Published in final edited form as:
Sociol Methods Res. 2011 May ; 40(2): 240–255. doi:10.1177/0049124111404821.

Sensitivity analysis for contagion effects in social networks

Tyler J. VanderWeele[1]

Abstract

Analyses of social network data have suggested that obesity, smoking, happiness and loneliness all travel through social networks. Individuals exert "contagion effects" on one another through social ties and association. These analyses have come under critique because of the possibility that homophily from unmeasured factors may explain these statistical associations and because similar findings can be obtained when the same methodology is applied to height, acne and head-aches, for which the conclusion of contagion effects seems somewhat less plausible. We use sensitivity analysis techniques to assess the extent to which supposed contagion effects for obesity, smoking, happiness and loneliness might be explained away by homophily or confounding and the extent to which the critique using analysis of data on height, acne and head-aches is relevant. Sensitivity analyses suggest that contagion effects for obesity and smoking cessation are reasonably robust to possible latent homophily or environmental confounding; those for happiness and loneliness are somewhat less so. Supposed effects for height, acne and head-aches are all easily explained away by latent homophily and confounding. The methodology that has been employed in past studies for contagion effects in social networks, when used in conjunction with sensitivity analysis, may prove useful in establishing social influence for various behaviors and states. The sensitivity analysis approach can be used to address the critique of latent homophily as a possible explanation of associations interpreted as contagion effects.

Introduction

Recent empirical analyses with social network data has suggested that social influence plays an important role in the spread of health-related attributes, behaviors and psychological states (Christakis and Fowler, 2007, 2008; Fowler and Christakis, 2008a; Cacioppo, Fowler and Christakis, 2009). Analyses conducted by Christakis and Fowler have suggested that such social influence plays a role in the spread of obesity, smoking, happiness and even loneliness. These states and behaviors, it is argued, travel through the network. An individual's becoming obese is thought to increase the likelihood that his or her friends become obese through social influence. Such effects, if present, are sometimes referred to as "peer effects," "contagion effects," or "induction." The analyses of Christakis and Fowler effectively suggest that obesity is "contagious," not so much by any sort of direct "passing on" but by changing weight-related behavior (diet, exercise, etc.) as well as norms, through one individual influencing another.

[1]Correspondence to: Harvard School of Public Health, Departments of Epidemiology and Biostatistics, 677 Huntington Avenue, Boston, MA, 02115; Phone: 617-432-7855; Fax: 617-432-1884; tvanderw@hsph.harvard.edu.

These analyses have come under critique. There was some skepticism about such states and behaviors traveling through networks when the findings were first published (Tamburlini and Cattaneo, 2007; Cohen-Cole and Fletcher, 2008a). Since then, additional analyses, employing the same methodology, have resulted in findings that would supposedly analogously indicate that other states and attributes, such as height, acne and head-aches, apparently also travel in networks (Cohen-Cole and Fletcher, 2008b). Irrespective of what one might think of obesity, smoking or happiness spreading through social influence, the case seems more implausible for factors such as height, acne and head-aches. These analyses have shed some doubt on the methodology employed in the contagion effect findings for obesity, smoking, happiness and loneliness.

In addition to these seemingly spurious findings, several further points of criticism have been raised with regard to the social network methodology employed including issues of model specification, estimation of standard errors and inadequate hypothesis testing (Cohen-Cole and Fletcher, 2008aa; Lyons, 2010); Christakis and Fowler have responded to some of these issues (Fowler and Christakis, 2008b; Christakis and Fowler, 2011). Perhaps the most important critique, however, concerns the possibility of latent (unmeasured) homophily (Shalizi and Thomas, 2011). Homophily refers to the tendency of individuals similar to one another to associate with each other i.e. to become friends. It may, for example, be the case that two friends simultaneously become obese not because one influences the other but because e.g. both enjoy excessive eating; this shared interest causes them to become friends and also causes them both to gradually become more and more obese, a change that might have occurred even if the two had never met. The association between the change in obesity status of one and that of the other may be principally due to such an unmeasured factor that influences both obesity status and the likelihood of their becoming friends. When control is not made for such variables in the analysis, it is difficult to attribute association to contagion effects (social influence) rather than homophily (Shalizi and Thomas, 2011).

