Social Interactions and Endogenous Association

March 2005

PRELIMINARY

Bruce A. Weinberg
Department of Economics
Ohio State University
1945 North High Street
Columbus, Ohio 43210
weinberg.27@osu.edu
http://econ.ohio-state.edu/weinberg

ABSTRACT

This paper develops a model of social interactions with endogenous associations. People are assumed to invest time to develop the relationships that maximize their utility. When associations are endogenous, the effects of groups on individual behavior will be non-linear even when the underlying behavioral model is linear; improving the performance of some members of the group may hurt others; and peer pressure emerges with people taking actions to make others want to interact with them. Using data on associations among high school students, we provide a range of evidence consistent with our model. Individuals associate with people whose behaviors and characteristics are similar to their own. This tendency is particularly strong when the pool of similar potential associates is larger and when the size of the group is larger.

I appreciate helpful comments from Larry Blume, Ed Glaeser, Kevin Lang, and Qingyan Shang and from participants at the 2005 AEA Meetings and the University of Alabama. Qingyan Shang provided excellent research assistance. I am also grateful for support from the Mershon Center at Ohio State University and the NBER, where much of the work on this paper was done. I am, of course, solely responsible for all errors.

Social Interactions and Endogenous Association

I. Introduction

Social scientists have increasingly turned to social interactions models, in which individuals' behavior is affected by their social groups, to understand large disparities in behaviors and outcomes, and especially low outcomes among underrepresented and economically disadvantaged groups.[1] An essential aspect of social interaction models is specifying who interacts with whom. Theoretical econometric analyses (*e.g.* Anselin [1988]; Lee [2001]; Brock and Durlauf [2001a, 2001b] require researchers to specify a fixed weighting matrix that gives association patterns. Empirical analyses typically assume that all members of some "macro-group," such as a school or neighborhood, interact equally with each other by relating an individual's behavior to the mean behavior in her macro-group.

Neither approach is satisfactory. While convenient, using macro-group means ignores that people associate more with some members of their macro-groups than with others.[2] Using existing intra-group interaction patterns to estimate a weighting matrix, while an improvement over virtually all empirical work,[3] ignores that interactions are

---

[1] See, for instance, Wilson [1987, 1996]; Massey and Denton [1993]; and Jargowsky [1997]. Early studies often contain weak controls for macro-group selection (Datcher [1982]; Corcoran, Gordon, Laren, and Solon [1992]). Recent attention has focused on controlling for macro-group selection. See surveys by Jencks and Mayer [1990]; Deitz [2002]; and Haurin, Deitz, and Weinberg [2002]. More recent studies seek to identify random variation in social groups. Such studies include Bayer, Ross, and Topa [2004]; Bertrand, Luttmer, and Mullainathan [2000]; Borjas [1995]; Case and Katz [1990]; Cutler and Glaeser [1997]; Evans, Oates, and Schwab [1992]; Glaeser, Scheinkman, and Sacerdote [1996, 2003]; Hoxby [2000]; Ioannides and Zabel [Forthcoming]; Solon, Page, and Duncan [2000]; Topa [2001]; Weinberg [2000] and studies in footnote 4.

[2] A number of papers show that people prefer to associate with others of the same racial or ethnic group (Moody [2001]; Bayer, McMillan, and Rueben [2002]; Marmaros and Sacerdote [2003]; see, however, Ross [2003]). People who are different from their macro-groups are less likely to be impacted by them (Duncan, Connell, and Klebanov [1997]; DiPasquale and Kahn [1999], Cummings, DiPasquale, and Kahn [2001]; Conley and Topa [2001]; and Hoxby [2000]).

[3] Exceptions are Conley and Udry [2001] and Bandiera and Rasul [2002]. Bertrand, Luttmer, and Mullainathan [2002] and Munschi [2002] use information to specify the group across interactions operate. Conley and Topa [2001] estimate propensities for racial groups to interact using a structural model. These papers do not explicitly study the association process, nor do they study how association-patterns are

determined by behaviors as much as behaviors are determined by interactions (Wiseman [2002] provides vivid examples). A person who neither smokes nor drinks, for instance, is unlikely to choose to associate with someone who does both heavily, and so the first person's behavior is unlikely to be much affected by the other person's substance use.

Many policies to address the effects of social interactions involve moving people across macro-groups. At an empirical level, many relocation studies often find small effects of social groups.[4] Our model suggests that individuals who are relocated into social groups that are very different from themselves, will likely segregate within these groups, attenuating any effects, even if finely defined groups do have effects. Consistent with this hypothesis, Angrist and Lang [Forthcoming] and Kling, Ludwig, and Katz [2005] find larger effects for girls than boys, who may integrate more into their new social groups. At a theoretical level, weighting-matrix studies provide no guidance as to who a newly added person will associate with, nor do they indicate how adding or removing people affects the associations of people who are already in the macro-group and remain in it. Sobel [2001] makes these point in a particularly striking form.

This paper develops a formal model of associations. We assume that individuals choose their behavior and their associations to maximize their utility, which depends on their own behavior and the behavior and characteristics of their associates. We specify a cost in terms of time and effort to an individual of associating with other members of their macro-groups. As is standard, preferences over actions depend linearly on a weighted average of the characteristics and actions of the other members of an individual's macro-group with the weights determined by the amount of their interaction.

Our model implies that people will associate the most with people whose

---

affected by changes in the population or in the behaviors of group members.

[4] Studies include Aaronson [1998]; Angrist and Lang [Forthcoming]; Gould, Lavy, and Passerman [2004a, 2004b]; Jacob [2004]; Katz, Kling, and Liebman [2001]; Kling; Ludwig; and Katz [2005]; Ladd and Ludwig [1997]; Ludwig, Duncan, and Hirschfield [2001]; Oreopoulos [2003]; Plotnick and Hoffman [1995]; Rosenbaum, DeLuca, and Miller [1999]; and Weinberg, Reagan, and Yankow [2004].

behaviors and characteristics are similar to their own. This tendency to associate with similar people will reinforce underlying behavioral tendencies. We apply our model to understand association patterns in schools. In that context, an increase in the number of students in a grade using substances, for example, will enable the students who are inclined toward substance use to associate with other substance users, reinforcing their original tendencies.

In our model changes in macro-group composition have non-linear effects on behavior. When there are few students inclined toward substance use, these students will associate with people who do not use substances, tending to discourage their own substance use. As the number of students inclined to substance use increases, these individuals will find like-minded associates accentuating their tendencies.[5] Thus, even with a linear behavioral equation, with endogenous associations, the effects of macro-groups on individual behavior will be non-linear and vary with individual characteristics and behaviors. At a policy level, these non-linearities and interactions imply that even in a linear-in-means model, social interactions are not zero-sum. These non-linearities also provide a potential solution to Manski's [1993] reflection problem, one that is quite different from Brock and Durlauf's [2001b] important results along these lines.

This study employs data from the National Longitudinal Study of Adolescent Health (Add Health), which surveyed all students in 132 schools, asking about the respondents' closest friends as well as their background and behaviors. We provide a range of evidence consistent with our model of endogenous association. We find that the majority of variation in friend's behavior arises within individual grades in schools. Individuals associate with others whose behaviors and characteristics are similar to their

[5] Krivo and Peterson [1996]; Galster, Quercia, and Cortes [2000]; Galster [2002]; and Weinberg, Reagan, and Yankow [2004] provide evidence for non-linear effects and interactions. Non-linear effects have been argued for at least since Crane [1991], although there are few formal micro-models of these effects (Quercia and Galster [Forthcoming] discuss theories).

own. An increase in the mean of a behavior or characteristic in a school-grade increases the mean of that behavior or characteristic among the associates of people with high levels of that behavior or characteristic more than those with low levels. There is a stronger relationship between own behavior and the behavior of associates in large macro-groups.

II. A Model

II.A. The Framework

This section develops a model of social interactions with endogenous association. Individuals are characterized by observable characteristics, *x*, and an unobservable characteristic, ε. They choose an action *y* and a set of people with whom they associate. An individual with characteristics *x* and ε has a utility function,

$$U(x,\varepsilon) = \beta xy + \varepsilon y - \frac{1}{2}y^2 + \int_X \int_{-\infty}^{\infty} \begin{matrix}\left[w(\tilde{x},\tilde{y};x,\varepsilon,y)S(\tilde{x},\tilde{y};x,y) - ct(\tilde{x},\tilde{y};x,\varepsilon,y)\right] \\ \varepsilon_y(\tilde{x},\tilde{y})n(\tilde{x},\varepsilon(\tilde{x},\tilde{y}))d\tilde{y}d\tilde{x}\end{matrix}.$$

Here $w(\tilde{x},\tilde{y};x,\varepsilon,y)$ gives the weight a person with (x,ε,y) places on people with characteristics $\tilde{x}$ and action $\tilde{y}$ and $S(\tilde{x},\tilde{y};x,y)$ gives the social utility or disutility he obtains from associating with a person with characteristics $\tilde{x}$ and action $\tilde{y}$. Both of these are discussed below. The time cost of associating with others is *c*, which could be negative if the opportunity cost of time is low and people inherently enjoy interacting. The amount of time that a person with (x,ε,y) associates with a person with characteristics $\tilde{x}$ and action $\tilde{y}$ is $t(\tilde{x},\tilde{y};x,\varepsilon,y)$.

The social utility is integrated over the population distribution of characteristics $\tilde{x} \in X$ and actions $\tilde{y} \in [-\infty,\infty]$. Integrating over $\tilde{x}$ and $\tilde{y}$ instead of $\tilde{x}$ and $\tilde{\varepsilon}$ simplifies the remaining analysis, but requires a change of variables to get back to the population distribution of $(\tilde{x},\tilde{\varepsilon})$, which is the primitive of the model. The function $\varepsilon(x,y)$ gives the value of the unobservable characteristic, ε, that leads a person with observable characteristic x to take action y in equilibrium. Let $n(\tilde{x},\tilde{\varepsilon})$ denote the measure of $(\tilde{x},\tilde{\varepsilon})$ people in the macro group. The size of the macro-group is

$N=\int_X \int_{-\infty}^{\infty} n(\tilde{x},\tilde{\varepsilon})d\tilde{\varepsilon}d\tilde{x}$.

The weight that a person with (x,ε,y) places on a person with $(\tilde{x},\tilde{\varepsilon})$ is

$w(\tilde{x},\tilde{y};x,\varepsilon,y)\equiv \dfrac{t(\tilde{x},\tilde{y};x,\varepsilon,y)}{T(x,\varepsilon)^{\alpha}}$ where

$T(x,\varepsilon,y)\equiv \int_X \int_{-\infty}^{\infty} t(\tilde{x},\tilde{y};x,\varepsilon,y)\varepsilon_y(\tilde{x},\tilde{y})n(\tilde{x},\varepsilon(\tilde{x},\tilde{y}))d\tilde{y}d\tilde{x}$. The weight is assumed to depend on time spent associating with the other person and total time spent associating with others, $T(x,\varepsilon,y)$.

Define $W(x,\varepsilon,y)\equiv \int_X \int_{-\infty}^{\infty} w(\tilde{x},\tilde{y};x,\varepsilon,y)\varepsilon_y(\tilde{x},\tilde{y})n(\tilde{x},\tilde{\varepsilon})d\tilde{y}d\tilde{x}$ as the total weight assigned to others. The parameter α governs crowding out of associations. If $\alpha=1$, then time associating with one person simply transfers weight to that person from others, with $W(x,\varepsilon,y)=1$. When $\alpha<1$ people who spend more time associating with others experience stronger social effects.

The social utility to someone with observable characteristics x and behavior y from interacting with someone with characteristics $\tilde{x}$ and action $\tilde{y}$ is given by,

$$S(\tilde{x},\tilde{y};x,y)\equiv \theta\tilde{y}y-\frac{\psi}{2}\tilde{y}^2+\omega\tilde{y}x+\Phi\tilde{x}x'+\gamma\tilde{x}y .$$

In Manski's [1994] terminology, θ gives endogenous social effect and γ gives the exogenous social effect. The parameters ψ, ω, and Φ have no counterparts in Manski's framework. They do not affect the action directly, but affect the utility of associations and actions through their effect on associations. For a given value of y, high values of ψ reduce the utility of association with people with extreme values of $\tilde{y}$. The parameter ω allows the utility of associating with someone to vary with the other person's behavior,

$\tilde{y}$, and for the effect to depend on a person's characteristics. For instance, people with behavioral problems, may produce disutility and the disutility may be particularly great for children whose parents are highly educated. The parameter Φ is analogous, but reflects the utility effects of associations as a function of associates' characteristics. For instance, associating with people who are attractive or athletic may raise utility and the effect may be strongest for people who are themselves attractive or athletic. People may gain more utility from associating with people from the same racial or ethnic group or of the same gender. These effects can also be captured in Φ.

