Identification and Estimation of Social Interaction-Based Models: A Changes-in-Changes Approach with an Application to Adolescent Substance Use

Muzhe Yang[*]

Department of Agricultural and Resource Economics

207 Giannini Hall #3310

University of California, Berkeley, CA 94720-3310

Abstract

This paper outlines a method for detecting and assessing the strength of social interactions through a changes-in-changes design. The proposed approach is based on a linear-in-means model and aims to resolve the "reflection problem", unobserved heterogeneities and endogenous group formation that plague identification of social interactions. Using longitudinal data from Add Health with rarely collected information on peer group's composition, we explore an exogenous variation in peer's drug use induced by a "mover friend" that occurs between Add Health's survey periods. This quasi-experiment shares a similar nature of a policy intervention of removing drug-user friends from a peer group. Such treatment-control group differences together with changes over time form the basis of our changes-in-changes design. Our study confirms a strong endogenous effect, which in turn motivates a "social multiplier", both of which are large enough to be relevant and are well worth attention to policy makers, researchers, health-care providers and educators for better understanding of how to protect young people and secure our future.

Key words: social interactions, linear-in-expectations, linear-in-means, difference-in-differences, changes-in-changes.

It is widely believed that adolescents' behaviors, particularly heath-related ones, are easily influenced by others, especially by their friends, during this rapid transition period from children to adults. How do peers affect one another, and why do such influences exist? Conceptual analysis of peer effects, or known as social or non-market interactions, has been thoroughly discussed by Manski (2000) and Glaeser and Scheinkman (2000). However, "there is little reason why a skeptic should be persuaded to change his mind by the statistical evidence [on social interactions] currently available." (Durlauf 2002). The weak state of empirical research on social interactions, to a large degree, is due to a lack of adequate data.

The National Longitudinal Study of Adolescent Health (Add Health), a school-based survey initiated in 1994 and by far the most comprehensive survey of adolescents ever undertaken in the United States, collects rich data on students in grades 7 through 12 and follows them through young adulthood (Chantala, Tabor et al. 1999; Chantala 2003). A key feature of Add Health is its precise measure of peer group composition on individual levels, which is rarely seen in other surveys targeting youth. Two consecutive waves of data (wave I and II) were collected between 1994 and 1996, which is followed by a third wave five years later. This longitudinal nature provides sources of identifying peer influences.

Using Add Health, I aim to identify and estimate peer effects of adolescents' health-related behaviors—substance use. To isolate peer effects from other factors, my identification strategy, different from leading alternatives, rests upon a spatial autoregressive model (SAR, Lee 2006) and exogenous variations in peers' behavior triggered by a quasi-experiment—a substance user friend moving away between wave I and II. Combining with two waves of data, I use a difference-in-differences (DID) design together with its generalized version—changes-in-changes (CIC, Athey and Imbens 2006)—to accommodate the possibility of sorting behavior—the group with

3

higher expected gain receiving the treatment. The goal is to identify peer effects and treatment effect under certain conditions, which can be seen correspond to an intervention of removing of a drug-user from his or her own peer group. Identification of these two parameters leads to constructing a "social multiplier", which bears rich policy implication.

The identification strategy, detailed below, to my knowledge, has never been used in identifying social interactions. It disentangles peer effects from sorting and unobserved heterogeneities through the idea of DID or CIC. Its implementation hinges upon measures of peer group's composition, sufficient variations in group sizes and data collected under a longitudinal design. All of these requirements are uniquely met by Add Health.

First, we place the research question back to where it comes from with a more proper framework. Adolescents spend more time with their closest friends, and these groups are usually not big. Considering substance use, it is more likely that observed actions of just being "cool" or actual "pursuits" of happiness seen from peers, rather than expected behaviors, that influence individual decision-making. In this sense, the linear-in-means model (Lee, 2006) is more appropriate and therefore adopted. This distinguishes my study from current empirical studies on social interactions, most of which are based on linear-in-expectations models (Manski, 1993).

Second, we utilize information of peer groups' composition provided by Add Health, which is rarely collected in similar longitudinal studies. We resolve the "reflection problem" (Manski, 1993) on the basis of linear-in-means model, thanks to the group size variations.

Third, we explore a source of exogenous changes in peers' drug use induced by a "mover friend" experiment occurring between Add Health's survey periods, which shares a similar nature of a policy intervention which removes drug-user friends from

a peer group. Such treatment-control group differences in combination with changes over time form the basis of a DID or CIC design.

Fourth, the "mover friend" experiment is shown to be randomly assigned conditional on peer groups' observable characteristics. This allows the use of treatment-control group difference to purge unobserved heterogeneities on the peer group's level. Disentangling endogenous effects from sorting and unobserved heterogeneities is resolved through the idea of DID or CIC.

Fifth, with the aid of this quasi-experiment, interacting the treatment group indicator with time gives an estimate of average treatment effect. We further relax the mean independence (Abadie 2005) restriction implied in DID and allow different quantiles in the distribution of group unobserved heterogeneities to interact with time. We then use CIC, and obtain estimates of average effects of treatment on both treatment and control groups, which imply a potential optimal policy intervention.

Sixth, combining CIC and DID treatment effect estimates with the estimate of endogenous effects, we next quantify a "social multiplier" (Glaeser, Sacerdote et al. 2003). This is also a case of treatment with "spillover effects". We hereby attempt to relax the stable-unit-treatment-value-assumption (SUTVA) (Rubin 1980) key to the current treatment effects literature.

Seventh, inferences on endogenous effects and various treatment effects are validated through bootstrap, which is currently underutilized (Cameron and Trivedi 2005). We correct the size and compute the power of testing these parameters of policy interest. We also utilize inverse power summary measure to facilitate interpretation and make valid inference when tests fail to reject the null hypothesis, which is widely neglected in applied work (Andrews 1989). This is an effort to address the weak state of inference in empirical studies (Ziliak and McCloskey 2004).

The rest of this article is organized as follows. We first give an overview of Add

Health study. Next, we discuss models of social interactions, identification strategies, estimation and inference procedures. Concluding remarks are provided in the end.

A Brief Overview of Add Health

As well noticed, the credibility of empirical studies on social interactions, to a large extent, depends on the availability of appropriate data. It is ideal to have a wealth of matched information on adolescents' characteristics, school information, family backgrounds and, most importantly, peer group composition from which we know who interacts with whom.

Not the best, but probably better than many other alternatives or nothing, Add Health is believed to enhance the quality of empirical research on social interactions through a detailed measure of friend network. Add Health is a school-based study of 80 high schools and 52 middle schools sampled from the United States.

Full details on Add Health are provided by the Carolina Population Center of the University of North Carolina at Chapel Hill, from which the following information is drawn.

> "(Add Health) is a nationally representative study that explores the causes of health-related behaviors of adolescents in grades 7 through 12 and their outcomes in young adulthood. Add Health seeks to examine how social context (families, friends, peers, schools, neighborhoods, and communities) influence adolescents' heath and risk behaviors."
>
> "Initiated in 1994 under a grant from the National Institute of Child Health and Human Development (NICHD) with co-funding from 17 other federal agencies, Add Health is the largest, most comprehensive survey of adolescents ever undertaken. Data at the individual, family, school, and

6

community levels were collected in two waves between 1994 and 1996. In 2001 and 2002, Add Health respondents, 18 to 26 years old, were re-interviewed in a third wave to investigate the influence that adolescence has on young adulthood."

Apart from the in-school interview conducted from September 1994 to April 1995, three waves of in-home survey data are available. A distinguishing feature of Add Health is its information collected on peers' network. In wave I, adolescents were asked to nominate up to five male friends and five female friends whose identification numbers make it possible to construct exact and meaningful reference groups where actual interactions occur. The wave I in-home survey collected detailed information on illegal drug use, based on the question "during the past 30 days, how many times did you use illegal drugs, including marijuana, cocaine, inhalants, LSD, PCP, ecstasy and heroin". About 14,000 adolescents were re-interviewed in wave II one year later with nearly identical survey questionnaires. There are about 5,000 adolescents who were interviewed in wave I not in wave II. Those are the ones either over-sampled in wave I (e.g., disabled or twins) or the ones who were in 12th grade in wave I and graduated from secondary school by wave II. From August 2001 to April 2002, wave III was conducted to re-interview all the wave I respondents who could be located, which resulted in a sample size of over 15,000. For our empirical study, we construct a longitudinal data set based on wave I and wave II in-home interviews including individual's self-reported reference group based on wave I's in-school and in-home interviews.

Issues on Identifying Social Interactions

Incisively pointed by Manski (2000), social interactions, not mediated via market, take effect through two different channels—expectation and preference. These two channels usually intertwine with each other. Herd behaviors occur usually when something becomes trendy or when it becomes a cheaper way to collect information. When there is less private or less asymmetric information, social interactions are more likely to be on the basis of aligned preferences. Inside the feedback process, and in this sense, the size of a peer group where interactions take place plays a key role in determining which channel will be muted. For a small group (among closest friends) information tends to be complete and outcomes within a group can be interpreted as Nash equilibria. In contrast, loss of information will increase as the group gets large, and expectation interactions become the dominating feature. This generates outcomes sharing the same nature as Bayesian Nash equilibria. Therefore, whether interactions take place on the basis of expectation or revealed preference via observed actions leads to two types of econometric modeling—linear-in-expectations (Manski 1993; Graham and Hahn 2005) and linear-in-means (Davezies, d'Haultfoeuille et al. 2006; Lee 2006).

The linear-in-expectations model at present is still the workhorse of many empirical studies. It is easy to estimate and interpret results. Apart from its emphasis on expectation interaction, such modeling becomes appropriate when the group size is large or the reference group can not be specified precisely. The linear-in-means model, instead, based on a spatial autoregressive model (SAR, Lee 2006), departs from the Manski (1993) model by measuring peer variables as spatially weighted averages of observed peer outcomes and characteristics instead of expectations. It suits small group and has some advantages over linear-in-expectation models in identifying

social interactions. It requires information on group size, and sufficient variations in the size.

Issues on identifying social interactions have been well framed, thanks to the seminal work by Manski (1993). Explanation of social interactions henceforth manages to proceed after theoretical descriptions or models. Such assessment requires differentiating three effects, which are defined by Manski (1993, pp. 532–533):

> "*Endogenous effects*, wherein the propensity of an individual to behave in some way varies with the prevalence of that behavior in the group;
>
> *Exogenous (contextual) effects*, wherein the propensity of an individual to behave in some way varies with the distribution of background characteristics in the group;
>
> *Correlated effects*, wherein individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments."

Identifying peer effects, especially disentangling endogenous effects from contextual and correlated effects, has become the main focus on empirical studies since then. Without direct intervention like controlled experiments, most observational studies rely on certain natural experiments to resolve confounding factors. Such confounding factors are pervasive on both group and individual level. The former creates the "reflection problem" coined by Manski (1993) and group level unobserved heterogeneities. The latter will induce sorting behavior and endogenous group formation. Identifying social interactions demands disentangling endogenous effects from individual sorting, contextual effects and unobserved group level heterogeneities (Sampson, Morenoff et al. 2002). In the case of social experiments, such as the Tennessee class size reduction experiment Project STAR, identification results can be robustified to

group level heterogeneity and sorting using second-order conditional moment restrictions (Graham 2006). However, contextual effects are still necessarily assumed away with this excess-variance-contrasts approach.

Contextual (exogenous) v.s. endogenous effects

Distinguishing contextual effects from endogenous effects actually have not been confronted directly in empirical studies. Such inseparability is due to the intrinsic collinearity between a group averaged choice and other covariates in individual behaviors, the so-called "reflection problem", because linear-in-expectations models, if adopted, by itself specifies a group averaged outcome as a linear transformation of group averaged characteristics. However, in discrete choice models with expectation interactions (Brock and Durlauf 2001; Durlauf 2001; Brock and Durlauf 2002; Brock and Durlauf 2004; Durlauf and Cohen-Cole 2004), the self-consistent expectation formation of a group choice, featuring fixed points derived from (Nash) equilibrium conditions, grants a critical nonlinearity between group averaged characteristics and group averaged behavior to circumvent this "reflection problem". However, in these discrete choice models with social interactions, group level heterogeneities, which confound endogenous effects as well, have not been dealt with directly. In addition, estimating such models which imply multiple equilibria is often intractable (Bisin, Moro et al. 2002; Cooper 2002).

The crux of identification problems in linear-in-expectations models partially lies in the model itself. Linear-in-expectations models treat within interactions symmetrically among group members. This makes it hard to simultaneously control for group heterogeneities and group expected outcomes, which is similar to the case that no more than one individual "fixed effects" can be identified in a panel setting. It

is noticeable that introducing between-group effects (Cohen-Cole 2006) may help to overcome these obstacles, but in the meantime this approach has shifted the estimand via including between-group complementarity, which differs from the usual observation that individuals interact with one another within a group, but not with members in other groups.