The analyses of Christakis and Fowler (Christakis and Fowler, 2007, 2008; Fowler and Christakis, 2008a; Cacioppo, Fowler and Christakis, 2009) consist of regressing one individual's (the ego's) obesity status on another's (the alter's) obesity status, along with the alter's lagged obesity status, the ego's lagged obesity status, and other covariates for the ego. Significant association between the ego's obesity status and the alter's obesity status when also controlling for the ego's and alter's lagged obesity and other variables was taken as evidence for a contagion effect. Christakis and Fowler argue that adjusting for the alter's lagged obesity status helps control for homophily (Christakis and Fowler, 2007; Carrington, Scott and Wasserman, 2005). The reasoning is that the latent factor giving rise to homophily would have to explain both the ego's obesity and the alter's obesity via pathways other than through the alter's lagged obesity. Although such control for alter's lagged obesity arguably does help somewhat, Shalizi and Thomas (2011) have persuasively argued that the problem of latent homophily still plagues such analyses. If the latent factor giving rising to the formation of friendship ties affects present obesity even when controlling for past obesity, associations can arise between the ego's and alter's current obesity even when the alter has no social influence on the ego. This critique has brought into question the whole range of supposed findings on social influence using observational social network data.

Interpreting social network associations as evidence for contagion effects is potentially problematic for another reason, namely, the possibility that a shared environmental factor might in fact affect both the ego and the alter. Christakis and Fowler argue against this as an explanation by noting that the effect estimates for ego-perceived friends are larger than those for alter-perceived friends, which would not occur if the associations were purely due to environmental confounding. While their argument carries some weight, Shalizi and Thomas (2011) show that in the presence of latent homophily, even if there is no unmeasured environmental confounding, the associations comparing ego-perceived friends and alter-perceived friends may differ in magnitude even when there is no social influence! The basic reasoning used by Christakis and Fowler (2007) to argue against environmental confounding in this case breaks down and we are then left with both latent homophily and environmental confounding as possible explanations.

This methodological problem is a real one. Fortunately, it is one that is arguably not insurmountable. It is not possible to analytically control for factors which are unmeasured, for which there is no data. Nevertheless it is possible to assess the extent to which such unmeasured factors would have to both influence the ego's obesity and the extent to which it would have to give rise to homophily in order to explain away supposed contagion effect findings. We use such a sensitivity analysis approach (VanderWeele and Arah, 2011) to examine the effect estimates in the current social networks literature and to assess the extent of the problem of latent homophily as potentially manifest in several recent studies.

Methods: Sensitivity Analysis for Contagion Effects

Christakis and Fowler (Christakis and Fowler, 2007, 2008; Fowler and Christakis, 2008a; Cacioppo, Fowler and Christakis, 2009) use longitudinal analyses of the characteristics of individuals in a social network to assess contagion effects. Suppose individual i names individual k as a friend. Let $Y_i(t)$ and $Y_i(t+1)$ denote the ego's outcome at times t and t+1 respectively. Let $Y_k(t)$ and $Y_k(t+1)$ denote the alter's outcome at times t and t+1 respectively. Let $Z_i(t+1)$ denote the ego's covariates at time t. Christakis and Fowler regress $Y_i(t+1)$ on $Y_i(t)$, $Y_k(t)$, $Y_k(t+1)$ and $Z_i(t+1)$ using either logistic regression (for binary outcomes) or linear regression (for continuous outcomes). The coefficient for $Y_k(t+1)$ on $Y_i(t+1)$ is taken as the contagion effect. Robust standard errors are computed using generalized estimating equations. There has been some debate concerning the appropriate estimation of standard errors in the analysis of social network data (Cohen-Cole and Fletcher, 2008a; Lyons, 2010; Fowler and Christakis, 2008b). Here we will use the standard errors and confidence intervals reported by Christakis and Fowler in their analyses; however, a similar approach could be used with other confidence intervals (Cohen-Cole and Fletcher, 2008a; Fowler and Christakis, 2008b).

Shalizi and Thomas (2011) note that $Y_i(t+1)$ may be associated with $Y_k(t+1)$ conditional on $Y_i(t)$, $Y_k(t)$, and $Z_i(t+1)$ even if there is no contagion effect if there is a latent factor (which we will denote by U_i for individual i and U_k for individual k) such that U_i and U_k affect both obesity and the likelihood that individuals i and k form a friendship tie. This gives rise to confounding due to latent homophily. By arguments from graphical networks (Shalizi and Thomas, 2011; Pearl, 2009; Spirtes, Glymour and Scheines, 2001), if it were possible to

control for U_i in the regression for $Y_i(t+1)$ on $Y_i(t)$, $Y_k(t)$, $Y_k(t+1)$ and $Z_i(t+1)$ the issue would be resolved in the sense that under the null hypothesis of no contagion effects (and assuming no further unmeasured confounding) we would have that $Y_i(t+1)$ would be independent of $Y_k(t+1)$ conditional on $Y_i(t)$, $Y_k(t)$, $Z_i(t+1)$ and U_i (Shalizi and Thomas, 2011).

