagen City Heart Study. *American Journal of Epidemiology, 161*(10), 960-967.

Bassett, E., & Moore, S. (2013). Social capital and depressive symptoms: The association of psychosocial and network dimensions of social capital with depressive symptoms in Montreal, Canada. *Social Science & Medicine, 86*, 96-102.

Berkman, L. F., & Krishna, A. (2014). Social network epidemiology. In L. F. Berkman, I. Kawachi, & M. Glymour (Eds.), *Social Epidemiology* (pp. 234-289).

Bojanowski, M., & Corten, R. (2014). Measuring segregation in social networks. *Social Networks, 39*, 14-32.

Bonilla-Silva, E. (2004). From bi-racial to tri-racial: Towards a new system of racial stratification in the USA. *Ethnic and Racial Studies, 27*(6), 931-950.

Bourdieu, P. ([1986] 2018). The Forms of Capital. In *The Sociology of Economic Life* (pp. 78-92): Routledge.

Burt, R. S. (2004). Structural holes and good ideas. *American journal of Sociology, 110*(2), 349-399.

Cattell, V. (2001). Poor people, poor places, and poor health: the mediating role of social networks and social capital. *Soc Sci Med, 52*(10), 1501-1516.

Child, S., Stewart, S., & Moore, S. (2017). Perceived control moderates the relationship between social capital and binge drinking: longitudinal findings from the Montreal Neighborhood Networks and Health Aging (MoNNET-HA) panel. *Annals of epidemiology, 27*(2), 128-134.

Choi, H. J., & Smith, R. A. (2013). Members, Isolates, and Liaisons: Meta-Analysis of Adolescents' Network Positions and Their Smoking Behavior. *Substance Use & Misuse, 48*(8), 612-622.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine, 357*(4), 370-379.

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M., Jr. (1997). Social ties and susceptibility to the common cold. *JAMA, 277*(24), 1940-1944.

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Cornwell, B. (2009). Good health and the bridging of structural holes. *Social Networks, 31*(1), 92-103.

Cornwell, B., & Laumann, E. O. (2015). The health benefits of network growth: New evidence from a national survey of older adults. *Social Science & Medicine, 125*, 94-106.

Dunbar, R. (2018). The anatomy of friendship. *Trends in cognitive sciences, 22*(1), 32-51.

Eagle, N., Macy, M., & Claxton, R. (2010). Network Diversity and Economic Development. *Science, 328*(5981), 1029-1031.

Ellwardt, L., Van Tilburg, T. G., & Aartsen, M. (2015). The mix matters: Complex personal networks relate to higher cognitive functioning in old age. *Social Science & Medicine, 125*, 107-115.

Erickson, B. H. (2003). Social networks: The value of variety. *Contexts, 2*(1), 25-31.

Escobar-Bravo, M. A., Puga-Gonzalez, D., & Martin-Baranera, M. (2012). Protective effects of social networks on disability among older adults in Spain. *Archives of Gerontology and Geriatrics, 54*(1), 109-116.

Feld, S. L. (1981). The Focused Organization of Social Ties. *American journal of Sociology, 86*(5), 1015-1035.

Feld, S. L. (1982). Social Structural Determinants of Similarity among Associates. *American Sociological Review, 47*(6), 797-801.

Fischer, C. S. (1977). *Networks and places: Social relations in the urban setting*: Free Press.

Gallup, I. (2014). *Gallup Panel Whitepaper Brief*. Retrieved from http://www.gallup.com:

Goldman, A. W., & Cornwell, B. (2015). Social network bridging potential and the use of complementary and alternative medicine in later life. *Social Science & Medicine, 140*, 69-80.

Goodreau, S. M., Kitts, J. A., & Morris, M. (2009). Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks. *Demography, 46*(1), 103-125.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American journal of Sociology, 78*(6), 1360-1380.

Haas, S. A., Schaefer, D. R., & Kornienko, O. (2010). Health and the Structure of Adolescent Social Networks. *Journal of Health and Social Behavior, 51*(4), 424-439.

Hodge, R. W., & Treiman, D. J. (1968). Class identification in the United States. *American journal of Sociology, 73*(5), 535-547.

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social Relationships and Mortality Risk: A Meta-analytic Review. *Plos Medicine, 7*(7).

Huisman, M. (2014). Imputation of missing network data: Some simple procedures. *Encyclopedia of social network analysis and mining*, 707-715.

Kawachi, I., & Berkman, L. F. (2001). Social ties and mental health. *Journal of Urban Health-Bulletin of the New York Academy of Medicine, 78*(3), 458-467.

Kelly, L., Patel, S. A., Narayan, K. M. V., Prabhakaran, D., & Cunningham, S. A. (2014). Measuring Social Networks for Medical Research in Lower-Income Settings. *Plos One, 9*(8).

Laumann, E. O., & Pappi, F. U. (1973). New directions in the study of community elites. *American Sociological Review*, 212-230.

Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society, 18*(1), 18-66.

Legh-Jones, H., & Moore, S. (2012). Network social capital, social participation, and physical inactivity in an urban adult population. *Social Science & Medicine, 74*(9), 1362-1367.

Lin, N. (1999). Building a network theory of social capital. *Connections, 22*(1), 28-51.

Lin, N. (2001). *Social capital : a theory of social structure and action*. Cambridge, UK ; New York: Cambridge University Press.

Lin, N., & Dumin, M. (1986). Access to Occupations through Social Ties. *Social Networks, 8*(4), 365-385.

Marsden, P. V. (1988). Homogeneity in confiding relations. *Social Networks, 10*(1), 57-76.

Marsden, P. V. (1990). Network Data and Measurement. *Annual Review of Sociology, 16*, 435-463.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 415-444.

Molesworth, T., Sheu, L. K., Cohen, S., Gianaros, P. J., & Verstynen, T. D. (2015). Social network diversity and white matter microstructural integrity in humans. *Social Cognitive and Affective Neuroscience, 10*(9), 1169-1176.

Moore, S., Daniel, M., Gauvin, L., & Dubé, L. (2009). Not all social capital is good capital. *Health & place, 15*(4), 1071-1077.

Moore, S., Daniel, M., Paquet, C., Dube, L., & Gauvin, L. (2009). Association of individual network social capital with abdominal adiposity, overweight and obesity. *Journal of Public Health, 31*(1), 175-183.

Moore, S., Teixeira, A., & Stewart, S. (2014). Effect of Network Social Capital on the Chances of Smoking Relapse: A Two-Year Follow-up Study of Urban-Dwelling Adults. *American Journal of Public Health, 104*(12), E72-E76.

Mowbray, O., Quinn, A., & Cranford, J. A. (2014). Social networks and alcohol use disorders: findings from a nationally representative sample. *American Journal of Drug and Alcohol Abuse, 40*(3), 181-186.

Newman, M. E. J. (2003). Mixing patterns in networks. *Physical Review E, 67*(2).

O'Malley, A. J., Arbesman, S., Steiger, D. M., Fowler, J. H., & Christakis, N. A. (2012). Egocentric Social Network Structure, Health, and Pro-Social Behaviors in a National Panel Study of Americans. *Plos One, 7*(5).

Offer, S., & Fischer, C. S. (2018). Difficult People: Who Is Perceived to Be Demanding in Personal Networks and Why Are They There? *American Sociological Review, 83*(1), 111-142.

Perry, B. L., Pescosolido, B. A., & Borgatti, S. P. (2018). *Egocentric network analysis : foundations, methods, and models*. Cambridge, UK ; New York, NY: Cambridge University Press.

Platt, J., Keyes, K. M., & Koenen, K. C. (2014). Size of the social network versus quality of social support: which is more protective against PTSD? *Social Psychiatry and Psychiatric Epidemiology, 49*(8), 1279-1286.

Putnam, R. (2001). Social capital: Measurement and consequences. *Canadian journal of policy research, 2*(1), 41-51.

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rice, E., Kurzban, S., & Ray, D. (2012). Homeless But Connected: The Role of Heterogeneous Social Network Ties and Social Networking Technology in the Mental Health Outcomes of Street-Living Adolescents. *Community Mental Health Journal, 48*(6), 692-698.

Santini, Z. I., Koyanagi, A., Tyrovolas, S., Haro, J. M., Fiori, K. L., Uwakwa, R., . . . Prina, A. M. (2015). Social network typologies and mortality risk among older people in China, India, and Latin America: A 10/66 Dementia Research Group population-based cohort study. *Social Science & Medicine, 147*, 134-143.

Schaefer, D. R., Kornienko, O., & Fox, A. M. (2011). Misery does not love company: Network selection mechanisms and depression homophily. *American Sociological Review, 76*(5), 764-785.

Shiovitz-Ezra, S., & Litwin, H. (2012). Social network type and health-related behaviors: Evidence from an American national survey. *Social Science & Medicine, 75*(5), 901-904.

Singh-Manoux, A., Marmot, M. G., & Adler, N. E. (2005). Does subjective social status predict health and change in health status better than objective status? *Psychosomatic Medicine, 67*(6), 855-861.

Small, M. L. (2018). *Someone to Talk to*: Oxford University Press.

Smith, J. A., McPherson, M., & Smith-Lovin, L. (2014). Social Distance in the United States Sex, Race, Religion, Age, and Education Homophily among Confidants, 1985 to 2004. *American Sociological Review, 79*(3), 432-456.

Smith, J. A., Moody, J., & Morgan, J. H. (2017). Network sampling coverage II: The effect of non-random missing data on network measurement. *Social Networks, 48*, 78-99.

Smith, K. P., & Christakis, N. A. (2008). Social networks and health. *Annual Review of Sociology, 34*, 405-429.

Song, L. (2011). Social Capital and Psychological Distress. *Journal of Health and Social Behavior, 52*(4), 478-492.

Song, L., Pettis, P. J., & Piya, B. (2017). Does Your Body Know Who You Know? Multiple Roles of Network Members' Socioeconomic Status for Body Weight Ratings. *Sociological Perspectives, 60*(6), 997-1018.

StataCorp. (2013). Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.

Tan, J. J. X., Kraus, M., Carpenter, N., & Adler, N. (2018). The Association Between Objective and Subjective Socioeconomic Standing and Subjective Well-being: A Meta-analytic Review.

Thoits, P. A. (2011). Mechanisms Linking Social Ties and Support to Physical and Mental Health. *Journal of Health and Social Behavior, 52*(2), 145-161.

VanderWeele, T. J., & Christakis, N. A. (2019). Network multipliers and public health. *International journal of epidemiology*.

Viruell-Fuentes, E. A., Morenoff, J. D., Williams, D. R., & House, J. S. (2013). Contextualizing nativity status, Latino social ties, and ethnic enclaves: an examination of the 'immigrant social ties hypothesis'. *Ethnicity & Health, 18*(6), 586-609.

Wang, C., Butts, C. T., Hipp, J. R., Jose, R., & Lakon, C. M. (2016). Multiple imputation for missing edge data: A predictive evaluation method with application to Add Health. *Social Networks, 45*, 89-98.

Wellman, B. (1979). The community question: The intimate networks of East Yorkers. *American journal of Sociology, 84*(5), 1201-1231.

Wu, Y.-H., Moore, S., & Dube, L. (2018). Social capital and obesity among adults: Longitudinal findings from the Montreal neighborhood networks and healthy aging panel. *Preventive medicine, 111*, 366-370.

Yang, Y. C., Boen, C., Gerken, K., Li, T., Schorpp, K., & Harris, K. M. (2016). Social relationships and physiological determinants of longevity across the human life span. *Proc Natl Acad Sci U S A, 113*(3), 578-583.

Zhang, T., Cao, W. H., Lv, J., Wang, N., Reilly, K. H., Zhu, Q., & Li, L. M. (2012). Size, Composition, and Strength of Ties of Personal Social Support Networks Among Adult People Living with HIV/AIDS in Henan and Beijing, China. *Aids and Behavior, 16*(4), 911-919.

