NBER WORKING PAPER SERIES

SOCIAL INTERACTIONS WITH ENDOGENOUS ASSOCIATIONS

Bruce A. Weinberg

Working Paper 13038
http://www.nber.org/papers/w13038

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2007

This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwistle for assistance in the original design. Persons interested in obtaining data from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu). I appreciate helpful comments from Larry Blume, Ed Glaeser, Bryan Graham, Larry Katz, Kevin Lang, Lung-fei Lee, Erzo Luttmer, Pat Reagan, and Qingyan Shang and from participants at the 2005 AEA Meetings, the University of Alabama, IZA, Harvard University, the London School of Economics, Ohio State University, the University of Chicago, and the University of Wisconsin, Madison. 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. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Social Interactions with Endogenous Associations
Bruce A. Weinberg
NBER Working Paper No. 13038
April 2007
JEL No. I21,I32,J15

ABSTRACT

This paper develops a model of social interactions with endogenous association. People are assumed to invest in relationships to maximize their utility. Even in a linear-in-means model, when associations are endogenous, the effect of macro-group composition on behavior is non-linear and varies across individuals. We also show that larger groups facilitate sorting. 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 stronger in large groups. We also show that behaviors vary within and between macro-groups in the way predicted by endogenous association.

Bruce A. Weinberg
Department of Economics
Ohio State University
410 Arps Hall
1945 North High Street
Columbus, OH 43210
and NBER
weinberg.27@osu.edu

Social Interactions and Endogenous Association

I. Introduction

Social scientists have increasingly turned to social interactions models, in which peoples' behaviors are affected by their social groups, to understand large disparities in behaviors and outcomes, and especially the low outcomes among underrepresented and economically disadvantaged groups.[1] Economists interested in social interactions have focused on controlling for the effect of selection into "macro-groups," such as schools or neighborhoods, on estimates of social interactions.[2] While analogous forces presumably operate within these macro-groups, to the best of our knowledge this paper is the first to systematically study the effects of endogenous association within groups. The absence of work on this question is troublesome for, while selection into groups almost surely biases estimates of the effects of groups up, endogenous association within groups limits the interpretation of conventional estimates and means that they are almost surely biased down. If associations are endogenous, social groups will have non-linear effects and individual characteristics and group composition will interact in determining behavior, markedly changing the policy implications of social interactions models.

This paper develops a formal model of within-group associations. We assume that individuals choose both their behavior and their associations to maximize their utility, which depends on both their own and their associates' behaviors and characteristics. We specify a cost in terms of time and effort to an individual of associating with the other

---

[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]).

[2] 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] and theoretical work by Bayer and Ross [2006]. 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.

members of his macro-group. We also assume and provide evidence that people obtain the most utility from associating with others whose behaviors and characteristics are similar to their own, known as homophily in sociology.

We specify a model that yields the standard implication that behaviors depend linearly on the (weighted) average of the characteristics and actions of the other members of an individual's macro-group, with the weights in our model determined by the amount of their interaction. Even with this linear behavioral model, when associations are endogenous, macro-group composition has a non-linear effect on behavior and the effect of macro-groups varies with individual characteristics and behaviors.[3] These non-linearities arise when associations are endogenous because associates' behaviors and characteristics will be non-linear functions of own characteristics. To the best of our knowledge, we are the first to generate these non-linearities endogenously. We apply our model to understand association patterns in schools. In that context, students who are inclined to substance use, for instance, in a school-grade where few other students are inclined toward substance use will be forced to associate with people who do not use substances, discouraging their own tendency to substance use. As the number of students inclined to substance use increases, these individuals will find more like-minded associates accentuating their tendencies.

At a policy level, these non-linearities and interactions imply that even in a linear-in-means model, social interactions are not zero-sum. To reduce the amount of a particular behavior, one would want to avoid placing people who are prone to that behavior in groups where there are many similar people – in these groups, there will be enough other people inclined toward the behavior with whom to associate to generate a

---

[3] Krivo and Peterson [1996]; Galster, Quercia, and Cortes [2000]; Galster [2002]; and Weinberg, Reagan, and Yankow [2004]; and Burke and Sass [2006] 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).

high level of the behavior. In terms of estimation, these non-linearities also provide a potential solution to Manski's [1993] reflection problem that is quite different from existing solutions (Brock and Durlauf [2001a,b]; Lee [2006]; Blume and Durlauf [2005]; Lin [2005]; and Shang [2006]).

Many studies of social interactions exploit relocations of individuals across macro-groups.[4] Most such studies involve relocating individuals who are at-risk along some dimensions from social groups where many individuals are at risk to social groups with a mix of people who are and are not at risk. These studies find surprisingly small effects of social groups. Our model provides an explanation for these small effects. It indicates that individuals relocated in this way will tend to re-segregate within their new groups, attenuating any benefits, even if associates do have effects. Consistent with this hypothesis, Angrist and Lang [2004] and Kling, Ludwig, and Katz [2005] find larger effects for girls than boys, who may integrate more into their new social groups, and Clampet-Lundquist, Edin, Kling, and Duncan [2005] provide evidence that the difference is due to differences in integration into new neighborhoods.

Emerging studies show that the associations that arise within macro-groups are not random, but empirical work on social interactions has not taken account of these results.[5] Most studies relate an individual's behavior to the mean in his macro-group, implicitly assuming that all members of the macro-group interact equally.[6] Some

---

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

[5] A number of papers show that people prefer to associate with others of the same racial, ethnic, or religious group (Moody [2001]; Bayer, McMillan, and Rueben [2002]; Marmaros and Sacerdote [2003]; Bisin, Topa, and Verdier [2004]; Mayer and Puller [2006] 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]).

[6] Exceptions are Conley and Udry [2001] and Bandiera and Rasul [2002] and Foster [2006]. Bertrand, Luttmer, and Mullainathan [2002], Munschi [2002], Grodner and Kniesner [2006]; and Hoxby and

theoretical analyses allow for variations in associations within macro-groups. Anselin [1988]; Lee [2001]; Brock and Durlauf [2001a, 2001b] specify a fixed weighting matrix that gives association patterns. This approach ignores that interactions are determined by behaviors as much as behaviors are determined by interactions (Coleman [1961]; Wiseman [2002]; and Eder [2003] provide vivid examples). A person who does not smoke, for instance, is unlikely to choose to associate with someone who smokes heavily, and so the first person's smoking is unlikely to be much affected by the other person's smoking. 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 points in a particularly striking form). This gap is troublesome given that most policies to address the effects of social interactions involve moving people across macro-groups.

This study employs data from the National Longitudinal Study of Adolescent Health (Add Health), which surveyed all students in 172 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 behaviors and characteristics arises within grades in schools. Individuals associate with others whose behaviors and characteristics are similar to their own. Consistent with our model, we find that so long as there are a few people with a particular characteristic in a group, people with that characteristic associate with each other, and that people avoid associating with people who are different from themselves until most of the group is comprised of people who are different. We show

---

Weingarth [2006] 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 affected by changes in the population or in the behaviors of group members.

that group size facilitates segregation and show that there is a stronger relationship between own characteristics and associates' characteristics in large macro-groups.

While estimating the effect of endogenous sorting on behaviors is complicated, we provide evidence that endogenous association affects behavior. As indicated, our model implies that the availability of some like-minded individuals permits people to associate with people with the same behavioral tendencies, accentuating their own behavioral tendencies. Consistent with this implication, we show the mean behavior among the associates of people who are inclined to a particular behavior (based on their exogenous characteristics) is higher (relative to people whose characteristics imply that they are unlikely to have high levels of the behavior) in groups where there some people inclined to the same behavior. We also show that the availability of people who are inclined to a particular behavior increases the behaviors of people who are inclined to that behavior (relative to people whose characteristics imply that they are unlikely to have high levels of the behavior). Lastly, we provide some two-stage least squares estimates of the effect of associates' behaviors on own behavior.

II. A Model

II.A. The Framework

This section develops a model of social interactions with endogenous association. People are characterized by observable characteristics, x, and an shock, ε. They choose an action, y, and a set of associations. An individual with characteristic x and shock ε who takes action y has a utility of,

$$U = \underbrace{\beta' x y + \varepsilon y - \frac{1}{2} y^2}_{\text{Private Utility}} - \underbrace{cT}_{\substack{\text{Cost of} \\ \text{Association}}} + \underbrace{\left(\theta \bar{y}^A y - \frac{\psi}{2} \overline{y^{A2}} + \omega' x \bar{y}^A + \bar{x}^{A\prime} \Phi x + \gamma' \bar{x}^A y \right) T^{\alpha}}_{\text{Social Utility}}$$

This utility function has three components: a private utility; a cost of interacting; and a social utility, each of which is described in turn. Variables with over-bars and A-superscripts denote means among a person's associates ($\overline{y^{A2}}$ denotes the mean of the squared values of associates' actions).

In the empirical analysis, we consider a variety of actions and outcomes. The actions include substance use and behavioral problems. We also study grades, which depend on own actions, whose utility depends on associates' actions, and on associates' grades (if there are human capital spillovers). The private utility is written to imply that behavior is a linear function of characteristics.

Interactions can be costly in terms of time and effort. To capture these costs parsimoniously, we assume that people incur a cost, c, per unit of time spent interacting with others, where $T(x, \varepsilon, y)$ denotes the total time that the person spends interacting with others.[7] We assume that associations are binary. Because of random encounters, people are assumed spend t_0/N "passively associating" with every other member of their macro-group. Here N gives the size of the macro-group, so total passive associations do not depend on group size. People can also "actively associate" with particular members of their macro-group, spending an additional unit of time for a total of $1+t_0/N$.

The social utility has both standard and novel components. In Manski's [1994] terminology, θ and γ give the effect of others' behaviors (the "endogenous effect") and characteristics (the "exogenous effect"). The parameters ψ, ω, and Φ have no counterparts in Manski's framework, because they do not affect the action directly. They affect the utility of associations and actions through their effect on associations. High values of ψ reduce the utility of associating with people with extreme values of y. The parameter ω allows the utility of associating with someone to vary with the other person's behavior, y, and for the effect to depend on a person's own characteristics. For instance, people with behavioral problems, may produce disutility. The parameter Φ is

[7] The cost could be negative if the opportunity cost of time is low and people like interacting.

analogous, but reflects the effect of associates' characteristics on the utility of associating. For instance, people likely obtain utility from associating with people who are attractive or athletic and people of the same race or ethnicity.

Once people are viewed as choosing to associate with a portion of their groups, it becomes important to consider how the number of associates affect the strength of interactions, which is governed by α. One view, which comes closest to the literature, is that new associations completely crowd out existing ones, leaving the total effect of associates fixed, that $\alpha = 0$. One might think that people who have many associates are affected by their associates (as a whole) more than people with few associates, that $\alpha > 0$. To simplify notation, it is convenient to define $\gamma = 1-\alpha$ and $\overset{=}{z}^{A} \equiv \bar{z}^{A} T^{a}$, which is the mean of some variable *z* among a person's associates weighted by the total amount of interactions, *T*.