We will let U denote a latent factor that may be taken as either the latent factor U_i responsible for homophily or alternatively an unobserved shared environmental factor. For simplicity, we will assume that U is binary but a more general approach is also possible (VanderWeele and Arah, 2011). We first consider the case of a binary outcome. Suppose that there is a binary U that increases the likelihood of the ego having the outcome present at time t by a factor of γ (conditional on the ego's lagged obesity, the alter's obesity and lagged obesity and the ego's covariate values) i.e.

$$\gamma = \frac{P\{Y_i(t+1)=1 | U=1, Y_k(t+1), Y_k(t), Y_i(t), Z_i(t+1)\}}{P\{Y_i(t+1)=1 | U=0, Y_k(t+1), Y_k(t), Y_i(t), Z_i(t+1)\}}.$$

Let the prevalence of U for the ego amongst the group with the alter having the outcome present or absent at time t+1 be π_1 and π_0 respectively (conditional on the ego's and the alter's lagged obesity and the ego's covariate values) i.e.

$$\pi_1 = P\{U=1 | Y_k(t+1)=1, Y_k(t), Y_i(t), Z_i(t+1)\}$$
$$\pi_0 = P\{U=1 | Y_k(t+1)=0, Y_k(t), Y_i(t), Z_i(t+1)\}.$$

The relation between π_1 and π_0 effectively captures the extent to which U is relevant for the formation of friendship ties. By arguments from graphical networks (cf. Shalizi and Thomas, 2011; Pearl, 2009), if the latent factor U had no effect on friendship ties then π_1 and π_0 would be the same. Under these assumptions, the ratio between (i) the estimate of the estimate of the effect of $Y_k(t+1)$ on the risk ratio scale in the regression of $Y_i(t+1)$ on $Y_i(t)$, $Y_k(t)$, $Y_k(t+1)$ and $Z_i(t+1)$ and (ii) the estimate that would have been obtained had it been possible to regress $Y_i(t+1)$ on U, $Y_i(t)$, $Y_k(t)$, $Y_k(t+1)$ and $Z_i(t+1)$ is given by (VanderWeele and Arah, 2011):

$$B = \frac{1+(\gamma-1)\pi_1}{1+(\gamma-1)\pi_0}.$$

The corrected estimate (i.e. what would have been obtained had adjustment been made for U) can then be computed by dividing the actual estimate by the bias factor B. Under this simple sensitivity analysis scenario, corrected confidence intervals can also be obtained by dividing both limits of the initial confidence interval by the factor B (VanderWeele and Arah, 2011). The parameters γ, π_1 and π_0 can be varied in a sensitivity analysis. The bias formula given above holds for risk ratios and holds approximately for odds ratios provided the outcome is rare.

Suppose now that the outcome is continuous and that linear regression is used to obtain estimates of the coefficient for $Y_k(t+1)$ when $Y_i(t+1)$ is regressed on $Y_i(t)$, $Y_k(t)$, $Y_k(t+1)$ and $Z_i(t+1)$. Suppose we wish to compare a value of $Y_k(t+1)$ of y_1 to that of y_0 e.g. for a one unit difference in $Y_k(t+1)$ we would have $y_1 - y_0 = 1$. Suppose again that U is binary and now let γ denote the mean outcome difference of $Y_i(t+1)$ comparing U=1 and U=0 (conditional on the ego's lagged obesity, the alter's obesity and lagged obesity and the ego's covariate values) i.e.

$$\gamma = E[Y_i(t+1)|U=1, Y_k(t+1), Y_k(t), Y_i(t), Z_i(t+1)] - E[Y_i(t+1)|U=0, Y_k(t+1), Y_k(t), Y_i(t), Z_i(t+1)].$$

and let δ be the difference in prevalence of U for the ego amongst the group with the alter having a value of $Y_k(t+1)$ of y_1 versus y_0 (conditional on the ego's and alter's lagged obesity and the ego's covariate values) i.e.

$$\delta = P\{U=1|Y_k(t+1)=y_1, Y_k(t), Y_i(t), Z_i(t+1)\} - P\{U=1|Y_k(t+1)=y_0, Y_k(t), Y_i(t), Z_i(t+1)\}.$$