II.B. First Order Conditions

Differentiation with respect to *y* gives the person's optimal action,

$$\begin{aligned} y &= \beta x + \int_X \int_{-\infty}^{\infty} w(\tilde{x},\tilde{y}) S_y(\tilde{x},\tilde{y}) \varepsilon_y(\tilde{x},\tilde{y}) n(\tilde{x},\varepsilon(\tilde{x},\tilde{y})) d\tilde{y} d\tilde{x} + \varepsilon \\ &= \beta x + \theta \bar{y}^A + \gamma \bar{x}^A + \varepsilon \end{aligned} . \qquad (*)$$

Here $\bar{y}^A \equiv \int_X \int_{-\infty}^{\infty} w(\tilde{x},\tilde{y}) \tilde{y} \varepsilon_y(\tilde{x},\tilde{y}) n(\tilde{x},\varepsilon(\tilde{x},\tilde{y})) d\tilde{y} d\tilde{x}$ and

$\bar{x}^A \equiv \int_X \int_{-\infty}^{\infty} w(\tilde{x},\tilde{y}) \tilde{x} \varepsilon_y(\tilde{x},\tilde{y}) n(\tilde{x},\varepsilon(\tilde{x},\tilde{y})) d\tilde{y} d\tilde{x}$ give the (interaction weighted) mean behaviors and observable characteristics of a person's associates. This is the standard linear-in-means behavioral equation, where the total social effect will vary across people if $\alpha \neq 1$.

Associations are assumed to have two parts. Because of random encounters people are assumed spend $\tau_0(N)$ passively associating with every other member of their macro-group. We assume that τ_0 declines with macro-group size. People can also actively associate with particular members of their macro-group by spending an additional unit of time beyond τ_0 with them, so $t(\tilde{x},\tilde{y}) \in \{t_0 N^{\sigma}, t_0 N^{\sigma} + 1\}$.[6]

---

[6] The assumption that the association decision is binary simplifies the analysis, but the assumption that active associations are for an additional unit of time is without loss of generality.

This structure implies the existence of thresholds, such that a person with (x, ε, y) actively associates with all people with characteristics $\tilde{x}$, whose actions are between $y^{-}(\tilde{x}; x, \varepsilon, y)$ and $y^{+}(\tilde{x}; x, \varepsilon, y)$. Differentiating the utility function with respect to the upper threshold for any given value of $\tilde{x}$ yields,

$$\frac{\partial U(x, \varepsilon)}{\partial y^{+}(\tilde{x})} = \frac{1}{T^{\alpha}} \left[S(\tilde{x}, y^{+}(\tilde{x})) - \alpha \bar{S}^{A} \right] - c = 0,$$

The first terms in the numerator reflect the utility from increased weight placed on $(\tilde{x}, y^{+}(\tilde{x}))$ people. The second term reflects the reduction in weight placed on all others, which depends on the amount of crowding out α and the social utility from the average interaction. An analogous expression can be obtained for the optimal value of $y^{-}(\tilde{x}; x, \varepsilon, y)$.

Figure 1 illustrates the association decision. We make a number of assumptions here and below that simplifying graphical analysis. First we assume that $\omega = 0$, $\Phi = 0$, and $\gamma = 0$, which implies that *social utility depends only on actions*. We assume $\alpha = 0$, or *no crowding out*. Under these assumptions, $S = \theta \tilde{y} y - \frac{\psi}{2} \tilde{y}^{2}$. The bliss-association – the association from which the person obtains the most utility – for someone whose action is *y* is with someone whose action is $\tilde{y} = \frac{\theta}{\psi} y$. To further simplify the graph, we assume that $\psi = \theta$, so that each person's bliss-association is with someone taking the same action, which we refer to as *social utility peaks at one's own action*. A person whose bliss association is with someone with action $\tilde{y}$ obtains utility of $\frac{\theta}{2} \tilde{y}$ from that association. For each value of $\tilde{y}$, the convex parabola shown in the figure gives the social utility from associating with someone with action $\tilde{y}$ for a person taking action $\tilde{y}$.

The two downward concave parabolas give the utility for associations for people with actions $y = \tilde{y}_L$ and $y = \tilde{y}_H$ as a function of associates' actions. As shown, the utility from an association declines as the difference between the associate's behavior and one's own action increases. People actively associate with everyone for whom their utility of associating exceeds the cost *c*. A person with action *y* chooses to associate with people with actions in the range $\left[y - \sqrt{(y)^2 - 2\frac{c}{\psi}}, y + \sqrt{(y)^2 - 2\frac{c}{\psi}} \right]$. The figure shows that as an individual's action increases, the range of their associates increases especially at the high end, but also at the low end.

Our model implies that people with extreme actions obtain more utility from interacting with others. Extrapolating to a model with multiple actions, we expect people to interact most heavily along the dimensions in which their actions are greatest in absolute value. This seems consistent with observation. There are groups for people who are deeply interested in a dizzying variety of music, arts, sports, political views, and religious activities, but few for people who are moderate along these dimensions. Thus, social interactions generate extremes in behavior rather than moderating them.

II.C. Interactions and Non-Linearities

The model endogenously generates interactions between individual and group characteristics. A person with action *y* actively associates with people taking actions in the range $\left[\frac{\theta y - \sqrt{(\theta y)^2 - 2c}}{\psi}, \frac{\theta y + \sqrt{(\theta y)^2 - 2c}}{\psi} \right]$. Consider, the effect on two people whose actions are y_L and $y_H > y_L$ of introducing as small measure of people with behaviors in the range $\left(\frac{\theta y_L + \sqrt{(\theta y_L)^2 - 2c}}{\psi}, \frac{\theta y_H + \sqrt{(\theta y_H)^2 - 2c}}{\psi} \right)$, above the upper threshold of associations for people with action y_L, but beneath the upper threshold for people taking action y_H. The y_H-people will actively associate with the new people, while the y_L-people will not, so the y_H-people will be affected more by the change in macro-group composition.

With endogenous association, changes in macro-group composition can have perverse effects on behavior. Consider a macro-group with two types, high and low, with characteristics x_L and x_H. As illustrated in Figure 2, if the high group is initially close to the upper bound of the low group's associations, increasing x_H to $x'_H > x_H$ may cause the high group to fall outside the low group's associations, causing the y_L-group's behavior to decline. A similar effect could arise from introducing a group of people with behavior somewhat above the initial level of the high group. In a standard model, the behaviors of every member of the macro-group are non-decreasing in mean behavior.

The model also endogenously generates a non-linear relationship between characteristics and behavior. The effects are most easily illustrated in the case where *c=0* and $\theta = \omega$. In this case, a person with action *y* associates with people whose actions $[0,2y]$. From (*), the effect of a change in *x* on behavior is given by $\frac{dy}{dx} = \beta + \theta \frac{d\bar{y}^A}{dy}\frac{dy}{dx}$.

Assuming that there are a constant measure $\varepsilon_y n$ of people with behaviors in the relevant range $\frac{d\bar{y}^A}{dy} = 4y\varepsilon_y n > 0$ and $\frac{dy}{dx} = \beta(1 - 4\theta y\varepsilon_y n)^{-1} > 0$. Taking the second derivative yields,

$$\frac{d^2 y}{dx^2} = \frac{4\beta\theta\varepsilon_y n}{(1 - 4\theta y\varepsilon_y n)^2}\frac{dy}{dx} > 0 .$$

Thus, people with higher actions choose to associate with others taking high actions and this endogenous association generates a convex relationship between *x* and *y* even though the behavioral equation (*) is linear (similar results for changes in ε).

If the densities of *x* and ε are normal or log-normal, the density of *y* will increase at low values of *y* and then decline for higher values. In this case, the relationship between ε or *x* and *y* will follow an S-curve.

With exogenous interactions, the linear in-means-model implies that movements of people across groups are zero-sum. The interactions and non-linearities that arise endogenously in our model break this zero-sum implication.

II.D. Group Size

In the present model, the size of the macro-group affects associations even holding the distribution of characteristics in the population fixed. The effects depend on how $\tau_0(N)$ varies with macro-group size. As indicated, we assume that $\tau_0{}'(N) < 0$. If $\frac{d \ln \tau_0}{d \ln N}(N) = -1$ then the total measure of passive associations is invariant to macro-group size. The effects of macro-group size are seen most easily in a partial-equilibrium analysis in which behaviors are fixed.

The simplest case is when there is no crowding out ($\alpha = 0$). In this case, changes in population size have no effect on peoples' association thresholds. Active associations increase as a share of total associations because active associations are proportional to macro-group size, but random associations increase at a slower rate than macro-group size (because $\tau_0'(N) < 0$).

If there is crowding out, it is possible to show that active associations increase unambiguously if $\frac{d \ln \tau_0}{d \ln N}(N) = -1$.[7] When $\frac{d \ln \tau_0}{d \ln N}(N) \in (-1,0]$, the effect is ambiguous. In a general equilibrium analysis, in which behaviors change, the results would be more complicated and, some groups might experience more or less sorting depending on how

[7] This can be seen by thinking about doubling the size of the macro-group. A person could completely neutralize the change in macro-group size by maintaining the same thresholds for active associations, but interacting with only one-half of the people within those bounds. With more people in the macro-group, there would be more people close to the individual's bliss-association, so the person gains by shrinking her association thresholds and actively associating with more people close to her bliss-association, from whom she derives more utility.

the behaviors of the people around them change.

II.E. Actions and Characteristics affect the Utility of Associations Directly

Some behaviors inherently generate utility for one's associates, while others generate disutility. For instance, it may be enjoyable to associate with people who party, but unpleasant to associate with people with behavioral problems. More people will associate with people taking actions that generate utility for their associates, causing those actions to proliferate if $\theta > 0$. The parameter ω governs the effect of associates' actions on utility. When $\omega > 0$ ($\omega < 0$), the action raises (lowers) associates' utilities. As shown in figure 3, people associate more with people taking high levels of pleasant actions, which will increase that action. By contrast, people tend to avoid people taking high levels of unpleasant actions, which will reduce the amount of those actions. It is well known that actions for which θ is high will spread, but these results imply that for a given value of θ, actions for which ω is high will spread too.

Similar effects arise if characteristics affect behavior. The behaviors of people whose characteristics generate utility for others will tend to proliferate. For instance, if athletic or attractive people generate utility for their associates, people will associate with athletes and attractive people and spend more time on athletic and, say, fashion.

People may obtain more utility from associates from the same racial, ethnic or gender groups, which will lower the cost of within-group associations (see Marmaros and Sacerdote [2003]). Figure 4 illustrates active associates for two costs of associating. People associate with a wider range of people for whom their cost of associating is low. Consequently, they will have more associates from low cost groups and, assuming the same distribution of behaviors and characteristics, their associates from low-cost groups will be more different from them and each other.[8]

---

[8] A similar result holds for associations with people who are attractive or athletic. People will tend to associate with more attractive or athletic people and the attractive or athletic people with whom they

II.F. Crowding Out

For much of the analysis we have assumed that there is no crowding out of friends, that $\alpha = 0$. Crowding out generates variations in the cost of association that depend on how similar a person is to the other members of her macro-group. When social utility depends only on actions ($\omega = 0$, $\Phi = 0$, and $\gamma = 0$), but $\alpha > 0$, the condition for an individual's optimal associations is,

$$\frac{\theta y^{+} y - \frac{\psi}{2} y^{+2}}{T^{\alpha}} = c + \alpha \frac{\theta \bar{y}^{A} y - \frac{\psi}{2} \overline{y^{A2}}}{T^{\alpha}}.$$

The same condition holds for y^{-}. When $\alpha > 0$, there are effectively diminishing returns to time spent associating with others, consequently the benefits of associating depend on *T*. Because additional associations crowd out existing associations, there is also a shadow cost of associating, which is reflected by the second term on the right hand side. People who are poorly matched to their macro-groups, in the sense that they receive less utility from their associations lose less by associating with others and actively associate with a broader range of people. Put differently, someone who obtains disutility from their passive associations has more active associates to diminish the weight on their passive associations.