II.B. Solving the Model

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

$$y = \beta' x + \left(\theta \bar{y}^{A} + \gamma \bar{x}^{A}\right) T^{\alpha} + \varepsilon . \qquad (*)$$

This is the standard linear-in-means behavioral equation, where the strength of the social effect varies across people if $\alpha \neq 0$. We consider a model with a binary, scalar characteristic. (Weinberg [2005] discusses a model with continuous characteristics.) We begin by eliminating the action to focus on associations on characteristics. We assume that the characteristic takes on values $x \in \{-1,1\}$, and refer to people with these characteristics as low and high types. Let π_i denote the share of the group with $x = i \in \{-1,1\}$. We assume $\Phi = 1$ and that the other social utility parameters (*θ, ψ, ω,* and *γ*) are all zero.

Let $\pi_{i|j}^{A} \in [0, \pi_i]$ denote the measure of people with characteristic x_i with whom a person with characteristic x_j associates, which must be non-negative and is constrained by group compostion. With $\Phi > 0$ and the values of the characteristic taking on opposite signs, people only associate with people of the same type, so that a type-i person chooses $\pi_{i|i}^{A} \in [0, \pi_i]$. The first order condition for active associations is,

$$\frac{\partial U}{\partial \pi_{i|i}^{A}} = \left(x_i{}' - (1-\alpha)\bar{x}^{A}{}' \right) \Phi x_i \left(t_0 + N\pi_{i|i}^{A} \right)^{\alpha - 1} - c = 0 \ \ subject\, to\ \pi_{i|i}^{A} \in [0, \pi_i]. \qquad (**)$$

Figure 1 illustrates the associations of a high (the results for lows are reversed). Panel A shows the "desired" value of $\pi_{H|H}^{A}$, which solves $\frac{\partial U}{\partial \pi_{H|H}^{A}} = 0$ ignoring the group composition constraint that $\pi_{H|H}^{A} \le \pi_H$, as a function of π_H. It is possible to show that the desired $\pi_{H|H}^{A}$ is non-increasing (and decreasing if $\alpha < 1$) in π_H. The figure also shows a 45º line, which gives the maximum feasible value of $\pi_{H|H}^{A}$. Up to three regions may be present. In region 1, the desired active associations exceed the maximum possible given the population composition. In the second region, the composition constraint does not bind and active associations are positive. In the third region, the person chooses not to actively associate. As the bold curve in the figure shows, active associations increase in region 1; decrease in region 2; and are zero in region 3. To fully characterize the model, the figures show all three regions, but the third may be less relevant, than the others especially with sorting on multiple dimensions.

Panel B shows $\overset{=}{x}{}_H^{A}$, where $\overset{=}{x}{}_i^{A}$ denotes the value of $\overset{=}{x}{}^{A}$ among type i people, as a function of π_H. The curve labeled $\frac{\partial U}{\partial \pi_{H|H}^{A}} = 0$ shows the desired value of $\overset{=}{x}{}^{A}$. It is possible to show that this curve is non-decreasing and concave (when $\alpha < 1$). Intuitively, when more of the group is high, a greater share of passive associations will be with highs, which decreases highs' motivation to actively associate with other highs (so long as $\alpha < 1$) partially, but not fully, offsetting the increase in $\overset{=}{x}{}^{A}$. Also shown are the maximum value of $\overset{=A}{x}$ the person can achieve by associating with all available highs, $\overset{=A\,Max}{x} = \frac{t_0\left[\pi_H x_H + (1-\pi_H)x_L\right] + \pi_H N x_H}{(t_0 + N\pi_H)^{1-\alpha}}$, which is concave if $\alpha < 1$, and the value that emerges in the absence of active associations, $\left[\pi_H x_H + (1-\pi_H)x_L\right]t_0^{\alpha}$, which is linear. The bold curve shows that $\overset{=A}{x}_H$ increases monotonically, with the greatest increase at low levels (in region 1).

Except on the left (and extreme right), changes in group composition have less effect on highs' associates than implied by the average characteristics in the group (given by the diagonal line from $(0, x_L)$ to $(1, x_H)$. This concavity, which is present except in region 3, is the key to our results.

Panel C shows that the difference in $\overset{=A}{x}$ between highs and lows, $\overset{=A}{x}_H - \overset{=A}{x}_L$, is hump-shaped in π_H. Intuitively, in groups that are largely composed of one of the groups, there is little scope for sorting, so the associations of both types are relatively similar. Segregation is greatest when there are many members of both groups.

II.C. Group Size

A novel implication of our model is that sorting should be greater in large groups than in small groups. Panel A of Figure 2 shows that increasing group size reduces the desired level of $\pi^A_{H|H}$. This can be seen from (**) where N and $\pi^A_{i|i}$ only appear multiplicatively, so as N increases, the same social influence can be obtained with a proportionately lower level of $\pi^A_{H|H}$. Panel B shows that $\overset{=A}{x}_H$ increases in region 1 and is otherwise unchanged (the increase in N corresponds to more associations in region 1, but is offset by the decrease in $\pi^A_{H|H}$ in region 2). Increasing group size shifts the $\overset{=A\,Max}{x}$ curve up because the amount of weight placed on passive associations does not change, but for any given level of π_H, there are more highs with whom other highs can associate.

The model implies more sorting in large groups when either group's associations are constrainted.

Glaeser, Sacerdote, and Scheinkman [1996] and Graham [2005] estimate the strength of social interactions using variations across large groups or variations in group size. Insofar as large groups facilitate sorting one might expect more variation in large groups. Endogenous association likely biases estimates using Glaeser, Sacerdote, and Scheinkman's method upward, but estimates using Graham's method downward.[8] At a policy level, our results on macro-group size indicate that attempts to integrate people from different groups will be more successful if macro-groups are kept small, limiting peoples' ability to re-segregate within groups.

II.D. Crowding Out

As indicated, once the number of associations is endogenous, one needs to think about how the number of associates someone has affects the strength of the social influence he experiences. By focusing on associates' means, existing models implicitly assume that additional associations fully crowd out existing ones ($\alpha = 0$), but it seems likely that people with more associates will be affected by them more.

It is possible to show that decreasing crowding out increases the difference in the mean associations between highs and lows. This effect is particularly large when most of the group is made up of one type or the other and in the middle. Thus, the assumption of complete crowding out, which is implicit in the literature, limits the scope for people to segregate within their macro-groups by active associating.

II.E. Endogenous Actions

So far we have considered models with sorting on exogenous characteristics. This section focuses on a binary action, $y \in \{-1,1\}$, assuming that people derive utility from associating with people taking the same action ($\theta > 0$). (For simplicity, we assume that the other social utility parameters, ψ, Φ, ω, and γ, are zero.) Conditional on the distribution of actions, sorting on an endogenous behavior is identical to sorting on an exogenous characteristic.

Characterizing the equilibrium involves solving for the share of the group taking an action as a function of the distribution of exogenous characteristics in the group. A person with characteristic x_i takes action y_H if,

$$y_H x_i{}' \beta + y_H \varepsilon - \frac{1}{2} y_H{}^2 + \bar{\bar{y}}_H^A \theta y_H > y_L x_i{}' \beta + y_L \varepsilon - \frac{1}{2} y_L{}^2 + \bar{\bar{y}}_L^A \theta y_L .$$

Let

$$\varepsilon^* \equiv \frac{\frac{1}{2}\left(y_H{}^2 - y_L{}^2\right) - \left(\bar{\bar{y}}_H^A \theta y_H - \bar{\bar{y}}_L^A \theta y_L\right)}{y_H - y_L} \text{ and } \varepsilon_i^* \equiv -x_i{}' \beta + \varepsilon^*$$

denote the value of ε at which a person with $x = 0$ or $x = x_i$ is indifferent between the actions. For simplicity, we again assume that $x_i \in \{-1,1\}$. The probability that a person with characteristic x_i takes the low (high) action is $\rho_{L|i} \equiv F(\varepsilon_i^*)$ ($\rho_{H|i} \equiv 1 - F(\varepsilon_i^*)$), where $F(.)$ denotes the cumulative distribution function of ε.

Let $\rho_H = \rho_{H|H} \pi_H + \rho_{H|L}(1 - \pi_H)$ denote the share of the group taking the high action, where π_H denotes the share of the group with the high characteristic. The effect of π_H on the mean action can be calculated by differentiating this expression with respect to π_H and rearranging. Formally,

$$\frac{\partial \rho_H}{\partial \pi_H} = \frac{F(\varepsilon_L^*) - F(\varepsilon_H^*)}{1 + \left(F'(\varepsilon_L^*)\pi_H + F'(\varepsilon_H^*)(1 - \pi_H)\right)\frac{\partial \varepsilon^*}{\partial \rho_H}},$$

where

[8] I am grateful to Bryan Graham for discussing these points.

$$\frac{\partial \varepsilon^*}{\partial \rho_H} = -\frac{\theta}{y_H - y_L}\left(\left.\frac{\partial \overset{=A}{y}_H}{\partial \rho_H}\right|_{\pi^A_{H|H}} y_H - \left.\frac{\partial \overset{=A}{y}_H}{\partial \rho_H}\right|_{\pi^A_{L|L}} y_L \right)$$

$$+\frac{1}{y_H - y_L}\left(cN^G - \theta \frac{\partial \overset{=A}{y}_H}{\partial \pi^A_{H|H}} y_H \right)\frac{\partial \pi^A_{H|H}}{\partial \rho_H} - \frac{1}{y_H - y_L}\left(cN^G - \theta \frac{\partial \overset{=A}{y}_L}{\partial \pi^A_{L|L}} y_L \right)\frac{\partial \pi^A_{L|L}}{\partial \rho_H} < 0 .$$

If there is some crowding out, if people choose to actively associate, and if $y_H > 0$ and $y_L < 0$, it is possible to show that $\frac{\partial \varepsilon^*}{\partial \rho_H}$ will be negative, so that as more of the group is high, more people take the high action, and that it will be more negative when ρ_H (and π_H) are either low or high.[9] At these extremes, $\frac{\partial \rho_H}{\partial \pi_H}$ will be large because its denominator will be small, so that changes in the share of highs in the group will have a particularly large effect on mean behavior.

Figure 3 shows the share of the group and people with the high and low characteristic taking the high action as a function of the share of the group with the high characteristic and the share of the group with the high action assuming that ε is distributed uniformly with bounds such that people are always in the interior of the support. This assumption is appealing because it implies that behaviors are linear in characteristics and associations. In both panels, the share of each type taking the high action is non-linear. The steeper sections at low and high levels emerge because at low levels, introducing a few highs makes it possible for highs to find like-minded associates, while at the high end, eliminating the few remaining highs makes it impossible for lows to find like-minded associates.