Under these assumptions, the difference between $(y_1 - y_0)$ times the coefficient for $Y_k(t+1)$ in the regression of $Y_i(t+1)$ on $Y_i(t)$, $Y_k(t)$, $Y_k(t+1)$ and $Z_i(t+1)$ and the $(y_1 - y_0)$ times the coefficient that would have been obtained had it been possible to regress $Y_i(t+1)$ on U, $Y_i(t)$, $Y_k(t)$, $Y_k(t+1)$ and $Z_i(t+1)$ is given by (VanderWeele and Arah, 2011):

$$B = \delta\gamma.$$

The corrected estimate can then be computed by subtracting the actual estimate by the bias factor B. Under this simple sensitivity analysis scenario, corrected confidence intervals can also be obtained by subtracting from both limits of the initial confidence interval the factor B (VanderWeele and Arah, 2011). The parameters δ and γ can be varied in a sensitivity analysis.

Results: Application to Christakis and Fowler Studies

We apply sensitivity analysis techniques (VanderWeele and Arah, 2011), adapted for social network data, to examine the sensitivity of contagion effect findings (Christakis and Fowler, 2007, 2008; Fowler and Christakis, 2008a; Cacioppo, Fowler and Christakis, 2009) to latent homophily. We examine contagion effects in four of the more prominent studies on social networks and social influence: namely those concerning the spread of obesity, smoking, happiness and loneliness (Christakis and Fowler, 2007, 2008; Fowler and Christakis, 2008a; Cacioppo, Fowler and Christakis, 2009). We apply the sensitivity analysis also to the presumably spurious contagion effect findings for height, acne and head-aches (Cohen-Cole and Fletcher, 2008b). Christakis and Fowler consider a number of different relationships including ego-perceived friends, alter-perceived friends, mutual friends, sibling, spouses and neighbors. Here, we will focus on assessing the effect of the state of an alter who the ego perceives as a friend on the ego's state, as these are the estimates that have attracted the

most attention in the literature. However, we will occasionally comment also on some of the other relationships.

Using data from the Framingham Heart Study (Dawber, 1980; Feinlib et al., 1975), Christakis and Fowler (2007) found that an individual's chances of being obese (body mass index>30) increased by 57% (95% CI: 6% – 123%) if he or she had a friend who was obese a given interval. They controlled for an ego's age, sex and education level, the ego's obesity status at the previous time point and the alter's obesity status at the previous time point. Adjustment for the alter's lagged obesity status was intended to control for homophily. As noted by Shalizi and Thomas (2011), if there is an unmeasured factor that explains both the formation of friendship ties and affects the likelihood of becoming obese at each period, then the analyses of Christakis and Fowler may be biased, even after control for the alter's lagged obesity. Such an unmeasured factor may effectively confound the relationship between the alter's current obesity and the ego's current obesity; estimates of the effect may be non-zero due to such confounding even if there is in fact no effect.

Using sensitivity analysis we can assess the extent to which such an unmeasured factor would have to be related to obesity in order to explain away the alleged effect of the alter's obesity on the ego's obesity status. Suppose that there were some unmeasured binary U denoting the presence or absence for the ego of some latent characteristic. If U increased the likelihood of an ego's being obese by 1.8-fold (conditional on the ego's lagged obesity, the alter's obesity and lagged obesity and the ego's covariate values) and if the prevalence of U for the ego amongst the group with the alter obese was 0.7 but only 0.3 amongst the group with the alter not obese (conditional on the ego's lagged obesity, the alter's lagged obesity and the ego's covariate values) then this would reduce the estimate of the effect of the alter's obesity on the ego's obesity from an increase of 57% (95% CI: 6% to 123%) to an increase of 25% (95 CI: −16% to 77%).

Such a sensitivity analysis can be employed more generally to consider a variety of values for (i) the effect of the latent factor on the ego's obesity (call it γ) and (ii) the prevalence of the latent characteristic for the ego amongst the groups with the alter obese (π_1) or not obese (π_0). Such a sensitivity analysis is given in Table 1. As can be seen, to completely explain away the estimate of a 57% increase in the likelihood of obesity, the latent factor would have to have a fairly strong effect on the ego's becoming obese and the latent factor would also have to be far more prevalent within those friendships in which the alter was obese. For example, to completely explain away the estimate of 57% the unmeasured factor U would have to have increase likelihood of obesity by 3-fold and have a prevalence of 0.7 in friendships in which the alter was obese and only 0.25 in friendships in which the alter was not obese. Although the latent factor U would have to be very strongly related to current obesity (even after controlling for the ego's covariates and the ego's and alter's lagged obesity), the latent U, as can also be seen in Table 1, would need to only be modestly related to obesity to bring down the confidence interval so as to include 0%. In the sensitivity analysis given here we have let the unmeasured factor U denote a characteristic that may affect the formation of friendship ties. The sensitivity analysis we have just described is, however, also applicable if U is taken to be a shared environmental factor for the ego and the alter (though Christakis and Fowler argue against this as an explanation of their findings

by noting the difference in the effect estimates for ego-perceived friends versus alter-perceived friends).