Unobserved group heterogeneities v.s. endogenous effects

As Manski (1993) examines this "reflection problem", it is found that identification of endogenous effects in liner-in-expectations models is not possible unless researchers have prior information which can specify the composition of reference groups. In this sense, a relatively tractable question would be directed to preference interactions within small groups where observed actions instead of their expectations generate feedback processes. This idea is corresponding to linear-in-means models pioneered by Lee (2006), which share the same nature as spatial autoregressive models (SAR). The identification is achieved in linear-in-means models through variations in known group sizes and asymmetric responses to peers within a group given that interaction is based on observed actions instead of expectation. The group size variation here provides the crucial nonlinearity to separately identify exogenous and endogenous effects, and it also allows for group fixed effects to pick up unobserved group heterogeneities, or termed correlated effects. Explicit conditions and identification results in this type of models have been established in Davezies, d'Haultfoeuille et al. (2006). And estimating these models has been discussed in Lee (2006), while the state of art is relatively limited to parametric approach (conditional maximum likelihood) or restricted by existence of valid instruments if semiparametric estimator (such as two-stage least squares) is preferred.

Although linear-in-means models enjoy certain advantages over linear-in-expectations models, selecting between these competing models still relies on the nature of interactions. If expectation interactions are the main form and even if exogenous effects can be properly excluded through quasi-experiments of sorts, identification in linear-in-expectations models will still be threatened by unobserved group heterogeneities. In cross-sectional settings, we may consider a quasi-panel approach (Graham and Hahn 2005) to isolating endogenous effects through excess between-group variations relative to within-group variations. But such identification strategy necessarily hinges upon valid instruments to account for measurement errors that come from the gap between expected and averaged group outcomes. The required number of instruments grows at the same rate of the number of groups, so identification, in addition, may suffer from the many-instruments problem. Compared with cross-sectional data, panel data may greatly facilitate identification by allowing of time-invariant factors and providing exclusion restrictions (Brock and Durlauf 2001) provided that within-group variations are sufficient.

Endogenous group memberships

However, the majority of studies, both cross-sectional (Alexander, Piazza et al. 2001) and longitudinal (Bauman, Carver et al. 2001; Haynie 2002) conclude that large endogenous effects are probably biased upward due to lack of control for self-selection and unobserved individual heterogeneities. It is arguably the best way to exclude confounding factors is through exogenous intervention, and random assignment hereby sets a "gold standard". With the aid of randomization, unobservables can be purged across groups, and sorting behavior can be avoided (Zimmerman 2003). But with observational studies of limited sources of exogenous variations, it is difficult to si-

12

multaneously deal with both group and individual unobserved heterogeneities with the latter generating endogenous group memberships.

Self-selection induces endogenous group formation (Evans, Oates et al. 1992). The endogenous choice of residential area plagues identification of neighborhood effects. Studies that attempt to measure the impact of neighborhoods on children's outcomes are susceptible to bias because families choose where to live. As a result, the effect of family unobservables, such as importance parents place on their children's welfare, and other unobservables that are common to geographically clustered households, may be mistakenly attributed to neighborhood influences. To combat such sorting behavior, panel data are often sought with various fixed effects having been explored to pick up key time-invariant factors. One approach is to use a sibling fixed effects model (Aaronson 1998), controlling for unobserved family characteristics that influence neighborhood choices. Such sibling fixed effects are expected to capture latent factors associated with neighborhood choice which do not vary across siblings. It is hoped that family residential differences provide a source of variations concerning neighborhood background that is free of the family-specific heterogeneity biases associated with neighborhood selection.

If a large set of observables is available, a practical way may be just to "kitchensink" regressions. It has been shown that standard regression models are also sensitive to the individual and family characteristics for which one controls, with strong effects when no individual and family characteristics are controlled and smaller and often non-significant effects when an extensive set of individual and family attributes are controlled (Ginther, Haveman et al. 2000). Returning to panel settings, some study (Mas and Moretti 2006) exploits fixed effects estimates as individual permanent characteristics and sorting can be presumably excluded conditional on these permanent characteristics. However, these nuisance parameters actually cannot be consistently

estimated unless the time span of panel data approaches infinity at a proper speed with the cross-sectional dimension.

A recent study on endogenous memberships (Zanella 2004) allows for equilibrium reference group formation and explores the driving force of stratified equilibria. However, such structural analyses of the reference group formation and social interactions require either the same agents or aligned preferences, and these requirements usually will be easily violated in parent-offspring contexts.

Another problem with group memberships is that many empirical studies make bold presumptions on group composition. Researchers know a priori the relevant social group within which meaningful (or hypothetical) interactions take place. Identification is essentially assumed to be obtained via diminished credibility.

Models of Social Interactions

Understanding the nature of interactions and defining social reference groups are prerequisites to analyzing peer effects. Just from an identification perspective, linear-in-means models are preferred because of the advantages over linear-in-expectations models in resolving the "reflection problem". However, such preference should not be reasoned only on the basis of pragmaticism. It can be shown that this choice is actually better justified for small group interactions based upon observed actions.

To fix ideas, we denote outcomes by y, individual characteristics by a vector $\mathbf{x}_1$, peer group characteristics by a vector $\mathbf{x}_2$, peer group by r ($r = 1, \cdots, R$), size of r-th peer group by m_r, individuals in r-th peer group by i ($i = 1, \cdots, m_r$), individual level heterogeneity by a single index α_{ri}, peer group level heterogeneity by a single index α_r and idiosyncratic disturbance by ϵ_{ri}.

First, consider the following individual payoff, additively separable in private and

social utilities, at a given time period:

(1)

$$\pi_{ri}(y_{ri}|\mathbf{x}_{1ri}, y_{r,-i}, \mathbf{x}_{2r,-i}, \alpha_{ri}, \alpha_r, m_r, \epsilon_{ri}) = u(y_{ri}|\mathbf{x}_{1ri}, \alpha_{ri}, \epsilon_{ri}) + v(y_{ri}|y_{r,-i}, \mathbf{x}_{2r,-i}, \alpha_{r,-i}, \alpha_r, m_r)$$

Impose some parametric assumptions, satisfying common regularity conditions:

(2)
$$u(y_{ri}|\mathbf{x}_{1ri}, \alpha_{ri}) = -\frac{1}{2}y_{ri}^2 + y_{ri}(\mathbf{x}'_{1ri}\lambda_1 + \alpha_{ri} + \epsilon_{ri})$$

(3)
$$v(y_{ri}|y_{r,-i}, \mathbf{x}_{2r,-i}, \alpha_{r,-i}, \alpha_r, m_r) = y_{ri}\left(\lambda_0\frac{1}{m_r-1}\sum_{j\neq i}^{m_r} y_{rj} + \frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\mathbf{x}'_{2rj}\lambda_2 + \frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\alpha_{rj} + \alpha_r\right)$$

Payoff maximization gives:

(4) $$y_{ri} = \lambda_0\frac{1}{m_r-1}\sum_{j\neq i}^{m_r} y_{rj} + \mathbf{x}'_{1ri}\lambda_1 + \alpha_{ri} + \frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\mathbf{x}'_{2rjt}\lambda_2 + \frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\alpha_{rj} + \alpha_r + \epsilon_{ri}$$

Next, consider employing this optimal decision rule under the following two situations.

Small group interactions with complete information

With complete information, α_{ri}, $\alpha_{r,-i}$ and α_r are "common knowledge" to each individual, so $\alpha_{r,-i}$ can be absorbed into α_r, denoted by $\widetilde{\alpha}_r$. A Nash Equilibrium gives:

(5)
$$\begin{aligned}
\mathbb{E}(y_{ri}|x_{ri}, y_{r,-i}, x_{r,-i}, \alpha_{ri}, \alpha_r, m_r) &= \lambda_0\frac{1}{m_r-1}\sum_{j\neq i}^{m_r} y_{rjt} + \mathbf{x}'_{1ri}\lambda_1 \\
&+ \frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\mathbf{x}'_{2rjt}\lambda_2 + \alpha_{ri} + \widetilde{\alpha}_r
\end{aligned}$$

where $\widetilde{\alpha}_r = (m_r-1)^{-1}\sum_{j\neq i}^{m_r}\alpha_{rj} + \alpha_r$

Equation (5) justifies a linear-in-means model:

$$
(6) \qquad y_{ri} = \lambda_0 \frac{1}{m_r-1} \sum_{j\neq i}^{m_r} y_{rjt} + \mathbf{x}'_{1ri}\lambda_1 + \frac{1}{m_r-1}\sum_{j\neq i}^{m_r} \mathbf{x}'_{2rj}\lambda_2 + \alpha_{ri} + \widetilde{\alpha}_r + \epsilon_{ri}
$$

Large group interactions with incomplete information

With incomplete information, $\alpha_{r,-i}$, peers' "type" can not be directly observed, though inferred, by each individual, while peer group's "type" is known to each member of the group so that α_r observed. Therefore, we have the following Bayesian Nash Equilibrium:

$$
(7) \qquad \mathbb{E}(y_{ri}|\mathbf{x}_{1ri},\alpha_{ri},\alpha_r) = \mathbb{E}_r[\mathbb{E}(y_{ri}|\mathbf{x}_{1ri}, y_{r,-i}, \mathbf{x}_{2r,-i}, \alpha_{ri}, \alpha_{r,-i}, \alpha_r, m_r)]
$$

$$
\begin{aligned}
(8)\,\mathbb{E}(y_{ri}|\mathbf{x}_{1ri},\alpha_{ri},\alpha_r) \;=\;& \lambda_0\frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\mathbb{E}_r(y_{rj}|\mathbf{x}_{2rj},\alpha_{rj},\alpha_r) + \mathbf{x}'_{1ri}\lambda_1 + \alpha_{ri}\\
&+\frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\mathbb{E}_r(\mathbf{x}_{2rj}|\alpha_r)'\lambda_2 + \frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\mathbb{E}_r(\alpha_{rj}|\alpha_r) + \alpha_r
\end{aligned}
$$

Under rational expectation or self-consistency, we have:

$$
(9) \qquad \mathbb{E}(y_{ri}|\mathbf{x}_{1ri},\alpha_{ri},\alpha_r) = \mathbb{E}_r(y_{rj}|\mathbf{x}_{2rj},\alpha_{rj},\alpha_r) = \mathbb{E}(y_r|\mathbf{x}_r,\alpha_r)
$$

If the group size (m_r) is big, then

$$
\begin{aligned}
(10)\frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\mathbb{E}_r(y_{rj}|\mathbf{x}_{2rj},\alpha_{rj},\alpha_r) \;\doteq\;& \frac{1}{m_r}\sum_{i=1}^{m_r}\mathbb{E}_r(y_{rj}|\mathbf{x}_{2rj},\alpha_{rj},\alpha_r) \equiv \mathbb{E}(y_r|\mathbf{x}_{2r},\alpha_r)\\
\frac{1}{m_r-1}\sum_{j\neq i}^{m_r}\mathbb{E}_r(\mathbf{x}_{2rj}|\alpha_r) \;\doteq\;& \frac{1}{m_r}\sum_{i=1}^{m_r}\mathbb{E}_r(\mathbf{x}_{2ri}|\alpha_r) \equiv \mathbb{E}(\mathbf{x}_{2r}|\alpha_r)
\end{aligned}
$$

So, (8) gives:

$$(11) \quad \mathbb{E}(y_{ri}|\mathbf{x}_{1ri}, \alpha_{ri}, \alpha_r) = \lambda_0 \mathbb{E}(y_r|\mathbf{x}_{2r}, \alpha_r) + \mathbf{x}'_{1ri}\lambda_1 + \alpha_{ri} + \mathbb{E}(\mathbf{x}_{2r}|\alpha_r)'\lambda_2 + \widetilde{\alpha}_r$$

$$\text{where } \widetilde{\alpha}_r = \frac{1}{m_r - 1}\sum_{j \neq i}^{m_r} \mathbb{E}_r(\alpha_{rj}|\alpha_r) + \alpha_r$$

Equation (11) corresponds to a linear-in-expectations model:

$$(12) \quad y_{ri} = \lambda_0 \mathbb{E}(y_r|\mathbf{x}_{2r}, \alpha_r) + \mathbf{x}'_{1ri}\lambda_1 + \mathbb{E}(\mathbf{x}_{2r}|\alpha_r)'\lambda_2 + \alpha_{ri} + \alpha_r + \epsilon_{ri}$$

Linear-in-expectations or Linear-in-means

Place the research question back to where it comes from: adolescents spend more time with their closest friends. Such groups are usually not big. Considering substance use, it is more likely that observed actions of "being cool" or actual pursuits of "happiness" from peers, rather than expected behaviors, that play key roles for individuals' decision-making. Besides, within a small group, among closest friends, each individual's "type" is probably "common knowledge". In this sense, the linear-in-means model better accords with the nature of interaction of our interest and therefore adopted. This distinguishes our study from current empirical studies on social interactions, most of which use linear-in-expectations models as the workhorse.