Figure 5 illustrates this effect. The macro-group population is assumed to be distributed across positive and negative values of *y*, with the y_H-group obtaining more social utility. Members of that group face a higher implicit cost of associating because the associations that will be crowded out generate more utility for them. As illustrated, the macro-group is sufficiently different from the y_L-group so that at an optimum they obtain an implicit benefit from reducing weight placed on their existing associations, which exceeds the direct time cost, *c*, of associating. They actively associate with a wider

---

associate will tend to be less like them.

range of people than they would otherwise.[9] If the average member of the macro-group were taking a very high action, the implicit costs of association would be reversed.

II.G. Multipliers and Reflection

The existing literature places particular emphasis on the multipliers generated when peoples' actions depend on the actions of the people in their social groups. The strength of these endogenous effects is not identified in the traditional linear-in-means model of social interactions because the expected behavior of associates is a linear function of associates' observable characteristics (see Manski [1994] and Brock and Durlauf [2001a,b]).

This section argues that endogenizing associations overturns these results in two ways. First, we expand on the previous results showing that endogenous associations generate a non-linear relationship between an individual's characteristics and her behaviors, making Manski's endogenous effect identifiable in principle. We also show that with endogenous associations, multipliers can arise even in a model without Manski's endogenous effect and illustrate this point in a simple model.

II.E.1. Reflection

In a general model, with no crowding out ($\alpha = 0$), the thresholds for active associations are,

$$y^{-}(\tilde{x}; x, \varepsilon, y) = \frac{\theta y + \omega x}{\psi} - Z^{\frac{1}{2}} \text{ and } y^{+}(\tilde{x}; x, \varepsilon, y) = \frac{\theta y + \omega x}{\psi} + Z^{\frac{1}{2}}$$

where $Z \equiv \left(\frac{\theta y + \omega x}{\psi}\right)^2 + \frac{2}{\psi}\left(\Phi \tilde{x} x' + \gamma \tilde{x} y - c\right)$. The effect of x on the thresholds is,

$$\frac{dy^{+}(\tilde{x})}{dx} = \frac{\theta}{\psi}\frac{dy}{dx} + \frac{\omega}{\psi} + \frac{1}{2}\frac{dZ}{dx} Z^{-\frac{1}{2}} \text{ and } \frac{dy^{-}(\tilde{x})}{dx} = \frac{\theta}{\psi}\frac{dy}{dx} + \frac{\omega}{\psi} - \frac{1}{2}\frac{dZ}{dx} Z^{-\frac{1}{2}}.$$

The effect of changing x on behavior is,

[9] As illustrated, crowding out can break the implication that the associates of people with low actions are a

$$\frac{dy}{dx} = \beta + \int_X \left\{ \begin{array}{l} \frac{\partial y^+(\tilde{x})}{\partial x}\left[\theta y^+(\tilde{x}) + \gamma\tilde{x}\right]\varepsilon_y\left(\tilde{x}, y^+(\tilde{x})\right)n\left(\tilde{x}, \varepsilon\left(\tilde{x}, y^+(\tilde{x})\right)\right) \\ \quad - \frac{\partial y^-(\tilde{x})}{\partial x}\left[\theta y^-(\tilde{x}) + \gamma\tilde{x}\right]\varepsilon_y\left(\tilde{x}, y^-(\tilde{x})\right)n\left(\tilde{x}, \varepsilon\left(\tilde{x}, y^-(\tilde{x})\right)\right) \end{array} \right\} d\tilde{x} .$$

In general, this (implicit) relationship between *x* and *y* will be non-linear.

This expression shows that the density of characteristics at the upper and lower thresholds of a person's associations affect the relationship between *x* and *y*. Given that the association thresholds are estimable and the density of *x* is observable, random variations in the distribution of *x* could be used to identify non-linearities in the relationship between *x* and *y*.[10]

II.E.2. Multipliers without Endogenous Effects

To illustrate how multipliers can arise even without Manski's endogenous effect consider a simple model. To highlight the effect of interest, assume

$$S(\tilde{x}, \tilde{\varepsilon}; x, y) \equiv \gamma y + \omega y(\tilde{x}, \tilde{\varepsilon}) - \frac{\psi}{2} y(\tilde{x}, \tilde{\varepsilon})^2 .$$

That is, we assume that the direct effect of $\tilde{y}$ on utility (given by ω) does not vary with own characteristics; that the marginal utility of the action depends on the number of associates, but not their characteristics (that the γ effect does not depend on $\tilde{x}$); and that there is no Manski endogenous effect ($\theta = 0$).

This model implies the behavioral equation:

$$y = \beta x + \int_X \int_{-\infty}^{\infty} w(\tilde{x}, \tilde{\varepsilon}) n(\tilde{x}, \tilde{\varepsilon}) d\tilde{\varepsilon} d\tilde{x} + \varepsilon .$$

Thus, the action depends on an individual's characteristics and the number of people with whom the person associates, which might be appropriate for thinking about partying. The thresholds for association are $y + \omega\tilde{y} - \frac{\psi}{2}\tilde{y}^2 = c$. The situation is illustrated in Figure 6.

---

subset of the associates of people with higher actions.

[10] One would have to make assumptions about the distribution of ε, such that it is independent of *x*.

Assume that initially *x* is distributed so that people's actions are concentrated around y_0, which lies slightly beneath the lower bound for associations for people taking action y_0. Raising *x* for some members of the group will cause these people to move above the lower threshold of association, leading others to associate with them. People who begin to associate with the people whose *x* increased, will themselves increase their action causing more people to associate with them, further increasing the actions.

Thus, multipliers can arise with a change in *x* for a small number of people leading many people to increase their action even when there is no endogenous effect in the model.

II.H. Peer Pressure

People often describe social interactions models as models of peer pressure. An important aspect of peer pressure is the withdraw of association when one person takes an action that produces disutility for another. To capture this idea, we assume that people derive utility directly from having people associate with them.

Define $y^{O-}(\tilde{x};x,\varepsilon,y)$) as the lowest action at which someone with characteristics $\tilde{x}$, will want interact with someone with (x,ε,y). Figure 7 is drawn assuming that social utility depends only on actions ($\omega=0$, $\Phi=0$, and $\gamma=0$) and that social utility peaks at one's own action ($\theta=\psi$). In it, $y^{O-}(\tilde{x};x,\varepsilon,y_2)=y_1 \ \forall\tilde{x},x,\varepsilon$, so that people will want to associate with someone taking action y_2 if and only if they are taking an action of y_1 or higher.

We modify the utility function to be,

$$U(x,\varepsilon)=\beta xy+\varepsilon y-\frac{1}{2}y^2+\int_X\int_{-\infty}^{\infty}\begin{matrix}[w(\tilde{x},\tilde{y};x,\varepsilon,y)S(\tilde{x},\tilde{y};x,y)-ct(\tilde{x},\tilde{y};x,\varepsilon,y)] \\ \varepsilon_y(\tilde{x},\tilde{y})n(\tilde{x},\varepsilon(\tilde{x},\tilde{y}))d\tilde{y}d\tilde{x}\end{matrix}$$
$$+\int_X\int_{y^{O-}(\tilde{x};x,\varepsilon,y)}^{\infty}\pi(\tilde{x},\tilde{y};x,\varepsilon,y)\varepsilon_y(\tilde{x},\tilde{y})n(\tilde{x},\varepsilon(\tilde{x},\tilde{y}))d\tilde{y}d\tilde{x}.$$

Here $[y^{O-}(\tilde{x};x,\varepsilon,y),\infty]$ gives the range of people who will choose to interact with someone with (x, ε, y) and $\pi(\tilde{x}, \tilde{y}, x, \varepsilon, y)$ gives the amount of utility that someone with (x, ε, y) obtains from having someone with $(\tilde{x}, \tilde{y})$ interact with them or strength of the peer pressure effect. The first order condition for the optimal action is now,

$$\begin{aligned} y &= \beta x + \int_X \int_{-\infty}^{\infty} w(\tilde{x}, \tilde{y}) S_y(\tilde{x}, \tilde{y}) \varepsilon_y(\tilde{x}, \tilde{y}) n(\tilde{x}, \varepsilon(\tilde{x}, \tilde{y})) d\tilde{y} d\tilde{x} + \varepsilon \\ &- \int_X \pi\left(\tilde{x}, y^{O-}(\tilde{x}; x, \varepsilon, y); x, \varepsilon, y\right) \frac{\partial y^{O-}(\tilde{x}; x, \varepsilon, y)}{\partial y} \varepsilon_y\left(\tilde{x}, y^{O-}(\tilde{x}; x, \varepsilon, y)\right) n\left(\tilde{x}, \varepsilon\left(\tilde{x}, y^{O-}(\tilde{x}; x, \varepsilon, y)\right)\right) d\tilde{x} \end{aligned}.$$

This condition is the same as (*) except for the last term, which gives the effect of an increase in the person's action on the measure of people who will associate with him ("weighted" by how much utility the person gets from having each person associate with them). In this model of peer pressure, actions do not only depend on the mean of associates behaviors and characteristics, but also on the density of people at the margin to interact with someone. In this sense, the standard linear-in-means model of interactions may be a poor representation of peer pressure.

Define $\underline{y^{O-}(\tilde{x}; x)}$ as the lowest action that someone with characteristics *x* can take that will generate positive net utility to people with characteristics $\tilde{x}$ to interact with. Among people with characteristics *x*, $\underline{y^{O-}(\tilde{x}; x)}$ is the action that induces the most people with characteristics $\tilde{x}$ to interact with a person. In the figure, where characteristics do not directly affect the utility of interacting and where people's bliss-interactions are with people taking the same action, $\underline{y^{O-}}$ is the lowest action at which the utility from the bliss-interaction exceeds the cost of association. Any deviation from $\underline{y^{O-}}$ reduces the number of people interacting with someone.

In the simple case where $\pi(\tilde{x}, \tilde{y}, x, \varepsilon, y) = \pi \ \forall \tilde{x}, \tilde{y}, x, \varepsilon, y$, a person with an action beneath $\underline{y^{O-}}$ experiences peer pressure to take a higher action in that doing so increases the number of people associating with him. Similarly a person with an action beneath $\underline{y^{O-}}$ experience peer pressure to take a lower action. Thus peer pressure can lead people to either increase or decrease their actions and the effect will depend on the distribution of people in the macro-group.

Characteristics that directly affect the utility of association, such as looks, affect the strength of "transmission" and "reception" of peer pressure. On the transmission-side, good looks will raise $\pi(\tilde{x}, \tilde{y}, x, \varepsilon, y)$ for others, so people will be particularly subject to peer pressure from attractive people. On the reception-side, insofar as most people in the macro-group will want to associate with them regardless of their actions, highly attractive people will be less subject to peer pressure than others. Interestingly, people whose characteristics make them very undesirable as associates may also be less subject to peer pressure insofar as few people will associate with them independent of their actions. Thus the model suggest that the extent to which people are subject to peer pressure as recipients follows an inverse-U in how appealing they are as associates. The strength of peer pressure from the transmission side will be strictly increasing in how appealing someone is an associate.

II.I. Two-Sided Associations

We have assumed that associations can be one-sided, and often one person invests much more in a relationship than the other. Investments by the people in a relationship are likely to be complementary. In this section we consider the extreme case where a person can only associate with another when the second person associates with him.

In Figure 7 a person with action y_2 would like to associate with people taking actions in the range $\left[y^{-}(\tilde{x}; x, \varepsilon, y_2), y^{+}(\tilde{x}; x, \varepsilon, y_2)\right] \forall \tilde{x}, x, \varepsilon$ and would do so if one-sided associations were possible. But people beneath $y_1 = y^{O-}(x_0; x, \varepsilon, y_2)$ do not associate with him, so if associations must be two-sided, he associates with people in the range $\left[y^{O-}(\tilde{x}; x, \varepsilon, y_2), y^{+}(\tilde{x}; x, \varepsilon, y_2)\right] \forall \tilde{x}, x, \varepsilon$.

