Peer Effects in Health Behaviors

JM Fletcher, Yale School of Public Health, New Haven, CT, USA

© 2014 Elsevier Inc. All rights reserved.

Introduction

Health economists have long been interested in examining the determinants of, and potential policies for, reducing unhealthy behaviors in the population. Although a main focus in this area has historically been on issues of policy involving taxation, access restrictions, advertising, etc., a shift toward evaluating the basic social or nonmarket determinants of unhealthy behaviors has occurred in the literature. This is perhaps most obvious in the research regarding children and adolescent behaviors, wherein peer pressure is often thought to play a substantial role in determining choices such as smoking decisions. Indeed, there is now a large and growing literature in health economics that asks variants of the question, “Do peers influence an individual’s health behavior decisions?” The areas of interest range from substance use to eating behaviors and weight outcomes whereas the peer group definitions range from best friends to classmates, residential neighbors, and beyond. Indeed, although there have been several recent reviews of the literature examining social effects on health behaviors (Fletcher, 2010a, 2011a), these papers are being updated because of the rapid expansion in research in this area. This new research is based on previous work; moreover, quasi-experimental methods as well as new identification strategies are utilized in the research.

One reason for the increasing interest in achieving these research milestones is the policy implications of the existence of peer effects in health behaviors. Specifically, peer effects often imply a ‘social multiplier’ for interventions – if the health of one individual is increased, the effect of the intervention may be multiplied through peers. This type of social effect is seen as an ‘endogenous social effect’ in the literature of economics. In contrast, peer effects that operate through the characteristics of peers are labeled ‘exogenous social effects or contextual effects’ (see Manski, 1993). The presence of such endogenous social effects could increase the potential benefits of intervention without increasing the costs. In contrast to these benefits, the presence of peer effects could also work to spread unhealthy behaviors (such as smoking). The awareness of a social multiplier operating in determining health can also help to inform whether targeted (e.g., based on influential individuals within networks) or broad-based policy is more effective. Also, peer effects imply that the composition of a person’s neighborhood and/or school could affect his/her health behavior; because many policies can reorganize peer groups, such as school ability grouping (tracking), busing, school grade-span configuration, and residential zoning, there are a host of potentially important policy domains that plays a role in reducing poor health behavior in the presence of peer effects.

Although research regarding health decisions of socially connected individuals may continue to expand along with the multiple growth of interaction within social ties, the empirical hurdles to credibly estimating peer influence remain relatively unchanged and difficult to overcome. This article discusses some of the general empirical issues with their brief history, current controversies, and future directions.

Empirical Issues

Just as the policy and health importance of peer effects is likely to be substantial, so too is the empirical difficulty of credibly estimating causal effects. There are (at least) four standard primary empirical issues that researchers face. In many empirical settings, some are generic problems of measurement and omitted variables, whereas others are somewhat specific to peer effects research.

First, researchers must define a relevant peer group. This step seems simple, but data limitations typically force researchers to define peers on the basis of convenience rather than on theory. This has created peer group definitions that range from state-based groups to nominated best friends, and everything in between. For example, Harding (2003) uses census tracts, Evans et al. (1992) use metropolitan level data, Case and Katz (1991) use city block level data, Fletcher (2010a) uses school grades, Fletcher (2010b) uses school classrooms, Mayer and Puller (2008) use ‘Facebook Friends,’ and Sacerdote (2001) uses roommates to create relevant reference groups for the outcomes to be examined. Although there are several data sets that include nominated friends and peers, the vast majority do not. New data sets may reduce this issue over time, particularly those collecting online social network data, but this will raise the issue of whether online social contacts represent an important and relevant peer group for the determination of health decisions, and if they do, what types of health decisions are relevant when considering online peers.

A second empirical difficulty is the endogeneity of peer groups. Does a person smoke because his friend smokes or did he choose his friend for the sake of smoking? Because individuals typically have some degree of choice over their interaction with others (schoolmates, neighbors, friends, etc.), separating peer selection from peer influence is a particularly difficult empirical problem, and peer selection effects would typically inflate standard estimates of ‘peer effects.’ In fact, there seems to be a ‘relevance-endogeneity’ trade-off between the first and the second empirical difficulties (Fletcher, 2010a). As the researcher broadens the definition of the peer group (such as pertaining to the state level), the endogeneity of the peer group probably diminishes, but the relevance of the peer group may weaken. In contrast, best friends are probably a relevant definition of a peer group for many health behaviors but the endogeneity of best friend is magnified.