Christakis and Fowler (2008) also examined whether smoking spread through social networks. Using analyses similar to those just described, they estimated that smoking cessation by a friend decreased the chances of smoking by 36% (95% CI: 12% to 55%). They also noted that smoking cessation by a spouse decreased a person's chances of smoking by 67% (95% CI: 59% to 73%). Let us first consider the supposed effects of the spouse on the ego. Once again the association might be explained away by homogamy (similar individuals marrying one another) or a common environmental factor. Table 2 gives a sensitivity analysis similar to that of Table 1 but for the effect of a spouse's smoking cessation on the ego's smoking status. As can be seen in Table 2, not only the estimate but also the entire confidence interval would require that the latent factor U had a very strong relationship with smoking to completely explain away the effect.

If we turn to the effect of a friend's smoking cessation on an ego's smoking status, the results of a sensitivity analysis are similar to that for obesity: a fairly strong relationship between the latent factor and smoking would need to be present to explain away the estimate of 36% (95% CI: 12% to 55%) but only a more modest association would be required to shift the confidence interval so that it included 0. In the study of obesity, the largest effect size was obtained for the effect of a mutual friend's (when both the ego and alter name one another as friends) obesity on the ego with an 171% (95% CI: 59% to 326%) increase in the likelihood of the ego being obese when the alter is. Like the estimate for the spouse in the setting of smoking cessation, the estimate for a mutual friend in the setting of obesity would require a very large amount of confounding to explain away.

For both obesity and smoking, the effect estimate for an ego-perceived friend requires a latent factor U that is very strongly related to the obesity or smoking state (even after controlling for the ego's covariates and the ego's and alter's lagged state), but only modest confounding by homophily would be required to shift the confidence interval to include 0. However, for both obesity and smoking, there are other contagion effect estimates (in the case of obesity, that for a mutual friend; in the case of smoking cessation, that for a spouse) for which both the estimate itself and the confidence interval would require very strong confounding by latent homophily to completely explain away.

We now turn to the analyses concerning happiness (Fowler and Christakis, 2008a) and loneliness (Cacioppo et al., 2009). In the social network analyses (Fowler and Christakis, 2008a), happiness was defined as having reported the maximum answers on all four questions about positive affect on the Center for Epidemiological Studies Depression scale (CES-D; Radloff, 1977). Loneliness was measured as the number of days an individual felt lonely the previous week as reported in a question on the CES-D (Radloff, 1977; Cacioppo et al., 2009). For happiness, Fowler and Christakis (2008a) find that a friend who lives within one mile who is happy increases the probability that the ego is happy by 25% (95% CI: 1% to 57%). Here, if a latent factor U increased the likelihood of an ego's being happy 1.8-fold and if the prevalence of U for the ego amongst the group with the alter was or was not happy were 0.7 and 0.3 (the same values initially used for obesity) this would reduce the

observed effect to −1% (95% CI: −20% to 25%), completely explaining it away. Since the initial confidence interval (95% CI: 1% to 57%) itself nearly contains 0%, very little confounding by homophily or environment would suffice to shift the confidence interval to include 0%. In fact, the contagion effect estimates obtained by Folwer and Christakis (2008) for happiness for all of the relationships (friend, spouse, sibling, neighbor, etc.) are all such that relatively modest confounding will shift the confidence interval to include 0%.

For loneliness, Cacioppo et al. (2009) estimate that for a friend who lives within one mile each extra day the friend is lonely increases the number of days the ego is lonely by 0.29 days (95% CI: 0.07 to 0.50). The estimate itself could be explained away if a latent factor U giving rise to homophily increased loneliness by 0.87 days a week on average with a prevalence difference in the ego of 0.33 for a one day difference in loneliness of the alter. We see that it would require moderate confounding to explain away the estimate. However, once again, more modest confounding would shift the confidence interval so as to include 0: a latent U that increased loneliness by 0.3 days a week on average with a prevalence difference in the ego of 0.23 for a one day difference in loneliness of the alter would suffice.