While the quantitative implications discussed above are altered when associations are required to be two-sided, most of the qualitative implications remain. Thus, people with higher levels of an action will associate with others with high levels of that action. The effects of groups will be non-linear, and there will in general be a non-linear relationship between own characteristics and behaviors.

II.J. Extensions

This section considers two extensions to the model – popularity and dislike. Popularity can be introduced simply by making one of the individual characteristics, x, be the measure of people in the macro-group who associate with a person.[11] Presumably, people obtain more utility from interacting with people who are popular. In the context of peer pressure, people will be particularly interested in having popular people associate with them.

When popularity *per se* has effects an intervention that increases the weight placed on one person or set of people will generate a feedback to other people in the macro group, which can generate multipliers – as one person begins to associate with a person that raises the utility that others derive from associating with him. Such effect explain the emphasis places on affecting how much weight people place on others, by making positive (negative) examples of people they want to do (not) want emulated, such as valedictorians (Moffitt [2001] discusses a variety of policies).

Lastly, one might allow people to invest time and energy in placing less weight on some members of their macro-group. People would optimally want to down-weight people who are most different from them.

III. Data

We use data from the National Longitudinal Study of Adolescent Health (Add Health). The Add Health data provide information on wide range of youth behaviors, including substance use, risky behaviors, and behavioral problems and data on family

[11] The fixed time that people spend with other members of their macro group may also depend on their popularity, insofar as popular people are likely to come up in conversations.

background. They contain nationally representative data on 90 thousand students enrolled in grades 7-12 in 132 schools. Students were asked about their 5 closest male and female friends and the confidential data provide identifiers so that it is possible to reconstruct the friendship networks within schools. These data have a number of advantages for our purposes. First, they provide information on both the macro groups – schools and grades – as well as the subgroupings that emerge within schools between individual students and all people in the macro-groups are surveyed symmetrically.[12]

We restrict attention to students enrolled in 9th through 12th grades in classes with at least 25 students. The data contain only a limited panel aspect, which we do not utilize in this study.

Appendix Table 1 lists the outcome variables used in this study and how they were constructed. In most cases variables were given as the frequency of a behavior. In these cases, we coded them so as to give a monthly frequency.

Table 1 reports descriptive statistics for the 44,760 people with data on associations. Slightly under half the sample is male; 16% identify as black (respondents were allowed to report multiple races); 14% identify as Hispanic. Just over three quarters of respondents live with their father, and their mothers have 13.5 years of schooling on average. On average the respondents smoked a bit more than once a week, they drink about once a week, and get drunk about half as frequently. Among the most common behaviors or behavior problems are having trouble paying attention and doing homework regularly and lying to parents. Skipping school and fighting are among the least common, occurring less than once every two weeks on average.

Also shown are the means and standard deviations of the associate averages of the

[12] Alternative surveys provide information about particular individuals and ask the respondents about the behaviors of their friends. While valuable, under this approach individuals and their friends are surveyed asymmetricly, so that less information is available on friends than on respondents. Reports of friends' behaviors may exhibit different reporting biases than reports of own behaviors.

behaviors and characteristics. The mean of the associates behaviors are consistently above the mean of the own behaviors, indicating that people with "worse" behaviors are listed as friends more often than people with better behavior. The mean of mother's education and, to a lesser extent, father present among associates is above the own means for these variables, suggesting that children with better family backgrounds may be more appealing because they are wealthier or more socially adept.

IV. Variance Decomposition

As indicated, most studies of social effects focus on variations in macro-group characteristics. We begin by investigating how much of the variation in friend's behaviors and characteristics arises within macro-groups as opposed to between macro-groups.

Meaningful macro-groups should be reasonably narrowly defined, yet contain a large portion of peoples' associates. School-grades meet this criterion with 74% of the surveyed friends being in the person's school-grade.[13]

Let z_{ij} denote a behavior or characteristic of the *j*th friend of person *i*. The mean value of *z* among person *i*'s friends is $\bar{z}_i = \frac{1}{N_i}\sum_j z_{ij}$, where N_i gives the number of people with whom *i* is friends. We define the macro-group as grades within schools, and are interested in decomposing the variance of $\bar{z}_i$ into within and between school-grade components. The total variance in $\bar{z}_i$ is $VAR[\bar{z}_i] = VAR[\bar{z}_i - E[\bar{z}_i | S_i, G_i]] + VAR[E[\bar{z}_i | S_i, G_i]]$. The former term represents the variance arising within school-grades and the later the variance arising across school-grades.

The share of the variance arising within school grades is the R^2 from an analysis

[13] This share is somewhat higher in grades nine (79%) and twelve (77%) and somewhat lower for grades

of variance with a full set of school-grade interactions, which is reported the first column of Table 2. The estimates indicate that for most behaviors about 90% of the variation in friends' behaviors arises within school-grades. Roughly 85% of the sorting on family background, measured by mother's education and the father's presence arises within school-grades. Considerably more of the variation in the racial and ethnic composition of friends can be accounted for by school-grade effects, often 40% or more, which is to be expected given racial and ethnic residential sorting.

Because people often form friendships with people who live near them, a portion of the within-school grade sorting likely arises from residential sorting within school-grades. The Add Health survey contains some data on the characteristics of the census block group of residence for a portion of respondents.[14] Because neighborhood characteristics are only available for one in five respondents, column 2 reports the R^2 from analysis of variance described above when applied to the smaller sample; column 3 reports results when the neighborhood variables interacted with school effects are added to the model. Switching to the smaller sample tends to increase the R^2 slightly because the sample size declines dramatically relative to the number of effects. Inclusion of neighborhood variables accounts for roughly 5% of the variance in friends' behavior, but the variance within school-grades and school-neighborhoods is still considerably larger. To the extent that people choose their friends based on their behaviors and there is an association between neighborhood characteristics and behavior, these estimates will overstate the effect of residential proximity.

While respondents were able to list 5 male and 5 female friends and most list fewer, respondents may have friends that they did not list either because they were

ten (71%) and eleven (69%).

[14] The neighborhood characteristics we use are whether the block group is urban, percent black, percent living in poverty, the percent of the adult population without high school degrees, the percent unemployed, and the log of the median family income.

constrained by the survey or because they chose not to list all their friends. Additionally, data was not available on friends outside of the sample. To see how much of the within school-grade variance in the mean friend characteristics and behaviors might be due to noise, we take people who list at least two friends, and randomly assign each friend to one of two groups, *A* and *B*, as close to equal in size as possible. The mean friend characteristics or behaviors in group *k* for person *i* are,

$$\bar{z}_i^k = \frac{1}{N_i^k}\sum_j z_{ij}^k = \mu_i + \frac{1}{N_i^k}\sum_k \varepsilon_{ij}^k \ ,$$

where z_{ij}^k denotes the characteristics or behavior of friend *j* in group *k* of person *i*'s friends; μ_i denotes the mean value of *z* among person *i*'s friends; and ε_{ij}^k gives sampling variation. The covariance of the means for the two groups conditional on the school, S_i, and grade, G_i, gives the within school-grade variance in μ_i,

$$COV\left[\bar{z}_i^A, \bar{z}_i^B \middle| S_i, G_i\right] = E\left[\left(\bar{z}_i^A - E\left[\bar{z}_i^A \middle| S_i, G_i\right]\right)\left(\bar{z}_i^B - E\left[\bar{z}_i^B \middle| S_i, G_i\right]\right)\right] = VAR\left[\mu_i \middle| S_i, G_i\right].$$

Column 4 reports these estimates of the within school-grade variance in mean friend behaviors and characteristics.

The variance in the mean of friends' behaviors and characteristics within school-grades is

$$\begin{aligned} VAR\left[\bar{z}_i \middle| S_i, G_i\right] &= E\left[\left(\bar{z}_i - E\left[\bar{z}_i \middle| S_i, G_i\right]\right)^2\right] = E\left[\left(\mu_i - E\left[\bar{z}_i \middle| S_i, G_i\right]\right)^2 + \left(\frac{1}{N_i}\sum_k \varepsilon_{ik}\right)^2\right] \\ &= VAR\left[\mu_i \middle| S_i, G_i\right] + VAR\left[\frac{1}{N_i}\sum_k \varepsilon_{ik}\right] \end{aligned}.$$

These figures are given in column 5 of the table. Under the assumption that each student has an infinite number of friends, the ratio of the within school-grade covariance and variance, $\frac{COV\left[\bar{z}_i^A, \bar{z}_i^B \middle| S_i, G_i\right]}{VAR\left[\bar{z}_i \middle| S_i, G_i\right]}$, gives the share of the within school-grade variation that is due to differences in the mean of friends' *z*'s within school-grades as opposed to noise.

Under the more reasonable assumption that each person has twice as many friends as they list, $\frac{COV\left[\bar{z}_i^A, \bar{z}_i^B \middle| S_i, G_i\right]}{2VAR\left[\bar{z}_i \middle| S_i, G_i\right]}$ gives the share of the within school-grade variance that is noise. The portion of the non-noise variation in mean friends' *z*'s that arises between macro-groups is

$$\frac{VAR\left[E\left[\bar{z}_i \middle| S_i, G_i\right]\right]}{\left(VAR\left[\bar{z}_i\right] - VAR\left[E\left[\bar{z}_i \middle| S_i, G_i\right]\right]\right)\frac{COV\left[\bar{z}_i^A, \bar{z}_i^B \middle| S_i, G_i\right]}{2VAR\left[\bar{z}_i \middle| S_i, G_i\right]} + VAR\left[E\left[\bar{z}_i \middle| S_i, G_i\right]\right]}.$$

These figures are reported in the last three columns of the table (for the models in columns (1), (2), and (3)).

In the estimates without neighborhood controls, between 20% and 90% of the variation in behaviors arises within-school grades, with the median being 70%. In the last column, the within school-grade and school-neighborhood variation is between 5% and 90% of the total, with the median being 55%. Thus, even after accounting for potential underreporting of associates, our estimates indicate that the majority of variation in associates behavior arises within macro-groups.

V. The Choice of Associates

V. A. Own Behavior and Characteristics and Mean Associate Behavior and Characteristics

Given that most of the variation in associates' behavior and characteristics arises within school-grades, this section studies the determinants of associates' mean behaviors and characteristics within school-grades. We begin with simple regressions of the mean of associates' behaviors on own behavior and characteristics. Let y_i denote the behavior or characteristic of person *i*. We regress the mean behavior of *i*'s associates, $\bar{y}_i^A$, on y_i We anticipate a positive relationship between $\bar{y}_i^A$ and y_i. Formally, our model is

$$\bar{y}_i^A = \beta y_i + \gamma X_i + \Pi \overrightarrow{SG}_i + \varepsilon_i.$$

Person *i*'s own observable characteristics, X_i, are included to control for differences in basic individual characteristics that correlate with behavior and affect the choice of associates.[15] Also included is a vector of dummy variables containing a complete set of interactions between the school and grade, $\overrightarrow{SG}_i$. With these school-grade effects, we estimate whether people with higher levels of *z* have associates with higher values of *z* compared to others in the same school-grade.

The estimates reported in Table 3 indicate a strong positive relationship between own behavior and characteristics and those of peoples' associates. Our goal is not to estimate the causal effect of a change in one persons' behavior on their choice of associates and there are a number of reasons not to interpret these estimates as such. First, if social interactions are important, a person's behavior will affect their friends' behavior and vice versa. Second, the estimated relationships may reflect sorting in other arenas such as neighborhoods or in classes or tracks within school-grades. Last, people selectively associate based on characteristics as well as behaviors. This form of sorting is, of course, in keeping with the model and is suggested by the estimates for individual characteristics. We do not attempt to disentangle selective sorting on behaviors from selective sorting on the characteristics that determine those behaviors.

To address neighborhood-based selection, we augment our earlier models by including neighborhood characteristics. (We do not have data on the classes that people take or on tracking.) The remaining columns of the table report estimates from the original specification for the sample with neighborhood characteristics and then estimates that include the neighborhood variables. The substantial reduction in the sample changes the estimates, often reducing their precision, but including the neighborhood variables has little effect.