A third empirical difficulty in peer effect research lies in its potential nature for omitted variable bias through shared influences. For example, smoking bans may reduce tobacco use in all members of a school-based or community-based
peer group. These shared factors can lead to inflated estimates of peer effects if sufficient control variables are not included.

A fourth empirical difficulty in peer effects research is the reflection problem (Manski, 1993), where the researcher may be unable to distinguish between whether Bill influences Ted or Ted influences Bill. Although it is not essential to disentangle these two influences in order to establish whether there is any social effect for determining health behaviors, it can be useful to separate these effects in order to understand the importance of the initial causal effect as against the feedback effects to further understand the processes of health spillovers. Although most researchers explicitly acknowledge each of these difficulties, they often adopt different approaches in attempting to overcome them.

Indeed, there is a two-decade-old history examining peer effects in many health behaviors, which can provide some examples of the difficulty with this research topic while outlining ways that other researchers have attempted to circumvent the empirical issues as outlined above. Typically, researchers have used neighborhood or school-based definition of ‘peers’ when examining health behaviors such as tobacco, alcohol, and drug use.

**A Brief History of Empirical Approaches**

Case and Katz (1991) provide a seminal look at the effects of neighborhood peers on risky behaviors and other outcomes, although they are unable to tackle many of the aforementioned empirical issues. In particular, the authors acknowledge that they are unable to control for all environmental confounders and their self-selection into neighborhoods. The authors use what has become a typical empirical framework in the literature:

\[ Y_{ig} = X_{ig}B + \bar{X}_{-ig}\delta + W_{ig}\theta + \gamma \bar{Y}_{-ig} + \epsilon_{ig} \]  

where, \( Y_{ig} \) is the health behavior choice of individual \( i \) in peer group \( g \) (e.g., neighborhood), individual and family characteristics are contained in a vector \( X \), and peer characteristics are measured as group-level averages of the \( X \) vector excluding the individual, labeled \( \bar{X}_{-ig} \). Unobserved factors are contained in the vector, \( W_{ig} \). Finally, \( \bar{Y}_{-ig} \) is the group-level average outcome excluding the individual (e.g., the proportion of individuals in the same neighborhood who report smoking). The main coefficient of interest is the endogenous effect \( \gamma \), which indicates the extent to which individuals are influenced by their peers’ choices. If \( \gamma = 0 \) positive, interventions that change the behavior of individuals (or subsets of individuals) within a reference group would be predicted to spillover on nontreated individuals in the same reference group. In addition to acknowledging the potential for omitted group-level variables as well as self-selection (where \( \epsilon_{ig} \) and \( \bar{Y}_{-ig} \) are correlated), the authors are also unable to resolve the simultaneity bias (this issue was not fully discussed until Manski, 1993). The authors find evidence of substantial correlation between own and neighborhood peer substance use, crime, and other behaviors.

Norton et al. (1998) focus on schoolmate peer effects in alcohol and tobacco use of teenagers, and they use an instrumental variables strategy to address the endogeneity of peer groups (see also Evans et al. (1992) for an analysis of teenage pregnancy). Although the focus on endogeneity is important, there is little scope to control for the shared environment due to both data limitations and the instruments (such as neighborhood drug availability and safety) being potentially invalid – in fact, the results have suggested that noninstrumented results are preferable for extremely large peer effects. The general approach of using schoolmates or grademates has been used by many subsequent studies (e.g., Gaviria and Raphael, 2001), wherein too the quality of the instruments are uncertain; specifically, all contextual effects are often assumed to not exist in order to use these variables as instruments.

More recently, Fletcher (2010a) has suggested this approach to be inappropriate and instead proposes a combined instrumental variables/fixed effects design with conceptually appealing diagnostic tests (following Bifulco et al., 2011; Lavy and Schlosser, 2007) in order to validate a preferred instrument set, although the validity of the instruments is still widely questioned. Specifically, Fletcher argues that the increasing proportion of the smoking grademates is due to smoking status of individuals in their households (which can be empirically demonstrated), which does not directly affect respondent smoking even when school-fixed effects are controlled (which is a maintained, untestable assumption). Although Fletcher shows the evidence that exposure of smoking grademates from households of smokers is conditionally random within school, there are ways by which this instrument could be invalidated because, for example, if mothers of grademates are smokers, it simply implies that there is access to tobacco for the respondent. See also Fletcher (2011b) for an examination of peer influences in alcohol consumption.