Empirical implication and relevance

To fix ideas, we define some useful matrices and a vector of 1's:

$$(13) \qquad \boldsymbol{\iota}_{m_r} = (1, \cdots, 1)'_{m_r \times 1}$$

$$(14) \qquad \underset{m_r \times m_r}{W_r} = \frac{1}{m_r - 1}\left(\boldsymbol{\iota}_{m_r}\boldsymbol{\iota}'_{m_r} - I_{m_r}\right) \text{ (spatial weighting matrix)}$$

$$(15) \qquad \underset{m_r \times m_r}{J_r} \equiv I_{m_r} - \frac{1}{m_r}\boldsymbol{\iota}_{m_r}\boldsymbol{\iota}'_{m_r} \text{ (within-group operator)}$$

And we derive the following properties:

$$(16) \qquad\qquad (1)\ \boldsymbol{\iota}'_{m_r} W_r = \boldsymbol{\iota}'_{m_r}$$

$$(17) \qquad\qquad (2)\ J_r \boldsymbol{\iota}_{m_r} = \mathbf{0}$$

$$(18) \qquad\qquad (3)\ J_r W_r = -\frac{1}{m_r - 1}J_r$$

A sample representation of (6) gives:

$$(19) \qquad\qquad \underset{m_r \times 1}{\mathbf{y}_r} = \lambda_0 W_r \mathbf{y}_r + X_{1r}\lambda_1 + W_r X_{2r}\lambda_2 + \boldsymbol{\alpha}_r + \boldsymbol{\iota}_{m_r}\widetilde{\alpha}_r + \boldsymbol{\epsilon}_r$$

Apply the within-group operator (15) to (19):

$$J_r \mathbf{y}_r = -\frac{\lambda_0}{m_r - 1}J_r \mathbf{y}_r + J_r X_{1r}\lambda_1 - \frac{1}{m_r - 1}J_r X_{2r}\lambda_2 + J_r \boldsymbol{\alpha}_r + J_r \boldsymbol{\epsilon}_r$$

$$(20)$$
$$J_r \mathbf{y}_r = \frac{m_r - 1}{m_r - 1 + \lambda_0}J_r X_{1r}\lambda_1 - \frac{1}{m_r - 1 + \lambda_0}J_r X_{2r}\lambda_2 + \frac{m_r - 1}{m_r - 1 + \lambda_0}J_r \boldsymbol{\alpha}_r + \frac{m_r - 1}{m_r - 1 + \lambda_0}J_r \boldsymbol{\epsilon}_r$$

If $\boldsymbol{\alpha}_r$ can be assumed away, then this within-group operator will help to separately identify endogenous effect (λ_0) and exogenous effect (λ_2) thanks to the nonlinearity

built-in through variations in group sizes (Davezies, d'Haultfoeuille et al. 2006; Lee 2006). The focus on identifying λ_0 is not solely for intellectual interest per se. It is because the feedback process depends on such endogenous effects that we have the "leverage" to manipulate aggregate equilibrium through exogenous interventions on individual levels. Unlike endogenous effects, neither exogenous nor correlated effects generate spillover effects and therefore they don't create this "social multiplier".

Consider applying a between-group operator:

$$
\begin{aligned}
(21) \quad \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\mathbf{y}_r &= \lambda_0 \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}W_r\mathbf{y}_r + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}X_{1r}\lambda_1 + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}W_rX_{2r}\lambda_2 \\
&\quad + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\boldsymbol{\alpha}_r + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\boldsymbol{\iota}_{m_r}\widetilde{\alpha}_r + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\boldsymbol{\epsilon}_r \\
\overline{y}_r &= \frac{1}{1-\lambda_0}\overline{\mathbf{x}}'_{1r}\lambda_1 + \frac{1}{1-\lambda_0}\overline{\mathbf{x}}'_{2r}\lambda_2 + \frac{1}{1-\lambda_0}(\overline{\alpha}_r + \widetilde{\alpha}_r) + \frac{1}{1-\lambda_0}\overline{\epsilon}_r
\end{aligned}
$$

On a peer group level, it is clearly seen that any exogenous interventions, whose effects will be magnified by $1/(1-\lambda_0)$, will possess multiplier effects as long as the endogenous effect is nonzero. Suppose that there is an exogenous treatment with effect τ. Such a hypothetical "treatment effect" in the presence of social interactions will be magnified to $\tau/(1-\lambda_0)$. We call this *social multiplier* of a particular intervention. It attempts to approximate a "general equilibrium" result instead of a partial one, which overlooks subsequent "ripple effects", and thus tends to underestimate the impacts from the intervention if positive spillovers persist.

Identification via linear-in-means models, although free of the "reflection problem", still need to deal with unobserved heterogeneities on both individual level and peer group level. The former induces sorting behavior, while the latter generates correlated effects (Manski 1993). Unlike studies resolving these issues through modeling, we aim to address them from the research design perspective. If unobserved

19

individual heterogeneities are time-invariant, then a two-period panel data can help to "difference out" such confounding factors. And, if certain "natural" perturbation occurs, generating a "treatment" group and a "control" group, then these two groups are likely to share similar characteristics. This implies that the peer group heterogeneities on this aggregate level can be reasonably "equalized" between treated and untreated groups, which therefore are likely to share a similar time trend as well. Differencing between the treated and untreated group therefore helps to purge unobserved group heterogeneities. In the end, this DID approach will sort out the endogenous effects. And, under certain circumstances, this DID estimate can be interpreted as the "treatment effect".

Add Health provides an excellent platform to implement this DID idea with its longitudinal design and a "mover friend" experiment. This quasi-experiment is of policy interest as well since it corresponds with an outside intervention of removing a substance user from a peer group. The treatment effect estimate can be further justified under weaker conditions using a CIC approach, which obtains counterfactual distributions of what would have happened in the absence or in the presence of treatment.

Identification

Similar to a spatial autoregressive (SAR) model, we use the following linear-in-means model already defined by (6) to specify structural parameters.

$$(22)\ y_{rit}\ =\ \lambda_0 \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} y_{rjt} + \mathbf{x}'_{1rit}\lambda_1 + \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \mathbf{x}'_{2rjt}\lambda_2 + \alpha_{ri} + \alpha_{rt} + \epsilon_{rit}$$
$$(i\ =\ 1, \cdots, m_r; t = 0, 1; r = 1, \cdots, R;\ \mathbb{E}(\epsilon_{rit}) = 0, \mathbb{E}(\epsilon_{rit}\epsilon_{rjt}) = 0, \forall i \neq j)$$

Note that the interdependence within a peer group is modeled in a "fixed effect" manner. After controlling for time-invariant individual heterogeneities and peer group time-varying heterogeneities, contemporaneous correlations across members in the same peer group vanish, namely no random effects. The trade-off between this second-order heterogeneity and the first-order one is currently ignored in this article.

Given the econometric model (22). We need to address two identification problems: (1) endogenous group membership due to personal preference α_{ri}, which is assumed time-invariant. This implies that individual preference for joining a certain group is stable over time; (2) peer group-level unobserved heterogeneity α_{rt}, which is allowed to be time-varying. This implies that the "correlated" effects can change over time.

Given the existence of time-invariant factors, a natural starting point is to difference out these nuisance parameters. Differencing (22) gives:

$$(23) \qquad \triangle y_{ri} = \lambda_0 \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \triangle y_{rj} + \triangle \mathbf{x}'_{1ri} \lambda_1 + \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \triangle \mathbf{x}'_{2rj} \lambda_2 + \triangle \alpha_r + \triangle \epsilon_{ri}$$

All structural parameters in this "differenced" model are generically identifiable under regularity conditions explicitly summarized by Davezies, d'Haultfoeuille et al. (2006). Estimation can be done by applying a "within-group" estimator to (23) with both parametric (conditional maximum likelihood, CML) and semiparametric (two-stage least squares, 2SLS) techniques. However, these proposals given by Lee (2006) still rely heavily on distributional assumptions in addition to the functional form restriction of the present econometric model and availability of valid instruments. The CML enjoys the efficiency in terms of the Cramer-Rao lower bound but less robust, while 2SLS essentially changes the estimand in the presence of heterogeneous responses to instruments.

Rather than seeking a mechanical way or an omnibus tool to solve this problem, we exploit a "natural" experiment which provides critical exclusion restrictions by way of a treatment-control comparison. These two "naturally" generated groups are likely to experience similar time trend if such an intervention comes in exogenously. In this sense, $\triangle \alpha_r$ can be purged out through the treatment-control difference under this "common trend" assumption. Therefore, we are just implementing a DID idea using time-differencing to exclude individual unobserved heterogeneity (a normalized single index) and treatment-control differencing to impose another exclusion restriction on peer group heterogeneities.

In the context of Add Health, we consider this quasi-experiment when a substance user friend moved away from his or her peer group between wave I and wave II, $g_{ri} \in G_{ri} = 1\{i$ has at least one substance user friend of peer group r moving away between two waves$\}$, and $t_{ri} \in T_{ri} = \{0, 1\}$. This includes the following cases: (1) moved to different tract but within the same county; (2) moved to different county but within the same state; (3) moved to different state. In this two-group-two-period setting, if there is no interaction effects between time and group, i.e., the change of outcomes in the control group over time can offer a counterfactual for the treated group in the absence of treatment, then the second-period "treatment" (relative to the first-period baseline treatment) can be properly defined as $I_{ri} = G_{ri} \cdot T_{ri}$, or equivalently, the treatment effect is the second-period treatment relative to the baseline treatment, $I_{ri} = 1\{G_{ri} = 1, T_{ri} = 1\}$. If the "common trend" assumption is plausible, we require that the group membership G_{ri} be at least randomly assigned on the peer-group level conditional on group-level observables, $\mathbf{x}_{2r}$. And, if this "mover" experiment is randomized at the peer-group level, conditional on peer group level observables, then peer-group level potential outcomes must be independent of this group membership. Equivalently, peer group level unobservables can be excluded through differencing

based upon this treatment group membership. The observed difference is therefore purely due to this membership assignment, and can be seen as the "normal difference". The "abnormal difference" arises over time will reveal the treatment effect since peer group level unobserved heterogeneities as confounders have been excluded.

Borrow notation from the potential outcome models (POM):

$$(24) \qquad (y_{0rt}, y_{1rt}) \perp G | \mathbf{x}_{2r}, T = t$$

$$(25) \qquad \Rightarrow \alpha_{rt} \perp G | \mathbf{x}_{2r}, T = t$$

Observed difference due to group membership assignment is thus not confounded by group level unobserved heterogeneities. This observed difference represents a "normal" or manipulable difference because it is cleaned from α_{rt}. Therefore,

$$y_{rit} | \alpha_{rt}, \mathbf{x}_{1ri}, \mathbf{x}_{2r}, G = g, T = t$$

$$(26) \qquad \sim y_{rit} | \mathbf{x}_{1ri}, \mathbf{x}_{2r}, G = g, T = t$$

$$\Rightarrow \alpha_{rt} \text{ properly excluded from the outcome}$$

We virtually argue that the difference between peer-group level unobserved heterogeneities averaged by the treatment-control group membership can be observed as or proxied by a "normal" difference induced by an exogenous manipulation which defines the treated and untreated group in the absence of treatment.

In summary, identification of social interactions in this article rests upon a linear-in-means model and a DID design. We first adopt the following linear specification,

and then extend it to a nonlinear setting where the treatment effect can be identified through CIC.

However, when the treated and untreated groups do not share a "common" trend, DID estimate doesn't have the meaning of treatment effects. This is likely even under random assignment since either treatment or control group may still experience different transient shocks over time which are not bound to be "common". In this case the interaction term between group membership and time indicator only helps to identify endogenous effects, and τ doesn't represent the effect from a manipulation and is not essentially of policy relevance. This limitation in some way results from the excessively restrictive linear setting. In contrast, CIC permits interactions taking places between time and different quantiles of individual unobservables across groups, so that counterfactuals in the absence of treatment are not necessarily restricted by using the expected time trend from the control group as what a conventional DID requires.

Endogenous effects

Identifying endogenous effects in linear-in-expectations model faces two main challenges: (1) unobserved heterogeneities; (2) the "reflection problem", namely the collinearity between exogenous effects and endogenous effects. Without imposing exclusion restrictions, (2) is impossible to circumvent unless discrete choice models are studied (Brock and Durlauf 2001). In our study, (2) is resolved by adopting a linear-in-means model, a model better capturing the nature of interactions. It has been shown that this model possesses crucial nonlinearities carried out by group sizes (Davezies, d'Haultfoeuille et al. 2006; Lee 2006). Unless all groups are of the same size or there is lack of variations in group sizes, where numerical problems arises due to

ill-conditioned observation matrices, disentangling endogenous effects from exogenous effects is generically feasible even when all members of the group are not observed.

Add Health provides valuable information in regard to friendship networks, sizes of peer groups and their variations. Given its study design, the composition of peer groups is directly surveyed, which, to a large extent, minimizes the measurement error and satisfies a prerequisite to linear-in-means models that the group size must be known. Table 2 provides details, from which we can see that these social reference groups are unanimously small in size. Nearly 90% of these groups consist of no more than three students. Considering the nature of these groups, interactions among closest friends are presumably influential and meaningful. In addition, small-group interactions are better suitable to linear-in-means models because with the case of large groups, "identification can be weak in the sense that the estimates converge in distribution at low rates." (Lee 2006).

Now consider how this DID approach assists identification of endogenous effects when treatment group assignment is independent of peer group-level unobserved heterogeneities conditional on certain peer group level observables:

$$(27) \quad y_{rit} = \lambda_0 \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} y_{rjt} + \mathbf{x}'_{1rit}\lambda_1 + \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \mathbf{x}'_{2rjt}\lambda_2 + \tau G_{ri}T_{ri} + \beta T_{ri} + \gamma G_{ri} + \epsilon_{rit}$$

This double differencing procedure will result in:

$$(28) \qquad \widetilde{y}_{ri} = \lambda_0 \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \widetilde{y}_{rj} + \widetilde{\mathbf{x}}'_{1ri}\lambda_1 + \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \widetilde{\mathbf{x}}'_{2rj}\lambda_2 + \tau + \widetilde{\epsilon}_{ri}$$

where $\widetilde{s} \equiv (s_{11} - s_{10}) - (s_{01} - s_{00})$ and $s_{gt} \equiv s|_{G=g,T=t}$.