Table 1. Sample characteristics (Baseline, Year 1)

Characteristics	n=	Cat % / Mean (SD)	Range	Characteristics	n=	Cat % / Mean (SD)	Range
Network covariates				*(Socio-demographic covars., cont'd)*			
Education assortativity	9108	-0.21 (0.23)	-1.0 - 0.5	Race			
Number of alters	10679	4.1 (1.89)	0 - 8	White	9,423	88.2%	
Personal network density	9108	0.87 (0.16)	0.22 - 1.0	Black	437	4.1%	
Average closeness to alters	9695	8.24 (1.12)	1 - 10	Asian	99	0.9%	
Average liking of alters	9696	8.65 (0.96)	1 - 10	Other	158	1.5%	
Health covariates				Multiple	471	4.4%	
Body-Mass Index	10177	27.8 (5.6)	15.3 - 59.6	NA (not available)	91	0.9%	
Exercised regularly				Ethnicity			
No	3,962	37.1%		Not Hispanic	9,896	92.7%	
Yes	6,464	60.5%		Hispanic	602	5.6%	
NA (not available)	253	2.4%		NA (not available)	181	1.7%	
Physical health				Employment status			
Poor	221	2.1%		Employed full-time	5076	47.50%	
Fair	1,740	16.3%		Employed part-time	1027	9.60%	
Good	6,364	59.6%		Other status	3998	37.40%	
Excellent	2,269	21.3%		NA (not available)	578	5.40%	
NA (not available)	85	0.8%		Education			
Mental health				<HS	98	0.9%	
Poor	114	1.1%		High school	946	8.9%	
Fair	1,011	9.5%		Some college	2471	23.1%	
Good	4,820	45.1%		College	2916	27.3%	
Excellent	4,583	42.9%		Postgraduate	4246	39.8%	
NA (not available)	151	1.4%		NA (not available)	2	0.02%	
Socio-demographic covariates				Income	9816	5.8 (1.8)	1 - 9
Gender				Household Assets	9241	3.9 (2.0)	1 - 8
Male	5,588	0.52		Subjective Social Status	10491	7.3 (1.6)	0 - 10
Female	5,091	0.48		Region			
Age	10679	58.1 (13.5)	18 - 95	Northeast	1,589	14.9%	
Marital Status				Midwest	2,838	26.6%	
Married/Living with a Partner	7602	71.2%		South	3,474	32.5%	
Other status	2496	22.4%		West	2,756	25.8%	
NA (not available)	581	6.4%		NA (not available)	22	0.2%	

Table 2. Associations with education network assortativity

	Year 1 OLS coeff.	Years 1-3 MLM coeff.
Key independent variables		
Educational Attainment (Ref: College)		
<HS	-0.067*	-0.012
	(-0.128,-0.006)	(-0.030,0.006)
HS	-0.001	-0.015*
	(-0.024,0.021)	(-0.028,-0.001)
Some college	-0.033***	-0.006
	(-0.047,-0.018)	(-0.017,0.004)
Postgraduate	0.016*	0.023***
	(0.004,0.029)	(0.011,0.035)
Network covariates		
Number of alters	0.032***	0.033***
	(0.028,0.036)	(0.031,0.035)
Personal network density	0.055**	0.040***
	(0.018,0.092)	(0.018,0.062)
Average closeness to alters	0.008	0.005
	(-0.001,0.018)	(-0.000,0.011)
Average liking of alters	-0.007	0.000
	(-0.018,0.004)	(-0.007,0.006)
Socio-demographic covariates		
Male (Ref: Female)	0.005	0.005
	(-0.005,0.016)	(-0.003,0.013)
Age	0.000	-0.001***
	(-0.001,0.000)	(-0.001,-0.000)
Race (Ref: White)		
Black	-0.017	-0.014
	(-0.044,0.010)	(-0.033,0.006)
Asian	0.013	0.022
	(-0.043,0.069)	(-0.018,0.062)
Other race	-0.003	0.008
	(-0.046,0.041)	(-0.024,0.039)
Multiracial	-0.006	0.001
	(-0.032,0.020)	(-0.018,0.020)
Hispanic	-0.018	-0.023**
	(-0.043,0.006)	(-0.041,-0.006)
Subjective social status	0.001	0.001
	(-0.003,0.004)	(-0.001,0.003)
N (observations)	-	21795
N groups (egos) groups (egos)	7308	9090
AIC	-1218.1	-4418.1
BIC	-956.0	-4074.6

Note: * p<0.05; ** p<0.01; *** p<0.001

Note: Both models adjust for covariates shown above, as well as categorical measures for region, marital status, household asset tiers, income tiers, and employment status. Multi-level model includes a continuous time measure.

Table 3. Network education assortativity and health (random-coefficient multilevel models)

	Exercise Regularly			BMI			Excellent Self-reported Physical Health (SRPH)			Excellent Self-reported Mental Health (SRMH)		
	(1) Baseline	(2) +Inc, Wealth	(3) +Ed x EdAssort	(1) Baseline	(2) +Inc, Wealth	(3) +Ed x EdAssort	(1) Baseline	(2) +Inc, Wealth	(3) +Ed x EdAssort	(1) Baseline	(2) +Inc, Wealth	(3) +Ed x EdAssort
	OR	OR	OR	coeff.	coeff.	coeff.	OR	OR	OR	OR	OR	OR
Key independent variables												
Educational Attainment (Ref: College)												
<HS	0.348***	0.411***	0.449***	0.230*	0.128	0.237	0.501**	0.607*	0.377**	1.049	1.159	1.218
	(0.254,0.475)	(0.299,0.564)	(0.297,0.678)	(0.023,0.437)	(-0.082,0.339)	(-0.033,0.507)	(0.314,0.798)	(0.378,0.973)	(0.191,0.745)	(0.768,1.433)	(0.843,1.594)	(0.795,1.866)
HS	0.557***	0.637***	0.762	0.137	0.061	0.229*	0.718*	0.821	0.732	1.046	1.131	1.077
	(0.445,0.698)	(0.507,0.801)	(0.564,1.029)	(-0.000,0.274)	(-0.079,0.201)	(0.045,0.414)	(0.524,0.985)	(0.595,1.133)	(0.478,1.121)	(0.834,1.312)	(0.899,1.425)	(0.792,1.464)
Some college	0.789**	0.842*	0.934	0.124**	0.088	0.124	0.962	1.031	0.816	1.016	1.059	1.184
	(0.665,0.936)	(0.709,1.000)	(0.735,1.186)	(0.036,0.212)	(-0.001,0.177)	(-0.004,0.252)	(0.762,1.216)	(0.814,1.305)	(0.590,1.130)	(0.854,1.209)	(0.888,1.263)	(0.928,1.511)
Postgraduate	1.439***	1.454***	1.632***	-0.775***	-0.773***	-0.703***	1.775***	1.721***	1.355	0.963	0.978	0.953
	(1.167,1.774)	(1.179,1.794)	(1.252,2.128)	(-1.020,-0.530)	(-1.019,-0.528)	(-0.967,-0.439)	(1.342,2.349)	(1.299,2.280)	(0.956,1.920)	(0.785,1.181)	(0.796,1.201)	(0.734,1.238)
Education Assortativity	1.283*	1.260*	0.76	0.012	0.009	-0.303	1.068	1.053	2.903*	1.155	1.152	1.096
	(1.020,1.615)	(1.002,1.585)	(0.387,1.490)	(-0.141,0.165)	(-0.143,0.162)	(-0.702,0.095)	(0.783,1.456)	(0.772,1.437)	(1.146,7.357)	(0.918,1.452)	(0.916,1.450)	(0.548,2.192)
Interactions (Ref: College)												
Education Assortativity x <HS	-	-	1.491	-	-	0.421	-	-	0.128*	-	-	1.24
			(0.479,4.642)			(-0.286,1.128)			(0.019,0.877)			(0.375,4.102)
Education Assortativity x HS	-	-	2.191	-	-	0.691**	-	-	0.529	-	-	0.831
			(0.931,5.159)			(0.182,1.199)			(0.151,1.860)			(0.341,2.028)
Education Assortativity x Some college	-	-	1.624	-	-	0.173	-	-	0.324*	-	-	1.643
			(1.000,1.000)			(-0.245,0.592)			(0.112,0.941)			(0.750,3.597)
Education Assortativity x Postgraduate	-	-	1.709	-	-	0.304	-	-	0.310*	-	-	0.883
			(0.804,3.634)			(-0.151,0.760)			(0.112,0.862)			(0.410,1.902)
Subjective social status	1.445***	1.407***	1.407***	-0.153***	-0.149***	-0.149***	2.892***	2.819***	2.821***	2.997***	2.959***	2.959***
	(1.386,1.506)	(1.349,1.467)	(1.349,1.467)	(-0.182,-0.124)	(-0.178,-0.120)	(-0.178,-0.120)	(2.672,3.131)	(2.604,3.053)	(2.605,3.055)	(2.837,3.165)	(2.800,3.126)	(2.800,3.127)
Network covariates												
Number of alters	1.023	1.023	1.023	-0.009	-0.009	-0.009	0.987	0.988	0.987	1.065**	1.065**	1.064**
	(0.983,1.065)	(0.983,1.065)	(0.983,1.065)	(-0.032,0.014)	(-0.033,0.014)	(-0.033,0.014)	(0.934,1.043)	(0.935,1.044)	(0.934,1.043)	(1.023,1.109)	(1.023,1.109)	(1.022,1.108)
Personal network density	0.551**	0.549**	0.547**	0.204	0.213	0.210	0.908	0.927	0.925	1.332	1.32	1.331
	(0.377,0.805)	(0.376,0.802)	(0.375,0.799)	(-0.016,0.425)	(-0.008,0.433)	(-0.010,0.431)	(0.538,1.533)	(0.548,1.566)	(0.547,1.563)	(0.908,1.954)	(0.899,1.938)	(0.906,1.955)
Average closeness to alters	0.945	0.944	0.945	-0.02	-0.018	-0.017	0.872*	0.866*	0.865*	1.021	1.022	1.021
	(0.859,1.040)	(0.858,1.038)	(0.859,1.040)	(-0.075,0.036)	(-0.074,0.037)	(-0.072,0.039)	(0.762,0.998)	(0.757,0.991)	(0.756,0.990)	(0.925,1.128)	(0.925,1.128)	(0.924,1.127)
Average liking of alters	1.03	1.036	1.034	0.000	-0.001	-0.002	1.258**	1.268**	1.274**	1.242***	1.246***	1.245***
	(0.926,1.147)	(0.931,1.152)	(0.929,1.151)	(-0.061,0.062)	(-0.062,0.060)	(-0.064,0.059)	(1.081,1.464)	(1.089,1.477)	(1.094,1.483)	(1.111,1.388)	(1.114,1.393)	(1.113,1.393)
Socio-demographic covariates												
Male (Ref: Female)	1.149	1.147	1.147	-0.996***	-0.986***	-0.987***	1.506***	1.511***	1.511***	0.669***	0.667***	0.668***
	(0.995,1.328)	(0.992,1.327)	(0.992,1.327)	(-1.224,-0.767)	(-1.214,-0.758)	(-1.215,-0.758)	(1.240,1.830)	(1.241,1.840)	(1.241,1.841)	(0.582,0.768)	(0.579,0.768)	(0.580,0.769)
Age	1.01	1.000	1.000	0.00	0	0	0.978***	0.977***	0.977***	1.020***	1.017***	1.018***
	(1.000,1.011)	(0.994,1.006)	(0.994,1.006)	(-0.011,0.006)	(-0.008,0.009)	(-0.008,0.009)	(0.971,0.985)	(0.969,0.985)	(0.969,0.985)	(1.015,1.025)	(1.011,1.023)	(1.012,1.024)
Race (Ref: White)												
Black	0.957	1.031	1.031	1.159***	1.119***	1.119***	0.338***	0.369***	0.368***	1.229	1.261	1.259
	(0.666,1.374)	(0.718,1.480)	(0.718,1.479)	(0.583,1.735)	(0.544,1.694)	(0.544,1.694)	(0.199,0.576)	(0.217,0.628)	(0.216,0.627)	(0.870,1.737)	(0.891,1.784)	(0.889,1.782)
Asian	2.379*	2.290*	2.292*	-3.679***	-3.651***	-3.648***	1.909	1.788	1.795	0.945	0.925	0.92
	(1.085,5.215)	(1.044,5.023)	(1.044,5.029)	(-4.870,-2.488)	(-4.840,-2.463)	(-4.836,-2.460)	(0.728,5.006)	(0.682,4.687)	(0.684,4.711)	(0.459,1.946)	(0.448,1.909)	(0.446,1.901)
Other race	1.144	1.129	1.127	0.859	0.859	0.857	0.598	0.594	0.597	0.881	0.885	0.886
	(0.629,2.082)	(0.622,2.049)	(0.621,2.047)	(-0.091,1.810)	(-0.089,1.807)	(-0.091,1.805)	(0.262,1.366)	(0.260,1.356)	(0.261,1.365)	(0.498,1.558)	(0.500,1.568)	(0.500,1.571)
Multiracial	1.580*	1.656**	1.656**	1.144***	1.117***	1.116***	0.449**	0.472**	0.472**	1.477*	1.507*	1.504*
	(1.110,2.250)	(1.164,2.355)	(1.164,2.355)	(0.578,1.709)	(0.553,1.681)	(0.552,1.680)	(0.268,0.753)	(0.281,0.792)	(0.281,0.792)	(1.051,2.076)	(1.071,2.121)	(1.069,2.118)
Hispanic	1.281	1.293	1.294	0.395	0.393	0.393	1.001	1.021	1.016	1.138	1.137	1.138
	(0.918,1.787)	(0.928,1.801)	(0.929,1.803)	(-0.129,0.919)	(-0.130,0.916)	(-0.130,0.915)	(0.642,1.560)	(0.654,1.593)	(0.651,1.585)	(0.830,1.560)	(0.828,1.561)	(0.829,1.562)
N (observations)	21525	21525	21525	21199	21199	21199	21701	21701	21701	21,600	21600	21600
N groups (egos) groups (egos)	9056	9056	9056	8967	8967	8967	9191	9191	9191	9065	9065	9065
AIC	24104.7	24073.7	24078.35	105676.9	105663	105662.5	16330.6	16300.4	16301.1	22243.1	22258.3	22260.2
BIC	24296.1	24416.7	24453.27	105883.9	106021.3	106052.6	16522.3	16643.8	16676.4	22434.6	22601.4	22635.3

Note: * p<0.05; ** p<0.01; *** p<0.001
Note: Model 1 adjusts for covariates shown, as well as region (categorical), marital status (categorical), and time (continuous). This model is arguably the best-fitting, with the lowest BIC across the four health models.
Note: Model 2 adds additional SES measures to prior model: household asset tiers (categorical), income tiers (categorical), employment status (categorical).
Note: Model 3 adds interaction between education attainment categories and education assortativity.