---

[9] Under these assumptions, the first term is negative and more negative whenever either of the groups has fewer active associations, which will be at extreme values of ρ_H (and π_H). The second and third terms are zero whenever people's active associations are unconstrained or zero and negative when the constraint $\pi^A_{i|i} \le \pi_i$ binds, which will be at extreme values of ρ_H (and π_H). As $\pi^A_{i|i}$ increases, these terms become

While the uniform case is useful for demonstrating that the model generates non-linearities endogenously, it is probably more realistic to assume that the shock, ε, is normally distributed. In this case, it is possible to show that the gap in both behaviors and associations between people with the high and low characteristic are hump-shaped in the share of the group with the high characteristic.[10]

Figure 4 shows results for normally distributed shocks. The difference in the action between people with the high and low characteristic is increasing in the share of the group that has the high characteristic and the share of the group with the high action. The relationship between the average behavior in the population and the share of the group with the high characteristic is non-linear, convex at low levels and concave at high levels. The difference in the mean behavior of the associates of people with the high and low characteristic are also hump-shaped in the share of the group with the high characteristic (and the high behavior). These implications are tested below.

---

less negative, so these terms are also most negative at extreme values of ρ_H (and π_H).

[10] Define $\Delta\rho_H = \rho_{H|H} - \rho_{H|L}$ as the difference in behavior between people with the high and low characteristic. The effect of the mean behavior in the group on this difference is given by

$$\frac{d\Delta\rho_H}{d\rho_H} = \left(F'\left(\varepsilon_L^*\right) - F'\left(\varepsilon_H^*\right)\right)\frac{d\varepsilon^*}{d\rho_H}.$$

As indicated above, $\frac{\partial\varepsilon^*}{\partial\rho_H}$ is negative and more negative at low and high values of ρ_H. Because $\varepsilon_H^* < \varepsilon_L^*$, $\left(F'\left(\varepsilon_L^*\right) - F'\left(\varepsilon_H^*\right)\right)$ is negative when ρ_H is low (and ε^* is high) and it is positive when ρ_H is high (and ε^* is low). Thus, $\Delta\rho_H$ is increasing when $\rho_H < .5$ and decreasing when $\rho_H > .5$ and the difference in behavior between people with the low and high characteristic is hump shaped in ρ_H. Define $\Delta\overset{=A}{y} = \left(\overset{=A}{y}_H - \overset{=A}{y}_L\right)\Delta\rho_H$ as the difference in the mean behavior of the associates of people with the low and high characteristic. From the results above on sorting on exogenous characteristics $\overset{=A}{y}_H - \overset{=A}{y}_L$ is hump-shaped in ρ_H. Given that $\Delta\rho_H$ is hump-shaped in ρ_H, so is $\Delta\overset{=A}{y}$. Note that these results hold for any distribution where $\left(F'\left(\varepsilon_L^*\right) - F'\left(\varepsilon_H^*\right)\right)$ is negative when ε^* is high positive when ε^* is low, including a uniform distribution where people move beyond the support.

II.F. Implications

A number of important implications arise immediately from the non-linear relationship between group composition and behavior. First, it breaks the zero-sum implication of the standard, linear-in-means model. A number of studies across a range of disciplines have argued for a non-linear relationship between group composition and behavior (see footnote 3 for references). To the best of our knowledge, our model is the first to generate non-linearities endogenously.

In the preceding uniform example, if one quarter of the population had the high characteristic and the population was large enough for three groups, if the highs were divided equally among the three groups, 30.9% of the population (68.4% of highs and 18.4% of lows) would take the high action. If the population were divided into 2 groups with no highs and one that was three quarter highs, 23.0% of the population (81.6% of highs; none of the lows with other lows, and 31.6% of the lows in the three-quarter high group) would take the high action. Thus, grouping the highs with enough other highs to allow them to heavily associate with each other generates high overall behavior.

The existing literature places particular emphasis on the multipliers generated when peoples' actions depend on the actions of the other 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]; Brock and Durlauf [2001a,b]; Blume and Durlauf [2006]; Lin [2005]; Lee [2006]; and Shang [2006]).

Endogenizing associations overturns these results in two ways. First, the non-linear relationship between the composition of a group and the mean behavior can, in principle, identify the model's parameters. Second, the distinction between traditional endogenous and exogenous effects itself becomes less interesting once associations are treated as endogenous because changes in the distribution of exogenous characteristics can generate multipliers through associations. Put differently, for multipliers to arise some endogenous variable has to respond to changes in the environment, and it can be associations instead of behaviors.

Lastly, endogenous association limits the interpretation of conventional estimates of the effects of macro-groups because the effect of macro-groups varies across people and with group composition. While one can accurately estimate the effect of moving people like those being studied between groups like those studied, it is impossible to estimate the effect of moving similar people between different types of groups or different types of people between the groups being studied. Put differently, one can estimate the effect of the groups studied on the people studied, but there is no single "effect" of groups to estimate because the effect of groups depends on associations which in turn depend on interactions between individual characteristics and group composition. As indicated, moving people who are at risk along some dimension from groups where many individuals are at risk to groups with a mix of people who are and are not at risk is likely to have less effect on the influences a person experiences than one would expect based on macro-group composition because people can find like-minded associates in their new groups.

II.G. Actions and Characteristics Directly Affect Interactions

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. People will associate with people taking actions that generate utility for their associates, causing those actions to proliferate if $\theta > 0$.

It can be shown that people associate more with people taking high levels of pleasant actions, causing those actions to spread. 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.

II.H. Peer Pressure

People often describe social interactions models as models of "peer pressure." Peer pressure involves the withdraw of association when one person takes an action that produces disutility for another. To model peer pressure, associations must be endogenous.

To capture this idea, we assume that people derive utility directly from having people associate with them and, in choosing their actions, take into consideration how their action will affect whether other people associate with them. Let the utility function be,

$$U = \underbrace{\beta' xy + \varepsilon y - \frac{1}{2} y^2}_{\textit{Private Utliity}} - \underbrace{cT}_{\substack{\textit{Cost of} \\ \textit{Associating}}} + \underbrace{\left(\theta \bar{y}^A y - \frac{\psi}{2} \overline{y^{A2}} + \omega' x \bar{y}^A + \bar{x}^{A\prime} \Phi x + \gamma' \bar{x}^A y \right) T^\alpha + \delta P}_{\textit{Social Utility}},$$

where P gives the person's popularity, the number of people associating with him. Returning to the case of a continuous action, the first order condition for the action is,

$$y = \beta x + \left(\theta \bar{y}^A + \gamma \bar{x}^A \right) T^\alpha + \delta \frac{\partial P}{\partial y} + \varepsilon . \qquad (*')$$

The condition now includes a term for the effect of a change in the action on the measure of people who will associate with him. Under peer pressure, a person's action depends on the density of people at the margin to interact with him as well as the mean of associates' behaviors and characteristics. In this sense, the standard linear-in-means model of interactions may be a poor representation of peer pressure.

The strength of "reception" and "transmission" of peer pressure will likely depend on characteristics. On the reception-side, insofar as most people in the macro-group will want to associate with them regardless of their actions, people whose characteristics make them appealing as associates will be less subject to peer pressure than others. 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 regardless of their actions. Thus the model suggests that the extent to which people are subject to peer pressure as recipients follows an inverse-U in how appealing they are as associates. On the transmission-side, people who are attractive as associates may get more weight in other people's *P*'s, so people will be particularly subject to peer pressure from attractive people. The strength of peer pressure from the transmission side will be strictly increasing in how appealing someone is as an associate.

II.I. Extensions

This section considers two further extensions to the model – endogenous popularity and dislike. One could extend the analysis of popularity by making one of the individual characteristics, x, be the measure of people in the macro-group who associate with a person. 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 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 an effect explain the emphasis placed on affecting how much weight people place on others, by making positive (or negative) examples of people, such as valedictorians (Moffitt [2001] discusses a variety of policies).

Lastly, one might allow people to invest time down-weighting some members of their macro-group. People would optimally want to down-weight people who are most different from them. Without crowding out, people who are most different from their macro-groups will invest the most to down-weight others. The possibility of down-weighting reduces the benefits from placing people with bad outcomes into better groups.

III. Data

We use data from the National Longitudinal Study of Adolescent Health, commonly known as Add Health. The Add Health data provide information on a wide range of youth behaviors, including substance use, risky behaviors, grades, and behavioral problems, and data on family background. They contain nationally representative data on 90,118 students enrolled in grades 7-12 in 172 schools. Students were asked about their 5 closest male and female friends, so information is available on both the macro groups – schools and grades – as well as the subgroupings that emerge within them. Also, all people are surveyed symmetrically.[11]

We restrict attention to students enrolled in 9th through 12th grades in classes with at least 10 students. The data contain a limited panel aspect, which we do not utilize.

Appendix Table 1 lists the behaviors and outcomes used and their construction. 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.[12] Slightly under half the sample is male; 16% identify as black (respondents were allowed to report multiple races); 15% identify as Hispanic. Just over three quarters of respondents live with their father, and their mothers have 13.6 years of schooling on average. Average grades vary from a B- to a B, with the lowest grades in math and the highest grades in history and social studies. On average the respondents smoked a bit more than once a week, they drink a couple of times a month, and get drunk about half as frequently. The most common behaviors or

---

[11] Alternative surveys provide information about particular individuals and ask the respondents about the behaviors of their friends. Under this approach less information is available on friends than on respondents and reports of friends' behavior may be biased differently than reports of own behavior.

[12] The original sample comprises 90,118 individuals. Once the sample is restricted to individuals in grades 9 through 12 with valid school, grade, and identification information, the sample is reduced to 62,203. Restricting the sample to people with some identifiable friends with valid information reduces the sample to 47,570. The remaining deletions are for having missing data for a variety of the other variables.

behavioral problems are having trouble paying attention and doing homework regularly. Skipping school and fighting are the least common, occurring about once a month.

Also shown are the means and standard deviations of the associate averages of the behaviors and characteristics. The mean of the associates' behaviors are above the mean of own behaviors, indicating that people with "worse" behavior are listed as friends more often than people with better behavior. On the other hand, the mean of associates' grades tend to be slightly higher than own grades, indicating that popularity is increasing in grades. Similarly, the mean of mother's education and father present among associates is above the own mean for these variables, indicating that children with "better" family backgrounds may be more appealing because they are wealthier or more socially adept.

IV. Within-Group Variations in Associations

As indicated, most studies of social effects focus on variations across macro-groups. 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.

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 $\bar{z}_i^A = \frac{1}{N_i}\sum_j z_{ij}$, where N_i gives the number of people with whom *i* is friends. (Here and below, by studying $\bar{z}_i^A$ we implicitly focus on the case where $\alpha = 0$; work that estimates α is in progress.) We decompose the variance of $\bar{z}_i$ into within and between school-grade components. The total variance in $\bar{z}_i^A$ is $VAR[\bar{z}_i^A] = VAR[\bar{z}_i^A - E[\bar{z}_i^A | S_i, G_i]] + VAR[E[\bar{z}_i^A | S_i, G_i]]$. The former term represents the within school-grade variance and the later the across school-grade variance.

The share of the variance arising within school-grades is the R^2 from an analysis 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 over 90% of the variation in friends' behaviors arises within school-grades. For grades (and television watching) about 80% of the variation in friends' behaviors and outcomes arises within school grades. Between 75% and 90% of the variation in family background, measured by mother's education, mother is a homemaker, and 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, frequently close to half, which is to be expected given the extent of racial and ethnic residential sorting.