Use a matrix representation:

$$(29) \qquad \widetilde{\mathbf{y}}_r = \lambda_0 W_r \widetilde{\mathbf{y}}_r + \widetilde{X}_{1r}\lambda_1 + W_r \widetilde{X}_{2r}\lambda_2 + \boldsymbol{\iota}_{m_r}\tau + \widetilde{\boldsymbol{\epsilon}}_r$$

A between-group estimator will reveal the "reflection problem":

$$\frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\widetilde{\mathbf{y}}_r = \lambda_0\frac{1}{m_r}\boldsymbol{\iota}'_{m_r}W_r\widetilde{\mathbf{y}}_r + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\widetilde{X}_{1r}\lambda_1 + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}W_r\widetilde{X}_{2r}\lambda_2 + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\tau + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\widetilde{\boldsymbol{\epsilon}}_r$$

$$\overline{\widetilde{y}}_r = \lambda_0\frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\widetilde{\mathbf{y}}_r + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\widetilde{X}_{1r}\lambda_1 + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\widetilde{X}_{2r}\lambda_2 + \tau + \frac{1}{m_r}\boldsymbol{\iota}'_{m_r}\widetilde{\boldsymbol{\epsilon}}_r$$

$$(30) \qquad \Rightarrow \overline{\widetilde{y}}_r = \overline{\widetilde{\mathbf{x}}}'_{1r}\frac{\lambda_1}{1-\lambda_0} + \overline{\widetilde{\mathbf{x}}}'_{2r}\frac{\lambda_2}{1-\lambda_0} + \frac{\tau}{1-\lambda_0} + \frac{\overline{\widetilde{\epsilon}}_r}{1-\lambda_0}$$

As (30) demonstrates, identifying λ_0 is not possible without additional exclusion restrictions, say $\lambda_2 = 0$, which is often imposed by empirical studies.

Instead, now consider a within-group estimator:

$$J_r\widetilde{\mathbf{y}}_r = \lambda_0 J_r W_r \widetilde{\mathbf{y}}_r + J_r X_{1r}\lambda_1 + J_r W_r X_{2r}\lambda_2 + J_r \boldsymbol{\iota}_{m_r}\tau + J_r\widetilde{\boldsymbol{\epsilon}}_r$$

$$J_r\widetilde{\mathbf{y}}_r = -\frac{\lambda_0}{m_r - 1}J_r\widetilde{\mathbf{y}}_r + J_r X_{1r}\lambda_1 - \frac{1}{m_r - 1}J_r X_{2r}\lambda_2 + J_r\widetilde{\boldsymbol{\epsilon}}_r$$

(31)
$$\Rightarrow \left(\frac{m_r - 1 + \lambda_0}{m_r - 1}\right)\left(\widetilde{y}_{ri} - \overline{\widetilde{y}}_r\right) = \left(\widetilde{\mathbf{x}}'_{1ri} - \overline{\widetilde{\mathbf{x}}}'_{1r}\right)\lambda_1 - \frac{1}{m_r - 1}\left(\widetilde{\mathbf{x}}'_{2ri} - \overline{\widetilde{\mathbf{x}}}'_{2r}\right)\lambda_2 + \left(\widetilde{\epsilon}_r - \overline{\widetilde{\epsilon}}_r\right)$$

(32)
$$\left(\widetilde{y}_{ri} - \overline{\widetilde{y}}_r\right) = \left(\widetilde{\mathbf{x}}'_{1ri} - \overline{\widetilde{\mathbf{x}}}'_{1r}\right)\frac{(m_r - 1)\lambda_1}{m_r - 1 + \lambda_0} - \left(\widetilde{\mathbf{x}}'_{2ri} - \overline{\widetilde{\mathbf{x}}}'_{2r}\right)\frac{\lambda_2}{m_r - 1 + \lambda_0} + \frac{m_r - 1}{m_r - 1 + \lambda_0}\left(\widetilde{\epsilon}_r - \overline{\widetilde{\epsilon}}_r\right)$$

It is clearly seen that the group size transforms a linear-in-means, which is linear in parameters, model into a nonlinear (in parameters) one that can be generically

identified under certain regularity conditions established in Davezies, d'Haultfoeuille et al. (2006). This model can be estimated by CML or 2SLS (Lee 2006) in principle.

We next focus on the role for the constant (intercept) τ, which "normalizes" the regression model (29). Some new notations are introduced.

$$
\begin{aligned}
\boldsymbol{\iota}_N &\equiv (1, 1, \cdots, 1)'_{N \times 1} = (\boldsymbol{\iota}'_{m_1}, \boldsymbol{\iota}'_{m_2}, \cdots, \boldsymbol{\iota}'_{m_R})'_{N \times 1} \\
\mathbf{y} &= (\mathbf{y}'_{m_1}, \mathbf{y}'_{m_2}, \cdots, \mathbf{y}'_{m_R})'_{N \times 1} \\
\underset{N \times N}{W} &\equiv
\begin{bmatrix}
\underset{m_1 \times m_1}{W_1} & 0 & \cdots & 0 \\
0 & \underset{m_2 \times m_2}{W_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \underset{m_R \times m_R}{W_1}
\end{bmatrix} \\
M_{0r} &\equiv I_{m_r} - \frac{1}{N} \boldsymbol{\iota}_{m_r} \boldsymbol{\iota}'_{m_r} \\
M_0 &\equiv I - \frac{1}{N} \boldsymbol{\iota}_N \boldsymbol{\iota}'_N \\
&=
\begin{bmatrix}
I_{m_1} & 0 & \cdots & 0 \\
0 & I_{m_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & I_{m_R}
\end{bmatrix}
- \frac{1}{N}
\begin{bmatrix}
\boldsymbol{\iota}_{m_1} \\
\boldsymbol{\iota}_{m_2} \\
\vdots \\
\boldsymbol{\iota}_{m_R}
\end{bmatrix}
\begin{bmatrix}
\boldsymbol{\iota}'_{m_1} & \boldsymbol{\iota}'_{m_2} & \cdots & \boldsymbol{\iota}'_{m_R}
\end{bmatrix}
\end{aligned}
$$

We have the following properties:

(33)
$$
(1) \ M_0 \boldsymbol{\iota}_N = \boldsymbol{\iota}'_N M_0 = 0
$$

(34)
$$
(2) \ M_{0r} W_r = W_r - \frac{1}{N} \boldsymbol{\iota}_{m_r} \boldsymbol{\iota}'_{m_r} \doteq W_r
$$

(35) $\qquad$ (3) $M_0 W\mathbf{y} = W\mathbf{y} - \boldsymbol{\iota}_N \overline{y}$

Now consider applying a "de-mean" operator to all peer groups:

$$
\begin{aligned}
\widetilde{\mathbf{y}} &= \lambda_0 W\widetilde{\mathbf{y}} + \widetilde{X}_1\lambda_1 + W\widetilde{X}_2\lambda_2 + \boldsymbol{\iota}_N\tau + \widetilde{\boldsymbol{\epsilon}} \\
M_0\widetilde{\mathbf{y}} &= \lambda_0 M_0 W\widetilde{\mathbf{y}} + M_0\widetilde{X}_1\lambda_1 + M_0 W\widetilde{X}_2\lambda_2 + M_0\boldsymbol{\iota}_N\tau + M_0\widetilde{\boldsymbol{\epsilon}} \\
\widetilde{\mathbf{y}} - \boldsymbol{\iota}_N\overline{\widetilde{y}} &= \lambda_0(W\widetilde{\mathbf{y}} - \boldsymbol{\iota}_N\overline{\widetilde{y}}) + (\widetilde{X}_1 - \overline{\widetilde{X}}_1)\lambda_1 + (W\widetilde{X}_2 - \boldsymbol{\iota}_N\overline{\widetilde{X}}_2)\lambda_2 + (\widetilde{\boldsymbol{\epsilon}} - \boldsymbol{\iota}_N\overline{\widetilde{\epsilon}})
\end{aligned}
$$

(36) $\quad \Rightarrow \ \widetilde{\mathbf{y}} = (1 - \lambda_0)(I - \lambda_0 W)^{-1}\boldsymbol{\iota}_N\overline{\widetilde{y}} + (I - \lambda_0 W)^{-1}(\widetilde{X}_1 - \overline{\widetilde{X}}_1)\lambda_1 +$

$\qquad\qquad (I - \lambda_0 W)^{-1}(W\widetilde{X}_2 - \boldsymbol{\iota}_N\overline{\widetilde{X}}_2)\lambda_2 + (I - \lambda_0 W)^{-1}(\widetilde{\boldsymbol{\epsilon}} - \boldsymbol{\iota}_N\overline{\widetilde{\epsilon}})$

This "de-mean" operator generating a nonlinearity in parameters, similar to Lee (2006), introduced by group sizes m_r's offers necessary conditions for identifying λ_0. Up to this stage, (36) is estimable via CML or 2SLS, which demands either distributional assumptions or valid instruments.

However, considering this DID design feature, if "white" noise exists, then we can estimate this linear-in-means model in the same spirit as we do in a spatial autoregressive model (SAR). Based on our previous arguments, confounding factors are controlled in a "fixed" effect manner, rather than resorting to "random" effects, which is essentially a trade-off between first-order and second-order heterogeneous responses that we wish to account for. Note that (27) is numerically equivalent to the following specification with individual fixed effects:

(37) $\quad y_{rit} = \lambda_0 \dfrac{1}{m_r - 1}\sum_{j \neq i}^{m_r} y_{rjt} + \mathbf{x}'_{1rit}\lambda_1 + \dfrac{1}{m_r - 1}\sum_{j \neq i}^{m_r} \mathbf{x}'_{2rjt}\lambda_2 + \tau I_{ri} + \alpha_{ri} + \alpha_{rt} + \epsilon_{rit}$

where $I_{ri} = G_{ri}T_{ri}$.

If we can accept the presumption that the dependence of y_{ri}'s over time is resolved by controlling for individual fixed effects (α_{ri}), and the contemporaneous independence of $y_{r \cdot t}$'s is achieved by conditioning on the peer group fixed effect (α_{rt}), then this model (37), or its equivalence (27), can be estimated without additional exclusion restrictions. We are hereby restricting the second-order heterogeneities embedded in the covariance structure in exchange for a flexible control for the first-order heterogeneities. The gain is less parametric assumptions for the data-generating process (DGP) at the cost of efficiency loss and changing estimand in the presence of heterogeneous responses. Acknowledging this trade-off, we estimate (37) through least squares to obtain consistent estimates. Given that the conditional variance-covariance of the disturbances may still possess a "cluster" feature, for robust and valid inferences, we use a block-bootstrap method to account for the clustering and panel data.

Social multipliers

Feedbacks arise from endogenous effects. Such potential "ripple effects" are of policy interest because the impact of certain intervention will be magnified, which bears rich implication for a benefit-cost evaluation. Rewrite the econometric specification (27) in a matrix representation.

$$(38) \qquad \mathbf{y} = \lambda_0 W \mathbf{y} + X_1 \lambda_1 + W X_2 \lambda_2 + \tau I + \beta T + \gamma G + \boldsymbol{\epsilon}$$

Take the mean, applying (16), and we will get the following:

$$(39) \qquad \begin{aligned} \frac{1}{N}\boldsymbol{\iota}'_N \mathbf{y} &= \lambda_0 \frac{1}{N}\boldsymbol{\iota}'_N \mathbf{y} + \frac{1}{N}\boldsymbol{\iota}'_N X_1 \lambda_1 + \frac{1}{N}\boldsymbol{\iota}'_N X_2 \lambda_2 \\ &\quad + \tau \frac{1}{N}\boldsymbol{\iota}'_N I + \beta \frac{1}{N}\boldsymbol{\iota}'_N T + \gamma \frac{1}{N}\boldsymbol{\iota}'_N G + \frac{1}{N}\boldsymbol{\iota}'_N \boldsymbol{\epsilon} \end{aligned}$$

Take the expectation condition on $I = 1$ and $I = 0$ respectively.

(40) $$\mathbb{E}(\overline{y}|I = 1) = \lambda_0\mathbb{E}(\overline{y}|I = 1) + \mathbb{E}(\overline{\mathbf{x}}_1)'\lambda_1 + \mathbb{E}(\overline{\mathbf{x}}_2)'\lambda_2 + \tau$$

(41) $$\mathbb{E}(\overline{y}|I = 0) = \lambda_0\mathbb{E}(\overline{y}|I = 0) + \mathbb{E}(\overline{\mathbf{x}}_1)'\lambda_1 + \mathbb{E}(\overline{\mathbf{x}}_2)'\lambda_2$$

Rearrange the difference between (40) and (41), we obtain a *social multiplier* defined as:

(42) $$\text{social multiplier} \equiv \mathbb{E}(\overline{y}|I = 1) - \mathbb{E}(\overline{y}|I = 0) = \frac{\tau}{1 - \lambda_0}$$

If "common trend" assumption in the DID literature between the treated and untreated group can be accepted, specifically, a "mover" intervention does not introduce additional confounding factors to the treatment group, and both groups experience similar changes between wave I and wave II of Add Health's survey periods, then τ bears the meaning of treatment effects which corresponds to removing a drug-user friend from a peer group. Furthermore, if this "mover" intervention is randomly assigned, then τ represents an average treatment effects (ATE); otherwise, τ can be only interpreted as the average effect of treatment on the treated group (ATT) under the "common trend" assumption and in the presence of heterogeneous treatment effects. If feedbacks are taken into account, then on an aggregate level, the expected mean outcomes will be changed by an overall treatment effect which is magnified by $1/(1 - \lambda_0)$ thanks to social interactions.