There have been several alternatives to the instrumental variable approach in the literature. Clark and Loheac (2007) use panel data and a lagged measure of peer behaviors that is combined with school-fixed effects in order to adjust for endogeneity, a large portion of the shared environment, and the reflection problem:

\[ Y_{ig} = X_{ig}B + \bar{X}_{-ig}\delta + W_{ig}\theta + \gamma \bar{Y}_{-ig-1} + \epsilon_{ig} \]  

The reflection problem is eliminated because current smoking decisions cannot affect past schoolmate smoking decisions. Although school-fixed effects reduce the issue of contextual effects, a maintained assumption is that, within schools, students choose friends randomly. A second weakness of this design is the need to assume a specific time structure where individual decision making and social influence processes are concerned (e.g., 1 day, 1 week, 1 month, 1 year, 2 years, etc.) (Manski, 1995). Specifically, Manski (1995, p. 136) states, "Of course, one cannot simply specify a dynamic model and claim that the problem of inference on social effects has been resolved. Dynamic analysis is meaningful only if one has reason to believe that the transmission of social effects follows the assumed temporal pattern."

An alternative to implementing a lag structure research design or an instrumental variables strategy is to focus on estimating contextual social effects instead of endogenous social effects. The most convincing work in this area uses random assignment of peers. For example, Kremer and Levy...
(2008) use data from a university that randomly assigns freshmen to shared dormitory rooms:
\[
Y_{ig} = X_{ig}B + \bar{X}_{ig}\delta + W_{ig}\theta + \epsilon_{igt} \tag{3}
\]
where, in this case, \(\bar{X}_{ig}\) could be thought of as a lagged endogenous social effect examined in some studies or roommate’s precollege alcohol consumption. What allows the estimate to produce a contextual effect rather than an endogenous one is that the individual is not exposed to the actual drinking behavior, but rather is being exposed to having a roommate who has the characteristic of being a past drinker. Additionally, the random assignment of roommates eliminates the concerns regarding the endogeneity of the peer group. Kremer and Levy show that a fresh student who is randomly assigned a roommate with alcoholic past during high school has lower college performance than the student who is assigned a nondrinking roommate. The focus on the roommate’s predetermined high school drinking behavior as the peer effect of interest also eliminates issues of simultaneity bias.

Because not all data sets are able to leverage the random assignment of ‘friends,’ several studies attempt to leverage quasi-random-variation in observational data. For example, Bifulco et al. (2011) use a cross-cohort, within-school design to link the outcomes of students to their (quasi-randomly assigned) classmates’ characteristics:
\[
Y_{ig} = X_{ig}B + \bar{X}_{ig}\delta + W_{ig}\theta + \epsilon_{igt} \tag{4}
\]
That is, the authors examine the ‘peer effects’ of having a higher share of grademates with educated mothers or a higher share of grademates who are racial/ethnic minorities. This focus on contextual effects sidelines the need for a solution to the reflection problem because student smoking cannot affect grademate race, but some of the important policy issues that are tied to a social multiplier through endogenous peer effects cannot be evaluated directly.

**Newer Approaches and Extension of Outcomes**

Although the more traditional literature examining peer effects and health behaviors has focused primarily on substance use outcomes and has used a range of empirical approaches, the more recent literature in this area has broadened research designs and has dramatically expanded the range of health outcomes under study – especially, weight and mental health outcomes.

Apart from the literature in health economics, the set of studies that has received the most media attention is from Nicholas Christakis and James Fowler, and a set of coauthors. Their first study has brought a new outcome of interest to the literature by examining whether obesity is ‘socially contagious.’ Specifically, the authors have found that the chances of an individual becoming obese increased by more than 50% when his/her friend has become obese. The authors have used the Framingham Heart Study data, which contain up to 32 years of longitudinal measures of BMI for individuals in one area of Massachusetts. To these data, the authors have matched information from the original respondents’ records, on which respondents have individually been asked to name a person who can be contacted in case the survey team does not reach them directly at follow up; this contact person is treated by Christakis and Fowler (2007) as a ‘friend.’ Thus, the first issue with this research is whether the contact person is truly a peer. The authors estimate regressions using the following parsimonious empirical model:
\[
\text{health}_{it} = \delta \text{health}_{jt} + \beta_1 \text{health}_{it-1} + \beta_2 \text{health}_{jt-1} + \beta_3 X_{it} + \epsilon_{it} \tag{5}
\]
where, the health (obesity) of person \(i\) is linked to person \(j\) and \(\delta\) is the coefficient of interest – the endogenous social effect or ‘social multiplier.’ A positive estimate on \(\delta\) suggests that an intervention which reduces the chances of an individual becoming obese will also reduce the chances of obesity in his/her peer.