Table 4. Tie strength moderation of network education assortativity and health (random-coefficient multilevel models)

	Exercise Regularly	BMI	Excellent Self-reported Physical Health (SRPH)	Excellent Self-reported Mental Health (SRMH)
	OR	coeff.	OR	OR
Education Assortativity	1.445	-0.026	18.795*	1.062
	(0.282,7.399)	(-1.069,1.018)	(1.518,232.7)	(0.171,6.582)
Average liking of alters	1.032	0.000	1.181*	1.249***
	(0.918,1.159)	(-0.069,0.068)	(1.003,1.391)	(1.104,1.412)
Interaction				
Education assortativity x Avg. alter liking	0.984	0.004	0.717*	1.009
	(0.815,1.188)	(-0.116,0.124)	(0.537,0.956)	(0.819,1.244)
Network covariates				
Number of alters	1.023	-0.009	0.989	1.065**
	(0.983,1.065)	(-0.033,0.014)	(0.936,1.045)	(1.023,1.109)
Personal network density	0.549**	0.213	0.928	1.32
	(0.376,0.802)	(-0.008,0.433)	(0.549,1.569)	(0.899,1.938)
Average closeness to alters	0.944	-0.018	0.866*	1.022
	(0.858,1.038)	(-0.074,0.037)	(0.757,0.991)	(0.925,1.128)
Socio-demographic covariates				
Male (Ref: Female)	1.147	-0.986***	1.510***	0.667***
	(0.992,1.327)	(-1.214,-0.758)	(1.240,1.840)	(0.579,0.767)
Age	1.000	0.000	0.977***	1.017***
	(0.994,1.006)	(-0.008,0.009)	(0.969,0.985)	(1.011,1.023)
Race (Ref: White)				
Black	1.031	1.119***	0.365***	1.261
	(0.718,1.479)	(0.544,1.694)	(0.215,0.622)	(0.891,1.784)
Asian	2.290*	-3.651***	1.776	0.925
	(1.044,5.023)	(-4.840,-2.463)	(0.677,4.660)	(0.448,1.910)
Other race	1.129	0.859	0.596	0.885
	(0.622,2.049)	(-0.089,1.807)	(0.261,1.361)	(0.500,1.568)
Multiracial	1.655**	1.117***	0.473**	1.507*
	(1.164,2.354)	(0.553,1.681)	(0.282,0.795)	(1.070,2.121)
Hispanic	1.293	0.393	1.017	1.137
	(0.928,1.801)	(-0.130,0.916)	(0.652,1.587)	(0.828,1.561)
Educational Attainment (Ref: College)				
<HS	0.411***	0.128	0.610*	1.159
	(0.299,0.564)	(-0.083,0.339)	(0.380,0.978)	(0.843,1.594)
HS	0.637***	0.061	0.823	1.131
	(0.507,0.801)	(-0.079,0.201)	(0.596,1.135)	(0.899,1.425)
Some college	0.842*	0.088	1.031	1.059
	(0.709,1.000)	(-0.001,0.177)	(0.814,1.305)	(0.888,1.263)
Postgraduate	1.454***	-0.773***	1.726***	0.978
	(1.179,1.794)	(-1.019,-0.528)	(1.302,2.287)	(0.796,1.201)
Subjective social status	1.407***	-0.149***	2.820***	2.959***
	(1.349,1.467)	(-0.178,-0.120)	(2.605,3.054)	(2.800,3.126)
N (observations)	21525	21199	21600	21701
N groups (egos)	9056	8967	9065	9191
AIC	24075.64	105665.0	16297.5	22260.3
BIC	24426.62	106031.2	16648.8	22611.4

Note: * p<0.05; ** p<0.01; *** p<0.001

Note: models adjust for income tiers, household asset tiers, region, employment status, marital status (all categorical), and time.

Note: independent covariance structure specified.

Figures

Figure 1. Egocentric education assortativity distribution. Note that lower values correspond with more dissortativity (greater attribute diversity), and higher values correspond with more assortativity (lower attribute diversity).

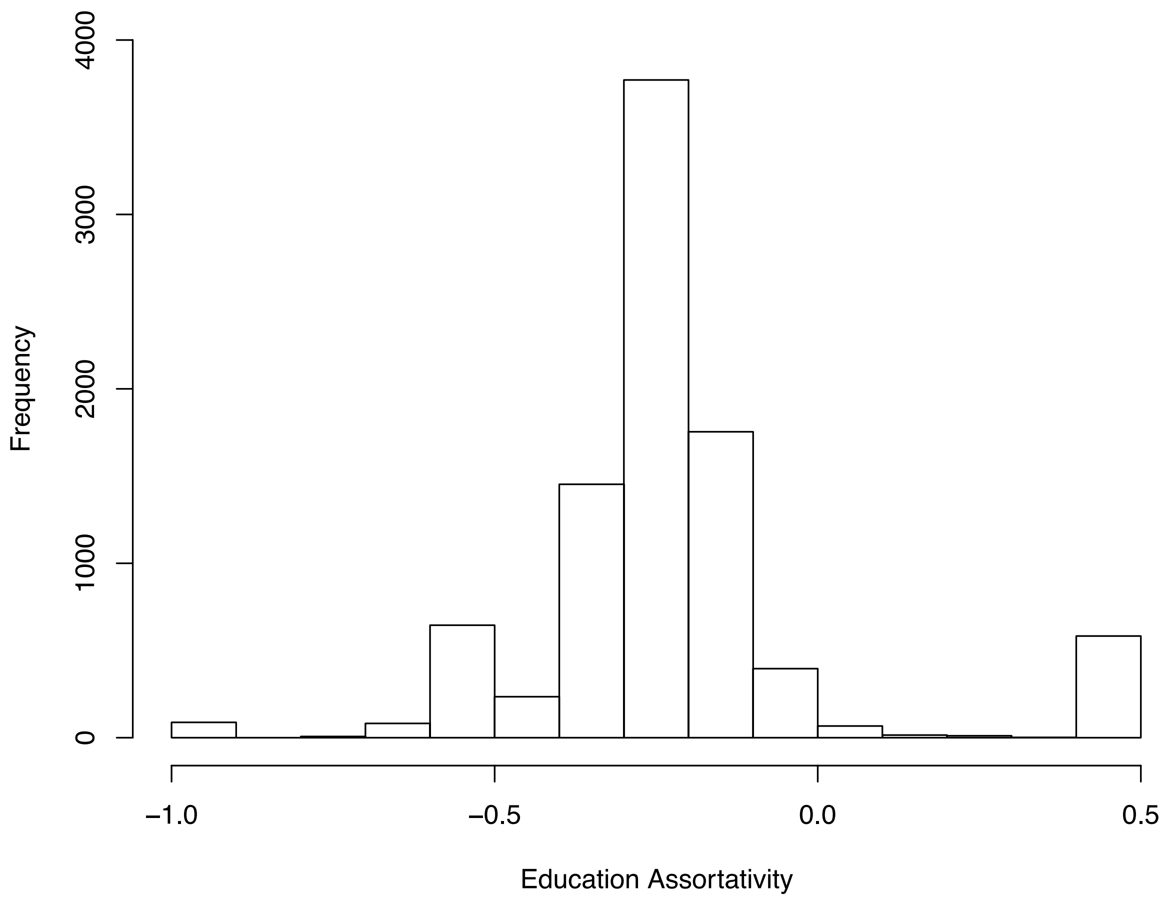


Figure 2. Mean egocentric network education assortativity, variation by participant characteristics. The assortativity scale is from -1.0 (most dissortative/maximally diverse education among alters) to 0.5 (most assortative/maximally homogeneous education among alters). There is a tendency for those who report frequent exercise, lower BMI, excellent self-rated physical health and mental health to have more educationally assortative egocentric networks.

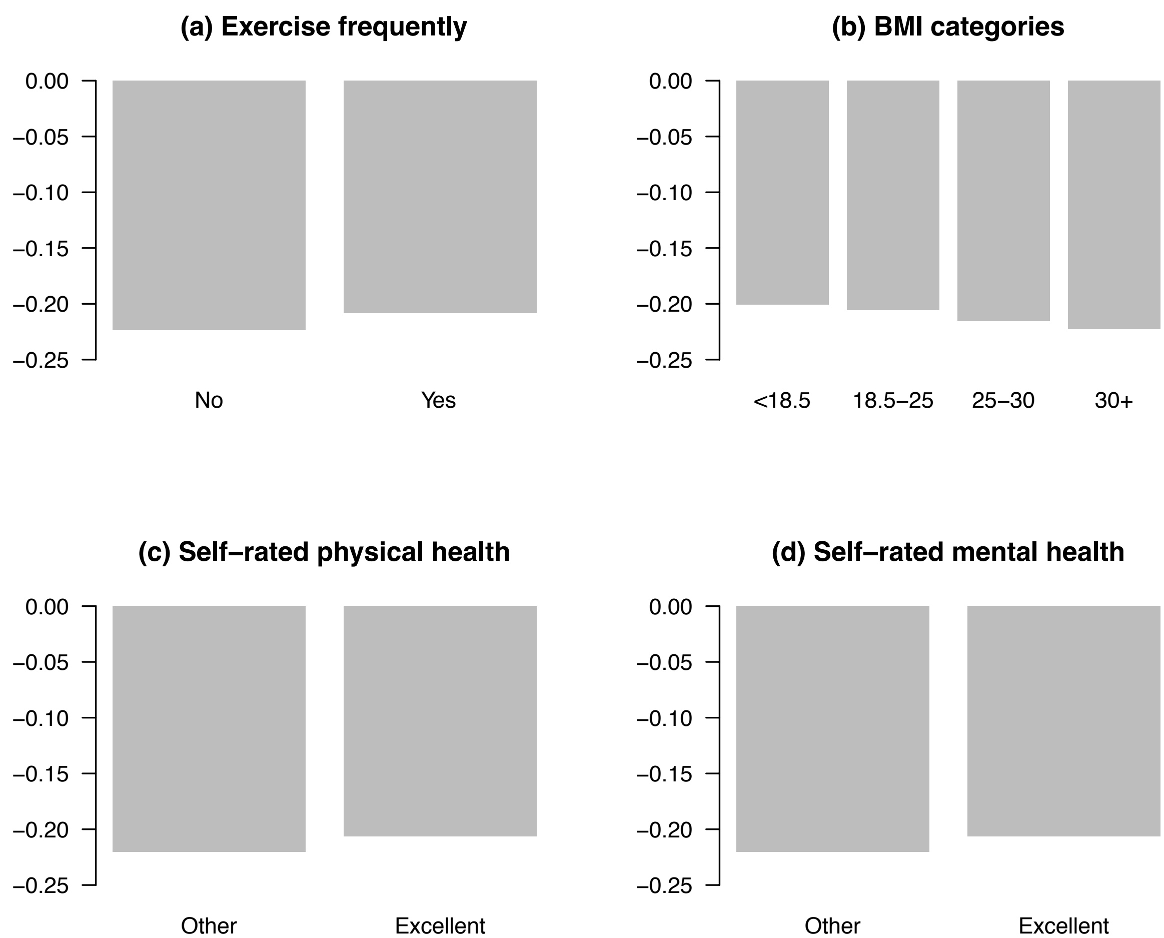


Figure 3. Mean egocentric network education assortativity, variation by socioeconomic status. The assortativity scale ranges from -1.0 (most dissortative/maximally diverse alter education) to 0.5 (most assortative alter education). These measures suggest a nearly linear gradient between those with lower SES having more educationally diverse networks, and those with higher SES having more educationally homogenous networks (the fewer number of responses in < High School and HS education categories suggest that if these categories were pooled, the pattern would be more linear; the same is the case with the three lowest subjective status categories).