Because people often form friendships with people who live near them, a portion of the within school-grade variation likely arises from residential sorting among students in the same school-grade. Tracking and selection into classes may also generate within school-grade variations in associations. The Add Health survey does not contain data on tracking or the classes taken, but it does contain characteristics of the census block group of residence for a portion of respondents. Because neighborhood characteristics are only available for one respondent in five, column 3 reports the R^2 from the analysis of variance described above when applied to the smaller sample; column 4 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 typically accounts for 5% of the variance in friends' behavior. Even in this smaller sample with neighborhood variables, within macro-group variations account for 80% or more of the variation in associates' behaviors; 30%-70% of the variations in associates' race and ethnicity; and 70% of the variation in associates' grades and family background. To the extent that people choose their friends based on their behaviors and there is an association between neighborhoods and behaviors, these estimates will overstate the effect of residential proximity.

The substantial variations in associates' behaviors can not be viewed as the effect of sorting insofar as exogenous variations in associations combined with social effects on behavior will generate variations in associates' behaviors. Nevertheless, we find substantial variations in associates' family backgrounds within macro-groups, and these variations are not caused by associations. Regardless of causality, our estimates indicate ignoring variations in associations with macro-groups leads to substantial mis-measuring the social influences to which people are exposed.

V. The Choice of Associates

The preceding analysis indicates that the vast majority of variation in the composition of social groups arises within narrowly-defined macro-groups. We begin by studying the determinants of associations on exogenous characteristics (race, ethnicity, and family background). It is easier to study sorting on exogenous characteristics than behaviors and outcomes because, as indicated, if social interactions operate, behaviors and outcomes are affected by others. That is, the relationship between a person's behavior and those of his associates may reflect the effect of his behavior on his associates as opposed to the effect of selective association. Our results on sorting on exogenous characteristics are of interest in their own right insofar as sorting on racial, ethnic, and socio-economic lines at the formative ages studied here likely affects attitudes toward other groups.

V. A. Own Characteristics and Associates' Characteristics

We begin by regressing the mean of associates' characteristics on own characteristics. Let x_i denote a characteristic of person *i*. We regress the mean *x* among *i*'s associates, $\bar{x}_i^A$, on x_i and anticipate a positive relationship. Our model is

$$\bar{x}_i^A = \phi x_i + \gamma X_i + \Pi \overrightarrow{SG}_i + \varepsilon_i \, .$$

Person *i*'s own observable characteristics, X_i (other than x_i), are included as controls.

Also included are fixed effects fully interacting school and grade, $\overrightarrow{SG}_i$. With these school-grade effects, we estimate whether people with higher levels of x have associates with higher values of x compared to others in the same school-grade.

The estimates reported in Table 3 indicate a strong positive relationship between own characteristics and those of peoples' associates. There are a number of reasons to be careful about interpreting these estimates as the causal effect of a person's characteristics on his associations. As discussed above, the estimated relationships may reflect sorting in other arenas such as neighborhoods. Moreover, our model implies that people will selectively associate based on behaviors as well as characteristics. If groups have different behaviors and this form of sorting is present, our estimates will reflect the effect of sorting on behaviors that are affected by the characteristics.

To address neighborhood-based selection, we augment our models by including neighborhood characteristics. 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 reduces the precision of the estimates, but neither the change in sample, nor the inclusion of the neighborhood variables substantially affects the point estimates.

Lastly our estimates here and below can not rule out the possibility of a hierarchy among groups. For instance, everyone may prefer to associate with white students, but because the white students associate with each other, students from other groups may be forced to associate with each other. Even in this context our estimates show active associations among white students.

V. B. Macro-Group Composition

Under the assumption that people prefer to associate with others who are similar to themselves, the model implies that for people who have a particular (binary) characteristic, there should be a concave relationship between the share of their associates that have that characteristic and the share of the group that has that characteristic. For people who do not have that characteristic, the model implies that there should be a convex relationship between the share of their associates that have that characteristic and the share of group that does. The difference in the share of associates who have the characteristic between people with and without it will be hump-shaped. Intuitively, once a small number of people with a characteristic are introduced, the people who also have that characteristic will be able to have associates who are similar to themselves; people without the characteristic will not be forced to associate with people with it until most of the group has the characteristic.

We test this hypothesis by regressing the share of peoples' associates who have some exogenous binary characteristic, $\bar{x}_i^A$, on cubics in the share of the group who have the characteristic, $\bar{x}_i^G$, and these cubics interacted with the person's value for the characteristic, x_i. We focus on race and ethnicity because of the strength of sorting on them (from Table 3).[13] Formally, we estimate,

$$\underbrace{\bar{x}_i^A}_{\substack{\text{Associate}\\\text{Mean}}} = \underbrace{\left(\gamma_1^0 \bar{x}_i^G + \gamma_2^0 \bar{x}_i^{G\,2} + \gamma_3^0 \bar{x}_i^{G\,3}\right)}_{\substack{\text{Cubic in Mean Characteristic}\\\text{in Grade}}} + \underbrace{\left(\gamma_1^1 \bar{x}_i^G + \gamma_2^1 \bar{x}_i^{G\,2} + \gamma_3^1 \bar{x}_i^{G\,3}\right) x_i}_{\substack{\text{Cubic in Mean Characteristic in}\\\text{Grade} \times \text{Own Characteristic}}}$$
$$+ \underbrace{\gamma X_i}_{\substack{\text{Own}\\\text{Characteristics}}} + \underbrace{\pi \bar{X}_i^{G,Race} \times X_i^{Race}}_{\substack{\text{Full Racial Composition}\\\text{Interacted with own Race}}} + \underbrace{\Pi \vec{S}_i \times x_i}_{\substack{\text{School Fixed Effects}\\\text{Interacted with}\\\text{Characteristic}}} + \varepsilon_i$$

The model also includes individual characteristics to control for the direct effect of characteristics on associations; interactions between the racial composition of the grade and the individual's race (in case, for instance, the share of the grade that is white affects interactions between Asians and blacks); and school fixed effects interacted with the characteristic. Our estimates are identified from differences in the associations across students in different grades in the same school.

The estimated polynomials, are reported in table 4. Figure 5 plots the implied share of people's associates with a given characteristic and the share of the group with that characteristic, with each point representing a single school-grade. (Echenique, Fryer, and Kaufman [2006] report similar results, but without school-grade fixed effects, obtained at the same time.) As expected, the difference in the relationship between people with and without the characteristics are globally concave. The curves for people with (own) and without (other) the characteristic generally indicate the expected relationships, especially in the regions with the greatest density.

V. C. Macro-Group Size

The model implies that large macro-groups facilitate sorting so that there should be a stronger relationship between individual characteristics and associates' characteristics in large macro-groups.[14] To assess how school-grade size affects selective association, we estimate

$$\bar{x}_i^A = \beta x_i * \log\left(N_i^{SG}\right) + \theta x_i + \gamma X_i + \pi x_i \bar{x}_i^G + \Pi \overrightarrow{SG}_i + \varepsilon_i .$$

Also included in the regression are individual characteristics (including the one being studied), interactions between the characteristic under investigation and the share of the school-grade with that characteristic, and a set of school-grade fixed effects. Our estimates are identified from differences in the strength of the relationship between own characteristics and those of associates across grades of varying sizes in the same school. Under the hypothesis that grade size facilitates selective sorting $\beta > 0$.

The estimates are reported in Table 5. For all of the characteristics (except father present), an increase in the size of the school-grade increases the relationship between own behavior or characteristics and those of associates. Thus, our estimates indicate that larger macro-groups do facilitate sorting.

---

[13] Results are not reported for Indians because less than 3% of school-grades are 15% or more Indian.
[14] The mean and standard deviation of grade size are 236 and 126 (4.28 and .672 in logs) when weighted by students and 148 and 118 (4.427 and 1.427 in logs) when the data are un-weighted.

VI. Behaviors

This section studies how endogenous association affects behavior. As indicated, one can not directly study the effect of sorting on endogenous variables in the same way as sorting on exogenous variables because endogenous variables may be affected by sorting. To address this concern, we employ a multi-step procedure in which we predict both individual and group behaviors using exogenous characteristics. We then study how changes in (predicted) behavior in a macro-group affects the behavior of people whose behavior is predicted to be high (or low) based on their characteristics.

Formally, we begin by estimating the relationship between behaviors and exogenous individual characteristics,

$$y_i = \gamma^1 X_i^1 + \gamma^2 X_i^2 + \Pi \overrightarrow{SG}_i + \varepsilon_i \,.$$

Here we have partitioned the matrix of characteristics, X_i, into two components, X_i^1 and X_i^2, where $X_i = \left[X_i^1\ X_i^2\right]$. The model also includes school-grade fixed effects to control for the effect of macro-groups on behaviors.

In our second step, we use the coefficients on the individual characteristics obtained from this model, to obtain two predictions of the behavior of each individual in the sample, $\hat{y}_i^j = \hat{\gamma}^j X_i^j$, $j \in \{1,2\}$. We then estimate whether student *i* is expected to be above the population mean of *y* based on each set of characteristics. Formally, define

$$\hat{y}_i^{Hj} = \begin{cases} 1 if\ \hat{y}_i^j > \bar{y}^j \\ 0 if\ \hat{y}_i^j \le \bar{y}^j \end{cases} \text{ for } j \in \{1,2\}.$$

We also estimate $\overline{\hat{y}_i^{GHj}}$, the share of *i*'s macro-group that is expected to be above the mean of *y* based on characteristic set X^j.

VI. A. Sorting on Behaviors

Endogenous sorting implies that the difference between people with high and low (predicted) behavior, in the share of associates with a high level of that behavior should be concave in the share of the group with a high level of that behavior. To test this hypothesis, we regress the mean behavior of person *i*'s associates on interactions between our prediction of whether i is above the mean for behavior y based on $X_i^j, \hat{y}_i^{Hj}$; and the share of his group that is expected to be above the mean of behavior y based on $X_i^{j'}$ (i.e. the other set of characteristics), $\overline{\hat{y}_i^{GHj'}}$, and its square. Formally, we estimate,

$$\underbrace{\overline{y}_i^A}_{\substack{\text{Mean of}\\ \text{Associates'}\\ \text{Behavior}}} = \underbrace{\beta_1 \hat{y}_i^1 \overline{\hat{y}_i^{GH2}}}_{\substack{\text{Own Behavior Above}\\ \text{Mean*Share of School}-\\ \text{Grade Above Mean}}} + \underbrace{\beta_2 \hat{y}_i^1 \left(\overline{\hat{y}_i^{GH2}}\right)^2}_{\substack{\text{Own Behavior Above}\\ \text{Mean*Share of School}-\\ \text{Grade Above Mean}^2}} + \underbrace{\gamma^1 X_i^1 + \gamma^2 X_i^2}_{\substack{\text{Individual}\\ \text{Characteristics}}} + \underbrace{\Pi \overrightarrow{SG}_i}_{\substack{\text{School}-\text{Grade}\\ \text{Fixed Effects}}} + \varepsilon_i .$$

As indicated, in the interactions between predicted individual and macro-group behaviors, we construct the predicted individual and macro-group components from the "opposite" components of characteristics. We do this because we have already shown that individuals sort on observable characteristics and we do not want our estimates of sorting on associations to be biased by sorting on exogenous characteristics.[15] We only include one set of interactions to avoid co-linearity. Our specification also includes school-grade fixed effects and individual characteristics, so our estimates are identified from differences in how the (predicted) behavior in a school grade affects people who (are predicted to) have high levels of an action relative to those who are not in the same school grade.