To overcome endogeneity, Christakis and Fowler (2007) have assumed that lagged health outcomes for the friend (\(\text{health}_{jt-1}\)) is a sufficient control, that is, after controlling for lagged obese status of a friend, they have assumed that there is no additional issue of friendship selection. Unfortunately, to the extent that this control variable does not completely eliminate selection effects, the estimated coefficient of interest (\(\delta\)) will probably be upwardly biased. The authors have controlled for own-lagged health in order to control for aspects of the individual’s genetic disposition or other time-invariant characteristics. The second issue is confounding due to shared influences. Without explicitly controlling for shared environmental factors, the authors have appealed to a comparison between mutually nominated friends and nonmutual friends (with unreciprocated nomination), arguing that directionality of nominations does not matter if environmental confounding is the primary explanation. Finally, the authors neither discussed nor attempted to overcome the empirical complications from the reflection problem. Unfortunately, each of those empirical issues listed above would probably lead to upwardly biased estimates of peer effects. So, what proportion of the 50% estimated peer effect is due to bias and what proportion is an actual peer effect? To address these empirical concerns, Cohen-Cole and Fletcher (2008a) have provided an examination focusing on one of the empirical issues in peer effects models – shared environmental factors that may bias upward the estimates. The authors have used the National Longitudinal Study of Adolescent Health (Add Health), which includes the nationwide longitudinal data on adolescents in the US over approximately seven years. Although the Framingham study has a much longer time horizon and focuses on adults, the Add Health data contain information on actual ‘best friends’ who are named by the respondent; this is arguably a more appropriate peer than the contact person in the Framingham data. Cohen-Cole and Fletcher have first estimated eqn [5] on the basis of the Add Health data in order to replicate the baseline findings of Christakis and Fowler (2007) that are based on the Framingham data. Interestingly, both papers, using different data of different age groups, arrive at point estimates for \(\delta\) from eqn [1] for the ‘peer effect’ of BMI of 0.05, meaning that a one unit increase in a friend’s BMI over time is correlated with a 0.05 unit increase in one’s own BMI. However, when
Cohen-Cole and Fletcher controlled for shared environmental factors such as school-fixed effects, the coefficient fell by approximately 40%, no longer being statistically significant. Thus, current evidence suggests that the empirical problems described above are problematic enough to reduce confidence in any peer effects in obesity resulting from this specific model.

In an attempt to further explore the potential upward bias in the Christakis/Fowler empirical model, Cohen-Cole and Fletcher (2008b) took an alternative approach. The authors asked the question: ‘Is the empirical model [5] so weak that it would produce estimates of peer effects in behaviors where the true peer effect should be zero?’, that is, the authors conducted a falsification test of the empirical model by showing that estimating eqn [5] with the Add Health data would also produce results suggesting ‘social contagion’ in outcomes that are unlikely to be contagious: acne, headaches, and height. Indeed, the estimates for peer effects in these health behaviors are in some case larger than the Christakis/Fowler estimates of peer effects in obesity. The results of the falsification exercise strongly suggest that the model is insufficiently specific to distinguish between true social effects and the alternative hypotheses as discussed above (e.g., endogeneity of friendships and exposure to shared environmental factors). As in previous work, Cohen-Cole and Fletcher (2008b) have shown that the magnitudes of the fictional social network effects are reduced and when shared environmental influences are controlled, these effects largely disappear.

Based on part on these findings, obesity and weight-related behaviors have been studied in several additional papers. Trogdon et al. (2008) use several empirical strategies to examine peer effects. They examine both grade-level peers, similar to the cross-cohort designs already discussed, as well as nominated friends. To control for shared environmental factors, the authors control for school-fixed effects. To address friendship selection and simultaneity bias, the authors use an instrumental variables strategy, where friend’s birthweight, weight of parents of friend, and other measures are used as instruments. The limitation with this approach is that it is unclear whether these variables are good instruments for friendship selection. It appears that the instruments have been mainly employed to reduce the importance of the simultaneity issue, though the instruments still need to be excludable from the equation determining one’s own weight. In addition to controlling for shared environmental influences, the authors use school-fixed effects to partially control for friendship selection. The implicit assumption with school-fixed effects is that within schools, friendships form randomly.

Like Trogdon et al. (2008), Renna et al. (2008) also use a single cross-section of the Add Health data to examine the correlation between own and friend’s weight outcomes; however, these two papers use different subsamples and Renna et al. focus only on nominated friends. Renna et al. use school-level fixed effects to control for shared environmental factors and also attempt to reduce the simultaneity issue with an instrumental variables approach. The authors also use the obesity status of parents of friends as instruments. To control for selection of friends, the authors include additional control variables and acknowledge that the estimates are likely to be biased upward. The authors find evidence for peer effects for both genders in the baseline models, whereas only females in the IV models, although the point estimates are very similar. Overall, these papers are suggestive of peer effects but are unable to control for the empirical issues necessary to make the evidence more conclusive.