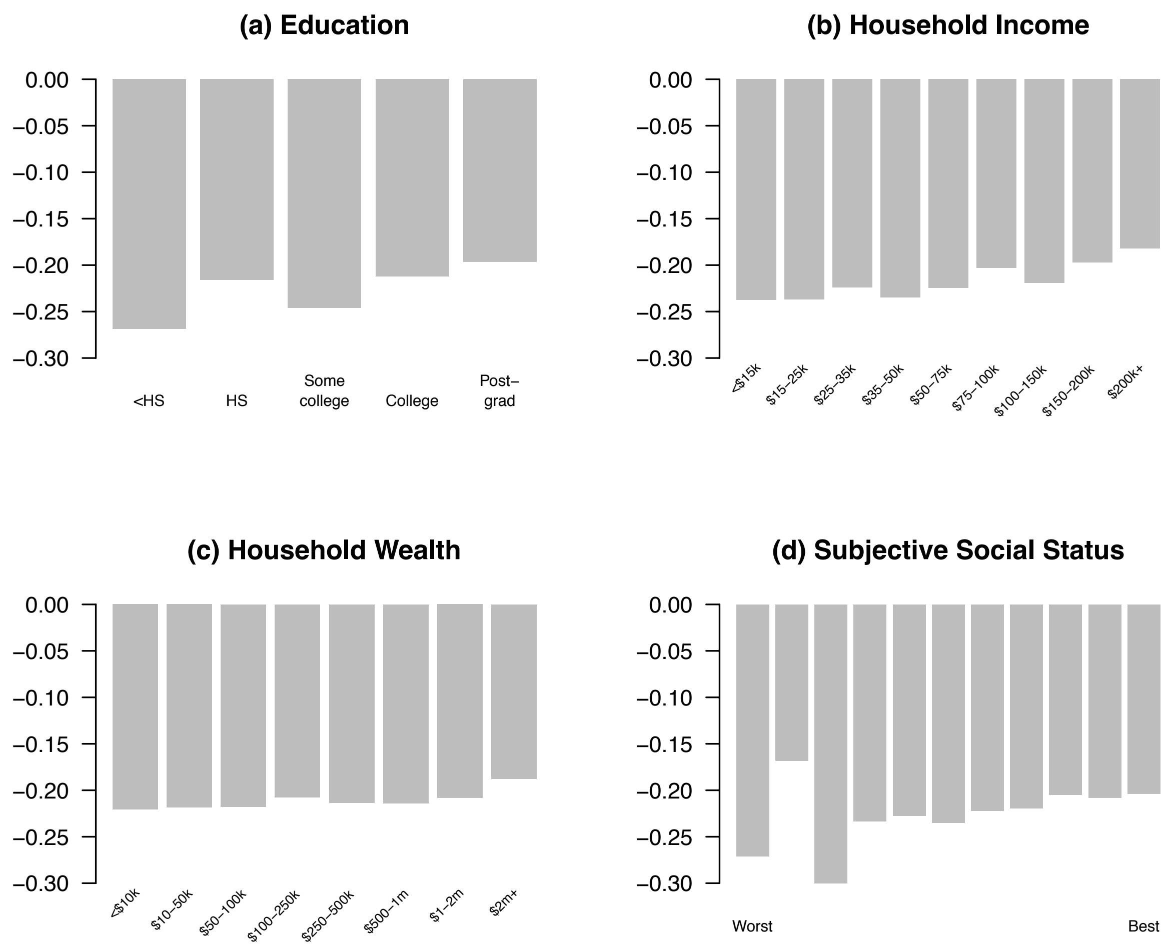
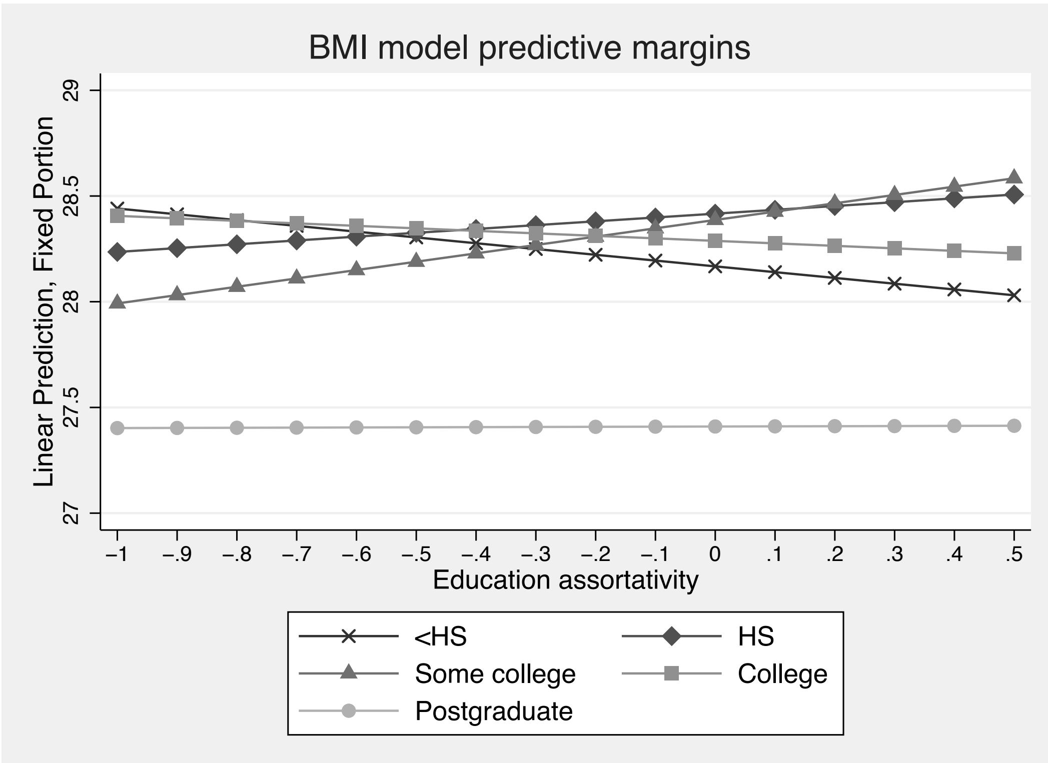


Figure 4. BMI model predictive margins from multi-level model (Table 3), showing that greater assortativity (lower network education diversity) is associated with higher BMI among individuals with a high school degree or some college.



Is Having an Educationally Diverse Social Network Good for Health?

Mark C. Pachucki, Diego F. Leal

Supplementary Information

In an effort to be maximally transparent with analysis of the relationship between education assortativity and health indicators, we report stepwise model estimation here. We observe that although the relationship between educational assortativity and our four outcomes remain essentially null, there is some evidence that there are differential returns to education assortativity depending upon one's level of educational attainment. This is notable given the abundance of role diversity research that has shown beneficial health returns to network diversity. Sensitivity Tables 1a-1d (below) report stepwise model progression, though only Models 3,4,5 are reported in the body text.

• When we begin with a baseline model (1) with *educational attainment as the only SES variable* among the other confounders, we see trends that individuals who have less education (especially HS only or "Some college" education) have worse health indicators (MH, PH, worse exercise, higher BMI), relative to individuals in the college-educated reference group.
• In models (2), we add in *education assortativity*. Across the four models, the direction of the coefficients suggests that greater assortativity (less diversity) is associated with better health indicators (though only propensity to exercise regularly is marginally significant). This is important because it suggests that a more precise measurement of diversity (by way of assortativity) reveals a different story than prior research which has found evidence of a diversity benefit to health using a role diversity measure.
• In models (3), we add *subjective social status*. Here, we observe that across all four health models, the fit is the best, with the lowest BIC for each model in the series. More importantly, as we add a term for subjective social status, we observe that the coefficient for education assortativity attenuates, providing support that perception of one's social status may be a mechanism linking network diversity and health.
• In models (4), we add *income tiers, wealth tiers, and employment status*. Here, although we do not see overall fit improving in any of the models, we believe it is critical to include key socioeconomic confounders different from education in order to better specify the relationship between education assortativity and health. Evidence in this regard comes from the fact that the coefficient for education assortativity decreases in size across all models.
• The fully-interacted models (5) are discussed at length in the manuscript.

Sensitivity Table 1a. Exercise Regularly (random-coefficient multilevel models)

	(1) Base OR	(2) +Ed Assort OR	(3) +SSS OR	(4) +Inc, Wealth OR	(5) +Ed x EdAssort OR
Key independent variables					
Educational Attainment (Ref: College)					
<HS	0.308***	0.309***	0.348***	0.411***	0.449***
	(0.225,0.422)	(0.226,0.424)	(0.254,0.475)	(0.299,0.564)	(0.297,0.678)
HS	0.516***	0.518***	0.557***	0.637***	0.762
	(0.411,0.647)	(0.413,0.650)	(0.445,0.698)	(0.507,0.801)	(0.564,1.029)
Some college	0.763**	0.764**	0.789**	0.842*	0.934
	(0.643,0.905)	(0.644,0.907)	(0.665,0.936)	(0.709,1.000)	(0.735,1.186)
Postgraduate	1.586***	1.575***	1.439***	1.454***	1.632***
	(1.283,1.961)	(1.274,1.947)	(1.167,1.774)	(1.179,1.794)	(1.252,2.128)
Education Assortativity	-	1.305*	1.283*	1.260*	0.76
		(1.036,1.644)	(1.020,1.615)	(1.002,1.585)	(0.387,1.490)
Interactions (Ref: College)					
Education Assortativity x <HS	-	-	-	-	1.491
					(0.479,4.642)
Education Assortativity x HS	-	-	-	-	2.191
					(0.931,5.159)
Education Assortativity x Some college	-	-	-	-	1.624
					(1.000,1.000)
Education Assortativity x Postgraduate	-	-	-	-	1.709
					(0.804,3.634)
Subjective social status	-	-	1.445***	1.407***	1.407***
			(1.386,1.506)	(1.349,1.467)	(1.349,1.467)
Network covariates					
Number of alters	1.050*	1.041	1.023	1.023	1.023
	(1.009,1.092)	(0.999,1.084)	(0.983,1.065)	(0.983,1.065)	(0.983,1.065)
Personal network density	0.568**	0.563**	0.551**	0.549**	0.547**
	(0.388,0.831)	(0.384,0.824)	(0.377,0.805)	(0.376,0.802)	(0.375,0.799)
Average closesness to alters	0.991	0.99	0.945	0.944	0.945
	(0.901,1.091)	(0.900,1.089)	(0.859,1.040)	(0.858,1.038)	(0.859,1.040)
Average liking of alters	1.054	1.053	1.03	1.036	1.034
	(0.947,1.173)	(0.946,1.172)	(0.926,1.147)	(0.931,1.152)	(0.929,1.151)
Socio-demographic covariates					
Male (Ref: Female)	1.167*	1.166*	1.149	1.147	1.147
	(1.008,1.352)	(1.007,1.350)	(0.995,1.328)	(0.992,1.327)	(0.992,1.327)
Age	1.012***	1.012***	1.01	1.000	1.000
	(1.007,1.018)	(1.007,1.018)	(1.000,1.011)	(0.994,1.006)	(0.994,1.006)
Race (Ref: White)					
Black	0.951	0.954	0.957	1.031	1.031
	(0.657,1.376)	(0.659,1.380)	(0.666,1.374)	(0.718,1.480)	(0.718,1.479)
Asian	2.520*	2.502*	2.379*	2.290*	2.292*
	(1.142,5.562)	(1.133,5.522)	(1.085,5.215)	(1.044,5.023)	(1.044,5.029)
Other race	1.139	1.136	1.144	1.129	1.127
	(0.620,2.093)	(0.618,2.087)	(0.629,2.082)	(0.622,2.049)	(0.621,2.047)
Multiracial	1.4	1.4	1.580*	1.656**	1.656**
	(0.978,2.004)	(0.978,2.004)	(1.110,2.250)	(1.164,2.355)	(1.164,2.355)
Hispanic	1.345	1.355	1.281	1.293	1.294
	(0.959,1.887)	(0.966,1.900)	(0.918,1.787)	(0.928,1.801)	(0.929,1.803)
N (observations)	21525	21525	21525	21525	21525
N groups (egos)	9056	9056	9056	9056	9056
AIC	24411.3	24408.4	24104.7	24073.7	24078.35
BIC	24586.8	24591.8	24296.1	24416.7	24453.27

Note: * p<0.05; ** p<0.01; *** p<0.001

Note: Model 1 adjusts for covariates shown, as well as region (categorical) , marital status (categorical), and time (continuous).

Note: Model 2 adds education assortativity measure to prior model, and a random coefficient for education assortativity; independent covariance structure specified.

Note: Model 3 adds subjective social status measure to prior model. (This model is arguably the best-fitting, with the lowest BIC across the four health models.

Note: Model 4 adds additional SES measures to prior model: household asset tiers (categorical), income tiers (categorical), employment status (categorical).

Note: Model 5 adds interaction between education categories and education assortativity.