Table 6 reports the estimates. As expected, the estimates show that the difference in the behaviors of the associates of people with high (predicted) behaviors relative to those with low (predicted) behaviors is concave in the share of the macro-group with high (predicted) behaviors. The linear term is positive and the quadratic is negative in 15 of 19 cases and the two are statistically significant in 12 of the cases. (None of the

---

[15] We have also regressed one set of predictions on the other and interacted the residuals. We split our characteristics into one set that comprises race and ethnicity and a second set that comprises family background and years at the school. The estimates reported use race and ethnicity to predict individual behavior and the other variables to predict macro-group behavior. Estimates that use race and ethnicity to

estimates where the signs are reversed are statistically significant.) Moreover, the two coefficients are frequently similar in magnitude, indicating that the difference in associations is greatest in groups where roughly equal numbers of people are predicted to be above and below the average behavior. As implied by endogenous association, we find that as the share of a macro-group with high (predicted) behavior increases, the mean behavior among the associates of people who (are predicted to) have a high level of that behavior increases relative to others initially and then declines as most of the macro-group (is predicted) to have a high level of that behavior.

VI. B. Reduced Form Behavioral Models

If associations are endogenous and social interactions operate, the addition of people with high (predicted) behavior to a group that has a low (predicted) behavior will raise the behavior among people with high predicted behavior more than those with low predicted behavior. Once most of the group has a high (predicted) behavior, adding more people with high (predicted) behavior will affect people with low predicted behavior more than those with high predicted behavior. Thus, the difference in behavior between people whose characteristics imply high behavior relative to those whose characteristics imply low behavior should be hump-shaped in the share of the group with high predicted behavior.

To test this hypothesis, we estimate a model that is similar to that above except we replace the mean behavior among associates as the dependent variable with the individual's own behavior. Formally we estimate,

$$\underbrace{y_i}_{Behavior} = \underbrace{\pi_1 \hat{y}_i^{H1} \overline{\hat{y}_i^{GH2}}}_{\substack{Own\ Behavior\ Above \\ Mean*Share\ of\ School- \\ Grade\ Above\ Mean}} + \underbrace{\pi_2 \hat{y}_i^{H1} \left(\hat{y}_i^{GH2}\right)^2}_{\substack{Own\ Behavior\ Above \\ Mean*Share\ of\ School- \\ Grade\ Above\ Mean^2}} + \underbrace{\gamma^1 X_i^1 + \gamma^2 X_i^2}_{\substack{Individual \\ Characteristics}} + \underbrace{\Pi \overrightarrow{SG}_i}_{\substack{School-Grade \\ Fixed\ Effects}} + \varepsilon_i .$$

As above, the interactions between predicted own behavior and predicted group behavior

---

predict group behavior and other variables to predict individual behavior are broadly comparable.

come from different sets of characteristics and the model includes both sets of individual characteristics and school-grade fixed effects. These estimates can be thought of as reduced form estimates in a two stage least squares regression of individual behaviors on associates' behaviors where interactions between predicted behaviors of the individual and the group are instruments for associates' behaviors.

Table 7 reports these estimates. As expected, we find that the difference in behaviors between those with high and low (predicted) behaviors is a concave function of and the share of the group with high (predicted) behaviors. Of the 19 estimates, 17 show the expected pattern (the other two are statistically insignificant) and 11 of the 17 are statistically significant. As above, the linear and quadratic terms are frequently similar in magnitude indicating that the difference in behavior peaks when roughly equal numbers of people (are predicted to) have high and low behaviors. Thus, our results for behaviors are consistent with social interactions with endogenous association.

VI. C. Two Stage Least Squares Estimates

As indicated, one can think of the two previous sets of estimates as the first stage and reduced form estimates of the effect of individual behaviors on associates' behaviors where interactions between predicted behaviors of the individual and group are instruments for associates' behaviors. The second stage equation in this system is

$$\underbrace{y_i}_{Behavior} = \underbrace{\theta \overline{y}_i^{A}}_{\substack{Mean\ of \\ Associates' \\ Behavior}} + \underbrace{\gamma^1 X_i^1 + \gamma^2 X_i^2}_{\substack{Individual \\ Characteristics}} + \underbrace{\Pi \overrightarrow{SG_i}}_{\substack{School-Grade \\ Fixed\ Effects}} + \varepsilon_i$$

and the first stage equation is

$$\underbrace{\overline{y}_i^{A}}_{\substack{Mean\ of \\ Associates' \\ Behavior}} = \underbrace{\beta^{12}\hat{y}_i^{H1}\overline{\hat{y}_i^{GH2}} + \beta^{21}\hat{y}_i^{H2}\overline{\hat{y}_i^{GH1}}}_{\substack{Own\ Behavior\ Above\ Mean*Share \\ of\ School-Grade\ Above\ Mean}} + \underbrace{\beta_2^{21}\hat{y}_i^{H1}\left(\overline{\hat{y}_i^{GH2}}\right)^2 + \beta_2^{21}\hat{y}_i^{H2}\left(\overline{\hat{y}_i^{GH1}}\right)^2}_{\substack{Own\ Behavior\ Above\ Mean*Share \\ of\ School-Grade\ Above\ Mean^2}} + \underbrace{\gamma^1 X_i^1 + \gamma^2 X_i^2}_{\substack{Individual \\ Characteristics}} + \underbrace{\Pi \overrightarrow{SG_i}}_{\substack{School-Grade \\ Fixed\ Effects}} + \varepsilon_i .$$

Here we include both sets of interactions between predicted own and group behavior in order to maximize the fit of our first stage equation. As above, the models include individual characteristics and school-grade fixed effects. Thus the estimates are identified from differences between individuals with high predicted behavior relative to those with low predicted behavior in the effect of group composition. Because the first-stage models generate hump-shaped differences in associates' behaviors, the identification comes from the non-monotonic relationship between group composition and the difference between people with high and low (predicted) behavior.

The results are presented in Table 8. All of the estimates but one are positive and 12 of the 19 are statistically significant. As expected the implied social effects are larger than those for the effect of macro-groups, but closer to estimates for people who are likely to associate with one another (Carrell, Fullerton, Gilchrist, and West [2006]).[16]

VII. Conclusion and Implications

We have developed a model of social interactions with endogenous association. Our theory implies the standard, linear-in-means model of behavior, but departs from the standard model by assuming that people invest in developing relationships and that they do so in order to maximize their utility.

At a policy level, the non-linearities and interactions that arise endogenously in the model imply that relocations of individuals can be positive or negative sum. Our finding of greater sorting in large macro-groups implies that relocations will be more effective when they are to small groups.

Our model provides a novel explanation for the weak effects of housing relocation experiments. We hypothesize that people who are relocated into groups where many (but not all) people are different from themselves choose to associate with the people who are like themselves. Thus, we are reluctant to conclude from these experiments that people's associates do not have large effects.

At a policy level, our results suggest that relocation programs may be less effective than policies that increase the incentives for desired behaviors (e.g. rewards for getting higher grades or more stiff penalties for crime). If endogenous effects are present, rewards will generate multipliers and because groups can be maintained, we expect less re-sorting within groups.

---

[16] The mean of the effects is .918 and the mean weighted by the inverse of the standard errors is .900.

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* 94 (No. 5, December 2004): 1613-1634.

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.

Bayer, Patrick and Stephen Ross. "Identifying Group Effects in the Presence of Sorting: A Neighborhood Effects Application." Working Paper 2006.

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.

Bisin, Alberto; Giorgio Topa; and Thierry Verdier. "Religious Intermarriage and Socialization in the United States." *Journal of Political Economy* (Vol. 112, No. 3, June 2004): 615-664.

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.

Blume, Lawrence and Steven N. Durlauf. "Identifying Social Interactions: A Review." Working Paper 2005.

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.

Burke, Mary A. and Tim R. Sass. "Classroom Peer Effects and Student Achievement." Working Paper 2006.

Carrell, Scott E.; Richard L. Fullerton; Robert N. Gilchrist; and James E. West. "Peer and Leadership Effects in Academic and Athletic Performance." Working Paper. 2006.

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.

Clampet-Lundquist, Susan; Kathryn Edin; Jeffrey R. Kling; and Greg J. Duncan. "Moving At-Risk Youth Out of High-Risk Neighborhoods: Why Girls Fare Better Then Boys." Princeton IRS Working Paper 509, March 2006.

Coleman, James S. (with John W. C. Johnstone and Kurt Jonassohn). 1961. *The Adolescent Society: The Social Life of the Teenager and its Impact on Education.* Westport, CT: Greenwood Press.

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.

Echenique, Federico; Roland G. Fryer, Jr.; and Alex Kaufman. “Is School Segregation Good or Bad?” *American Economic Review* 96 (No. 2, May 2006): 265-269.

Eder, Donna (with Catherine Colleen Evans and Stephen Parker). 2003. *School Talk: Gender and Adolescent Culture*. New Brunswick, NJ: Rutgers University Press.

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.

Foster, Jennifer. “It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism.” *Journal of Public Economics* 90 (August 2006): 1455-1475.

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.

Graham, Bryan S. 2005. “Identifying Social Interactions through Excess Variance Contrasts.” Working Paper.

Grodner, Andrew and Thomas Kniesner. “An Empirical Model of Labor Supply with Social Interactions: Econometric Issues and Tax Policy Implications.” Working Paper. 2006.

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.

Hoxby, Caroline M. and Gretchen Weingarth. “Taking Race out of the Equation: School Reassignment and the Structure of Peer Effects.” Working Paper. 2006.

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.

Lee, Lung-fei. “Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects.” *Journal of Econometrics.* (2006), doi:10.1016/j.jeconom.2006.07.001.

Lin, Xu. "Peer Effects and Student Academic Achievement: An Application of Spatial Autoregressive Model with Group Unobservables”. Working Paper. 2005.

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.

Mayer, Adalbert and Steven L. Puller. “The Old Boy (and Girl) Network: Social Network Formation on University Campuses.” Working Paper 2006.

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.

Sacerdote, Bruce. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *Quarterly Journal of Economics* (May 2001): 681-704.

Shang, Qingyan. “Estimating Endogenous Neighborhood Effects from Variations in Price and Price-Responsiveness,” Working paper. 2006.

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.

Zimmerman, David J. “Peer Effects in Academic Outcomes: Evidence from A Natural Experiment.” *Review of Economics and Statistics* 85 (No 1, February 2003): 9-23.