Cohen-Cole and Fletcher (2008b) pursue similar analyses for contagion effects for head-aches, acne and height. After adjustment using an approach similar to that of Christakis and Fowler (2007, 2008), Cohen-Cole and Fletcher (2008b) obtain an odds ratio effect estimate of a friend's head-aches on the ego's head aches of 1.14 (standard error 0.27) i.e. 14% increase in odds. The confidence interval contains 0%; the estimate itself could be explained away if a latent factor U giving rise to homophily increased the likelihood of head-aches by 1.4 fold with a prevalence in the ego of 0.7 versus 0.3 when a friend did or did not have head-aches. Cohen-Cole and Fletcher likewise obtain an odds ratio effect estimate of a friend's acne on the ego's acne of 1.23 (standard error 0.35) i.e. 23% increased odds. The confidence interval contains 0%; the estimate itself could be explained away if a latent factor U giving rise to homophily increased the likelihood of acne by 1.7 fold with a prevalence in the ego of 0.7 versus 0.3 when a friend did or did not have acne. Finally, Cohen-Cole and Fletcher obtain an estimate of the effect of a friend's height on an ego's of a 0.21 (95% CI: 0.19 to 0.23). This could be explained away if a latent factor U giving rise to homophily increased the height by 3 inches with a prevalence average with a prevalence difference in the ego of 0.07 for every increased inch of height of the alter.

Discussion

Social network analyses can be informative in assessing peer or "contagion" effects. These analyses are, however, as we have seen, somewhat problematic due to the possibility of homophily and shared environmental factors. Sensitivity analyses can be a useful method by which to assess the evidence for the conclusion of actual social influence. We have applied sensitivity analysis techniques to four social network analyses undertaken by Christakis and Fowler (2007, 2008; cf. Fowler and Christakis, 2008a; Cacioppo, Fowler and Christakis, 2009). The sensitivity analysis suggested that at least some of the findings indicating contagion effects for obesity and smoking (mutual friend for obesity, spouse for smoking) were reasonably robust to latent homophily or environmental factors for which control was not made. In contrast, the effect estimates for the supposed spread of happiness and

loneliness were much more subject to latent homophily or shared environmental factors as a possible explanation.

The analyses of Christakis and Fowler had also come under critique because the application of similar methodology to height, acne and head-aches suggested that these too spread through social networks (Cohen-Cole and Fletcher, 2008b). We have seen that all of the estimates for the supposed contagion effects of height, acne and head-aches are easily explained away in sensitivity analysis. Other critiques (Lyons, 2010) of the methodology employed by Christakis and Fowler may carry some weight but the supposed findings for height, acne and head-aches do not on their own constitute grounds for dismissing the findings for e.g. obesity and smoking. The approach employed by Christakis and Fowler, when used in conjunction with sensitivity analysis, may still be useful in assessing contagion effects. A similar approach might also prove useful in other studies of peer/contagion effects (Aral, Muchnik and Sundararajan, 2009; Bakshy, Barrer, Adamic, 2009). For example, Shalizi and Thomas (2011) proposed addressing a critique of Lyons (2010) by further lagging the alter's outcome by an additional period. The sensitivity analysis technique described here could be used in a similar manner with the proposal of Shalizi and Thomas (2011) simply by replacing $Y_k(t)$ and $Y_k(t+1)$ in the discussion above with $Y_k(t-1)$ and $Y_k(t)$.

The approach we have described here is, however, subject to various limitations. Perhaps most importantly, in attempting to reason about whether latent homophily might explain away a contagion effect estimate we have had to make decisions regarding whether homophily of a particular magnitude is or is not plausible. Without studies providing precise measures of such homophily it is difficult to know what levels might be sufficiently large to be considered implausible. It is nonetheless informative to compare the sensitivity parameters required to explain away an effect across different traits. In interpreting these sensitivity analyses it is, moreover, important to bear in mind that the sensitivity parameters governing homophily are conditional on the alter's lagged trait thereby eliminating some of the association that would otherwise be attributable to homophily. Indeed the very purpose of conditioning on the alter's lagged trait was to accomplish this (Christakis and Fowler, 2007; Carrington et al., 2005). Nonetheless, further empirical data regarding the extent of homophily for various traits and how this changes over time would unquestionably help inform the interpretation of and the conclusions drawn from sensitivity analyses such as the ones presented here.