Sensitivity Table 1b. BMI (random-coefficient multilevel models)

	(1) Base coeff.	(2) +Ed Assort coeff.	(3) +SSS coeff.	(4) +Inc, Wealth coeff.	(5) +Ed x EdAssort coeff.
Key independent variables					
Educational Attainment (Ref: College)					
<HS	0.252*	0.254*	0.230*	0.128	0.237
	(0.044,0.460)	(0.048,0.461)	(0.023,0.437)	(-0.082,0.339)	(-0.033,0.507)
HS	0.146*	0.144*	0.137	0.061	0.229*
	(0.008,0.284)	(0.007,0.281)	(-0.000,0.274)	(-0.079,0.201)	(0.045,0.414)
Some college	0.125**	0.128**	0.124**	0.088	0.124
	(0.037,0.214)	(0.040,0.216)	(0.036,0.212)	(-0.001,0.177)	(-0.004,0.252)
Postgraduate	-0.824***	-0.825***	-0.775***	-0.773***	-0.703***
	(-1.071,-0.577)	(-1.071,-0.578)	(-1.020,-0.530)	(-1.019,-0.528)	(-0.967,-0.439)
Education Assortativity	-	0.012	0.012	0.009	-0.303
		(-0.140,0.164)	(-0.141,0.165)	(-0.143,0.162)	(-0.702,0.095)
Interactions (Ref: College)					
Education Assortativity x <HS	-	-	-	-	0.421
					(-0.286,1.128)
Education Assortativity x HS	-	-	-	-	0.691**
					(0.182,1.199)
Education Assortativity x Some college	-	-	-	-	0.173
					(-0.245,0.592)
Education Assortativity x Postgraduate	-	-	-	-	0.304
					(-0.151,0.760)
Subjective social status	-	-	-0.153***	-0.149***	-0.149***
			(-0.182,-0.124)	(-0.178,-0.120)	(-0.178,-0.120)
Network covariates					
Number of alters	-0.009	-0.009	-0.009	-0.009	-0.009
	(-0.031,0.014)	(-0.033,0.014)	(-0.032,0.014)	(-0.033,0.014)	(-0.033,0.014)
Personal network density	0.228*	0.221*	0.204	0.213	0.210
	(0.008,0.448)	(0.000,0.441)	(-0.016,0.425)	(-0.008,0.433)	(-0.010,0.431)
Average closesness to alters	-0.029	-0.028	-0.02	-0.018	-0.017
	(-0.085,0.026)	(-0.084,0.027)	(-0.075,0.036)	(-0.074,0.037)	(-0.072,0.039)
Average liking of alters	0.001	-0.001	0.000	-0.001	-0.002
	(-0.060,0.062)	(-0.062,0.060)	(-0.061,0.062)	(-0.062,0.060)	(-0.064,0.059)
Socio-demographic covariates					
Male (Ref: Female)	-1.003***	-1.004***	-0.996***	-0.986***	-0.987***
	(-1.233,-0.774)	(-1.234,-0.774)	(-1.224,-0.767)	(-1.214,-0.758)	(-1.215,-0.758)
Age	-0.005	-0.005	0.00	0	0
	(-0.014,0.003)	(-0.014,0.003)	(-0.011,0.006)	(-0.008,0.009)	(-0.008,0.009)
Race (Ref: White)					
Black	1.176***	1.170***	1.159***	1.119***	1.119***
	(0.596,1.756)	(0.589,1.750)	(0.583,1.735)	(0.544,1.694)	(0.544,1.694)
Asian	-3.698***	-3.704***	-3.679***	-3.651***	-3.648***
	(-4.898,-2.499)	(-4.903,-2.504)	(-4.870,-2.488)	(-4.840,-2.463)	(-4.836,-2.460)
Other race	0.863	0.869	0.859	0.859	0.857
	(-0.094,1.821)	(-0.088,1.826)	(-0.091,1.810)	(-0.089,1.807)	(-0.091,1.805)
Multiracial	1.192***	1.195***	1.144***	1.117***	1.116***
	(0.623,1.762)	(0.625,1.764)	(0.578,1.709)	(0.553,1.681)	(0.552,1.680)
Hispanic	0.378	0.373	0.395	0.393	0.393
	(-0.149,0.906)	(-0.155,0.900)	(-0.129,0.919)	(-0.130,0.916)	(-0.130,0.915)
N (observations)	21199	21199	21199	21199	21199
N groups (egos)	8967	8967	8967	8967	8967
AIC	105865.7	105780	105676.9	105663	105662.5
BIC	106048.8	105979.1	105883.9	106021.3	106052.6

Note: * p<0.05; ** p<0.01; *** p<0.001

Note: Model 1 adjusts for covariates shown, as well as region (categorical) , marital status (categorical), and time (continuous).

Note: Model 2 adds education assortativity measure to prior model, and a random coefficient for education assortativity; independent covariance structure specified.

Note: Model 3 adds subjective social status measure to prior model. (This model is arguably the best-fitting, with the lowest BIC across the four health models.

Note: Model 4 adds additional SES measures to prior model: household asset tiers (categorical), income tiers (categorical), employment status (categorical).

Note: Model 5 adds interaction between education categories and education assortativity.

Sensitivity Table 1c. Excellent Self-reported Physical Health (MLM)

	(1) Base OR	(2) +Ed Assort OR	(3) +SSS OR	(4) +Inc, Wealth OR	(5) +Ed x EdAssort OR
Key independent variables					
Educational Attainment (Ref: College)					
<HS	0.403***	0.405***	0.501**	0.607*	0.377**
	(0.247,0.658)	(0.248,0.661)	(0.314,0.798)	(0.378,0.973)	(0.191,0.745)
HS	0.639**	0.640**	0.718*	0.821	0.732
	(0.462,0.883)	(0.463,0.884)	(0.524,0.985)	(0.595,1.133)	(0.478,1.121)
Some college	0.914	0.915	0.962	1.031	0.816
	(0.720,1.159)	(0.721,1.160)	(0.762,1.216)	(0.814,1.305)	(0.590,1.130)
Postgraduate	2.415***	2.408***	1.775***	1.721***	1.355
	(1.796,3.247)	(1.790,3.237)	(1.342,2.349)	(1.299,2.280)	(0.956,1.920)
Education Assortativity	-	1.125	1.068	1.053	2.903*
		(0.819,1.544)	(0.783,1.456)	(0.772,1.437)	(1.146,7.357)
Interactions (Ref: College)					
Education Assortativity x <HS	-		-	-	0.128*
					(0.019,0.877)
Education Assortativity x HS	-		-	-	0.529
					(0.151,1.860)
Education Assortativity x Some college	-		-	-	0.324*
					(0.112,0.941)
Education Assortativity x Postgraduate	-		-	-	0.310*
					(0.112,0.862)
Subjective social status	-	-	2.892***	2.819***	2.821***
			(2.672,3.131)	(2.604,3.053)	(2.605,3.055)
Network covariates					
Number of alters	1.015	1.011	0.987	0.988	0.987
	(0.960,1.073)	(0.956,1.070)	(0.934,1.043)	(0.935,1.044)	(0.934,1.043)
Personal network density	1.09	1.086	0.908	0.927	0.925
	(0.638,1.862)	(0.636,1.856)	(0.538,1.533)	(0.548,1.566)	(0.547,1.563)
Average closesness to alters	0.946	0.946	0.872*	0.866*	0.865*
	(0.827,1.082)	(0.827,1.082)	(0.762,0.998)	(0.757,0.991)	(0.756,0.990)
Average liking of alters	1.335***	1.335***	1.258**	1.268**	1.274**
	(1.146,1.556)	(1.146,1.556)	(1.081,1.464)	(1.089,1.477)	(1.094,1.483)
Socio-demographic covariates					
Male (Ref: Female)	1.652***	1.651***	1.506***	1.511***	1.511***
	(1.345,2.028)	(1.344,2.027)	(1.240,1.830)	(1.241,1.840)	(1.241,1.841)
Age	0.998	0.998	0.978***	0.977***	0.977***
	(0.990,1.005)	(0.990,1.005)	(0.971,0.985)	(0.969,0.985)	(0.969,0.985)
Race (Ref: White)					
Black	0.330***	0.330***	0.338***	0.369***	0.368***
	(0.189,0.576)	(0.189,0.577)	(0.199,0.576)	(0.217,0.628)	(0.216,0.627)
Asian	2.466	2.455	1.909	1.788	1.795
	(0.869,6.996)	(0.865,6.966)	(0.728,5.006)	(0.682,4.687)	(0.684,4.711)
Other race	0.555	0.555	0.598	0.594	0.597
	(0.231,1.337)	(0.230,1.335)	(0.262,1.366)	(0.260,1.356)	(0.261,1.365)
Multiracial	0.337***	0.337***	0.449**	0.472**	0.472**
	(0.196,0.579)	(0.196,0.579)	(0.268,0.753)	(0.281,0.792)	(0.281,0.792)
Hispanic	1.136	1.139	1.001	1.021	1.016
	(0.711,1.814)	(0.713,1.819)	(0.642,1.560)	(0.654,1.593)	(0.651,1.585)
N (observations)	21701	21701	21701	21701	21701
N groups (egos)	9191	9191	9191	9191	9191
AIC	17159.1	17160.7	16330.6	16300.4	16301.1
BIC	17334.8	17344.3	16522.3	16643.8	16676.4

Note: * p<0.05; ** p<0.01; *** p<0.001

Note: Model 1 adjusts for covariates shown, as well as region (categorical) , marital status (categorical), and time (continuous).

Note: Model 2 adds education assortativity measure to prior model, and a random coefficient for education assortativity; independent covariance structure specified.

Note: Model 3 adds subjective social status measure to prior model. (This model is arguably the best-fitting, with the lowest BIC across the four health models.

Note: Model 4 adds additional SES measures to prior model: household asset tiers (categorical), income tiers (categorical), employment status (categorical).

Note: Model 5 adds interaction between education categories and education assortativity.

Sensitivity Table 1d. Excellent Self-reported Mental Health (MLM)

	(1) Base OR	(2) +Ed Assort OR	(3) +SSS OR	(4) +Inc, Wealth OR	(5) +Ed x EdAssort OR
Key independent variables					
Educational Attainment (Ref: College)					
<HS	0.764	0.767	1.049	1.159	1.218
	(0.552,1.056)	(0.555,1.061)	(0.768,1.433)	(0.843,1.594)	(0.795,1.866)
HS	0.847	0.85	1.046	1.131	1.077
	(0.671,1.069)	(0.674,1.073)	(0.834,1.312)	(0.899,1.425)	(0.792,1.464)
Some college	0.954	0.955	1.016	1.059	1.184
	(0.801,1.135)	(0.802,1.137)	(0.854,1.209)	(0.888,1.263)	(0.928,1.511)
Postgraduate	1.269*	1.262*	0.963	0.978	0.953
	(1.021,1.577)	(1.015,1.569)	(0.785,1.181)	(0.796,1.201)	(0.734,1.238)
Education Assortativity	-	1.241	1.155	1.152	1.096
		(0.981,1.570)	(0.918,1.452)	(0.916,1.450)	(0.548,2.192)
Interactions (Ref: College)					
Education Assortativity x <HS	-	-	-	-	1.24
					(0.375,4.102)
Education Assortativity x HS	-	-	-	-	0.831
					(0.341,2.028)
Education Assortativity x Some college	-	-	-	-	1.643
					(0.750,3.597)
Education Assortativity x Postgraduate	-	-	-	-	0.883
					(0.410,1.902)
Subjective social status	-	-	2.997***	2.959***	2.959***
			(2.837,3.165)	(2.800,3.126)	(2.800,3.127)
Network covariates					
Number of alters	1.111***	1.103***	1.065**	1.065**	1.064**
	(1.067,1.157)	(1.059,1.150)	(1.023,1.109)	(1.023,1.109)	(1.022,1.108)
Personal network density	1.398	1.39	1.332	1.32	1.331
	(0.945,2.069)	(0.939,2.057)	(0.908,1.954)	(0.899,1.938)	(0.906,1.955)
Average closesness to alters	1.150**	1.149**	1.021	1.022	1.021
	(1.041,1.270)	(1.040,1.269)	(0.925,1.128)	(0.925,1.128)	(0.924,1.127)
Average liking of alters	1.305***	1.305***	1.242***	1.246***	1.245***
	(1.167,1.459)	(1.167,1.459)	(1.111,1.388)	(1.114,1.393)	(1.113,1.393)
Socio-demographic covariates					
Male (Ref: Female)	0.712***	0.712***	0.669***	0.667***	0.668***
	(0.611,0.830)	(0.611,0.830)	(0.582,0.768)	(0.579,0.768)	(0.580,0.769)
Age	1.043***	1.043***	1.020***	1.017***	1.018***
	(1.037,1.049)	(1.037,1.049)	(1.015,1.025)	(1.011,1.023)	(1.012,1.024)
Race (Ref: White)					
Black	1.163	1.168	1.229	1.261	1.259
	(0.789,1.713)	(0.793,1.720)	(0.870,1.737)	(0.891,1.784)	(0.889,1.782)
Asian	1.133	1.126	0.945	0.925	0.92
	(0.507,2.534)	(0.503,2.518)	(0.459,1.946)	(0.448,1.909)	(0.446,1.901)
Other race	0.872	0.871	0.881	0.885	0.886
	(0.465,1.635)	(0.465,1.633)	(0.498,1.558)	(0.500,1.568)	(0.500,1.571)
Multiracial	1.1	1.1	1.477*	1.507*	1.504*
	(0.757,1.599)	(0.757,1.598)	(1.051,2.076)	(1.071,2.121)	(1.069,2.118)
Hispanic	1.338	1.344	1.138	1.137	1.138
	(0.942,1.899)	(0.947,1.908)	(0.830,1.560)	(0.828,1.561)	(0.829,1.562)
N (observations)	21,600	21,600	21,600	21600	21600
N groups (egos)	9065	9065	9065	9065	9065
AIC	24436.4	24435.3	22243.1	22258.3	22260.2
BIC	24612.0	24618.8	22434.6	22601.4	22635.3

Note: * p<0.05; ** p<0.01; *** p<0.001

Note: Model 1 adjusts for covariates shown, as well as region (categorical) , marital status (categorical), and time (continuous).