In order to obtain the full strength of such a social multiplier, we need to identify τ apart from λ_0.

Difference-in-differences (DID) v.s. Changes-in-changes (CIC)

The essence of these two double differencing approaches is to extract information about what would have happened to treatment group in the absence of the treatment from the control group. DID fails in the situation when there is an interaction between the treated group and the time period during which the treatment takes place (Meyer 1995). One way out of this threat is to select the control group as similar as possible to the treated group so that such an interaction is unlikely to happen.

Standard or conventional DID model treats time and group symmetrically. CIC allows for asymmetric impact from time and group, and therefore can deal with heterogeneous treatment effects which remain under randomization. In contrast to DID, CIC focuses on the distributions of outcomes, and is able to identify counterfactual distributions of what could have happened in absence of treatment and in presence of treatment (Athey and Imbens 2006). In this sense, CIC takes into account endogenous treatment selection based on anticipated incremental benefit of policy, which also provides a method to evaluate "optimal" policy adoption. In this article, we first obtain a DID estimate as a baseline, since standard DID is still able to identify homogeneous treatment effect. And then we apply the CIC estimator proposed by Athey and Imbens (2006) to address the following limitation inherited in the DID approach.

First, DID accommodates the first-order heterogeneities, but it ignores the effects of unobservables which can change over time, and therefore the mean-variance trade-off is neglected as well. Instead, CIC can separately identify the effects of treatment on the treated and the control group, which allows for much richer policy implications.

Second, if expected (mean) time trend differs between the treated and untreated groups at the time of pre-treatment or post-treatment, standard DID doesn't work. In

contrast, CIC relaxes the common trend assumption by only requiring the distribution of unobserved heterogeneities stays the same over time in the absence of treatment and by accommodating the interactions between time and quantiles of distributions of unobservables in both treated and untreated groups.

Third, conventional DID rules out certain heterogeneous effect due to its linearity and additivity restrictions. CIC works out in nonlinear settings by taking advantage of a monotonicity assumption of the latent production functions mapping unobservables to potential outcomes. It utilizes the relative ranking of pre-treatment outcomes in the treated group in reference to outcomes of the control group, and calculates the time effect according to that relative ranking. Therefore CIC allows different individuals in treated group have different time effects, and exploits the distribution of time effects to construct counterfactual distributions. In this sense, CIC considers the heterogeneities in time effects which is ignored by DID. And, DID assumes a homogeneous time effects that everyone, on average, experiences the same time trend.

In addition, CIC fits application with short time periods where compositions of treatment-control groups stay the same. Add Health again provides an excellent environment thanks to its longitudinal design and two waves of surveys conducted successively.

In the following section, we first estimate τ under conventional DID assumptions. Next we apply the CIC approach to identify τ under weaker conditions, which is also an effort to robustify the DID estimate. In the end, we compare the inference properties regarding these two estimators in terms of size and power.

Adjusting for covariates can be done either parametrically or nonparametrically. We hereby follows the procedure proposed by Athey and Imbens (2006), which can be seen as a semiparametric approach—adjusting for covariates parametrically and identifying the treatment effect nonparametrically. The idea is illustrated in the

following.

Given the econometric model (27):

$$y_{rit} = \lambda_0 \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} y_{rjt} + \mathbf{x}'_{1rit}\lambda_1 + \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \mathbf{x}'_{2rjt}\lambda_2 + \tau G_{ri}T_{ri} + \beta T_{ri} + \gamma G_{ri} + \epsilon_{rit}$$

We uses its equivalence as the parametric specification.

$$(43) \qquad y_{rit} = \lambda_0 \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} y_{rjt} + \mathbf{x}'_{1rit}\lambda_1 + \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \mathbf{x}'_{2rjt}\lambda_2 + D'_{ri}\delta + \epsilon_{rit}$$

$$D_{ri} = [(1 - T_{ri})(1 - G_{ri}), T_{ri}(1 - G_{ri}), (1 - T_{ri})G_{ri}, T_{ri}G_{ri}]'$$

The "residualized" outcome of an individual in the absence of intervention satisfies the following:

$$(44) \qquad \widetilde{y}^N_{rigt} \equiv h(u_{ri}, t) \equiv D'_{ri}\delta + \epsilon_{rit}$$

$$
(45) \qquad
\begin{aligned}
\widetilde{y}^N_{rigt} &\stackrel{d}{\sim} y^N_{ri}|G_{ri} &=& g, T_{ri} = t \\
\widetilde{y}^I_{rigt} &\stackrel{d}{\sim} y^I_{ri}|G_{ri} &=& g, T_{ri} = t \\
\widetilde{y}_{rigt} &\stackrel{d}{\sim} y_{ri}|G_{ri} &=& g, T_{ri} = t \\
u_{rig} &\stackrel{d}{\sim} u_{ri}|G_{ri} &=& g \\
g &\in& \{0,1\}, t \in \{0,1\}
\end{aligned}
$$

So (43) can be rewritten as:

$$(46) \qquad y_{rigt} \equiv h(u_{ri}, t) + \lambda_0 \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} y_{rjt} + \mathbf{x}'_{1rit}\lambda_1 + \frac{1}{m_r - 1} \sum_{j \neq i}^{m_r} \mathbf{x}'_{2rjt}\lambda_2$$

33

DID identifies τ as τ_{ATE}^{DID} under regularity conditions and in the presence of heterogeneous treatment effects.

Under CIC's regularity conditions, the effect of treatment on the treated group is identified by CIC as:

$$(47) \qquad \tau_{ATT}^{CIC} = \mathbb{E}(\widetilde{y}_{11} - \widetilde{y}_{11}^N) = \mathbb{E}(\widetilde{y}_{11}) - \mathbb{E}[F_{\widetilde{y},01}^{-1}(F_{\widetilde{y},00}(\widetilde{y}_{10}))]$$

Similarly, the effect of treatment on the control group is identified by the following:

$$(48) \qquad \tau_{ATC}^{CIC} = \mathbb{E}(\widetilde{y}_{01}^I - \widetilde{y}_{01}) = \mathbb{E}[F_{\widetilde{y},11}^{-1}(F_{\widetilde{y},10}(\widetilde{y}_{00}))] - \mathbb{E}(\widetilde{y}_{01})$$

The consistency and asymptotic normality of this covariance-adjusted estimator is established by Athey and Imbens (2006).

Estimation

Our empirical study is focused on adolescents' substance use, namely illegal drugs, where peer's influence is believed to play a key role in shaping individual's behavior. To explore peer effects among health-related behaviors, Add Health data have been deemed well-suited in many respects due to its longitudinal design and specific measures of peer groups. Several studies have examined adolescents' risky behavior, such as substance use, in relation to their peers' behavior (Alexander, Piazza et al. 2001; Bauman, Carver et al. 2001; Haynie 2002; Eisenberg 2004; Clark and Loheac 2005) and have found endogenous effects to be compelling or "large enough to be policy-relevant" (Eisenberg 2004).

Following our brief overview of Add Health, we construct a sample based on multiple datasets from wave I and wave II: in-home interview, parent interview, in-school

interview, in-school and in-home friendship nominations, friends network, neighborhood and contextual databases.

Data

Add Health Study is a nationally representative, probability-based survey of adolescents in grades 7 through 12 in the United States. Wave I interviews were conducted in the 1994-95 school year. A total number of 14,396 adolescents was interviewed. Wave II data collection includes follow-up in-home interviews with adolescents and follow-up school administrator interviews conducted in 1996. The second wave surveyed 13,568 students one year after wave I. The sample for the wave II in-home interview comprised the respondents to the wave I in-home interview, with the following exceptions (cited from "Add Health Design Focus", UNC Carolina Population Center):

- "Respondents who were in the 12th grade at wave I and who were not part of the genetic sample were not interviewed at wave II."

- "Respondents who were in only the wave I disabled sample were not re-interviewed."

- "An additional 65 adolescents who were members of the genetic sample and who had not been interviewed at wave I were recruited at wave II."

Wave II interviews took place from April through August 1996. The interviews were generally identical to the ones conducted in wave I, which allows us to use a subsample from these two-wave surveys and construct a two-period panel dataset that is relevant to our empirical study.

Studies on children and adolescents' problem behaviors have reported indicators for perception to be important explanatory variables (Smetters and Gravelle 2001).

I choose not to include such variables because of the potential endogeneity. Table 1 gives summary statistics and variable definitions. It is noticeable that the times of illegal drug use per month on average have gone up by nearly 30% between wave I and wave II as students are getting older. Meanwhile, the variance increases by about 358%! So, a simple conjecture of a "mean" shift of drug use between these two survey periods will not be consistent with this giant increase in the variations of drug use. One possible explanation to this rests upon social interactions, which can generate "excess" variance of between-group variations relative to within-group variations (Graham 2006). Peer's influence also appear to be more prominent as adolescents age in this longitudinal setting.

A, probably, most rare feature of Add Health compared to other similar longitudinal studies on youth's behavior, is that adolescents (respondents) were asked to nominate up to five close male and five close female friends from the school roster. Friends' identification numbers make it possible to link a respondent's information to his or her friends' and therefore construct a relevant and meaningful peer group. This gives us potential advantages over many empirical studies in which peer effects are merely evaluated based upon either broadly or hypothetically defined reference group where actual interactions are likely to be overstated. Table 2 lists the variations of these peer groups' sizes. As we discussed later, such small group sizes with sufficient variations fit right in linear-in-means models (Lee 2006).

Mover experiment

Identification of the endogenous effect in this article hinges upon an exogenous perturbation of peer groups, which generates an *observed* difference between the treated and untreated group. This difference is manipulable in the sense of group level unobserved heterogeneities being purged of conditional on certain observables which are

not directly affected by such an intervention, and the distribution of which stays invariant in both pre-treatment and post-treatment periods. This is in line with the idea of "selection on observables".

Add Health provides a plausibly exogenous changes, a quasi-experiment, in peer environment given its longitudinal design—a friend who is a substance user moving away in between wave I and wave II. This brings up the following essential question: if in period 0 ($T = 0$) A and B are identical in observable characteristics and both of them have a friend who uses (illegal) drugs, then why does A use fewer (illegal) drugs than B in period 1 ($T = 1$) if A's friend who is a substance user moves away in between these two periods ($G = 1$) while B's friend doesn't ($G = 0$)? This quasi-experiment shares the same idea of a policy intervention which removes a substance user from his or her peer group. If this "mover friend" experiment takes place exogenously on the peer group level conditional on group level observables, then the difference in outcomes between the treated group ($G = 1$) and the untreated group ($G = 0$) in the pre-treatment period ($T = 0$) can be interpreted as a "normal" difference. In the post-treatment period ($T = 1$), the difference between the treatment ($G = 1$) and control ($G = 0$) group will include this "normal difference" plus the treatment effect, which is essentially the effect of the second-period treatment ($I \equiv G \cdot T = 1$) relative to the baseline treatment ($I = 0$). Therefore, a random intervention, conditional on group level observables, will eliminate unobserved heterogeneities which confound endogenous effect and provides a DID estimate. Whether this DID estimate can be interpreted as a treatment effect estimate depends upon the randomness of the intervention. We can also look at this DID approach from the other perspective. For the untreated group ($G = 0$), its difference between period 0 ($T = 0$) and period 1 ($T = 1$) represents a "common trend" whose validity, still, depends on the (conditional) random assignment of the treatment. Note that such a "common trend" is

not testable with two periods, which can be tested with multiple periods. Considering the treated group ($G = 1$), its counterfactual state—outcomes in the absence of treatment—will be constructed by this "common trend", so that the difference between period 0 ($T = 0$) and period 1 ($T = 1$) subtracting the "common trend" will reveal the treatment effect. In this way, DID eliminates individual heterogeneities, such as preference and ability, which are possibly time-invariant, through the differencing over time and purges group heterogeneities via a second differencing across groups. This DID idea, later extended to CIC, is illustrated in table 3.

Table 3 provides a primitive description of how things have changed over two periods. We observe that for the treated group drug use on average dropped dramatically in comparison with the control group. And, the "normal difference" between these two experimental groups has been even reversed. Regression analysis of this primitive findings help to disentangle effects of certain attributes, which *do not* bear causal meanings (Holland 1986), and effects due to things which are manipulable and whose causal interpretation is therefore valid. Of these, we are interested in separately identifying endogenous and contextual effects. From the observation of table 1 and table 3, the role for endogenous interactions is emerging. We notice that observed characteristics stay stable over time while outcomes have changed, relatively, a lot. Through a primitive DID analysis, we find that there seems to be some important "unknown factors" that have played a key role in changing the outcomes. Whether such "unknown factors" can be identified depends critically on this "mover-friend" experiment.