---

[15] Our controls are age, mother's education, and dummy variables for, gender, race (white, black, Asian, Indian, and other), Hispanic background, and mother and father present. Dummy variables are included for

V. B. The Effects of Mean School-Grade Behavior and Characteristics

The model implies that people who are in macro-groups where other members have similar behaviors or characteristics, will tend to interact particularly heavily with those people. To test this hypothesis, we augment the previous model by including an interaction between a behavior or characteristics, z_i, and the mean of that behavior or characteristic in person *i*'s school-grade, $\bar{z}_i^{SG}$. (For variables that are not binary, we also include the square of z_i.) The theory implies a positive relationship between $\bar{z}_i^A$ and the interaction $z_i * \bar{z}_i^{SG}$. The model is,

$$\bar{z}_i^A = \phi_1 z_i + \phi_2 z_i^{\,2} + \beta z_i * \bar{z}_i^{SG} + \gamma X_i + \Pi \overrightarrow{SG}_i + \varepsilon_i .$$

Our estimates indicate how an increase in a behavior or characteristic in a school-grade affects the prevalence of that behavior or characteristic among the associates of people with high values of it relative to others in the school-grade.

The estimates, reported in the first two column of table 4, are positive for all but 3 behaviors, and are positive in the 8 cases that are statistically significant. Among the background characteristics 9 of the 10 are statistically significant and 8 of these are positive. Again, we address neighborhood based sorting, by including neighborhood characteristics. Columns 3 and 4 report estimates from the original specification for the sample with neighborhood characteristics; columns 5 and 6 include these controls. The 80% reduction in the sample makes the estimates less precise, but inclusion of the neighborhood variables has a negligible impact on the estimates.

These estimates may be biased up if students in school-grades with a high value of a behavior or characteristics are different in unobservable ways from those in school-grades with lower levels of that behavior or characteristic. For instance, students in school-grades with a high percentage from their racial or ethnic group may come from

---

Hispanic background missing and mother's education missing.

families that particularly identify with that group (although Austen-Smith and Fryer [Forthcoming] suggest that student responses might compensate).

In an approach that is similar to Hoxby [2000], we focus on the component of the cross-grade variation within schools that is most likely to be random. To identify this variation, we regress the mean behavior in each school grade on a school dummy variables and school-specific quadratics. The residuals from these regressions are likely to be dominated by random variations in the composition of classes (which may be amplified by social interactions). We instrument for the interaction between own behavior and the mean in the school grade, $y_i * \bar{y}_i^{SG}$, with the interaction between own behavior and the residuals from these regressions and their square. The estimates are reported in the last column of the table. Again, the reduction in the variation makes the estimates are less precise, but they remain positive and statistically significant for 6 of the variables and are of roughly similar magnitude.

V. C. School-Grade Size

The model implies that large macro-groups facilitate sorting so that there should be a stronger relationship between individual behavior and associates behavior in large macro-groups. Before discussing our methods and results, we assess how grade size relates to observable characteristics. To do this, we estimate the model,

$$\ln\left(N^{SG}\right)_{sg} = \beta\bar{X}_{sg}^{SG} + \theta\overline{Neig}_{sg}^{SG} + \sum_{j=9}^{11}\gamma_j I\left(Grade_{sg} = j\right) + \varepsilon_{sg} .$$

Here $\ln\left(N_i^{SG}\right)_{sg}$ denotes the natural logarithm of the number of students in school-grade *sg*, $\bar{X}_{sg}^{SG}$ denotes their mean observable characteristics, and $\overline{Neig}_{sg}^{SG}$ denotes the mean of their neighborhood characteristics. The specification includes dummy variables, γ_j, to control for cross-grade variation in size. The standard errors are clustered within schools. The estimates in Table 5 indicate few systematic relationships between student or neighborhood characteristics and grade size

To assess how school-grade size affects selective association, we estimate

$$\bar{z}_i^A = \phi_1 z_i + \phi_2 z_i^{\ 2} + \beta z_i * \log\left(N_i^{SG}\right) + \gamma X_i + \Pi \overrightarrow{SG}_i + \varepsilon_i .$$

Under the hypothesis that grade size facilitates selective sorting $\beta > 0$. The estimates are reported in the first two columns of Table 6. In all but two case, an increase in the size of the school-grade increases the relationship between own behavior or characteristics and those of associates and the 14 statistically significant relationships are all positive.

Our first estimates do not suggest that grade size is unrelated to observable characteristics of students. To further control for heterogeneity, we instrument for the interaction between the behaviors or characteristics and log grade size with the interaction between the behaviors or characteristics and the residual of log grade size from school-specific quadratics in grade. These estimates are reported in columns 3 and 4. Here 7 of the 9 statistically significant relationships are positive.

VI. Conclusion

We have developed a model of social interactions with endogenous association. Our theory implies the standard linear-in-means social interactions model of behavior, but departs from the standard model in its assumptions about interactions. We assume that people are able to affect who they associate with within their groups by investing time and effort in developing relationship and that they do so in order to maximize their utility. People optimally choose to associate with people whose characteristics and behaviors are similar to their own, so the effects of social groups on behavior are non-linear and interact with individual characteristics. Our model predicts more selective association in large macro groups. When associations are endogenous "improvements" in the composition of a macro-group may have adverse consequences for portions of the macro-group.

We investigate a range of the model's empirical implications. Consistent with the model, people with higher levels of a behavior associate with others with high levels of that behavior. An increase in the mean of a behavior in a macro-group raises the mean of that behavior more for people heavily engaged in that behavior. We show that increases in macro-group size also increases sorting as measured by the strength of the relationship between own behavior and the behavior of associates.

References

Aaronson, Daniel, "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *Journal of Human Resources* 33 (no. 4, Fall 1998), 915-946.

Angrist, Joshua D. and Kevin Lang. "How Important are Classroom Peer Effects? Evidence from Boston's Metco Program." *American Economic Review*, Forthcoming.

Anselin, Luc. *Spatial Econometrics: Methods and Applications.* Dordrecht: Kluwer Academic Publishers. 1988.

Austen-Smith, David and Roland G. Fryer Jr. "An Economic Analysis of 'Acting White'" *Quarterly Journal of Economics*, Forthcoming.

Bandiera, Oriana, and Imran Rasul. "Social Networks and Technology Adoption in Northern Mozambique." Working Paper. London School of Economics. 2002.

Bayer, Patrick; Robert McMillan; and Kim Rueben. "An Equilibrium Model of Sorting in and Urban Housing Market: A Study of the Causes and Consequences of Residential Segregation." Working Paper. Yale University. 2002.

Bayer, Patrick; Stephen L. Ross; and Giorgio Topa. "Place of Work, Place of Residence: Informal Hiring Networks and Labor Market Outcomes." Working Paper. University of Connecticut. 2004.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics* 115 (no. 3 August 2000): 1019-55.

Borghans, Lex; Bas ter Weel; and Bruce A. Weinberg. "People People." In progress.

Borjas, George J. "Ethnicity, Neighborhoods, and Human Capital Externalities." *American Economic Review* 85 (no. 3, June 1995): 365-390.

Brock, William A. and Steven N. Durlauf. "Discrete Choice with Social Interactions." *Review of Economic Studies* 68 (2001a): 235-260.

Brock, William A. and Steven N. Durlauf. "Interactions-Based Models." *Handbook of Econometrics, vol 5.* James J. Heckman and Edward E. Leamer, eds. Amsterdam: North-Holland Press. 2001b.

Case, Anne C., and Lawrence Katz. "The Company you Keep: The Effects of Family and Neighborhoods on Disadvantaged Youth." NBER Working Paper no. 3705. Cambridge, MA. 1990.

Conley, Timothy G. and Giorgio Topa. "Socio-Economic Distance and Spatial Patterns in Unemployment." *Journal of Applied Econometrics* 17 (no. 4, July-August 2002): 303-27.

Conley, Timothy G. and Christopher R. Udry. Learning About a New Technology: Pineapple in Ghana. Working Paper 817. Economic Growth Center, Yale University. 2001.

Corcoran, Mary, Roger Gordon, Deborah Laren, and Gary Solon. "The Association between Men's Economic Status and Their Family and Community Origins." *Journal of Human Resources* 27 (no. 4, Fall 1992): 575-601.

Crane, Jonathan. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." *American Journal of Sociology* 96 (no. 5, March 1991): 1226-1259.

Cummings, Jean and Denise DiPasquale, and Matthew Cummings. Measuring the Consequences of Promoting Inner City Homeownership. Working Paper. Tufts University. 2001.

Cutler, David M., and Edward L. Glaeser. "Are Ghettos Good or Bad?" *Quarterly Journal of Economics* 112 (no. 3, August 1997): 827-872.

Datcher, Linda. "Effects of Community and Family Background on Achievement." *Review of Economics and Statistics* 64 (no. 1, Feb., 1982): 32-41.

Deitz, Robert. "Estimation of Neighborhood Effects in the Social Sciences: An Interdisciplinary Literature Review." *Social Science Research* 31 (No. 4, December 2002): 539-75).

DiPasquale, Denise and Matthew Kahn. "Measuring Neighborhood Investments: An Examination of Community Choice." *Real Estate Economics* 27 (1999): 389-424.

Duncan, Greg J., James P. Connell, and Pamela K. Klebanov. "Conceptual and Methodological Issues in Estimating Causal Effects of Neighborhoods and Family Conditions on Individual Development." In *Neighborhood Poverty, Volume I: Context and Consequences for Children*, edited by Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber. New York: Russell Sage Foundation, 1997.

Evans, William N., Oates, Wallace E., and Schwab, Robert M. "Measuring Peer Group Effects: A Study of Teenage Behavior" *Journal of Political Economy* 100 (no. 5, October 1992): 966-991.

Galster, George C.; Roberto G. Quercia; and Alvaro Cortes. "Identifying Neighborhood Thresholds: An Empirical Exploration." *Housing Policy Debate* 11(no. 3, 2000): 701-32.

Galster, George. "An Economic Efficiency Analysis of Deconcentrating Poverty Populations." *Journal of Housing Economics* 11 (2002): 303-329.

Glaeser, Edward L.; Bruce Sacerdote; and Jose A. Scheinkman. "Crime and Social Interactions." *Quarterly Journal of Economics* 111 (No. 2, May 1996): 507-48.

Glaeser, Edward L.; Bruce Sacerdote; and Jose A. Scheinkman. “The Social Multiplier.” *Journal of the European Economic Association* 1 (No. 2-3, April-May 2003): 345-53.

Gould, Eric D; Victor Lavy; and M. Daniele Paserman. “Immigrating to Opportunity: Estimating the Effect of School Quality Using a Natural Experiment on Ethiopians in Israel.” *Quarterly Journal of Economics* 119 (No. 2, May 2004a): 489-526.

Gould, Eric D; Victor Lavy; and M. Daniele Paserman. “Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence.” Working Paper, Hebrew University. 2004b.

Haurin, Donald R., Robert Deitz, and Bruce A. Weinberg. “The Impact of Neighborhood Homeownership Rates: A Review of the Theoretical and Empirical Literature.” *Journal of Housing Research* 13 (no. 2, 2002): 119-52.

Hoxby, Caroline. “Peer Effects in the Classroom: Learning from Gender and Race Variation.” National Bureau of Economic Research Working Paper 7867. 2000.

Ioannides, Yannis M. and Jeffrey E. Zabel. “Neighborhood Effects and Housing Demand.” *Journal of Applied Econometrics*, Forthcoming.

Jacob, Brian. “Public Housing, Housing Vouchers and Student Achievement: Evidence from the Public Housing Demolitions in Chicago.” *American Economic Review* 94 (No. 1, March 2004): 233-58.

Jargowsky, Paul A. *Poverty and Place: Ghettos, Barrios, and the American City*. New York: Russell Sage Foundation, 1997.

Jencks, Christopher, and Mayer, Susan E. “The Social Consequences of Growing Up in a Poor Neighborhood.” In *Inner-City Poverty in the United States*, edited by Laurence E. Lynn, Jr. and Michael G. H. McGeary. Washington: National Academy Press, 1990.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. “Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment.” *Quarterly Journal of Economics* 116 (May 2001): 607-654.

Kling, Jeffrey R.; Jens Ludwig; and Lawrence F. Katz. “Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Mobility Experiment.” *Quarterly Journal of Economics* 120 (February 2005).

Krauth, Brian. “Social Interactions, Thresholds, and Unemployment in Neighborhoods.” Working Paper. 2000.

Ladd, Helen F. and Jens Ludwig. “Federal Housing Assistance, Residential Relocation, and Educational Opportunities: Evidence from Baltimore.” *American Economic Review* 87 (no. 2, May 1997): 272-277.