Another limitation of the sensitivity analyses here were the assumptions that the latent factor U was binary and that its effects on the ego's being obese were the same irrespective of whether or not the alter was obese (essentially no interaction between the effect of U and the effect of the alter's obesity status). Such assumptions are likely unrealistic. More general sensitivity analysis techniques that do not make these assumptions (VanderWeele and Arah, 2011), could similarly be employed. However, one need not believe any specific sensitivity analysis model or the particular values used for the sensitivity parameters for such techniques to be informative. Even under the simplifying assumptions that were employed here the techniques can to a certain extent help assess whether the various effect estimates could be explained away by homophily or unmeasured confounding. If very modest effect

sizes for the unmeasured factor suffice to completely explain away the contagion effect estimate, it is reasonable to be skeptical that the effect is real. Sensitivity analysis under weaker assumptions (allowing for interactions) would still explain away the effect estimate since the sensitivity parameter values that explained away the estimate under the simplified technique would still be within the space of sensitivity parameters values under the more general technique. Certainly, none of the specific corrected estimates should be believed but when one considers a broad range of plausible sensitivity analysis parameters and all of these give rise to the same qualitative conclusion, one can be somewhat confident about the presence of an effect.

It should be noted that an estimate that can easily be explained away by an unmeasured factor does not mean that there is no social influence but only that one probably ought not draw that conclusion from the data at hand. In such settings, where effect sizes are likely more modest, alternative approaches perhaps employing experiments (Fowler and Christakis, 2010) may prove to be more effective in empirically assessing and establishing the presence of social influence.

Acknowledgments

The author thanks Nicholas Christakis and Felix Elwert, the editor, and two referees for helpful comments. The research was funded by NIH grants ES017876 and HD060696.

References

Aral, Siman; Muchnik, Lev; Sundararajan, Arum. Distinguishing intfluence based contagion from homophily driven diffusion in dynamic networks. Proceedings of the National Academy of Sciences (USA). 2009; 106:21544–21549.

Bakshy, Eytan; Karrer, Brian; Adamic, Lada A. Social influence and the diffusion of user-created content. Proceedings of the Tenth ACM Conference on Electronic Commerce; 2009. p. 325-334.

Cacioppo, John T.; Fowler, James H.; Christakis, Nicholas A. Alone in the crowd: the structure and spread of loneliness in a large social network. Journal of Personality and Social Psychology. 2009; 97(6):977–991. [PubMed: 19968414]

Carrington, Peter J.; Scott, John; Wasserman, Stanley. Model and methods in social network analysis. New York: Cambridge University Press; 2005.

Christakis, Nicholas A.; Fowler, Jason H. The spread of obesity in a large social network over 32 years. New England Journal of Medicine. 2007; 357:370–379. [PubMed: 17652652]

Christakis, Nicholas A.; Fowler, Jason H. The collective dynamics of smoking in a large social network. New England Journal of Medicine. 2008; 358:2249–2258. [PubMed: 18499567]

Christakis, Nicholas A.; Fowler, Jason H. Examining dynamic social networks and human behavior. Working Paper. 2011

Cohen-Cole, Ethan; Fletcher, Jason M. Is obesity contagious? Social network vs. environmental factors in the obesity epidemic. Journal of Health Economics. 2008; 27(5):1382–1387. [PubMed: 18571258]

Cohen-Cole, Ethan; Fletcher, Jason M. Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis. British Medical Journal. 2008; 337:a2533. [PubMed: 19056789]

Dawber, Thomas R. The Framingham Study: the Epidemiology of Atherosclerotic Disease. Cambridge, MA: Harvard University Press; 1980.

Feinleib, Manning; Kannel, William B.; Garrison, Robert J.; McNamara, Patricia M.; Castelli, William P. The Framingham Offspring Study: design and preliminary data. Preventive Medicine. 1975; 4:518–525. [PubMed: 1208363]

Fowler, James H.; Christakis, Nicholas A. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. British Medical Journal. 2008a; 337:a2338. [PubMed: 19056788]

Fowler, James H.; Christakis, Nicholas A. Estimating peer effects on health in social networks. Journal of Health Economics. 2008b; 27(5):1386–1391.

Fowler, James H.; Christakis, Nicholas A. Cooperative behavior cascades in human social networks. PNAS: Proceedings of the National Academy of Sciences. 2010; 107(9):5334–5338.

Lyons, Russell. The spread of evidence-poor medicine via flawed social network analysis. 2010. arXiv: http://arxiv.org/abs/1007.2876

Pearl, Judea. Causality. 2. Cambridge: Cambridge University Press; 2009.

Radloff, Lenore S. The CES-D Scale: “A self-report depression scale for research in the general population”. Applied Psychological Measurement. 1977; 1:385–401.