Table 4 provides mean-comparison results of evaluating this mover-friend experiment. For this first-moment comparison, we cannot find sufficient "evidence" to conclude that these peer group level characteristics are *not* similar between the treated and control group in both pre-treatment and post-treatment periods. This, to a cer-

tain extent, will validate our assumptions, though still not testable, of "selection on observables". In addition, these "balanced" characteristics, to some degrees, invalidate the concern of potential interactions between time and either the treatment or the control group. Group assignment per se is, empirically shown, to be less likely to introduce unobserved heterogeneities.

Admitting that such a standard check, i.e. group similarity on observables, certainly has its own deficiency, we next provide further evidence to show that such exogeneity is hard to refute even though it is still not perfectly believable. Thanks to Add Health, we obtain information on parents' intent to move between wave I and II, which is likely to be correlated with actual "removals" of drug-user friends from their peer groups. Table 5A and 5B offer statistical evidence for the independence between the treatment assignment and parents' intent to move. Two survey questions answered by parents, who make actual moving decisions, are of particular importance. One is "how much would you like to move away from this neighborhood?", and the other one is "in this neighborhood, how big a problem are drug dealers and drug users?" Both table 5A and 5B provide no (sufficient) evidence to reject the hypothesis that the treatment assignment is independent of parents' intent to move in terms of Pearson χ^2 and Kendall's rank correlation coefficient.

Table 5A reveals the fact that for those who wanted "very much" to move out of their current neighborhood at wave I, 1.9% of them were actually "assigned" to the treatment group. For those who didn't want to move "at all", 1% of them were assigned to receive the "treatment". In addition, within the treatment group, there are 44.9% of the teenagers having parents who didn't want to move "at all", 34.7% having parents who had "some" intent to move and only 20.4% of the students whose parents wanted to move "very much".

Similarly, table 5B shows that for those who regard the drug dealer and drug user

"a big problem", none of them were actually "assigned" to the treatment group. For those who didn't care "at all", 1.3% of them were assigned to receive the "treatment". In addition, within the treatment group, there are 61.2% of the teenagers having parents who didn't think drug dealer and drug user a problem "at all" and 38.8% of the students whose parents only believed "some" drug problems existing in current neighborhoods.

Table 6 further reveals the relationship between families' actual residential choice and their moving pattern. We observe that primary reasons for choosing current neighborhoods, in descending order, are: "affordable good housing", "close to friends or relatives", "better schools" and "less crime". For those who place priority on "better school", 85.7% of those families didn't move, and for those who care about "less crime", 89.3% stay put. It is also shown that the primary residential changes were made by families who placed "affordable good housing" on the first place. Therefore, considering the facts in table 5A, 5B and table 6, parental choice of residential neighborhood, in ideal situations possibly depending upon peer groups' behavior, is unlikely to be correlated with such a specific "mover friend" experiment during this particular period between wave I and wave II. Parents can have good intent to choose best neighborhoods for children, but there are real world constraints that could make "accidental" or unplanned or even undesirable moving happen, which essentially provides valuable exogenous changes from a research perspective.

Estimator

Based upon our discussion about the advantages of CIC over DID, we use the following estimators proposed by Athey and Imbens (2006) for the effect of treatment on the treated and on the control groups respectively. The estimate from DID is used as a baseline, or can be interpreted as an average treatment effect under DID's regularity

conditions.

Observations obtained from (44) group g and time period t are denoted by $\widetilde{y}_{gt,k(ri)}$, for $r = 1, \cdots, R$, $i = 1, \cdots, m_r$ and $k = 1, \cdots, N_{gt}$. We use the empirical distribution as an estimator for the distribution function.

$$(49) \qquad \widehat{F}_{\widetilde{y},gt}(y) = \frac{1}{N_{gt}} \sum_{k=1}^{N_{gt}} 1\{\widetilde{y}_{gt,k(ri)} \leqslant y\}$$

The estimator for the inverse of the distribution function is given by:

$$(50) \qquad \widehat{F}_{\widetilde{y},gt}^{-1}(q) = \inf\{y \in \widetilde{Y}_{gt,k(ri)} | \widehat{F}_{\widetilde{y},gt}(y) \geqslant q\}$$

The sample version of the CIC estimators corresponding to (47) and (48) are:

$$(51) \qquad \widehat{\tau}_{ATT}^{CIC} = \frac{1}{N_{11}} \sum_{n=1}^{N_{11}} \widetilde{y}_{11} - \frac{1}{N_{10}} \sum_{n=1}^{N_{10}} [\widehat{F}_{\widetilde{y},01}^{-1}(\widehat{F}_{\widetilde{y},00}(\widetilde{y}_{10}))]$$

$$(52) \qquad \widehat{\tau}_{ATC}^{CIC} = \frac{1}{N_{00}} \sum_{n=1}^{N_{00}} [\widehat{F}_{\widetilde{y},11}^{-1}(\widehat{F}_{\widetilde{y},10}(\widetilde{y}_{00}))] - \frac{1}{N_{01}} \sum_{n=1}^{N_{01}} \widetilde{y}_{01}$$

Policy implications

Estimation results on adolescents' drug use under a social-interactions based model, given by table 8, are in line with our intuition. Endogenous effects due to peer's influence predominate other effects. It is shown that 61.42% of the increment of peer's drug use will be "transferred" to individual's incremental usage of illegal drugs. This effect is singled out from contextual effects and is statistically significant at 1% level. Given its magnitude, we would expect a large multiplier to come into play when outside intervention occurs. Less importantly, going to a private school can alleviate

this drug use problem. This effect is significant at 5% level but it doesn't contribute to a multiplier effect. It is also noticeable that a male primary care-giver can lower the frequency of drug use and male students, on average, have higher potential use of illegal drugs.

Since treatment and control groups differ in the distribution of individual heterogeneities such as preferences, which affects drug use, we may wonder whether the considered "mover-friend" treatment assignment is "need-based". An optimal or ideal policy intervention should be conducted to the group in which adolescents have higher use of illegal drugs and their parents are worried about such drug problems. Table 7 includes information identical to table 5B except that percentages are calculated. In both treated and untreated groups, the distribution of parents' attitude toward drug problems in current residential neighborhood are similar. As we have previously argued, such a "mover-friend" experiment is assigned not according to parents' need. Since parents make actual moving decision, from individual student's perspective, this removal of "substance-user" intervention is randomly assigned. This intervention will be optimal in the sense that it is conducted directly to the group where high drug usage is pervasive. Table 3 has verified this. It is the group in which drug use nearly four times greater than the other group that did adopt the "treatment". Under this optimal policy adoption presumption, we would expect that magnitudes of various treatment effects identified by CIC and DID obey the following relationship:

$$(53) \qquad\qquad |\widehat{\tau}^{CIC}_{ATC}| < |\widehat{\tau}^{DID}_{ATE}| < |\widehat{\tau}^{CIC}_{ATT}|$$

Table 9 verifies (53). The point estimates are in accord with our conjecture of this

optimal policy adoption.

$$(54) \qquad\qquad |-0.240| < |-1.439| < |-2.064|$$

Inference

We make heavy use of bootstrap for inference given that exact finite-sample results are unavailable for estimators implemented in this article. Employing bootstrap methods permit us to draw statistical inference when analytical results, based on limiting distributions, for standard errors are difficult to compute or asymptotic methods work poorly in finite samples. When asymptotically pivotal statistic exists, we use bootstrap to implement asymptotic refinement, which essentially provides a numerical method to implement the Edgeworth expansion reducing asymptotic errors from $O(N^{-1/2})$ to $O(N^{-1})$ for one-sided tests and from $O(N^{-1})$ to $O(N^{-3/2})$ for two-sided hypothesis tests.

Finite sample size correction

Finite sample inference based on asymptotic approximation tends to understate the actual size of testing parameters obtained from a consistent and asymptotically normal estimator. Taking into account this potential size distortion, we bootstrap standard errors of estimates of the endogenous effects and various treatment effects. We use two—percentile and biased-corrected—methods to construct 95% confidence intervals for these estimates.

The estimate of the endogenous effect remains highly significant from zero with smallest finite sample bias. Although point estimates of various treatment effects give correct signs, they can not be distinguished from zero statistically at 5% significance

level. However, for the DID estimate, its asymptotically pivotal statistic exists. We use bootstrap to implement an asymptotic refinement, and obtain its p-value 0.038 in table 10. Therefore, the null hypothesis $\tau_{ATE}^{DID} \geqslant 0$ can be rejected at 5% significance level. Implementing a mandatory policy that removes a substance-user from a peer group will lower the average frequencies of adolescents' drug use. Unlike the DID estimate, CIC estimates of the average effects of treatment on both the treated and untreated groups do not offer asymptotic refinements. It can be shown that in general CIC will be less efficient if outcomes in the counterfactual states are more volatile compared with DID estimates (Athey and Imbens 2006). Nonetheless, both CIC estimates point to the correct direction. We next bootstrap t-ratios for each treatment effect, and then use the limiting (normal) distribution to obtain associated p-values. The result on τ_{ATC}^{CIC} is encouraging. The null hypothesis $\tau_{ATC}^{CIC} \geqslant 0$ can now be rejected with high significance level ($\alpha \leqslant 1\%$). Combining the endogenous effect estimate, we may construct different "social multipliers" corresponding to different treatment effects on the basis of (42). We then obtain p-values for these various "social multipliers", and the multiplier for the average effect of treatment on the control group remains highly significant ($\alpha \leqslant 1\%$). Under the case of optimal policy adoption, such a multiplier effect for the untreated group which hypothetically received the treatment may serve as the lower bound of the actual multiplier effect. In our example, this multiplier effect arising from a hypothetical treatment on the control group, in magnitude, is 159.07% times greater than the pure average effect of treatment on the untreated group per se. This lower bound is also statistically significant from zero. Given the magnitude order of different treatment effects established under optimal policy adoption, we may hypothesize that multipliers generated by the average treatment effect and the average effect of treatment on the treated shall be even more prominent. However, due to limitation of the sample size, there is no enough

power for us to distinguish these two multipliers from zero. Detailed testing results are reported in table 10.

Power calculation and inverse power functions

Powers of testing key parameters are reported in table 11. It is only the test for the endogenous effect that does not fail to reject at least 50% of the time, which is actually better than flipping a coin to decide the result. None of the tests for treatment effects possess sufficient power. To make valid inference of these hypothesis tests when the tests fail to reject the null hypothesis, we compute inverse power (IP) functions to yield summary measures, proposed by Andrews (1989), that can facilitate the interpretation of the test results.

We use inner IP function to answer questions such as: which deviations from the null hypothesis have "a good chance" of being *undetected* by the test. Here "a good chance" is represented by flipping a coin (with its weight evenly distributed). Within the range of this inner IP, the power of a test is less than 50%. We also use outer IP function to answer questions like: which deviations from the null hypothesis have a very good chance of being *detected* by the test. Here "a very good chance" is determined by $1 - \alpha$. Given a test of size equal to 5%, the outer IP gives the range of deviations from the null for which the test rejects at least 95% of the time. Loosely speaking, within the range defined by outer IP, probabilities of making type II errors are no greater than type I errors, which is the case that we can control type II errors by limiting type I errors.

Table 11 gives the summary measures of power, inner IP and outer IP. Figure 1 and 2 and 3 illustrate approximated power and inverse power functions. The DID test maintains superior power to the CIC test for average effect of treatment on the

45

treated, but its power is inferior to the CIC test for average effect of treatment on the control group. Given the directions set by CIC estimates, the inner IP function of the test for the effect of treatment on the untreated tells us that such a test can not outperform a coin flip in the parameter region $[-0.7899, 0]$, but it is able to reject the null hypothesis more than 95% of the time once the magnitude of the leftward deviation from zero reaches 1.4529 given a 5% significance level. The low power of the test for average effect of treatment on the treated is mostly due to the modest sample size, which limits meaningful inference. However, under the case of optimal policy adoption, the average effect of the treatment on the untreated group sets a lower bound with adequate power, which still conveys useful information that such a "social multiplier" generated by hypothetically removing a substance-user from peer groups is still too large to be overlooked and it is well worth attention to policy makers, researchers, health-care providers and educators for better understanding of how to protect young people and secure our future.

Conclusion

It is widely believed that adolescents' behaviors, particularly heath-related ones, are easily influenced by others, especially by their friends, during this rapid transition period from children to adults. It is also acknowledged that teenage behaviors and their health outcomes will not be well understood without considering social interactions. Finding a strong correlation in behavior among peers cannot justify the existence of peer effects. The impact of friends on an individual's behavior will be confounded by their mutual influences, individual's self-selection into peer groups and peer's shared unobserved environmental factors. I confront these problems by modeling the nature of interactions, exploring exogenous variations of peer groups and taking advantage

of the longitudinal design provided by Add Health.

Using Add Health, I identify and estimate peer effects of adolescents' health-related behaviors—substance use—through difference-in-differences (DID), using time-differencing to exclude individual "fixed effect" and treatment-control-differencing to exclude peer group level unobserved heterogeneities. To accommodate possible behaviors of sorting into treatment—the group with higher expected gain receiving the treatment—I estimate such heterogeneous treatment effects based on changes-in-changes (CIC), a generalized version of DID.