Note: Model 2 adds education assortativity measure to prior model, and a random coefficient for education assortativity; independent covariance structure specified.

Note: Model 3 adds subjective social status measure to prior model. (This model is arguably the best-fitting, with the lowest BIC across the four health models.

Note: Model 4 adds additional SES measures to prior model: household asset tiers (categorical), income tiers (categorical), employment status (categorical).

Note: Model 5 adds interaction between education categories and education assortativity.

Sensitivity Analysis: Comparison of Education Assortativity & Education Distance

In their 2014 study, Smith and colleagues found *greater education homophily among less-educated individuals* in the GSS using both continuous and categorical versions of an education distance measure. As a sensitivity analysis, we replicated Smith and colleagues' categorical variant of educational distance using the Gallup data (here, educational attainment was not measured in continuous years for egos, as it was in the GSS). Following Smith et al. (2014) we measured social distance by education between ego and her alters (i.e., education network diversity) as the average (absolute) difference in education (measured categorically) between ego and her alters.

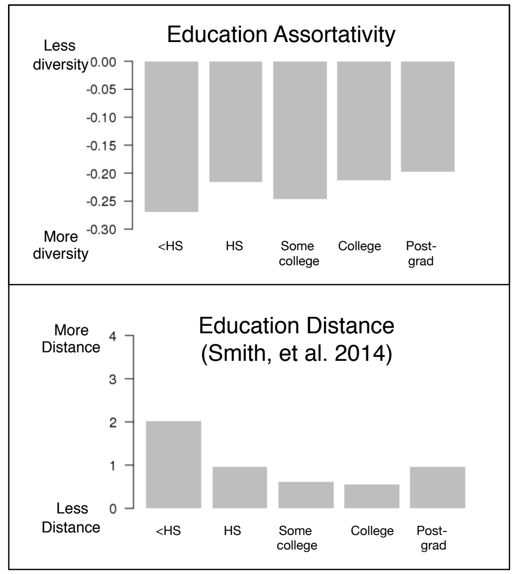


Figure S1. Distribution of network diversity measures by education categories.

Fig S1 at right shows how, in the Gallup sample, *education assortativity* and *education distance* share some degree of consistency at their extremes in terms of how they vary by ego education. More specifically, low-education egos have greater network diversity (as well as more dissimilar levels of education among alters), compared with high-education individuals, who have less diverse networks (and alters with more similar education levels). We note that this is somewhat different than Smith, et al. (2014) insofar as here, less-educated individuals appear to have *less* education homophily.

Next, we examined how educational attainment and network characteristics are associated with both education assortativity and education distance in a multi-level model that adjusts for the full range of socio-demographic characteristics (Sensitivity Table 2, below). We find that less-educated individuals (<HS) and the most-educated individuals (postgraduate) have the most education distance (i.e. the least education homophily). Though a more thorough comparison and interpretation of measures is beyond the scope of this study, future work may productively interrogate this direction.

References

Smith, J. A., McPherson, M., & Smith-Lovin, L. (2014). Social distance in the United States: Sex, race, religion, age, and education homophily among confidants, 1985 to 2004. *American Sociological Review*, *79*(3), 432-456.

Sensitivity Table 2. Comparative associations with education diversity measures

	Education Assortativity	Education Distance
	Years 1-3	Years 1-3
	MLM	MLM
	coeff.	coeff.
Key independent variables		
Educational Attainment (Ref: College)		
<HS	-0.012	0.215***
	(-0.030,0.006)	(0.174,0.256)
HS	-0.015*	0.028
	(-0.028,-0.001)	(-0.001,0.057)
Some college	-0.006	-0.017
	(-0.017,0.004)	(-0.037,0.003)
Postgraduate	0.023***	0.339***
	(0.011,0.035)	(0.310,0.368)
Network covariates		
Number of alters	0.033***	-0.008**
	(0.031,0.035)	(-0.013,-0.003)
Personal network density	0.040***	0.092***
	(0.018,0.062)	(0.045,0.139)
Average closesness to alters	0.005	0.019**
	(-0.000,0.011)	(0.007,0.031)
Average liking of alters	0.000	-0.013
	(-0.007,0.006)	(-0.026,0.000)
Socio-demographic covariates		
Male (Ref: Female)	0.005	-0.026*
	(-0.003,0.013)	(-0.048,-0.003)
Age	-0.001***	0.001*
	(-0.001,-0.000)	(0.000,0.002)
Race (Ref: White)		
Black	-0.014	0.046
	(-0.033,0.006)	(-0.010,0.101)
Asian	0.022	-0.056
	(-0.018,0.062)	(-0.171,0.059)
Other race	0.008	0.06
	(-0.024,0.039)	(-0.032,0.151)
Multiracial	0.001	0.04
	(-0.018,0.020)	(-0.014,0.094)
Hispanic	-0.023**	0.085***
	(-0.041,-0.006)	(0.035,0.136)
Subjective social status	0.001	-0.007*
	(-0.001,0.003)	(-0.012,-0.001)
N (observations)	21795	21795
N groups (egos)	9090	9090
AIC	-4418.1	30140.5
BIC	-4074.6	30484.0

Note: * p<0.05; ** p<0.01; *** p<0.001
Note: both models adjust for covariates shown here, as well as categorical measures for region, marital status, household asset tiers, income tiers, and employment status. Multi-level model includes a continuous time measure.