Shalizi, Cosma R.; Thomas, Andrew C. Homophily and contagion are generically confounded in observational social network studies. Sociological Methods and Research. 2011 in press.

Spirtes, Peter; Glymour, Clark; Scheines, Richard. Causation, Prediction and Search. 2. Cambridge, MA: MIT Press; 2001.

Tamburlini, Giorgio; Cattaneo, Adriano. The spread of obesity in a social network. New England Journal of Medicine. 2007; 357:1866. [PubMed: 17978297]

VanderWeele, Tyler J.; Arah, Onyebuchi A. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments and confounders. Epidemiology. 2011; 22:42–52. [PubMed: 21052008]

Table 1

Sensitivity Analysis for the Effect of a Friend's Obesity Status on an Ego's Obesity Corrected Contagion Effect Estimates Under Latent Homophily for Obesity: Percent Increase (95% CI)

Prevalence of U	**Effect of latent factor U on Obesity (γ)**						
	1.2	**1.5**	**1.8**	**2**	**2.5**	**3**	**4**
π_1 =0.6, π_0=0.4	51 (2,115)	45 (−2,105)	40 (−5,99)	37 (−7,95)	32 (−10,88)	28 (−13,82)	23 (−17,75)
π_1 =0.65, π_0=0.35	48 (0,111)	39 (−1,98)	32 (−11,88)	28 (−13,82)	21 (−18,72)	16 (−21,64)	9 (−26,55)
π_1 =0.7, π_0=0.3	46 (−1,107)	34 (−10,90)	25 (−16,77)	20 (−19,70)	11 (−25,58)	5 (−29,49)	−4 (−35,37)
π_1 =0.7, π_0=0.25	45 (−2,105)	31 (−11,86)	21 (−18,72)	15 (−22,64)	5 (−29,50)	−2 (−38,39)	−11 (−40,26)
π_1 =0.75, π_0=0.25	43 (−3,104)	28 (−13,82)	18 (−21,67)	12 (−24,59)	2 (−31,44)	−6 (−36,34)	−15 (−43,20)
π_1 =0.8, π_0=0.25	42 (−4,101)	26 (−15,79)	15 (−22,63)	9 (−26,55)	−2 (−34,39)	−9 (−39,29)	−19 (−45,14)
π_1 =0.8 π_0=0.2	41 (−5,100)	23 (−17,75)	11 (−25,58)	5 (−29,49)	−7 (−37,31)	−15 (43,20)	−26 (−50,5)

Settings with $\pi_1 < \pi_0$ all result with percent increase and 95% CI large than the observed estimate of 57 (6, 123)

Table 2

Sensitivity Analysis for the Effect of a Spouse's Smoking Cessation Status on an Ego's Smoking Cessation: Corrected Contagion Effect Estimate Under Latent Homophily for Smoking Cessation: Percent Increase (95% CI)

Prevalence of U	**Effect of latent factor U on Smoking Cessation (γ)**						
	1.2	**1.5**	**1.8**	**2**	**2.5**	**3**	**4**
π_1 =0.6, π_0=0.4	61 (53,67)	54 (47,60)	49 (42,54)	46 (39,51)	41 (34,46)	37 (30,42)	31 (25,36)
π_1 =0.65, π_0=0.35	58 (51,64)	48 (41,53)	41 (34,46)	37 (30,42)	29 (23,34)	23 (18,28)	16 (10,20)
π_1 =0.7, π_0=0.3	55 (48,61)	42 (35,47)	32 (26,38)	28 (22,32)	18 (12,22)	11 (5,15)	2 (−2,6)
π_1 =0.7, π_0=0.25	54 (46,59)	39 (33,44)	28 (22,33)	23 (17,27)	12 (6,16)	4 (−1,8)	−6 (−10,−2)
π_1 =0.75, π_0=0.25	52 (45,58)	37 (30,42)	25 (19,30)	19 (13,23)	8 (2,12)	0 (−5,4)	−10 (−14,−7)
π_1 =0.8, π_0=0.25	51 (44,57)	34 (28,39)	22 (16,27)	16 (10,20)	4 (−1,8)	−4 (−8,0)	−14 (−18,−10)
π_1 =0.8 π_0=0.2	50 (43,55)	31 (25,36)	18 (12,22)	11 (5,15)	−1 (−6,2)	−10 (−14,−6)	−21 (−25,−19)

Settings with $\pi_1 < \pi_0$ all result with percent increase and 95% CI large than the observed estimate of 67 (59, 73)