To isolate peer effects from other factors, my linear-in-means modeling strategy rests upon a spatial autoregressive model. Notice that adolescents spend more time with their closest friends. As to substance use, it is more likely that observed actions of just being "cool" or actual "pursuits" of happiness seen from peers, rather than expected behaviors, that influence individual decision-making. In this sense, the linear-in-means model is more appropriate and therefore adopted. This distinguishes my study from current empirical studies on social interactions, most of which are based on linear-in-expectations models. Resolving the "reflection problem" (Manski 1993) under linear-in-means models has been discussed by Lee (2006). The necessary condition for variations in peer group sizes is well satisfied by Add Health data.

To combat unobserved group heterogeneities, I explore a source of exogenous variations in peers' drug use, a "treatment" induced by a friend, who is a substance user, moving away between wave I and II from his or her own peer group. Such a removal, due to, for example, parents' job change, is likely to be independent of peer group level unobserved heterogeneities conditional on a sufficient number of group-level observable characteristics (generously available in Add Health). This quasi-experimental like change shares a similar nature of a policy intervention which removes drug user friends from a peer group. Since individual preference for peer group

formation can be reasonably assumed time-invariant in the short period between Add Health's first two survey periods, this sorting behavior is explicitly controlled for in a "fixed effect" manner under Add Health's longitudinal design. This strategy allows for identifying not only peer effects, but also the treatment effect corresponding to a hypothetical policy intervention of removing a drug-user friend from his or her own peer group. Identification of these two parameters ultimately leads to constructing a "social multiplier", which bears rich policy implication, especially seen from an economic perspective.

Inferences on the endogenous effect and various treatment effects are conducted through bootstrap, which is currently underutilized. We correct the size and compute the power of testing these parameters of policy interest. We also utilize inverse power summary measure to facilitate interpretation and make valid inference when tests fail to reject the null hypothesis, as an effort to address the weak state of inference in empirical studies.

Exploring the nature of peer effects and measuring the associated "social multiplier" is very relevant. It is well worth attention to policy makers, researchers, health-care providers and educators for better understanding of how to protect young people and thereby secure our future.

References

Aaronson, D. 1998. "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *Journal of Human Resources* 33: 915-946.

Abadie, A. 2005. "Semiparametric Difference-in-Differences Estimators." *Review of Economic Studies* 72(1): 1-19.

Alexander, C., M. Piazza, D. Mekos, and T. Valente. 2001. "Peers, schools, and adolescent cigarette smoking." *Journal of Adolescent Health* 29(1): 22-30.

Andrews, D.W.K. 1989. "Power in Econometric Applications." *Econometrica* 57: 1059-1090.

Athey, S. and G.W. Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* 74: 431-497.

Bauman, K.E., K. Carver, and K. Gleiter. 2001. "Trends in parent and friend influence during adolescence: The case of adolescent cigarette smoking." *Addictive Behaviors* 26: 349-361.

Bisin, A., A. Moro, and G. Topa. 2002. "The Empirical Content of Models with Multiple Equilibria." Unpublished, New York University.

Brock, W. and S. Durlauf. 2001. "Interactions-Based Models." In J. Heckman and E. E. Leamer, ed. *Handbook of Econometrics*: Elsivier Science, pp. 3297-3380.

——. 2001. "Discrete Choice with Social Interactions." *Review of Economic Studies* 68(2): 235-260.

——. 2002. "Market Implications of Peer and Neighborhood Effects: A Multinomial-Choice Model of Neighborhood Effects." *American Economic Review* 92(2): 298-303.

——. 2004. "Identification of Binary Choice Models with Social Interactions." Unpublished, University of Wisconsin, Madison.

Cameron, A.C. and P.K. Trivedi. 2005. *Microeconometrics: Methods and Applications*, Cambridge University Press.

Chantala, K. 2003. "Introduction to analyzing Add Health data." Carolina Populaton Center, University of North Carolina, Chapel Hill.

Chantala, K., J. Tabor. 1999. "Strategies to Perform a Design-Based Analysis Using the Add Health Data." Carolina Populaton Center, University of North Carolina at Chapel Hill.

Clark, A.E. and Y. Loheac. 2005. "'It Wasn't Me, It Was Them!' Social Influence in Risky Behavior by Adolescents." Working paper, Institute for the Study of Labor (IZA).

Cohen-Cole, E. 2006. "Multiple groups identification in the linear-in-means model." *Economics Letters* 92(2): 157-162.

Cooper, R. 2002. "Estimation and Identification of Structural Parameters in the Presence of Multiple Equilibria." Working Paper, NBER.

Davezies, L., X. d'Haultfoeuille, and F. Denis. 2006. "Identification of Peer Effects Using Group Size Variation." Working paper, Institute for the Study of Labor (IZA).

Durlauf, S. 2001. "A Framework for the Study of Individual Behavior and Social Interactions." *Sociological Methodology* 31: 47-87.

——. 2002. "Groups, social influences and inequality: a memberships theory perspective on poverty traps." Unpublished, University of Wisconsin, Madison.

Durlauf, S. and E. Cohen-Cole. 2004. "Social interaction models." Unpublished, University of Wisconsin, Madison.

Eisenberg, D. 2004. "Peer Effects for Adolescent Substance Use: Do They Really Exist?" Unpublished, University of California, Berkeley.

Evans, W.N., W.E. Oates, and R.M. Schwab. 1992. "Measuring Peer Group Effects: A Study of Teenage Behavior." *Journal of Political Economy* 100: 966-991.

Ginther, D., R. Haveman, and B. Wolfe. 2000. "Neighborhood Attributes as Determinants of Children's Outcomes: How Robust Are the Relationships?" *Journal of Human Resources* 35: 603-642.

Glaeser, E.L., B.I. Sacerdote, and J.A. Scheinkman. 2003. "The Social Multiplier." *Journal of the European Economic Association* 1(2-3): 345-353.

Graham, B.S. and J. Hahn. 2005. "Identification and Estimation of the Linear-in-Means Model of Social Interactions." *Economics Letters* 88(1): 1-6.

Graham, B.S. 2006. "Identifying Social Interactions through Conditional Variance Restrictions." Unpublished, University of California, Berkeley.

Haynie, D.L. 2002. "Friendship Networks and Delinquency: The Relative Nature of Peer Delinquency." *Journal of Quantitative Criminology* 18(2): 99-134.

Lee, L. Forthcoming. "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects." *Journal of Econometrics*, in Press.

Manski, C.F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60: 531-542.

——. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives* 14(3): 115-136.

Mas, A. and E. Moretti. 2006. "Peers at Work." Working paper, NBER.

Meyer, B.D. 1995. "Natural and Quasi-Experiments in Economics." *Journal of Business and Economic Statistics* 13(2): 151-161.

Rubin, D.B. 1980. "Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment." *Journal of the American Statistical Association* 75: 591-593.

Sampson, R.J., J.D. Morenoff, and T. G-Rowley. 2002. "Assessing 'Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28: 443-478.

Smetters, K. and J. Gravelle. 2001. "The Exchange Theory of Teenage Smoking and the Counterproductiveness of Moderate Regulation." Working Paper, NBER.

Zanella, G. 2004. "Discrete Choice with Social Interactions and Endogenous Memberships." Unpublished, University of Siena.

Ziliak, S. T. and D. N. McCloskey. 2004. "Size Matters: The Standard Error of Regressions in the American Economic Review." *Journal of Socio-Economics* 33: 527-546.

Zimmerman, D.J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics* 85(1): 9-23.

Table 1. Summary Statistics and Variable Definitions

Variable	Wave I			Wave II		
	mean	std. deviation	N	mean	std. deviation	N
Individual behavior						
Drug use	0.7208	6.900	3922	0.9296	14.7686	3922
Individual characteristics						
Age	15.2404	1.5313	3922	16.1629	1.5568	3922
Male	0.4776	0.4996	3922	0.4776	0.4996	3922
White	0.6311	0.4826	3922	0.6311	0.4826	3922
Black	0.2035	0.4026	3922	0.2035	0.4026	3922
Grade	9.3383	1.4548	3922	10.2906	1.4571	3878
Private School	0.0750	0.2634	3920	0.0750	0.2634	3920
Church	0.5681	0.4954	3855	0.5263	0.4994	3825
Allowance	7.4447	10.5153	3922	8.2449	12.4101	3920
Family factors						
Household income	45.7944	50.8590	3706	45.7944	50.8590	3706
PCG's education	0.1399	0.3469	3868	0.1399	0.3469	3868
Spouse's education	0.1330	0.3397	3405	0.1330	0.3397	3405
PCG (male)	0.0734	0.2608	3842	0.0734	0.2608	3842
Contextual factors						
Median income (block)	31797.09	14784.62	3905	31755.62	14934.85	3899
Education (block)	0.2190	0.1388	3922	0.2196	0.1397	3917
Unemployment rate (block)	0.0761	0.0516	3913	0.0758	0.0510	3903
Urban	0.4850	0.4998	3922	0.4881	0.4999	3917
Friends' averaged characteristics						
Grade	9.3383	1.4541	3922	10.3238	1.5779	3890
Church	0.5706	0.4934	3876	0.5258	0.4976	3850
Household income	46.7665	52.4584	3767	467665	52.4584	3767
PCG's education	0.1402	0.3441	3875	0.1402	0.3441	3875
Spouse's education	0.1356	0.3334	3531	0.1356	0.3334	3531
Moved between wave I and II						
Mover	0.1119	0.3153	3922	0.1119	0.3153	3922

Note:

1. "drug use": during the past 30 days, how many times have you used illegal (marijuana, cocaine, inhalants, LSD, PCP, ecstasy and heroin) drugs (0-900 times)

2. "grade": what grade are/were you in (grade 7-12, 13 beyond high school)

3. "church": ever go (=1) to church or not (=0)

4. "allowance": how much is your allowance each week ($0-$95)

5. "income": about how much total income before taxes did your family receive in 1994 (in thousands)

6. "PCG's education": primary care-giver graduated from college/university

7. "median income (block)": block median household income ($49,999-$148,752)

8. "education (block)": within block proportion of population aged 25 and above with college degree or higher (0.000 to 0.944)

9. "unemployment rate": within block unemployment rate (0.000 to 0.593)

10. "mover": respondent moved to a different census tract between wave I and II. This includes the following cases: (1) moved to different tract but within the same county; (2) moved to different county but within the same state; (3) moved to different state.

Table 2. Variations of Peer-Group Sizes

Sizes of Peer-Groups	Frequency	Percentage
2	1088	69.21
3	324	20.61
4	85	5.41
5	37	2.35
6	21	1.34
7	13	0.83
8	4	0.25
Total	1572	100

Table 3. A Difference-in-Differences Analysis without Covariates

drug use (mean)	$T = 0$	$T = 1$	difference
$G = 0$	0.6808 ($N_{00} = 3872$)	0.9396 ($N_{01} = 3872$)	0.2588 ("common trend")
$G = 1$	3.8200 ($N_{10} = 50$)	0.1600 ($N_{11} = 50$)	-3.6600
difference	3.1392 ("normal difference")	-0.7796	-3.9188 ("treatment effect")

Note:

1. T: Wave I ($T = 0$), Wave II ($T = 1$).

2. G: binary indicator (=1) for the "mover-treatment" group.

3. "mover-treatment": a substance user friend moved away between Wave I and II.

Table 4. Evaluating "Mover-Friend" Quasi-Experiment – Mean Comparison

Treatment-Control	Pre-treatment			Post-treatment		
Comparison	mean	std.error	t	mean	std.error	t
Individual characteristics						
White	0.0638	0.1275	0.50	0.0644	0.1268	0.51
Black	0.0182	0.1049	0.17	0.0303	0.1042	0.29
private school	−0.0508	0.0654	−0.78	−0.0533	0.0657	−0.81
allowance	3.4888	2.8199	1.24	−5.6031	3.3772	−1.66
Contextual factors						
median income (block)	−1053.872	3653.922	−0.29	−2078.014	3621.189	−0.57
college education (block)	0.0087	0.0342	0.25	0.0501	0.0344	1.46
unemployment rate (block)	−0.0114	0.0133	−0.86	0.0020	0.0130	0.15
urban	−0.0381	0.1325	−0.29	−0.0146	0.1321	−0.11

Note: The "mover-friend" quasi-experiment is randomly assigned conditional on the following peer-group level averaged characteristics: age, sex, grade, religion, household income, primary care-giver's education, spouse's education.