Lee, Lung-Fei. "Generalized Method of Moments Estimation of Spatial Autoregressive Processes." Working Paper, Ohio State University. 2001.

Ludwig, Jens, Greg J. Duncan, and Paul Hirschfield. "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment." *Quarterly Journal of Economics* 116 (May 2001):655-680.

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

Marmaros, David; and Bruce Sacerdote. "How Friendships Form." Working Paper, Dartmouth College. 2003.

Massey, Douglas S., and Nancy A. Denton. *American Apartheid: Segregation and the Making of the Underclass.* Cambridge, MA and London: Harvard University Press. 1993.

Moffitt, Robert A. "Policy Interventions, Low-Level Equilibria, and Social Interactions." *Social Dynamics*, Stven N. Durlauf and H. Peyton Young, eds. Cambridge, MA: MIT Press. 2001.

Moody, James. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology* 107 (No. 3, November 2001): 679-716.

Munschi, Kaivan. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." Working Paper. University of Pennsylvania. 2002.

Oreopolous, Philip. "The Long-run Consequences of Living in a Poor Neighborhood." *Quarterly Journal of Economics* 118 (No. 4, November 2003): 1533-75.

Plotnick, Robert D., and Saul D. Hoffman. "Fixed Effect Estimates of Neighborhood Effects." Working Paper. University of Delaware. 1995.

Quercia, Roberto G. and George C. Galster. "Threshold Effects and Neighborhood Change." *Journal of Planning Education and Research*, Forthcoming.

Rosenbaum, James E., Stefanie DeLuca, and Shazia Miller. "The Long-Term Effects of Residential Mobility on AFDC Receipt: Studying the Gautreaux Program with Administrative Data." Manuscript. Northwestern University. 1999.

Ross, Stephen L. "Segregation and Racial Preferences: New Theoretical and Empirical Approaches." *Annales d'Economie et de Statistique* 71-72 (No. 0, Special Issue July-Dec. 2003): 143-72.

Solon, Gary, Marianne E. Page and Greg J. Duncan. "Correlations between Neighboring Children in Their Subsequent Educational Attainment." *Review of Economics and Statistics* 82(no. 3, August 2000): 383-92.

Sobel, Michael E. "What are Neighborhood Effects?" Manuscript. Columbia University. 2001.

Topa, Giorgio. “Social Interactions, Local Spillovers, and Unemployment.” *Review of Economic Studies* 68 (No. 2, April 2001): 261-95

Weinberg, Bruce A. “Black Residential Centralization and the Spatial Mismatch Hypothesis.” *Journal of Urban Economics* 48 (2000): 110-34.

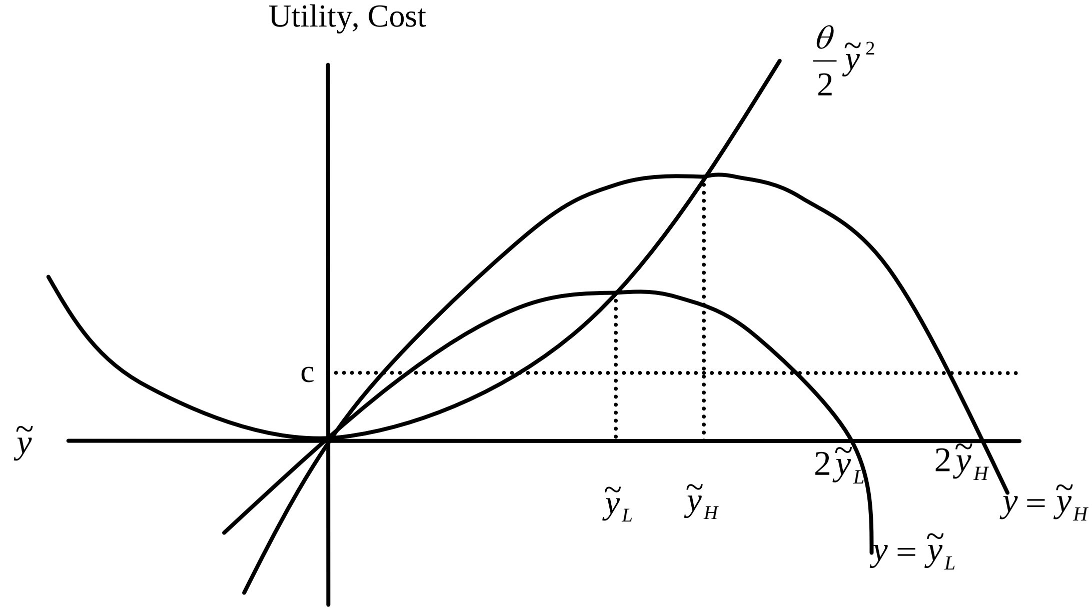
Weinberg, Bruce A., Patricia B. Reagan, and Jeffrey J. Yankow. “Do Neighborhoods Affect Work Behavior? Evidence from the NLSY79.” *Journal of Labor Economics*, Forthcoming, 2004.

Wilson, William Julius. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press, 1987.

Wilson, William Julius. *When Work Disappears: The World of the New Urban Poor*. New York: Alfred A. Knopf, 1996.

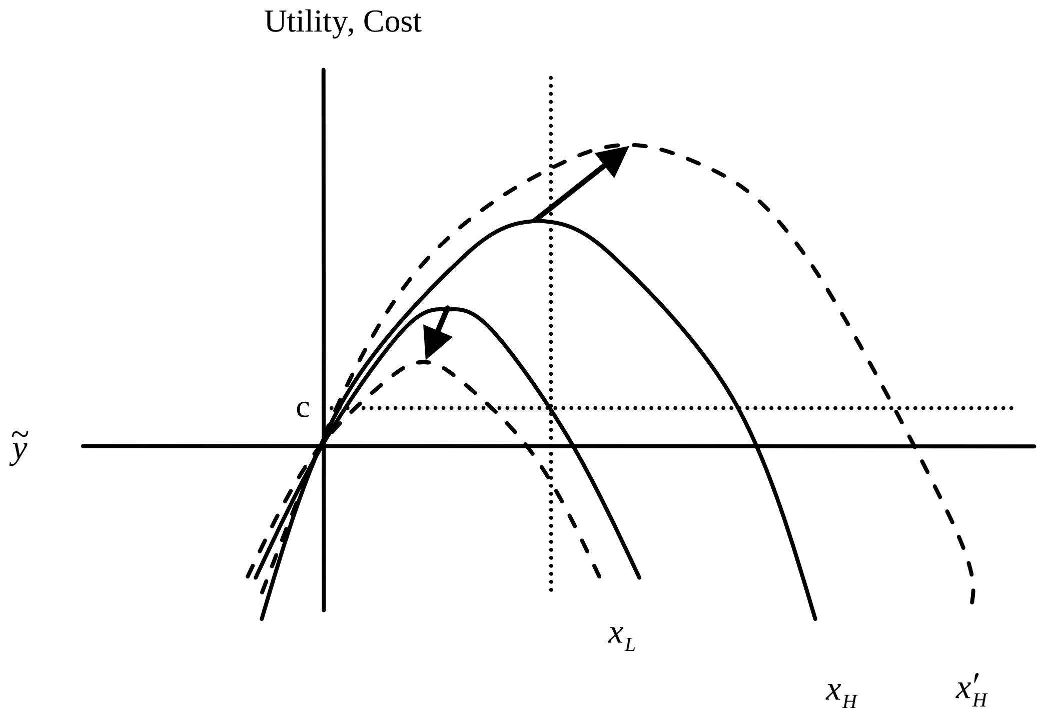
Wiseman, Rosalind. *Queen Bees and Wannabes*. New York: Crown Publishers. 2002.

Figure 1. Optimal associations.



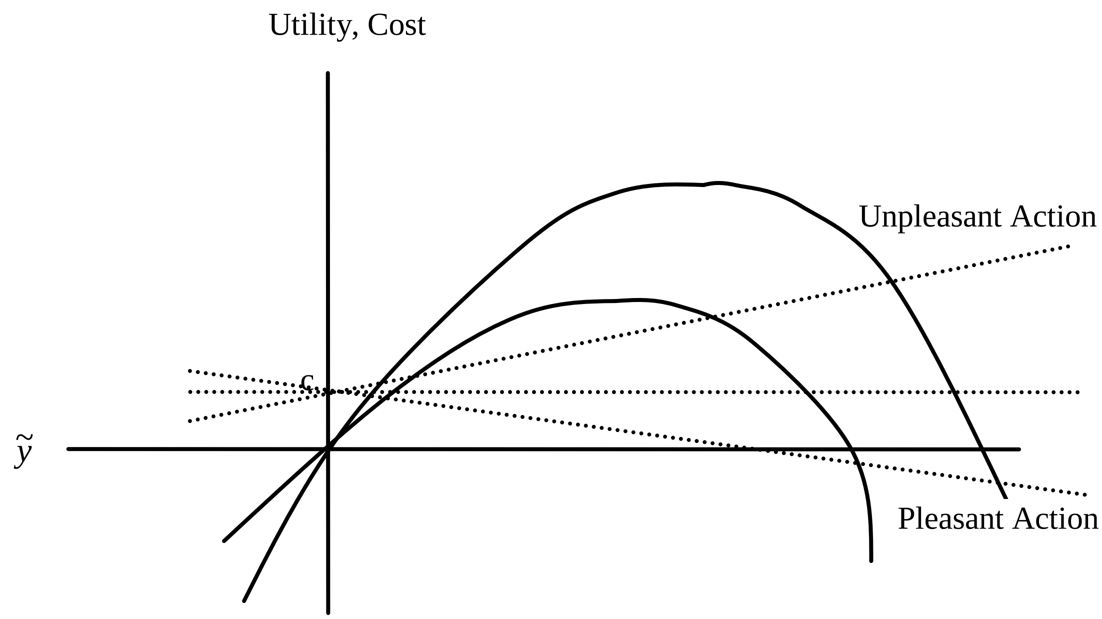
Note. Figure gives utility and cost of associating with others as a function of the other person's behavior.

Figure 2. Improvements in macro-groups can hurt some members of the macro-group.



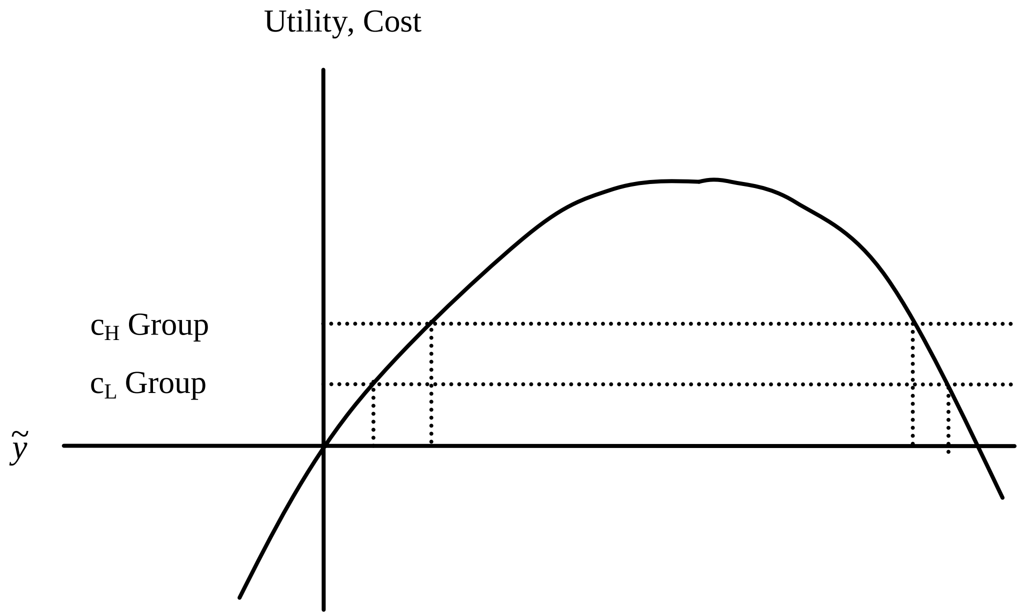
Note. The solid curves gives utility of associating with others as a function of the other person's behavior for two values of characteristics, $x_L < x_H$. The dashed curves give the associations assuming that x_H is raised to x'_H above the upper association threshold for the x_L-group.