Figure 1. Sorting on Characteristics.

A. Active Associations

B. (Weighted) Mean Characteristics of Associates

C. High-Low Difference in (Weighted) Mean Characteristics of Associates

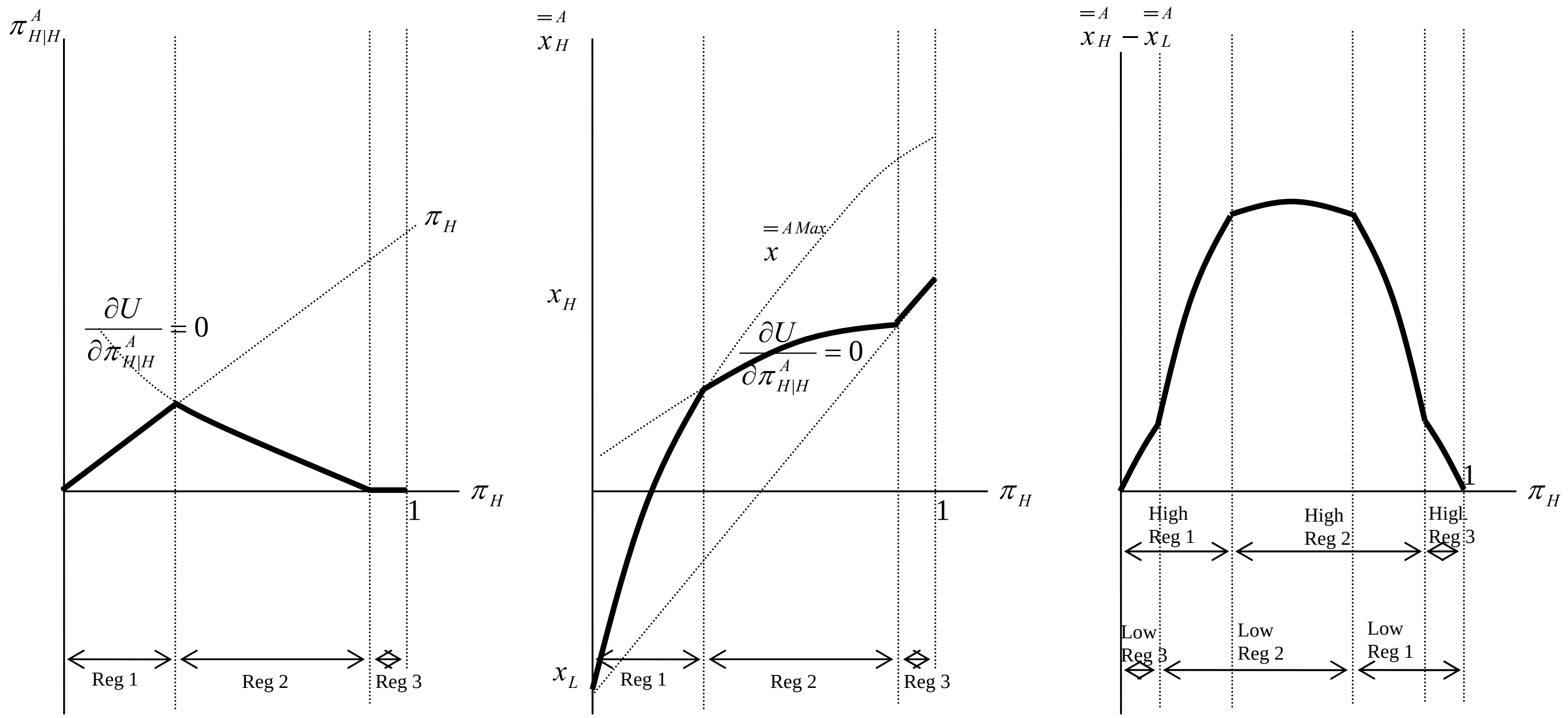


Figure 2. The Effect of Increasing Group Size.

A. Active Associations

B. (Weighted) Mean Characteristics of Associates

C. High-Low Difference in (Weighted) Mean Characteristics of Associates

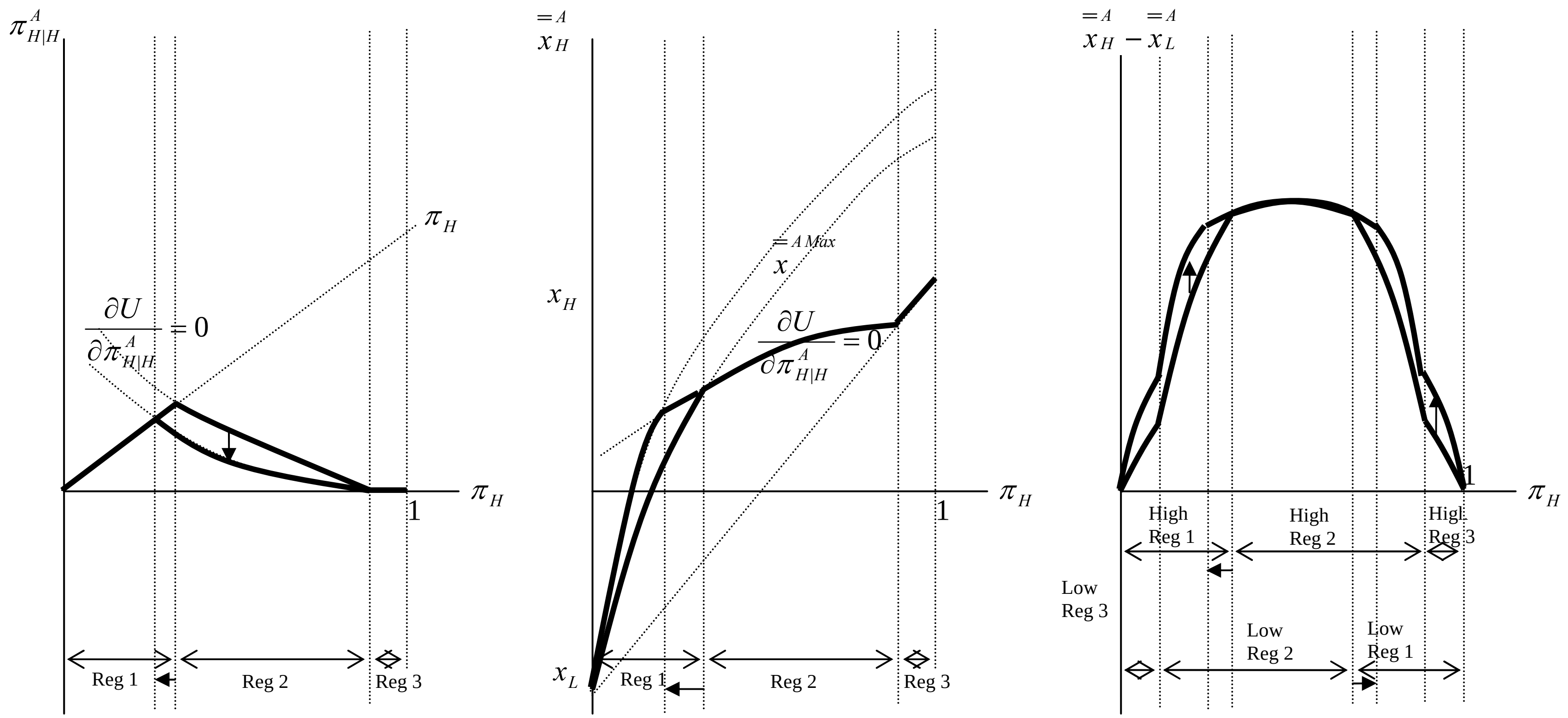
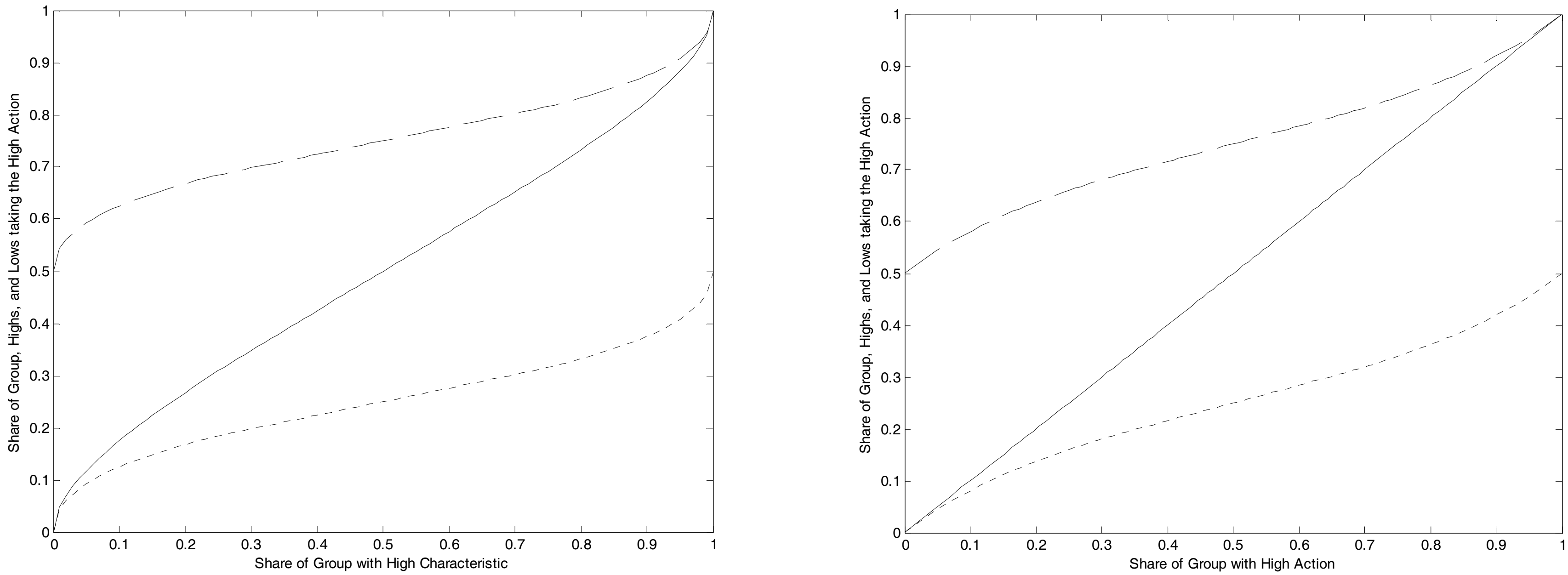
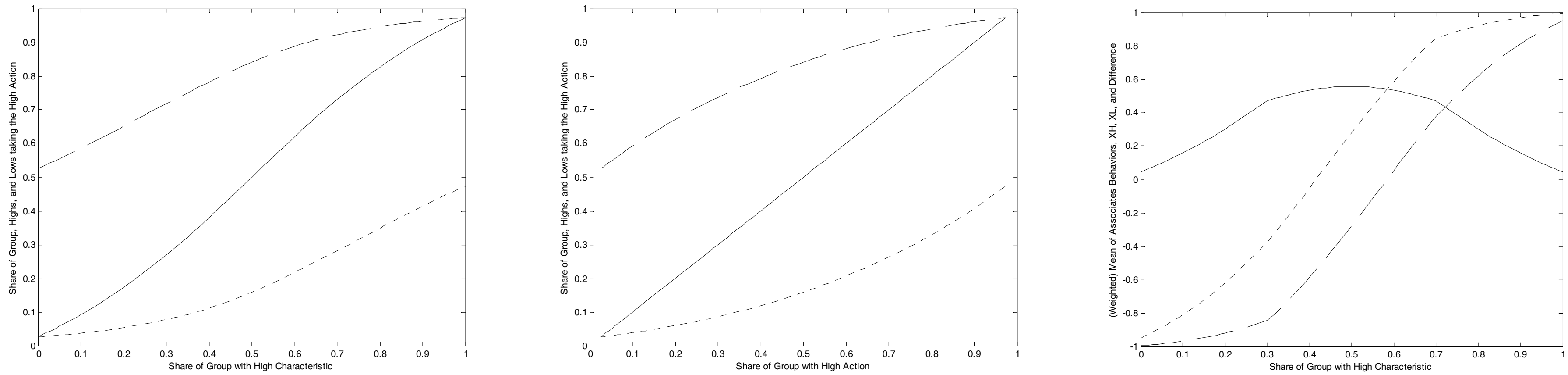


Figure 3. The Effect of Group Composition on Actions with Uniformly Distributed Preference Shocks.
A. Normally distributed unobserved preference shock.