Example Code # 1. R Script to Compute Educational Assortativity

```r
## Is Having an Educationally Diverse Social Network Good for Health?
## Pachucki, Mark C. & Diego F. Leal.
## Code written by Diego F. Leal (www.diegoleal.info)
## Last revision: 10/14/19 by DFL/MCP
## Purpose: This is a simplified version of the code to clean wave 1 Gallup data and
## to compute (1) educational assortativity and (2) education distance both based on ego net data
## For access to the full script, which includes other measures and robustness checks, please contact the authors.


## clear all
rm(list=ls())

library(reshape)
library(reshape2)
library(igraph)
library(readstata13)
library(foreign)
library("RColorBrewer")

# ## info for full replicability
# R version 3.3.2 (2016-10-31)
# Platform: x86_64-redhat-linux-gnu (64-bit)
# Running under: Red Hat Enterprise Linux Server release 6.8 (Santiago)
#
# locale:
# [1] LC_CTYPE=en_US.UTF-8 LC_NUMERIC=C LC_TIME=en_US.UTF-8 LC_COLLATE=en_US.UTF-8 LC_MONETARY=en_US.UTF-8 LC_MESSAGES=en_US.UTF-8 LC_PAPER=en_US.UTF-8
# [8] LC_NAME=C LC_ADDRESS=C LC_TELEPHONE=C LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
#
# attached base packages:
# [1] stats graphics grDevices utils datasets methods base
#
# other attached packages:
# [1] RColorBrewer_1.1-2 brew_1.0-6 readstata13_0.9.0 igraph_1.1.2 reshape2_1.4.2 reshape_0.8.7 bindrcpp_0.2 dplyr_0.7.3 plyr_1.8.4
#
# loaded via a namespace (and not attached):
# [1] Rcpp_0.12.12 assertthat_0.2.0 R6_2.2.2 magrittr_1.5 rlang_0.1.2 stringi_1.1.5 tools_3.3.2 stringr_1.2.0 glue_1.1.1 pkgconfig_2.0.1 bindr_0.1
# [12] tibble_1.3.4
########################### DATA PREPARATION ####################################

## Recording initiation time
g.time <-Sys.time()

#Import data
setwd("/RELEVANT PATH HERE")
```

```
covars_w1<-read.dta13("/RELEVANT PATH HERE")
covars_w1$RESPONDENT_ID<-as.character(covars_w1$RESPONDENT_ID)

## generate a covars list for W1
covars_trunc <- covars_w1
rownames(data) <- NULL

#This observation gives R a crash, but just in Wave 2 for some reason - drop from all waves' analyses.
(covars_trunc <- covars_trunc[covars_trunc$EMPLOYEE_KEY_VALUE != "4083317012_78914_01",
])

## list of all variables names in the full data set
allColLabels<-colnames(covars_trunc[,])

## locate position of the the variable "Q18_1_1",
## the first variable (from left to right) where ego reports an alter's name

for (i in 1:length(allColLabels))
{
if (allColLabels[i]=="Q18_1_1")
{
firstAlter<-i
}
}

## replace empty strings " " in alters' names by NAs
for (i in (firstAlter):(firstAlter+7))
{
for (j in 1:nrow(covars_trunc))
{
if (covars_trunc[j,i] == " ")
{
covars_trunc[j,i] <-NA
}
}
}


#restricted data set: Ego ID and Alters names, "covars_trunc" is the full data set
mat <-covars_trunc[,c("Q18_1_1", "Q18_1_2",
"Q18_1_3", "Q18_1_4", "Q18_1_5", "Q18_1_6",
"Q18_1_7", "Q18_1_8")]

#replace anything different from 0 for a 1 (0 represents an absent alter)
mat <-ifelse(is.na(mat[,]),NA,1)

#sum alters to get the size of each egonet
alters <-(rowSums(mat[],na.rm=T))
alters <-as.data.frame(alters)
mat <-cbind(covars_trunc$RESPONDENT_ID,mat,alters) ## covars_trunc is the full data set

## rename "RESPONDENT_ID" column
names(mat)[names(mat) == "covars_trunc$RESPONDENT_ID"] <- "RESPONDENT_ID"
```

```r
## merge full data set with the count of alters of each ego
mat2 <-as.data.frame(mat$RESPONDENT_ID)
mat3 <-as.data.frame(mat$alters)
mat2 <-cbind(mat2,mat3)
colnames(mat2) <-c("RESPONDENT_ID","alters")
covars_trunc <-merge(covars_trunc,mat2,by="RESPONDENT_ID",sort=F)
## list of all variables names in the full data set
allColLabels<-colnames(covars_trunc[,])
## locate position of the the variable "Q31A_YR1", "how much do you like alter A"
## the first variable (from left to right) where ego reports an alter's name
for (i in 1:length(allColLabels))
{
if (allColLabels[i]=="Q31A_YR1")
{
firstAlter<-i
}
}
## replace empty strings " " by NAs
for (i in (firstAlter):(firstAlter+7))
{
for (j in 1:nrow(covars_trunc))
{
if (covars_trunc[j,i] == " ")
{
covars_trunc[j,i] <-NA
}
}
}
#restricted data set: Ego ID and Alters names, "covars_trunc" is the full data set
mat <-covars_trunc[,c("Q31A_YR1", "Q31B_YR1",
"Q31C_YR1", "Q31D_YR1",
"Q31E_YR1", "Q31F_YR1",
"Q31G_YR1", "Q31H_YR1")]
#make sure mat is a data.frame object
mat<-as.data.frame(mat)
avg.like <-(rowMeans(mat[],na.rm=T))
avg.like <-as.data.frame(avg.like)
mat <-cbind(covars_trunc$RESPONDENT_ID,mat,avg.like) ## covars_trunc is the full data set
names(mat)[names(mat) == "covars_trunc$RESPONDENT_ID"] <- "RESPONDENT_ID"
mat2 <-as.data.frame(mat$RESPONDENT_ID)
mat3 <-as.data.frame(mat$avg.like)
mat2 <-cbind(mat2,mat3)
```

```
colnames(mat2) <-c("RESPONDENT_ID","avg_like")
covars_trunc <-merge(covars_trunc,mat2,by="RESPONDENT_ID",sort=F)

############## MEASURE 1: EDUCATION ASSORTATIVITY #######################
## Recording initiation time
e.time <-Sys.time()

## create (super)lists to store all egonets in their different forms

allEgoNetsFullEdu <-vector("list",nrow(mat)) ## egonets in matrix format w/relationship type & including NAs
allEgoNetsEdu <-vector("list",nrow(mat)) ## egonets in matrix format w/relationship type NOT including NAs
allEgoNetsBinaryEdu <-vector("list",nrow(mat)) ## egonets in matrix format, relationship type is binarized
allEgoNetsIgraphEdu <-vector("list",nrow(mat)) ## allEgoNetsBinary in igraph format

## list of all variables' names in the full data set, find the position of  variables "Q32A" and "alter_ed_5cat_1",
## relationship type between ego and alter and educational attainment of alter, respectively
allColLabels<-colnames(covars_trunc[,])

for (i in 1:length(allColLabels))
{
if (allColLabels[i]=="Q32A")
{
colNumber<-i
}
if (allColLabels[i]=="alter_ed_5cat_1")
{
colNumber3<-i
}
}

name_no_education <-as.data.frame(matrix(ncol=3,nrow=(nrow(covars_trunc)*8))) ## create a df to store alters with names and with no education info
colnames(name_no_education) <-c("RESPONDENT_ID","ALTER_NAME","ALTER_#")

########## recovering missing values in the education variable

count <- 0
for(a in 1:nrow(covars_trunc)) ## for each ego
{
for (b in 1:8) ## for each of egos's alters
{
count <- count + 1
string<-(covars_trunc[a,firstAlter + b - 1]) ## get the name of the bth alter
if (is.na(covars_trunc[a,firstAlter + b - 1]) == FALSE) ## if alter does have a name
{
if (((nchar(covars_trunc[a,firstAlter + b - 1]) >= 2)) & (is.na(covars_trunc[a,colNumber3 + b - 1]))) ## if alters' name is a string of at least two characters
{
name_no_education[count,1]<-covars_trunc[a,"RESPONDENT_ID"] ## store egos's ID in the name_no_education object
```

```r
name_no_education[count,2]<-covars_trunc[a, firstAlter + b -1] ## store alters' name in the name_no_education object
name_no_education[count,3]<-colnames(covars_trunc)[firstAlter + b - 1] ## store alters' number (e.e. alter_2 or alter_3) in the name_no_education object
}
}
}
}
name_no_education<-name_no_education[complete.cases(name_no_education),] ## get rid of NAs
#see the resulting data
View(covars_trunc)
#creste variables to store missing data
covars_trunc$missing_education <-NA
covars_trunc$missing_rel_type_edu <-NA
## create egonets based on the right number of dimensions
for (j in 1:nrow(mat))
{
egoNetName <-as.character(mat[j,1]) # store ego's ID ("RESPONDENT_ID")
egoNetDim <-max(mat[,"alters"]) + 1 # store the egonets dims (all are 9 by 9 matrices)
egoNet <-matrix(nrow=egoNetDim,ncol=egoNetDim) # create the matrix object
labels <-c(egoNetName,colnames(mat)[2:egoNetDim]) # label the matrix first row and first columns are the ego's ID ("RESPONDENT_ID")
colnames(egoNet) <-labels
rownames(egoNet) <-labels
#star-like egonet (i.e. everyone is connected to ego, no connections between alters)
egoNet <-as.matrix(egoNet) #save egonet as a matrix object
egoNet[,] <-9999 #all cells == 9999
egoNet[,1] <-10 #all cells in first column = 10
egoNet[1,] <-10 #all cells in first row = 10
diag(egoNet) <-10 #all cells in main diagonal = 10
egoNet[upper.tri(egoNet)] <- 0 #make cells in the upper triangle = 0
allEgoNetsEdu[[j]]<-egoNet #store the egonet in allEgoNets
}
################## M2.Education assortativity MASTER LOOP###################
# This loop populates the egonets based on the info of covars_trunc and calculates assortativity coeffcients based on education
```

```
for (j in 1:nrow(covars_trunc))
{
X<-allEgoNetsEdu[[j]] ## select the jth egonet in allEgoNetsEdu, that is, select the egonet of the ego in row j in the main data set (covars_trunc)
Y<-melt(X) ## transform the egonet from matrix format to edgelist format
colnames(Y) <- c("X1","X2","value") ## rename columns of the edgelist. "value" = relationship type between alters
count<-0
for (i in 1:(9*9)) ## at the beginning, all sociomatrices are 9 by 9 matrices. Equivalently, all edgelists have 81 (9 * 9) rows
{
if(Y[i,3]==9999) ## using the cells in the lower triangle only:
{
Y[i,3]<-covars_trunc[j,colNumber+count] ## replace the 9999 in the ith row of edgelist with value (i.e.relationship type) reported by ego between a given pair of his/her alters
count <- count + 1 ## go to next alter
}
}
Y<-acast(Y, X1~X2,value="value") ## from edgelist to sociomatrix format
Y[upper.tri(Y)]<- t(Y)[upper.tri(Y)] ## symmetrize the egonet
eduEgo <-covars_trunc$DEMO_EDUC_5CAT ## store all egos' education in the object educationEgo
eduEgo <-eduEgo[j] ## store the jth ego's education in the object educationEgo
eduAlters <-covars_trunc[j,(colNumber3):(colNumber3+7)] ## retrieve the education of ego's alters from the main data set (covars_trunc)
education <-cbind(eduEgo,eduAlters) ## bind ego's education and alters education
education <-t(education) ## transpose education to make it a vertical vector
colnames(education) <-covars_trunc[j,"RESPONDENT_ID"] ## rename "education" object with ego's unique ID
Y <-cbind(Y,education) ## bind the education vector to the egonet
#### This section deals with missing values in the relationship type between alters
### the code assumes that if there is an NA is because one of the alters did not exist
U<-as.data.frame(matrix(ncol=1,nrow=8)) ## create an empty vector to store inexistent alters
V<-as.data.frame(matrix(ncol=1,nrow=8)) ## create an empty vector to store alters with missing info in their education identity
S<-as.data.frame(matrix(ncol=1,nrow=8)) ## create an empty vector to store alters with missing info in their (RELATIONSHIP TYPE?)
for (q in 2:10) ## for each alter (i.e. for each columns in Y)
{
if(q<=9) ## alters are in columns 2 to 9
{
I<-unlist(Y[,q]) ## list all values of column q (i.e. the relationship type between alter q and all other alters)
```

```
if(sum(is.na(I))>=7) ## if alter q has no ties to other alters (i.e. if alter q has 7 NAs)
{
U[q-1,1]<-q ## store the number q in the object U
}
}
if(q==10) ## go to the education identity info of the alters
{
I<-unlist(Y[,q]) ## list of values, that is, all education identities of alters
for (g in 2:length(I)) ## go through the list of alters' education identities
{
if (is.na(I[g])) ## if a given alter has a missing value, store its position (i.e.,its row number) in the ego network
{
V[g-1,1]<-g ## store that info in the object V
}
}
}
}
for (q in 2:9) ## for each alter (i.e. for each columns in Y)
{
I<-unlist(Y[,q]) ## list all values of column q (i.e. the relationship type between alter q and all other alters)
if ((sum(is.na(I)) < 7) & (sum(is.na(I)) >0)) ## if alter q is indeed present in the egonet (if it has at least one relationship with another alter in the egonet)
{
if (sum(!(is.na(I))) + length(U) == 9) ## if alter q is indeed present in the ego net, its relationships to alter + the number of "fully" missing alters (alters that ego do not report at all) must be = 7
{
S[q-1,1]<-q ## store the number q in the object U. If q >= 1, it means that some alters reported by ego have a NA in their relationship with other alters
}
}
}
########### M2.CREATE LISTs W/ ISOLATED ALTERS (U) OR ALTERS W/NO EDUCATION INFO (V) #########
## V and U are then merged and the final set of alters to calculate EDUCATION assortativity (W) excludes the alters in U or V
U<-U[complete.cases(U),] ## get rid of NAs in object U (keep alters with no conection to any other alter)
U<-unlist(U) ## make U a list
V<-V[complete.cases(V),] ## get rid of NAs in object V (keep alters with no education info)
V<-unlist(V) ## make V a list
U<-c(U,V) ## concatenate U and V
U<-unique(U) ## keep unique element of U (i.e., keep column with missing alters or keep rows with alters with missing education info)
S<-S[complete.cases(S),] ## get rid of NAs in object V (keep alters with no education info)
S<-unlist(S) ## make S a list
covars_trunc[j,"missing_rel_type_edu"] <- length(S)
if (length(U)<8) ## if there is at least one alter with education info:
{
```

```
ifelse (length(U)>=1,W<-Y[-c(U),-c(U)],W<-Y) ## delete columns and rows with misssing info


#######M2.CREATE EGONETS BASED ON ALTERS WITH FULL EDUCATION & RELATIONAL INFO #########

attributes <-W[,(ncol(W))] ## create an node-attributes data set with the education info of the nodes
attributes <-as.data.frame(attributes) ## save the attributes object as data frame
colnames(attributes) <-"education" ## rename the column with education info with the label "education"
Y <-Y[,-(ncol(Y))] ## make Y the egonet with NAs
W <-W[,-(ncol(W))] ## make W the egonet with no NAs
Z <-ifelse(W[,]>1,1,0) ## make Z the binarized version of W
ZZ<-Z
ZZ
W[,1] <-1 ## populate the first column with 1s (ego's degree)
W[1,] <-1 ## populate the first row with 1s (ego's degree)
diag(W) <-0 ## populate main diagonal with 0s
diag(Z) <-0 ## populate main diagonal with 0s
Y[,1] <-1 ## technically, we should replace 1s w/numbers that represent relation type between ego & alters
Y[1,] <-1 ## technically, we should replace 1s by numbers that represent relation type between ego & alters
diag(Y) <-0 ## populate main diagonal with 0s

net<-graph_from_adjacency_matrix(Z, mode = "undirected") %>% ## creating the igraph object based on Z (binarized "egonet")
set_vertex_attr("education", value = attributes$education) ## setting education attribute


educationAssort <-assortativity_nominal(net, as.numeric(as.factor(V(net)$education)), directed = F) ## calculate education assortativity
covars_trunc[j,"educationAssortativity"] <-educationAssort ## store egos's education assortativity in the main data set (covars_trunc)
covars_trunc[j,"missing_education"] <-covars_trunc[j,"alters"] - (nrow(Z) - 1)

allEgoNetsFullEdu[[j]] <-Y ##store the full egonet (egonet with NAs) in the allEgoNetsFull list
allEgoNetsEdu[[j]] <-W ##store the egonet (egonet with NO NAs) in the allEgoNets list
allEgoNetsBinaryEdu[[j]] <-Z ##store the binarized egonet in the allEgoNetsBinary list
allEgoNetsIgraphEdu[[j]] <-net ##store the egonet as an igraph object in the allEgoNetsIgraph list
}
if (length(U)==8)
{
covars_trunc[j,"educationAssortativity"] <-"no_info"
covars_trunc[j,"missing_education"] <-length(U)
}
}


#generate data frame for education inspection
mat6 <-covars_trunc[,c("RESPONDENT_ID" ,"DEMO_EDUC_5CAT" ,"alter_ed_5cat_1","alter_ed_5cat_2","alter_ed_5cat_3",
"alter_ed_5cat_4" ,"alter_ed_5cat_5","alter_ed_5cat_6",
"alter_ed_5cat_7" ,"alter_ed_5cat_8","missing_education",
"missing_rel_type_edu", "educationAssortativity")]