Figure 3. The effect of actions that directly generate utility or disutility on optimal associations.



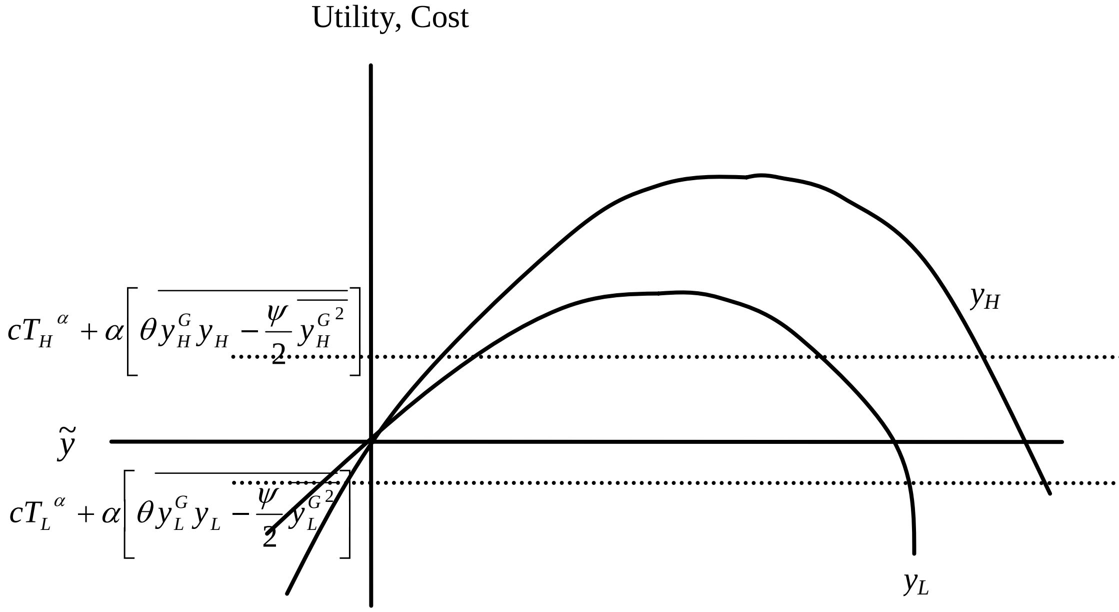
Note. Figure gives utility and cost of associating with others as a function of the other person's behavior. People associate with people taking higher levels of the pleasant action, from which they derive utility, than the unpleasant action, from which they derive disutility.

Figure 4. Optimal within and across group associations.



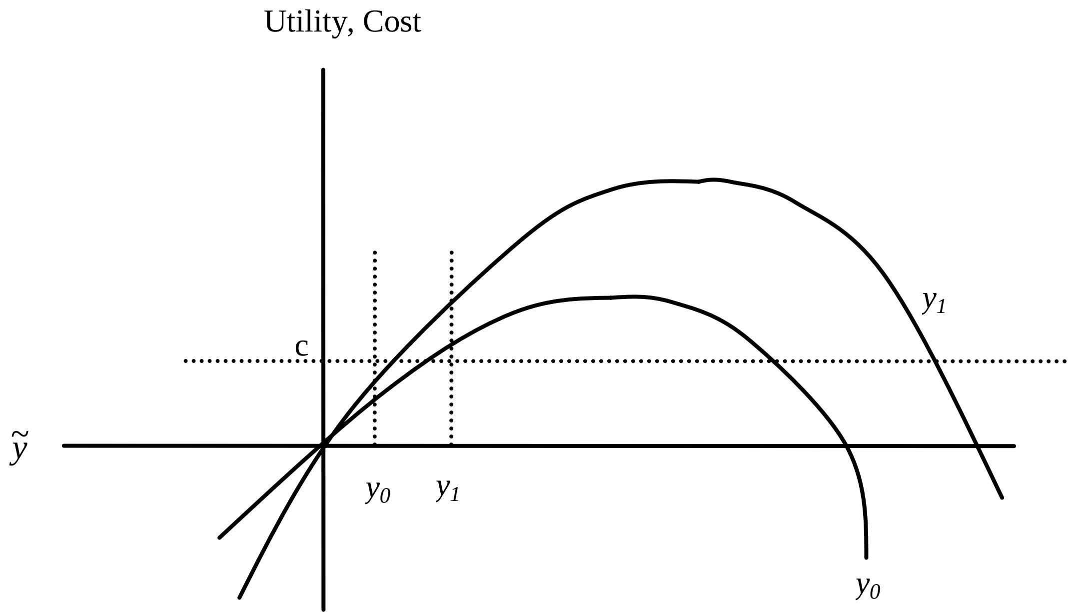
Note. Figure gives utility and cost of associating with others as a function of the other person's behavior for two costs of associating. The cost of associating with a member of ones' own group are reduced if associations with people from one's own group generate utility.

Figure 5. Effect of action on crowding out on associations.



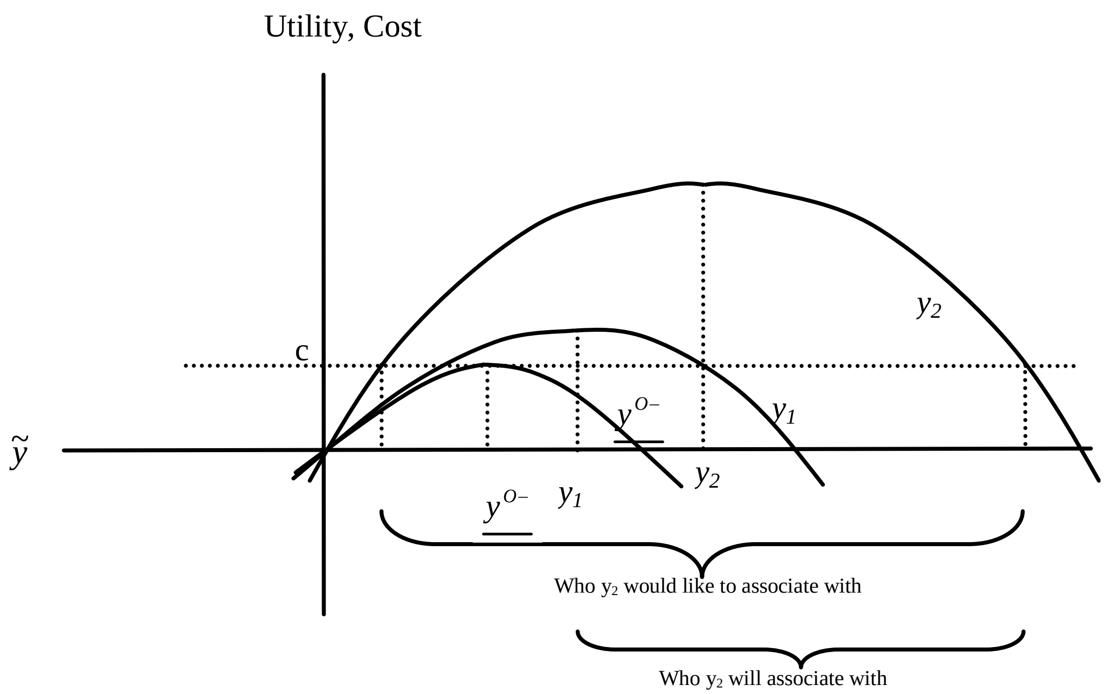
Note. Figure gives utility and cost of associating with others as a function of the other person's behavior. The y_H -group is assumed to be better matched to the macro-group in the sense that people in that group obtain more utility from their associations than people in the y_L -group.

Figure 6. Non-endogenous effect mutipliers.



Note. Figure gives utility and cost of associating with others as a function of the other person's behavior. The benefit curve labeled y_0 is for someone taking action y_0. The benefit curve labeled y_1 is for someone taking action y_1.

Figure 7. Two-sided associations.



Note. Figure gives utility and cost of associating with others as a function of the other person's behavior. The benefit curve labeled y_1 is for someone taking action y_1. The benefit curve labeled y_2 is for someone taking action y_2.

Table 1. Descriptive Statistics.

	Own Behavior or Characteristic		Associate Mean	
Drink many times	0.638	(0.481)	0.658	(0.310)
Smoke	5.349	(10.937)	5.525	(7.625)
Drink	3.811	(8.937)	4.120	(5.935)
Get Drunk	2.153	(6.981)	2.400	(4.684)
Race	4.760	(10.065)	4.710	(6.050)
Does Dangerous Things	1.962	(6.636)	2.088	(4.118)
Lies	7.205	(11.533)	7.292	(6.881)
Skips School	1.570	(5.898)	1.632	(3.962)
Fights	1.597	(3.097)	1.606	(1.926)
Effort Studying (reverse coded)	1.814	(0.653)	1.833	(0.401)
Trouble with Teacher	17.379	(14.201)	17.603	(8.394)
Trouble paying Attention	22.598	(11.012)	22.824	(6.326)
Trouble doing Homework	22.233	(11.503)	22.270	(6.700)
Trouble with Students	18.621	(13.767)	18.712	(7.952)
Male	0.460	(0.498)	0.453	(0.306)
Hispanic	0.137	(0.344)	0.140	(0.278)
Hispanic Unknown	0.076	(0.264)	0.078	(0.158)
White	0.669	(0.470)	0.666	(0.403)
Black	0.162	(0.368)	0.164	(0.336)
Asian	0.070	(0.256)	0.070	(0.205)
Indian	0.046	(0.209)	0.043	(0.123)
Other Race	0.077	(0.266)	0.074	(0.178)
With Mom	0.929	(0.257)	0.928	(0.159)
Mother’s Education (37,320 Obs.)	13.586	(2.499)	13.675	(1.790)
Mother's Education Missing	0.147	(0.354)	0.146	(0.218)
With Dad	0.778	(0.415)	0.783	(0.263)
N	44,760		44,760	

Note. Standard deviations of means in parentheses.

Table 2. Decomposition of Variance in Associates Behavior.

	Between Share	Between Share	Between Share	Variance of Individual Means	Total Individual Variance	Between Share, Double Friends	Between Share, Double Friends	Between Share, Double Friends
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drink many times	0.176	0.152	0.194	0.022	0.062	0.231	0.202	0.253
Smoke	0.115	0.154	0.205	16.471	33.455	0.117	0.156	0.208
Drink	0.111	0.126	0.172	5.726	35.661	0.280	0.310	0.393
Get Drunk	0.097	0.110	0.161	3.260	11.393	0.158	0.178	0.251
Race	0.091	0.087	0.135	1.923	22.943	0.374	0.362	0.482
Does Dangerous Things	0.065	0.092	0.139	0.745	10.449	0.328	0.415	0.531
Lies	0.055	0.085	0.129	2.492	27.510	0.243	0.339	0.450
Skips School	0.111	0.123	0.166	1.022	6.306	0.278	0.302	0.380
Fights	0.081	0.096	0.142	0.324	2.253	0.235	0.270	0.365
Effort Studying (reverse coded)	0.157	0.139	0.181	0.011	0.089	0.430	0.395	0.472
Trouble with Teacher	0.090	0.137	0.171	2.773	40.635	0.420	0.538	0.602
Trouble paying Attention	0.080	0.111	0.156	0.235	24.115	0.817	0.865	0.905
Trouble doing Homework	0.084	0.117	0.159	1.008	27.856	0.559	0.647	0.723
Trouble with Students	0.057	0.091	0.132	1.278	36.992	0.467	0.592	0.688
Hispanic	0.418	0.514	0.558	0.013	0.027	0.427	0.523	0.567
White	0.490	0.600	0.668	0.042	0.061	0.411	0.521	0.594
Black	0.468	0.447	0.620	0.040	0.048	0.345	0.327	0.495
Asian	0.361	0.396	0.462	0.010	0.017	0.324	0.358	0.422
Indian	0.089	0.171	0.236	0.000	0.009	0.524	0.699	0.777
Other Race	0.187	0.211	0.260	0.003	0.015	0.365	0.401	0.468
With Dad	0.131	0.143	0.210	0.006	0.036	0.311	0.334	0.444
Mother's Education	0.164	0.178	0.231	1.314	6.646	0.332	0.354	0.432
Observations	53,763	7,186	7,186					
Effects	498	303	600					
Full Sample	Yes					Yes		
Neighborhood Sample		Yes	Yes				Yes	Yes
School-Neighborhood Effects			Yes					Yes

Note. Between share gives the share of variance in the mean behavior or characteristics of associates arising between school-grades (in column (1)), or between school grades and school-neighborhoods (in columns (2) and (3)). Columns (6), (7), and (8) give the between share based on estimates in columns (1), (2), and (3) estimated assuming that each person has twice as many friends as they report.