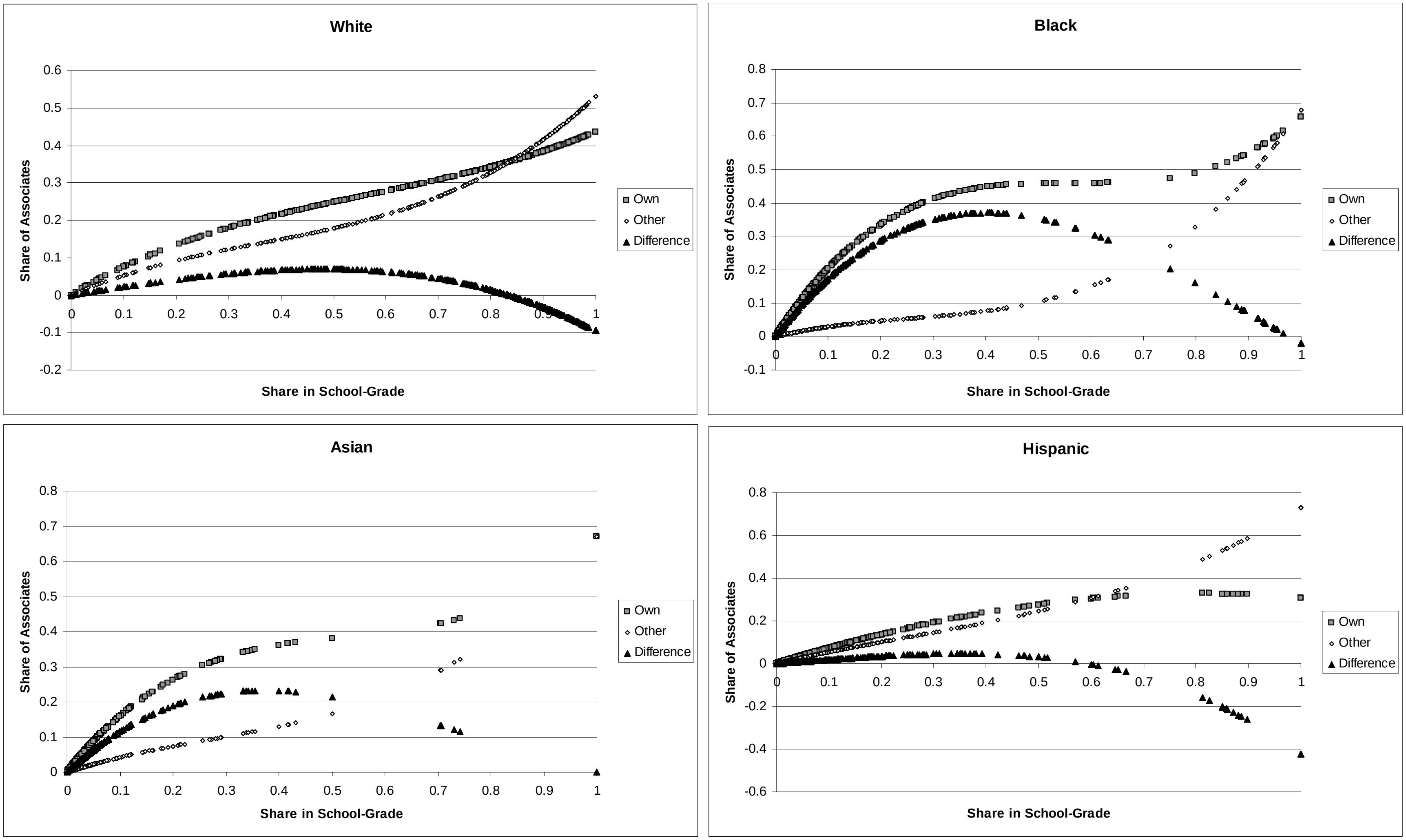
Note. The curves give mean behavior in the group (solid), among people with $x=-1$ (dotted), and $x=1$ (dashed). Solutions for $\beta=1$, $\theta=1$, $t_0=1$, $\alpha=.5$, $x\in\{-1,1\}$, $y\in\{-1,1\}$, $c=.5$, and $N=4$. Bounds of shocks are -2 and 2.

Figure 4. The Effect of Group Composition on Actions with Normally Distributed Preference Shocks.



Note. In the first two figures, the curves give mean behavior in the group (solid), among people with $x = -1$ (dotted), and $x = 1$ (dashed). In the third figure the curves give weighted mean of associates behavior for people with the high characteristic (dotted) the low characteristic (dashed) and the difference (solid). Solutions for $\beta = 1$, $\theta = 1$, $t_0 = 1$, $\alpha = .5$, $x \in \{-1,1\}$, $y \in \{-1,1\}$, $c = .5$, and $N = 4$. Shocks are distributed with mean of 0 and variance of 1.

Figure 5. The Effect of Group Composition on Associates' Characteristics, by the Individual's Characteristics.



Note. The own curve gives the share of the associates in the group for people who are in the group. The other curve gives the share of the associates in the group for people who are not in the group. The difference gives the difference between the own and other curves. Each point represents a school-grade. Plots generated from estimates in Table 4.

Table 1. Descriptive Statistics.

	Own Behavior or Characteristic		Associates' Behavior or Characteristic		
	Mean	Std. Dev.	Mean	Std. Dev.	N
Grade - English	2.858	(0.962)	2.885	(0.641)	36,384
Grade - Math	2.766	(1.002)	2.769	(0.667)	34,816
Grade - History, Social Studies	2.926	(0.978)	2.940	(0.682)	31,551
Grade - Science	2.860	(0.989)	2.877	(0.666)	32,379
Smoke	4.368	(9.688)	4.538	(6.801)	41,555
Drink Many Times	0.637	(0.481)	0.658	(0.301)	41,555
Drink	1.764	(4.215)	1.959	(2.805)	41,555
Get Drunk	0.993	(3.283)	1.152	(2.220)	41,555
Trouble with Teachers	4.911	(8.777)	4.977	(5.111)	41,555
Trouble Paying Attention	9.056	(10.405)	9.272	(5.964)	41,555
Touble with Homework	8.730	(10.468)	8.865	(6.071)	41,555
Trouble with Students	7.107	(10.664)	7.204	(6.366)	41,555
Effort Studying (reverse coded)	1.817	(0.655)	1.837	(0.388)	41,555
Race Bicycle	3.163	(7.530)	3.228	(4.475)	41,555
Do Dangerous Things	1.282	(4.675)	1.437	(2.968)	41,555
Lie	4.239	(7.758)	4.336	(4.620)	41,555
Skip School	0.921	(3.607)	0.990	(2.389)	41,555
Fight	1.181	(2.221)	1.215	(1.374)	41,555
TV	1.878	(1.283)	1.816	(0.826)	41,555
White	0.673	(0.469)	0.664	(0.396)	41,555
Black	0.161	(0.367)	0.165	(0.332)	41,555
Asian	0.068	(0.251)	0.068	(0.196)	41,555
Indian	0.045	(0.208)	0.045	(0.118)	41,555
Other Race	0.078	(0.268)	0.075	(0.172)	41,555
Hispanic	0.149	(0.356)	0.152	(0.285)	41,247
Mother's Education	13.641	(2.464)	16.679	(5.399)	34,721
Mother Homemaker	0.178	(0.382)	0.176	(0.225)	36,943
With Dad	0.785	(0.411)	0.791	(0.248)	41,059
Age	15.730	(1.217)	15.745	(1.019)	41,543
Male	0.460	(0.498)	0.465	(0.295)	41,539

Note. Standard deviations of means in parentheses.

Table 2. Decomposition of Variance in Associates' Behavior and Characteristics.

	Full Sample		Neighborhood Sample		
	Between Share	N	Between Share	Between Share	N
Grade - English	0.181	45,451	0.210	0.270	7,842
Grade - Math	0.207	45,153	0.240	0.297	7,781
Grade - History, Social Studies	0.166	44,388	0.181	0.242	7,666
Grade - Science	0.170	44,625	0.220	0.278	7,664
Smoke	0.086	46,314	0.148	0.212	7,966
Drink many times	0.092	46,314	0.147	0.199	7,966
Drink	0.056	46,314	0.096	0.154	7,966
Get Drunk	0.050	46,314	0.083	0.139	7,966
Trouble with Teacher	0.070	46,314	0.096	0.144	7,966
Trouble paying Attention	0.046	46,314	0.089	0.141	7,966
Trouble doing Homework	0.059	46,314	0.097	0.149	7,966
Trouble with Students	0.100	46,314	0.117	0.173	7,966
Effort Studying (reverse coded)	0.091	46,314	0.138	0.193	7,966
Race	0.056	46,314	0.094	0.153	7,966
Does Dangerous Things	0.046	46,314	0.079	0.150	7,966
Lies	0.045	46,314	0.075	0.126	7,966
Skips School	0.075	46,314	0.103	0.146	7,966
Fights	0.063	46,314	0.087	0.150	7,966
TV	0.233	46,314	0.288	0.349	7,966
White	0.491	46,314	0.612	0.685	7,966
Black	0.431	46,314	0.448	0.630	7,966
Asian	0.350	46,314	0.387	0.461	7,966
Indian	0.077	46,314	0.170	0.246	7,966
Other Race	0.206	46,314	0.223	0.280	7,966
Hispanic	0.459	45,955	0.508	0.561	7,910
Mother's Education	0.243	45,278	0.251	0.323	7,797
Mother Homemaker	0.081	45,659	0.114	0.177	7,867
With Dad	0.130	46,240	0.148	0.227	7,951
With Neighborhood Controls	No		No	Yes	
Effects (Main Models)	324		294	700	

Note. Between share gives the share of variance in associates between school-grades (or school-grades and school-neighborhoods in the last column).