```

```
##### Sensitivity analysis (i.e., computing dyadic edu distance). See the supplementary info of the published paper for more info ##

## R&R.2 (9/17/19). Creating a Distance Measure following Smith, McPherson, and Smith-Lovin's Social Distance in the United States (ASR, 2014)
## Based on Table 1 of that article, ASR 2014 measures social distance as the "absolute education difference between respondent and confidant"
## Here we assume that what they did was to compute the absolute differences between ego and her alters

#extract the labels of the variables in the mat6 object
mat6.labels<-colnames(mat6)

#find the position of the alter_ed_5cat_1 variable (the edu level of the first alter) in mat6
for (i in 1:length(mat6.labels))
{
if (mat6.labels[i]=="alter_ed_5cat_1")
{
mat6.ego1<-i
}
}

#make sure mat6 is a data frame
mat6<-as.data.frame(mat6)

#create a new variable called 'eduDistanceSmith' This variable will contain the edu distance
#following smith et al. Initially, the variable is populated with 99999s
mat6$eduDistanceSmith<-99999

#Loop to compute the absolute value of the average distance between ego and alter
for (iii in 1:nrow(mat6)) #for each ego
{
ego.alter.distance<-(matrix(99999,1,8)) #create a vector to store the dyadic distances
for (jjj in 1:ncol(ego.alter.distance)) #for each possible alter
{
if ((mat6[iii,"educationAssortativity"]!= "no_info") & (mat6[iii,"educationAssortativity"]!= "NaN")) #if ego has at least two alters (this will make the samples between assortativity and dyadic distance comparable)
{
ego.alter.distance[jjj]<-mat6[iii,"DEMO_EDUC_5CAT"] - mat6[iii,mat6.ego1+jjj-1] #subtract egos and alter's level of education
}
if (mat6[iii,"educationAssortativity"]== "no_info") #if ego does not have two or more alters
{
ego.alter.distance[jjj]<-99999 #simply put an 99999 as the distance between ego and each of her alters
}
if (mat6[iii,"educationAssortativity"]== "NaN") #if ego and all her alters have the same level of education (i.e., if there is perfect assortativity)
{
```

```
ego.alter.distance[jjj]<-0 #simply put a 0 as the distance between ego and each of her alters
}
}
mat6[iii,"eduDistanceSmith"]<-abs(rowMeans(ego.alter.distance,na.rm=T)) #compute the average distance between ego and her alters, then take the absolute value of the distance
}


for (iii in 1:nrow(mat6)) #this loops replaces 99999 for "no_info" in the education distance (i.e., egos with less than two alters)
{
if(mat6[iii,"eduDistanceSmith"] == 99999)
{
mat6[iii,"eduDistanceSmith"] <- "no_info"
}
}

#bind the new edu distance variable (eduDistanceSmith) to the main data set (covars_trunc)
covars_trunc<-cbind(covars_trunc,mat6$eduDistanceSmith)

#add a label to the new variable (eduDistanceSmith) in the context of the main data set (covars_trunc)
colnames(covars_trunc)[ncol(covars_trunc)]<-"eduDistanceSmith"

##### end of changes related to R&R.2 (i.e., computing dyadic edu distance) ##############

#count how many "no info" instances there are (answer: 5942). Remember, "no_info" entries mean that ego does not have 2 or more alters with education info.
count_noinfoedu <-as.data.frame(table(covars_trunc$educationAssortativity, useNA = "always"))
sum(count_noinfoedu[, "Freq"][1:nrow(count_noinfoedu)]) ## check that all egos (20366) have a gender_assortativity value

#check how much time did the R script take to run
print(Sys.time() - e.time)


###### M2.TEST ASSORTATIVITY EDUCATION CODE #########

#tabulate distribution
# When this new variable is > 0,it means that, for a given alter,
#the sum of its relationships + the known number of absent alters in the egonet
#(i.e. the # of "structural NAs") reported by ego is different from 7. In other words,
#if a given ego has "missing_rel_type" > 0 that's an indication of inconsistencies in the data.
count_missingreltypeedu<-as.data.frame(table(covars_trunc$missing_rel_type_edu, useNA = "always"))
count_missingeducation<-as.data.frame(table(covars_trunc$missing_education, useNA = "always")) # 8 missing alters = # no info in the count_noinfoedu object = 5942


#### TEST EDUCATION ASSORTATIVITY CODE for ego 20365, (assortativity = -0.2)
#show education info of ego and alters
as.factor(V(net)$education)
#manually check education values
```

```
covars_trunc[20365,"DEMO_EDUC_5CAT"] # 4
covars_trunc[20365,"alter_ed_5cat_1"] # 4
covars_trunc[20365,"alter_ed_5cat_2"] # 3
covars_trunc[20365,"alter_ed_5cat_3"] # 3
covars_trunc[20365,"alter_ed_5cat_4"] # 3
covars_trunc[20365,"alter_ed_5cat_5"] # 3


attributes2edu <-matrix(c(4,4,3,3,3,3),nrow=6,ncol=1)
attributes2edu <-as.data.frame(attributes2edu)
colnames(attributes2edu) <-"education"

net2edu<-graph_from_adjacency_matrix(allEgoNetsBinaryEdu[[20365]], mode = "undirected") %>% ## creating the igraph object based on Z (binarized "egonet")
set_vertex_attr("education", value=attributes2edu$education) ##setting education attribute


assortativity_nominal(net, as.numeric(as.factor(V(net)$education)), directed = F) ## calculate education assortativity

# END TEST ASSORTATIVITY EDUCATION CODE

#save data
save.image("WAVE 1.Rdata")
#export analytic data frame to Stata version #
write.dta(covars_trunc1, WAVE 1.dta')
```

Example Code # 2. Stata Do-file to Compute Multilevel Models

```
// Pachucki, Mark C. & Diego F. Leal
// Is Having an Educationally Diverse Social Network Good for Health?
// Code written by Mark Pachucki
// Multilevel model specifications of data from waves 1-3
// Last revision: 11/17/19 by MCP
//
// R data from Waves 1,2,3 with assortativity vars have been imported from R,
// merged on ID, transformed to long format, w/xtset denoting panel data.

use "/WAVE123_long_20191117.dta", replace

***************************************************************
*0. Table 2 - Education Assortativity as outcome
***************************************************************
//OLS - assortativity as outcome (baseline Wave 1 only)
regress edu_ass_n_0_5 ib4.DEMO_EDUC_5CAT  ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME Q14 ib3.Q15 ///
ib3.demo_region if time==1
```

```
est store edass_rc_OLS_v0
est save edass_rc_OLS_v0, replace
save, replace

//MLM - assortativity as outcome
mixed edu_ass_n_0_5 ib4.DEMO_EDUC_5CAT  ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME Q14 ib3.Q15 ///
ib3.demo_region time ///
|| ID:, covariance(independent)
est store edass_rc_v1
est save edass_rc_v1, replace
save, replace

***********************************************************
*SI. Sensitivity Tables 1a-d, Network ed. assortativity & health (Model 1)
***********************************************************
//A. BMI
quietly mixed BMIcorr ib4.DEMO_EDUC_5CAT ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q14!=. & Q15!=. ///
|| ID: , covariance(independent)
est store bmi_rc_v0
est save bmi_rc_v0, replace
save, replace

//B. Exercise Regularly
quietly melogit exer_bin ib4.DEMO_EDUC_5CAT ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q14!=. & Q15!=. ///
|| ID:, covariance(independent) or noheader
est store exer_rc_v0
est save exer_rc_v0, replace
save, replace

//C. Excellent Self-reported Mental Health
quietly melogit Q10_bin ib4.DEMO_EDUC_5CAT ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q14!=. & Q15!=. ///
|| ID:, covariance(independent) or noheader
est store mh_rc_v0
est save mh_rc_v0, replace
save, replace

//D. Excellent Self-reported Physical Health
quietly melogit Q8_bin ib4.DEMO_EDUC_5CAT ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q14!=. & Q15!=. ///
```

```
|| ID:, covariance(independent) or noheader
est store ph_rc_v0
est save ph_rc_v0, replace
save, replace

************************************************************
*SI. Sensitivity Tables 1a-d, Network ed. assortativity & health (Model 2)
*          This model adds education assortativity
************************************************************

//A. BMI
quietly mixed BMIcorr ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q14!=. & Q15!=. ///
|| ID:edu_ass_n_0_5, covariance(independent)
est store bmi_rc_v1
est save bmi_rc_v1, replace
save, replace

//B. Exercise Regularly
quietly melogit exer_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q14!=. & Q15!=. ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store exer_rc_v1
est save exer_rc_v1, replace
save, replace

//C. Excellent Self-reported Mental Health
quietly melogit Q10_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q14!=. & Q15!=. ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store mh_rc_v1
est save mh_rc_v1, replace
save, replace

//D. Excellent Self-reported Physical Health
quietly melogit Q8_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q14!=. & Q15!=. ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store ph_rc_v1
est save ph_rc_v1, replace
save, replace
```

```
************************************************************
*Table 3. Network education assortativity and health (Model 1)
*Also reported in:
*SI. Sensitivity Tables 1a-d, Network ed. assortativity & health (Model 3)
*           This model adds subjective social status
************************************************************
//A. BMI
quietly mixed BMIcorr ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP Q14 ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=.  & Q15!=. ///
|| ID:edu_ass_n_0_5, covariance(independent)
est store bmi_rc_v1_5
est save bmi_rc_v1_5, replace
save, replace

//B. Exercise Regularly
quietly melogit exer_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP Q14 ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q15!=. ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store exer_rc_v1_5
est save exer_rc_v1_5, replace
save, replace

//C. Excellent Self-reported Mental Health
quietly melogit Q10_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP Q14 ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q15!=. ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store mh_rc_v1_5
est save mh_rc_v1_5, replace
save, replace

//D. Excellent Self-reported Physical Health
quietly melogit Q8_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP Q14 ///
ib3.demo_region time if employ!=. & DEMO_INCOME!=. & Q15!=. ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store ph_rc_v1_5
est save ph_rc_v1_5, replace
save, replace
```

```
***********************************************************
*Table 3. Network education assortativity and health (Model 2)
*Also reported in:
*SI. Sensitivity Tables 1a-d, Network ed. assortativity & health (Model 4)
*           This model adds income, wealth, employment status
***********************************************************

//A. BMI
quietly mixed BMIcorr ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME Q14 ib3.Q15 ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent)
est store bmi_rc_v2
est save bmi_rc_v2, replace
save, replace

//B. Exercise Regularly
quietly melogit exer_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME Q14 ib3.Q15 ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store exer_rc_v2
est save exer_rc_v2, replace
save, replace

//C. Excellent Self-reported Mental Health
quietly melogit Q10_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME Q14 ib3.Q15 ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store mh_rc_v2
est save mh_rc_v2, replace
save, replace

//D. Excellent Self-reported Physical Health
quietly melogit Q8_bin ib4.DEMO_EDUC_5CAT edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME Q14 ib3.Q15 ib3.demo_region time ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or noheader
est store ph_rc_v2
est save ph_rc_v2, replace
save, replace
```

```
************************************************************
*Table 3. Network education assortativity and health (Model 3)
*Also reported in:
*SI. Sensitivity Tables 1a-d, Network ed. assortativity & health (Model 5)
*		This model adds interaction education x assortativity
************************************************************
//A. BMI
quietly mixed BMIcorr ib4.DEMO_EDUC_5CAT##c.edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME b4.DEMO_EDUC_5CAT ib3.Q15 Q14 ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent)
quietly margins DEMO_EDUC_5CAT, at(edu_ass_n_0_5=(-1(0.1)0.5))
marginsplot
graph save Graph "BMI_marginsplot_assort.gph", replace
est store bmi_rc_v3
est save bmi_rc_v3, replace
save, replace

//B. Exercise Regularly
quietly melogit exer_bin ib4.DEMO_EDUC_5CAT##c.edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME b4.DEMO_EDUC_5CAT ib3.Q15 Q14 ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or
est store exer_rc_v3
est save exer_rc_v3, replace
save, replace
quietly margins DEMO_EDUC_5CAT, at(edu_ass_n_0_5=(-1(0.1)0.5)) ///
predict(mu fixedonly) vsquish
marginsplot
graph save Graph "exer_marginsplot_assort.gph", replace
est store exer_rc_v3
est save exer_rc_v3, replace
save, replace

//C. Excellent Self-reported Mental Health
quietly melogit Q10_bin ib4.DEMO_EDUC_5CAT##c.edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME b4.DEMO_EDUC_5CAT ib3.Q15 Q14 ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or
quietly margins DEMO_EDUC_5CAT, at(edu_ass_n_0_5=(-1(0.1)0.5)) ///
predict(mu fixedonly) vsquish
marginsplot
graph save Graph "MH_marginsplot_assort.gph", replace
est store mh_rc_v3
est save mh_rc_v3, replace
save, replace
```

```
//D. Excellent Self-reported Physical Health
quietly melogit Q8_bin ib4.DEMO_EDUC_5CAT##c.edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME b4.DEMO_EDUC_5CAT ib3.Q15 Q14 ///
ib3.demo_region time ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or
quietly margins DEMO_EDUC_5CAT, at(edu_ass_n_0_5=(-1(0.1)0.5)) ///
predict(mu fixedonly) vsquish
est store ph_rc_v3
est save ph_rc_v3, replace
save, replace
marginsplot
graph save Graph "PH_marginsplot_assort.gph", replace
est store ph_rc_v3
est save ph_rc_v3, replace
save, replace

***********************************************************
* Table 4: Tie strength moderation
***********************************************************

//A. BMI
quietly mixed BMIcorr c.avg_like##c.edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME ib4.DEMO_EDUC_5CAT ib3.Q15 Q14 ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent)
est store bmi_rc_table4
est save bmi_rc_table4, replace
save, replace

//B. Exercise Regularly
quietly melogit exer_bin c.avg_like##c.edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME b4.DEMO_EDUC_5CAT Q14 ib3.Q15 ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or
est store exer_rc_table4
est save exer_rc_table4, replace
save, replace

//C. Excellent Self-reported Mental health
quietly melogit Q10_bin c.avg_like##c.edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME b4.DEMO_EDUC_5CAT Q14 ib3.Q15 ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or
est store mh_rc_table4
```

```
est save mh_rc_table4, replace
save, replace

//D. Excellent Self-reported Physical Health
quietly melogit Q8_bin c.avg_like##c.edu_ass_n_0_5 ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME b4.DEMO_EDUC_5CAT ib3.Q15 Q14 ib3.demo_region time ///
ib3.demo_region time ///
|| ID:edu_ass_n_0_5, covariance(independent) or
est store ph_rc_table4
est save ph_rc_table4, replace
save, replace

************************************************************
*SI. Sensitivity Table 2 - Education Assortativity  & Education Distance
************************************************************
//MLM - assortativity as outcome
mixed edu_ass_n_0_5 ib4.DEMO_EDUC_5CAT  ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME Q14 ib3.Q15 ///
ib3.demo_region time ///
|| ID:, covariance(independent)
est store edass_rc_v1
est save edass_rc_v1, replace
save, replace

//MLM - education distance as outcome
mixed eduDistanceSmith_num ib4.DEMO_EDUC_5CAT  ///
graphdensity alters avg_close avg_like ///
DEMO_GENDER DEMO_AGE marital ib1.DEMO_RACE HISP ///
ib2.employ ib4.DEMO_INCOME Q14 ib3.Q15 ///
ib3.demo_region time ///
|| ID:, covariance(independent)
est store edass_rc_v1ssmith
est save edass_rc_v1ssmith, replace
save, replace
```