However, three recent papers have attempted to overcome the methodological issues with the above papers by pursuing alternative research designs. Yakusheva et al. (2011) use the roommate design described above with females from a private Midwestern university. The authors show negative correlations between having a heavy roommate and own weight outcomes. Carrell et al. (2011) stretch the literature further into the outcome of physical fitness by using random assignment to squadrons in the US Air Force Academy in order to show that squadronmates’ level of physical fitness is highly correlated with one’s own fitness. Finally, using a new instrumental variable strategy (characteristics of friend of peer), a so-called ‘friend of friend’ instrument pioneered by Bramoulle et al. (2009) and Fortin and Yaceck (2011) show some evidence of peer effects in fast-food consumption. Although these papers have considerably strengthened the research designs from past work and have extended the set of health behaviors, additional work is needed to further understand the potential for whether obesity is indeed ‘socially contagious.’ This work requires different (and hopefully more representative) samples and further replication.

In addition to weight-related outcomes, the literature examining peer influences in health outcomes has also begun to examine the realm of mental health. Although some older papers have attempted to examine social influences on suicidal behaviors, this literature is yet to incorporate newer and more rigorous research designs. Hence, the existence of peer effects is still uncertain. However, other measures of mental health have been explored recently. Eisenberg et al. (2011) have applied the roommate design to a variety of anxiety and depressive symptoms using a sample of freshman college students from two universities. The authors find no evidence of peer influence in measures of happiness. However, symptoms of anxiety appear to be correlated between roommates and there is some suggestive evidence of depressive symptoms being correlated between male roommates. See also Fletcher (2010c) for evidence that classmate mental health may reduce school performance.

Considering that research has expanded the domain of health behavior under study, new directions have been adopted in empirical methods on the basis of nonexperimental data. For example, a new direction in the study of social networks with implications for the study of health is the analysis of interdependent duration decisions. Because many health outcomes and behaviors have important time components such as smoking and drinking histories, utilizing new methods in this area could prove useful. The current state of the art includes the theoretical framework as outlined in Brock and Durlauf (2008) as well as the empirical applications of de Paula (2009) and de Paula and Honore (2010). Likewise, Fletcher and Ross (2012) have attempted to combine a control function approach with a cross-cohort design (as outlined above) to estimate the effects of best friend’s smoking and drinking behaviors on individual health choices. Work by Yves Zenou and colleagues have accumulated a set of papers that build a game theoretic model of network formation with interesting empirical implications (e.g., Calvó-Armengol et al., 2009).
In addition to these new research methods, there are also new data opportunities as well as research design opportunities emerging. Mayer and Puller (2008) leverage data from the social networking website Facebook.com in order to examine correlations between friends’ health behaviors, but they do not focus on causal inference. Mapping friendship networks through the use of cell phone usage information may also transform our ability to construct and track social networks in the future (Eagle et al., 2009). However, these new data sources will not alleviate the need to confront the difficulties of estimating empirical models of social influence.

Conclusions

This article briefly outlines some of the history, empirical challenges, current research and controversies, and directions for the future of this research area. Indeed, it is important to point out that this article being necessarily unexhaustive, does not cover important research areas that share many of the same issues described herein. Perhaps most obvious are the exclusion of the neighborhood effects literature that focuses on health outcomes and the emerging literature that examines potential peer effects in other health related areas such as doctors’ prescribing patterns and medical technology adoption. These important areas are beyond the scope of this article, which focuses only on the peer effects in health behaviors.

Research in the economics literature examining the effects of peers on health behaviors is now more than two decades old. There has been impressive progress as well as a stable set of challenges still not fully resolved. There has been a broadening of the set of behaviors and outcomes under consideration including weight and mental health besides the use of quasi-experimental research designs for additional outcomes of interest. In contrast, the growing volume of data on peer influence, especially online, has not been met by new research designs and methodologies that can produce entirely convincing results. This will be an important challenge to the researchers’ work of expanding the literature on peer effects in health behaviors.

Acknowledgments

This article extends similar summaries of current work on peer effects/social network effects in health behaviors found in Fletcher (2010a, 2011a). The author thanks the Robert Wood Johnson Foundation Health & Society Scholars program for its financial support.

See also: Smoking, Economics of

References


**Further Reading**