Table 5A. Evaluating the Independence between "Mover-Friend" and Parents' Intent to Move

		"Want to Move" at Wave I			
		not at all (=1)	some (=2)	very much (=3)	Total
"Mover-Friend" between	$G=0$	2106	1195	518	3819
Wave I and Wave II	$G=1$	22	17	10	49
	Total	2128	1212	528	3868

Note:

1. The survey question is: "how much would you like to move away from this neighborhood?"

2. Pearson $\chi^2 = 2.7625$, Prob. $= 0.251$; Kendall's τ-b $= 0.0248$, Asy.std.error $= 0.016$, $z = 1.55$

Table 5B. Evaluating the Independence between "Mover-Friend" and Parents' Intent to Move

		"How big the drug problem is" at Wave I			
		not at all (=1)	small (=2)	a big problem (=3)	Total
"Mover Friend" between	$G=0$	2295	1201	313	3809
Wave I and Wave II	$G=1$	30	19	0	49
	Total	2325	1220	313	3858

Note:

1. The survey question is: "in this neighborhood, how big a problem are drug dealers and drug users?"

2. Pearson $\chi^2 = 4.8371$, Prob. $= 0.089$; Kendall's τ-b $= -0.0090$, Asy.std.error $= 0.014$, $z = -0.64$

Table 6. Family Residential Choice and Moving Pattern

"Why live here?"	no	same block	same tract	same county	same state	diff. state	unknown	Total
1	51	0	0	0	0	0	1	52
2	294	8	0	20	6	0	7	335
3	302	0	2	25	7	1	8	345
4	621	7	4	54	22	1	11	720
5	465	0	0	34	7	10	5	521
6	185	0	0	10	1	5	5	206
7	510	6	7	32	15	1	11	582
8	481	2	3	40	19	6	10	561
9	49	0	0	0	1	0	0	50
10	224	0	0	4	2	5	1	236
Total	3182	23	16	219	80	29	59	3608

The top spanning header reads "Moved between Wave I and Wave II".

Notes on "why live here":

1 = near old workplace; 2 = near current workplace; 3 = had outgrown previous housing; 4 = affordable good housing; 5 = less crime; 6 = less illegal activity; 7 = close to friends or relatives; 8 = better schools; 9 = children of appropriate ages; 10 = born here.

Table 7. Optimal Policy Implication

		"How big the drug problem is" at Wave I			
		not at all (=1)	small (=2)	a big problem (=3)	Total
"Mover-friend" between	$G = 0$	2295	1201	313	3809
		60.25%	31.53%	8.22%	100%
Wave I and II ("treatment)	$G = 1$	30	19	0	49
		61.22%	38.78%	0%	100%
	Total	2325	1220	313	3858
		60.26%	31.62%	8.11%	100%

Table 8. Difference-in-Differences (DID) Estimates

drug use	coefficient	std. error	t
"mover friend" treatment			
$G \cdot T = 1$	-1.4385	1.4763	-0.97
peer's average behavior			
drug use	0.6142^{***}	0.2035	3.02
individual characteristics			
age	0.0639	0.0908	0.70
male	0.4497^{*}	0.2477	1.82
White	-1.1125	0.4287	-0.26
Black	-0.3568	0.3169	-1.13
grade	-0.1139	1.0479	-0.11
private school	-0.3356^{**}	0.1629	-2.06
church	1.3052	0.9573	1.36
allowance	-0.0005	0.0121	-0.04
family factors			
household income	0.0006	0.0007	0.85
PCG's college education	0.1457	0.2197	0.66
spouse's college education	0.2155	0.2720	0.79
PCG (male)	-0.3097^{*}	0.1786	-1.73
contextual factors			
median income (block)	0.00001	$8.77e-06$	1.18
college education (block)	-2.9744^{**}	1.4391	-2.07
unemployment rate (block)	-4.0280	2.9738	-1.35
urban	0.4491^{*}	0.2343	1.92
peer's average characteristics			
grade	-0.1320	1.0598	-0.12
church	-0.9518	0.9705	-0.98
household income	0.0007	0.0012	0.61
PCG's college education	-0.0383	0.1895	-0.20
spouse's college education	-0.2739	0.2850	-0.96

$F(25, 3230) = 5.76$

$\Pr > F = 0.0000$

$R^2 = 0.4257$

number of observations $= 6335$

number of clusters $= 3231$

Notes: Asterisks (e.g., *, **, ***) denote significant at 10%, 5% and 1% level, respectively. Standard errors are robust to heteroskedasticity and clusters. Cluster-robust standard errors are calculated based on random effects in the panel setting.

Table 9. Changes-in-Changes (CIC) and Difference-in-Differences (DID) Estimates

N=6335, cluster=3231, rep.=1000	$\widehat{\tau}_{ATT}^{CIC}$	$\widehat{\tau}_{ATC}^{CIC}$	$\widehat{\tau}_{ATE}^{DID}$	$\widehat{\lambda}_0$
observed coefficient	-2.064	-0.240	-1.4385	0.614
bias	-1.050	0.043	0.1112	0.023
bootstrap standard error	4.789	0.285	1.5308	0.198
95% confidence interval (percentile)	$[-16.47, 1.771]$	$[-0.836, 0.315]$	$[-4.2929, 1.7931]$	$[0.247, 0.936]$
95% confidence interval (bias-corrected)	$[-18.35, 1.490]$	$[-1.117, 0.156]$	$[-4.6069, 1.3087]$	$[0.206, 0.931]$

Notes: Block-bootstraps are implemented in this panel setting with clustered data. Resample is done over blocks, jointly defined by individual student and his or her peer group.

Table 10. Testing Treatment Effects and Social Multipliers

Hypothesis testing		bootstrap w/ asy. refinement		bootstrap w/o asy. refinement	
treatment (τ)		t	percentile-t p-value	t	normal p-value
$H_0 : \tau \geqslant 0$	τ_{ATE}^{DID}	-0.970	0.038	-0.940	0.174
$H_1 : \tau < 0$	τ_{ATT}^{CIC}	$-$	$-$	-0.300	0.382
	τ_{ATC}^{CIC}	$-$	$-$	-5.052	0.000
multiplier (M_τ)		t	percentile-t p-value	t	normal p-value
$H_0 : M_\tau \geqslant 0$	$M_{\tau_{ATE}^{DID}}$	$-$	$-$	-0.467	0.320
$H_1 : M_\tau < 0$	$M_{\tau_{ATT}^{CIC}}$	$-$	$-$	-0.147	0.442
	$M_{\tau_{ATC}^{CIC}}$	$-$	$-$	-2.533	0.006

Notes: Block-bootstraps are implemented in this panel setting with clustered data. Resample is done over blocks, jointly defined by individual student and his or her peer group.

Table 11. Approximated Powers and Inverse Powers

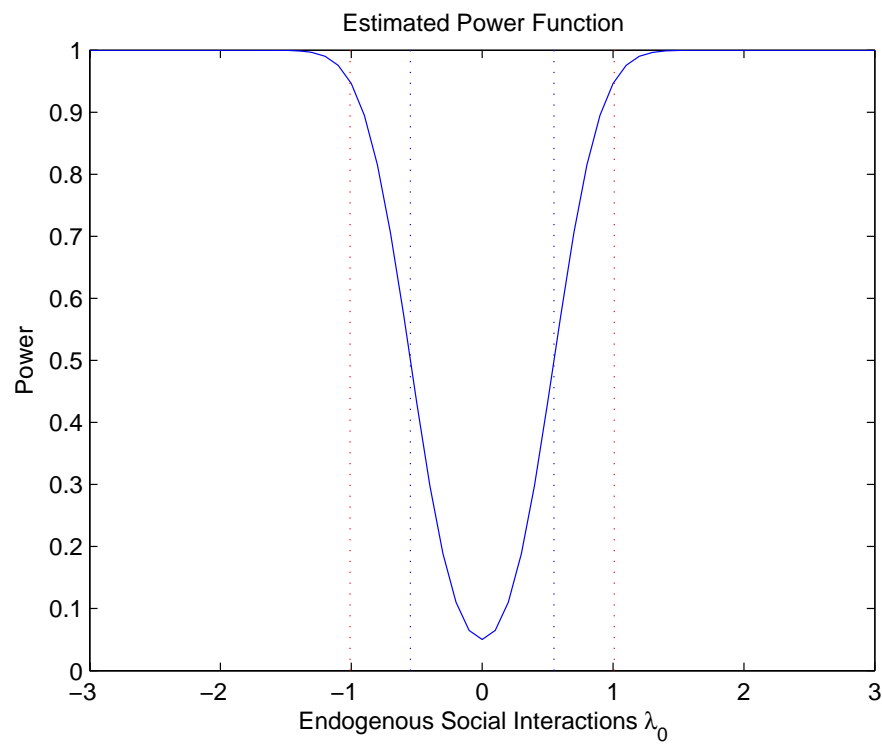| $\alpha = 0.05$	Power	Inner Inverse Power	Outer Inverse Power
$H_0 : \lambda_0 = 0$	0.5921	$[-0.5488, 0.5488]$	$(-\infty, -1.0094] \cup [1.0094, \infty)$
$H_0 : \tau_{ATE}^{DID} = 0$	0.1019	$[-4.2428, 4.2428]$	$(-\infty, -7.8040] \cup [7.8040, \infty)$
$H_0 : \tau_{ATT}^{CIC} = 0$	0.0610	$[-13.273, 13.273]$	$(-\infty, -24.414] \cup [24.414, \infty)$
$H_0 : \tau_{ATC}^{CIC} = 0$	0.0915	$[-0.7899, 0.7899]$	$(-\infty, -1.4529] \cup [1.4529, \infty)$

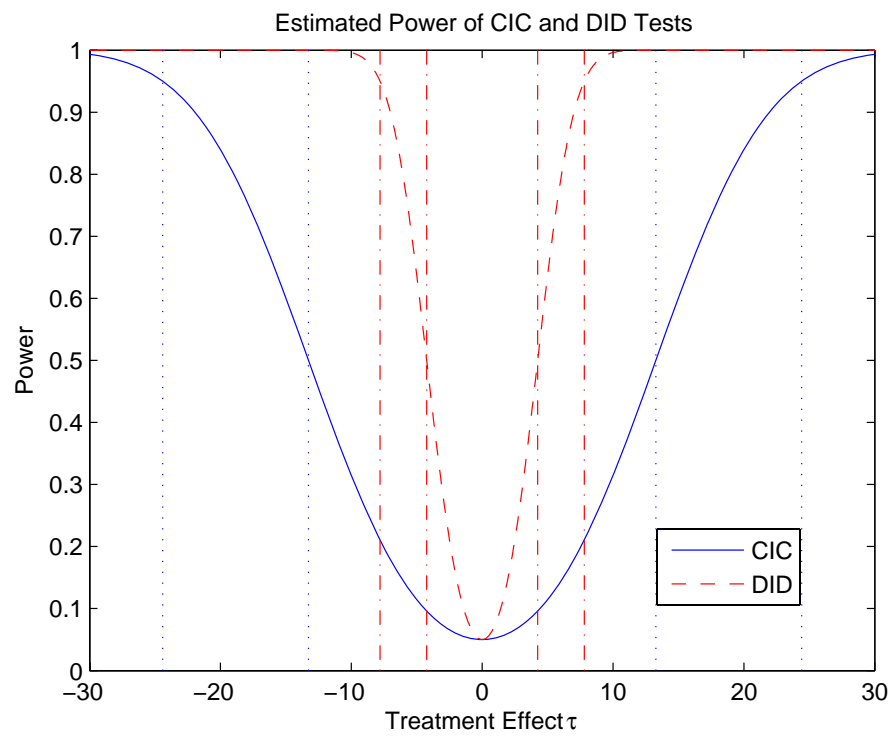**Figure 1. Approximated power and inverse power functions of testing the endogenous effect** (λ_0)

Figure 2. Approximated power and inverse power functions of testing the average effect of treatment on the treated group(τ_{ATT})
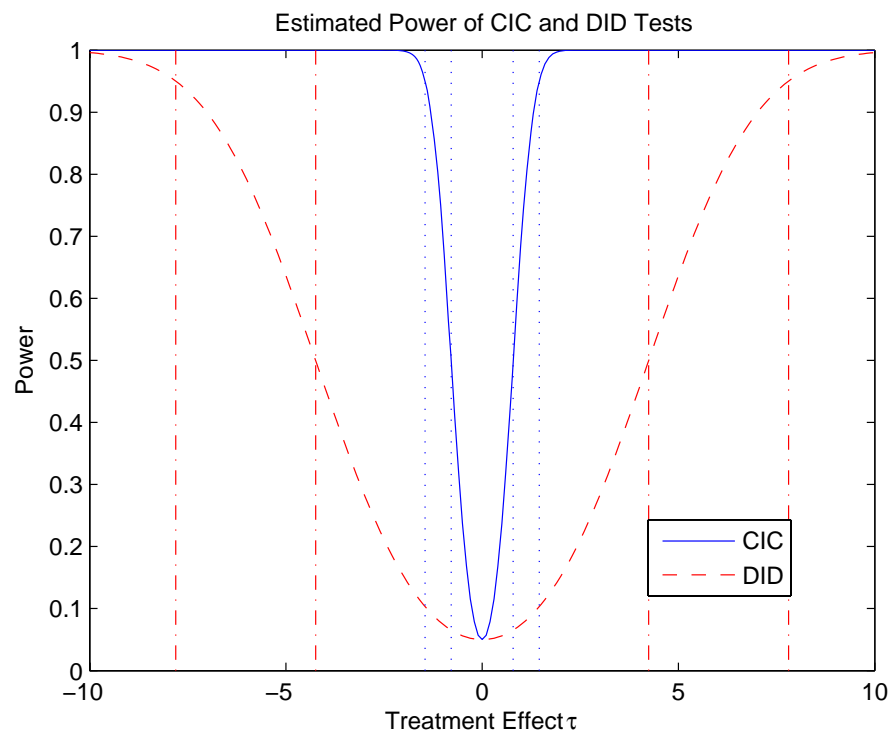
Figure 3. Approximated power and inverse power functions of testing the average effect of treatment on the control group (τ_{ATC})