Table 3. Mean Associate Behavior or Characteristics as Determined by Own Behavior or Characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)
Drink many times	0.159	(0.003)	0.151	(0.006)	0.151	(0.006)
Smoke	0.272	(0.003)	0.258	(0.007)	0.258	(0.007)
Drink	0.149	(0.003)	0.136	(0.007)	0.135	(0.007)
Get Drunk	0.121	(0.003)	0.095	(0.007)	0.094	(0.007)
Race	0.046	(0.003)	0.035	(0.006)	0.035	(0.006)
Does Dangerous Things	0.042	(0.003)	0.041	(0.006)	0.041	(0.006)
Lies	0.041	(0.003)	0.053	(0.006)	0.053	(0.006)
Skips School	0.089	(0.003)	0.101	(0.007)	0.102	(0.007)
Fights	0.066	(0.003)	0.070	(0.006)	0.070	(0.006)
Effort Studying (reverse coded)	0.059	(0.003)	0.058	(0.006)	0.058	(0.006)
Hispanic	0.219	(0.004)	0.235	(0.008)	0.227	(0.008)
White	0.228	(0.004)	0.222	(0.009)	0.214	(0.009)
Black	0.577	(0.003)	0.591	(0.007)	0.538	(0.008)
Asian	0.316	(0.003)	0.392	(0.007)	0.384	(0.007)
Indian	0.028	(0.003)	0.034	(0.006)	0.033	(0.006)
Other Race	0.049	(0.003)	0.070	(0.007)	0.069	(0.007)
With Dad	0.033	(0.003)	0.030	(0.006)	0.027	(0.006)
Mother's Education	0.124	(0.003)	0.123	(0.008)	0.114	(0.008)
Male	0.233	(0.003)	0.233	(0.006)	0.233	(0.006)
N	44,760		8,893		8,893	
Full Sample	Yes					
Neighborhood Sample			Yes		Yes	
With Neighborhood Controls					Yes	

Note. Dependent variable is mean behavior among associates. Each estimate is from a separate regression of the mean behavior or characteristics of associates on the individual's own value of that behavior or characteristic. Standard errors reported in parentheses. Estimates include school-grade fixed effects.

Table 4. Associate Behaviors (and Chars.) Related to own Behavior (Char.) Interacted with Mean in School-Grade.

	OLS		OLS		OLS		2SLS	
Drink many times	0.1038	(0.0314)	-0.0075	(0.0654)	-0.0017	(0.0654)	-0.0583	(0.1220)
Smoke	0.0129	(0.0013)	0.0101	(0.0024)	0.0101	(0.0024)	0.0101	(0.0036)
Drink	0.0224	(0.0015)	0.0108	(0.0029)	0.0110	(0.0029)	0.0109	(0.0063)
Get Drunk	0.0251	(0.0020)	0.0079	(0.0042)	0.0080	(0.0042)	-0.0209	(0.0069)
Race	0.0090	(0.0017)	0.0153	(0.0038)	0.0152	(0.0038)	0.0074	(0.0063)
Does Dangerous Things	0.0009	(0.0028)	-0.0041	(0.0050)	-0.0040	(0.0050)	-0.0096	(0.0083)
Lies	0.0029	(0.0018)	0.0044	(0.0036)	0.0046	(0.0036)	-0.0013	(0.0052)
Skips School	0.0158	(0.0021)	0.0194	(0.0047)	0.0193	(0.0047)	0.0148	(0.0057)
Fights	0.0103	(0.0051)	0.0103	(0.0100)	0.0098	(0.0100)	0.0014	(0.0138)
Effort Studying (reverse coded)	0.0574	(0.0227)	0.0314	(0.0465)	0.0297	(0.0466)	0.0180	(0.1046)
Trouble with Teacher	-0.0001	(0.0011)	0.0011	(0.0023)	0.0014	(0.0023)	0.0066	(0.0039)
Trouble paying Attention	-0.0011	(0.0015)	-0.0002	(0.0027)	-0.0002	(0.0027)	-0.0008	(0.0064)
Trouble doing Homework	-0.0002	(0.0014)	0.0015	(0.0026)	0.0014	(0.0026)	0.0096	(0.0035)
Trouble with Students	0.0018	(0.0015)	0.0040	(0.0029)	0.0040	(0.0029)	0.0025	(0.0039)
Male	0.2519	(0.0469)	0.2766	(0.0968)	0.2782	(0.0969)	-0.0507	(0.1178)
Hispanic	0.5795	(0.0156)	0.5876	(0.0386)	0.5614	(0.0387)	0.4740	(0.1473)
White	-0.1701	(0.0122)	-0.1655	(0.0247)	-0.1668	(0.0247)	1.1393	(0.4123)
Black	0.1718	(0.0124)	0.2783	(0.0298)	0.2123	(0.0297)	-0.1106	(0.1082)
Asian	0.8115	(0.0165)	1.1488	(0.0361)	1.1321	(0.0360)	0.8280	(0.2742)
Indian	0.6611	(0.0498)	1.0617	(0.0733)	1.0589	(0.0735)	-0.5464	(0.3191)
Other Race	0.2981	(0.0300)	0.3658	(0.0619)	0.3642	(0.0619)	0.1947	(0.1889)
Mother's Education (37,320 / 7,450 Obs.)	0.0097	(0.0041)	0.0107	(0.0098)	0.0067	(0.0098)	-0.0382	(0.0311)
With Dad	-0.0373	(0.0280)	-0.1279	(0.0654)	-0.1198	(0.0654)	-0.0381	(0.1963)
N	44,760		8,893		8,893		44,760	
Neighborhood Sample			Yes		Yes			
Neighborhood Controls					Yes			

Note. The dependent variable is the mean of each behavior (or characteristic) among associates. Estimates are the coefficient on the interaction between own behavior (or characteristic) and the mean of that behavior (or characteristic) in the school-grade from separate models. Standard errors reported in parentheses. 2SLS estimates instrument with the product between the own behavior (or characteristic) and the deviation of that behavior (or characteristic) from a school-specific quadratic. Estimates include individual characteristics (age, gender, Hispanic background, Hispanic background unknown, race dummy variables, years at school, mother's education, mother's education missing, and dad present), the behavior (or characteristic), and its square and school-grade dummy variables.

Table 5. Relationship between Observable Characteristics and Log Grade Size.

	Log Grade Size		Log Grade Size Residual	
Mother's Education	0.118	(0.117)	-0.004	(0.005)
Mother's Education Missing	1.381	(1.946)	-0.011	(0.113)
With Data	0.492	(0.991)	0.040	(0.062)
Male	-0.606	(0.583)	-0.013	(0.018)
Hispanic	1.608	(0.626)	0.078	(0.048)
Hispanic Unkown	-0.187	(1.194)	-0.006	(0.081)
White	0.149	(1.004)	-0.024	(0.056)
Black	0.600	(1.179)	0.007	(0.078)
Asian	0.279	(1.020)	0.015	(0.049)
Indian	0.251	(1.256)	0.106	(0.099)
Other Race	-1.725	(1.440)	-0.258	(0.123)
Years at School	-0.425	(0.064)	0.001	(0.002)
Age	0.544	(0.324)	0.001	(0.011)
Neighborhood - % Urban	0.150	(0.152)	0.004	(0.004)
Neighborhood - % Black	-0.323	(0.839)	-0.021	(0.052)
Neighborhood - Ln(Median Family Inc.)	-0.230	(0.289)	0.018	(0.018)
Neighborhood - % Poverty	-2.505	(1.175)	0.008	(0.050)
Neighborhood - % High School Dropout	0.263	(0.910)	0.012	(0.034)
Neighborhood - % Unemployed	0.162	(3.119)	-0.038	(0.112)
N	276		276	
R^2	0.479		0.066	

Note. Standard errors in parentheses. Regressions include grade dummy variables. Log grade size residual is residual from school dummy variables and school-specific quadratics in grade. Regressions weighted by square root of number of observations in school-grade. Standard errors clustered at the school-level.

Table 6. Mean Associate Behavior (and Chars.) Related to own Behavior (Char.) Interacted with and Log School-Grade Size.

	OLS		2SLS	
Drink many times	0.0218	(0.0044)	0.0615	(0.0159)
Smoke	0.0169	(0.0045)	-0.0785	(0.0345)
Drink	0.0050	(0.0046)	-0.1345	(0.0428)
Get Drunk	0.0101	(0.0047)	0.0511	(0.0173)
Race	0.0144	(0.0043)	0.0881	(0.0379)
Does Dangerous Things	-0.0041	(0.0045)	0.0290	(0.0146)
Lies	0.0012	(0.0043)	-0.0186	(0.0344)
Skips School	0.0223	(0.0050)	0.0250	(0.0159)
Fights	0.0085	(0.0042)	-0.0192	(0.0465)
Effort Studying (reverse coded)	0.0180	(0.0043)	-0.0088	(0.0435)
Trouble with Teacher	0.0161	(0.0042)	0.0860	(0.0414)
Trouble paying Attention	0.0033	(0.0039)	-0.0107	(0.0138)
Trouble doing Homework	0.0035	(0.0041)	0.0171	(0.0438)
Trouble with Students	0.0038	(0.0041)	-0.0786	(0.0386)
Male	0.0106	(0.0040)	0.0170	(0.0143)
Hispanic	0.1467	(0.0049)	0.1486	(0.0149)
White	0.0092	(0.0046)	-0.0188	(0.0493)
Black	0.0730	(0.0042)	0.0682	(0.0231)
Asian	0.1243	(0.0051)	0.0366	(0.0382)
Indian	-0.0057	(0.0041)	-0.0315	(0.0128)
Other Race	0.0081	(0.0046)	0.0023	(0.0147)
Mother's Education	0.0357	(0.0045)	0.0569	(0.0341)
With Dad	0.0009	(0.0044)	-0.0221	(0.0164)
N	44,760		44,760	

Note. The dependent variable is the mean of each behavior among associates. Estimates are the coefficient on the interaction between own behavior and the log of grade size. 2SLS estimates instrument with the product between the own behavior (or characteristic) and the deviation of log grade size from a school-specific quadratic. Estimates include individual characteristics (age, gender, Hispanic background, Hispanic background unknown, race dummy variables, years at school, mother's education, mother's education missing, and dad present), the behavior (or characteristic), and its square and school-grade dummy variables.

Appendix Table 1. Variable Descriptions.

Variable	Description	Coding
Drink many times	Have you had a drink of beer, wine, or liquor—not just a sip or a taste of someone else's drink—more than two or three times in your life?	0 (No); 1 (Yes)
	During the past twelve months, how often did you:	
Smoke	...smoke cigarettes?	0 (never); .5 (once or twice); 1(once a month or less); 2.5 (2 or 3 days a month); 16 (3 to 5 days aweek); 30 (nearly every day)
Drink	...drink beer, wine, or liquor	As above
Get Drunk	...get drunk?	As above
Race	...race on a bike, on a skateboard or roller blades, or in a boat or car?	As above
Does Dangerous Things	...do something dangerous because you were dared to?	As above
Lies	...lie to your parents or guardians?	As above
Skips School	...skip school without an excuse?	As above
Fights	In the past year, how often have you gotten into a physical fight?	0 (never); 1.5 (1 or 2 times); 4 (3 or 5 times); 6.5 (6 or 7 times); 10 (more than 7 times)
Effort Studying (reverse coded)	In general, how hard do you try to do your school work well?	1 (I try very hard to do my best) to 4 (I never try at all)
	Since school started this year, how often have you had trouble:	
Trouble with Teacher	...getting along with your teachers?	0 (never); 1 (just a few times); 4 (about once a week); 3.5 (almost everyday); 30 (every day)
Trouble paying Attention	...paying attention in school?	As above
Trouble doing Homework	...getting your homework done?	As above
Trouble with Students	...getting along with other students?	As above