The neighborhood characteristics 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.

Table 3. Mean Associate Characteristics Related to Own Characteristics.

	Full Sample		Neighborhood Sample			
White	0.235	(0.004)	0.217	(0.009)	0.210	(0.009)
Black	0.572	(0.003)	0.577	(0.008)	0.522	(0.008)
Asian	0.316	(0.003)	0.397	(0.008)	0.390	(0.008)
Indian	0.033	(0.003)	0.044	(0.006)	0.043	(0.006)
Hispanic	0.231	(0.004)	0.237	(0.009)	0.228	(0.009)
Mother has some College	0.094	(0.003)	0.086	(0.008)	0.080	(0.008)
Mother Homemaker	0.018	(0.003)	0.013	(0.007)	0.013	(0.007)
With Dad	0.026	(0.003)	0.028	(0.007)	0.025	(0.007)
N (Race Variables)	46,990		8,080		8,080	
N (Hispanic)	42,822		7,371		7,371	
N (Family Background)	36,942		6,350		6,350	
Full Sample	Yes					
Neighborhood Sample			Yes		Yes	
With Own*Neighborhood Interactions					Yes	

Note. Standard errors reported in parentheses. Dependent variable is mean characteristic among associates. Each estimate is from a separate regression of the mean characteristics of associates on the individual’s own value of that characteristic. Estimates include individual characteristics (age, gender, Hispanic background, Hispanic background unknown, race dummy variables, years at school, years at school missing, mother’s education, mother’s education missing, mother is a homemaker, mother is a homemaker missing, dad present, and dad present missing) and school-grade fixed effects. The neighborhood characteristics 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.

Table 4. Associate Characteristics Related to own Characteristic Interacted with Mean in School-Grade.

	Group Share of School-Grade	Group Share of School-Grade^2	Group Share of School-Grade^3	Person in Group * Group Share of School-Grade	Person in Group * Group Share of School-Grade^2	Person in Group * Group Share of School-Grade^3
White	0.608	-0.926	0.849	0.220	-0.011	-0.305
	(0.230)	(0.476)	(0.303)	(0.303)	(0.512)	(0.325)
Black	0.380	-0.986	1.284	2.042	-3.269	1.207
	(0.100)	(0.316)	(0.271)	(0.271)	(0.480)	(0.359)
Asian	0.496	-0.826	0.999	1.392	-2.462	1.069
	(0.080)	(0.336)	(0.388)	(0.388)	(0.769)	(0.693)
Hispanic	0.580	-0.507	0.656	0.170	0.157	-0.748
	(0.102)	(0.454)	(0.473)	(0.473)	(0.869)	(0.731)

Note. Standard errors reported in parentheses. Each row represents a separate regression. Regressions for whites, blacks, and Asians contain 46,690 observations. The regression for Hispanics contains 42,822 observations because people with missing Hispanic background were excluded. The dependent variable is the mean of each characteristic among associates. Estimates include individual characteristics (age, gender, Hispanic background, Hispanic background unknown, race dummy variables, years at school, years at school missing, mother’s education, mother’s education missing, mother is a homemaker, mother is a homemaker missing, dad present, and dad present missing), the racial (or ethnic) composition of the school-grade and interactions between race (or ethnic background) and the racial (or ethnic) composition of the school-grade and school dummy variables interacted with race (or Hispanic background).

Table 5. The Effect of Macro-Group Size on Sorting: Associates' Characteristics Related to own Characteristics Interacted with Macro-Group Size.

	Own Characteristic * Log(School-Grade Size)	
White	0.0040	(0.0046)
Black	0.0672	(0.0042)
Asian	0.1382	(0.0049)
Indian	0.0093	(0.0041)
Hispanic	0.1089	(0.0054)
Mother has some College	0.0056	(0.0022)
Mother Homemaker	0.0230	(0.0045)
With Dad	-0.0021	(0.0042)
N (Race Variables)	46,990	
N (Hispanic)	42,822	
N (Family Background)	36,942	
School-Grade Effects	Yes	

Note. Standard errors reported in parentheses. The dependent variable is the mean of each characteristic among associates. Estimates are the coefficient on the interaction between own characteristics and the log of school-grade size. Estimates include individual characteristics (age, gender, Hispanic background, Hispanic background unknown, race dummy variables, years at school, years at school missing, mother's education, mother's education missing, mother is a homemaker, mother is a homemaker missing, dad present, and dad present missing), interactions between the composition of the school-grade and own characteristics, and school-grade fixed effects.

Table 6. Sorting on Behaviors: Associates' Behaviors Related to Predicted own Behavior Interacted with Predicted Mean in School-Grade.

Independent Variable:	Own Behavior Predicted to be High * Share of School-Grade with Behavior Predicted to be High		Own Behavior Predicted to be High * Share of School-Grade with Behavior Predicted to be High^2		Observations
Grade - English	2.387	(0.398)	-2.032	(0.354)	33,758
Grade - Math	1.203	(0.348)	-1.056	(0.315)	32,246
Grade - History, Social Studies	2.121	(0.371)	-1.806	(0.333)	29,217
Grade - Science	1.830	(0.404)	-1.579	(0.354)	30,001
Smoke	8.525	(2.905)	-5.331	(3.273)	38,422
Drink many times	-0.200	(0.154)	0.276	(0.173)	38,422
Drink	3.055	(2.726)	-2.440	(2.531)	38,422
Get Drunk	0.933	(1.903)	-1.603	(1.697)	38,422
Trouble with Teacher	17.589	(3.009)	-18.766	(3.477)	38,422
Trouble paying Attention	12.105	(3.771)	-12.166	(4.250)	38,422
Trouble doing Homework	11.543	(3.436)	-10.263	(3.936)	38,422
Trouble with Students	10.474	(3.222)	-9.778	(3.745)	38,422
Effort Studying (reverse coded)	-0.299	(0.278)	0.190	(0.241)	38,422
Race	6.594	(3.249)	-6.575	(2.826)	38,422
Does Dangerous Things	-2.520	(2.377)	1.516	(1.976)	38,422
Lies	5.831	(3.614)	-6.979	(3.825)	38,422
Skips School	7.899	(1.916)	-9.928	(2.113)	38,422
Fights	-1.132	(0.945)	0.616	(0.794)	38,422
TV	0.880	(0.376)	-1.007	(0.371)	38,422

Note. Standard errors reported in parentheses. The dependent variable is the mean of each behavior among associates. Estimates are the coefficient on whether the person's behavior is predicted to be above the population mean interacted with the share of the school-grade whose behavior is predicted to be above the population mean and its square. Own behaviors predicted from regressions of behaviors on race and ethnicity dummy variables; school-grade behaviors predicted from years at school, years at school missing, mother's education, mother's education missing, dad present, and dad present missing. Estimates include individual characteristics (age, gender, Hispanic background, Hispanic background unknown, race dummy variables, years at school, years at school missing, mother's education, mother's education missing, mother is a homemaker, mother is a homemaker missing, dad present, and dad present missing), and school-grade fixed-effects variables.

Table 7. Effect of Endogenous Association on Behavior: Own Behaviors Related to Predicted own Behavior Interacted with Predicted Mean in School-Grade.

Independent Variable:	Own Behavior Predicted to be High * Share of School-Grade with Behavior Predicted to be High		Own Behavior Predicted to be High * Share of School-Grade with Behavior Predicted to be High^2		Observations
Grade - English	2.363	(0.626)	-1.892	(0.557)	33,758
Grade - Math	1.051	(0.559)	-0.752	(0.507)	32,246
Grade - History, Social Studies	2.067	(0.555)	-1.541	(0.498)	29,217
Grade - Science	0.965	(0.628)	-0.728	(0.551)	30,001
Smoke	9.553	(4.257)	-4.991	(4.797)	38,422
Drink many times	-0.563	(0.254)	0.541	(0.285)	38,422
Drink	3.117	(4.080)	-1.490	(3.788)	38,422
Get Drunk	3.129	(2.756)	-2.747	(2.457)	38,422
Trouble with Teacher	23.548	(5.227)	-25.951	(6.038)	38,422
Trouble paying Attention	18.231	(6.648)	-19.414	(7.493)	38,422
Trouble doing Homework	13.209	(6.008)	-12.292	(6.882)	38,422
Trouble with Students	15.307	(5.543)	-16.948	(6.443)	38,422
Effort Studying (reverse coded)	1.169	(0.479)	-1.090	(0.415)	38,422
Race	7.282	(5.490)	-8.496	(4.776)	38,422
Does Dangerous Things	6.437	(3.686)	-4.227	(3.064)	38,422
Lies	11.706	(6.136)	-12.510	(6.494)	38,422
Skips School	1.630	(2.880)	-2.086	(3.177)	38,422
Fights	-0.209	(1.527)	0.949	(1.283)	38,422
TV	1.026	(0.638)	-1.269	(0.630)	38,422

Note. Standard errors reported in parentheses. The dependent variable is the individual's behavior. Estimates are the coefficient on whether the person's behavior is predicted to be above the population mean interacted with the share of the school-grade whose behavior is predicted to be above the population mean and its square. Own behaviors predicted from regressions of behaviors on race and ethnicity dummy variables; school-grade behaviors predicted from years at school, years at school missing, mother's education, mother's education missing, dad present, and dad present missing. Estimates include individual characteristics (age, gender, Hispanic background, Hispanic background unknown, race dummy variables, years at school, years at school missing, mother's education, mother's education missing, mother is a homemaker, mother is a homemaker missing, dad present, and dad present missing), and school-grade fixed effects variables.

Table 8. Two Stage Least Squares Estimate of Associates' Social Effect.

	Estimate	(Std. Err.)	Observations
Grade - English	1.209	(0.234)	33,758
Grade - Math	0.791	(0.314)	32,246
Grade - History, Social Studies	1.044	(0.184)	29,217
Grade - Science	0.752	(0.235)	30,001
Smoke	1.239	(0.099)	38,422
Drink many times	0.716	(0.353)	38,422
Drink	2.659	(1.353)	38,422
Get Drunk	0.562	(0.598)	38,422
Trouble with Teacher	1.204	(0.336)	38,422
Trouble paying Attention	0.671	(0.396)	38,422
Trouble doing Homework	1.009	(0.389)	38,422
Trouble with Students	1.106	(0.420)	38,422
Effort Studying (reverse coded)	-1.389	(1.390)	38,422
Race	1.970	(0.780)	38,422
Does Dangerous Things	0.527	(0.511)	38,422
Lies	1.386	(0.750)	38,422
Skips School	0.328	(0.252)	38,422
Fights	0.364	(0.561)	38,422
TV	1.285	(0.433)	38,422

Note. Standard errors reported in parentheses. The dependent variable is the individual's behavior. First-stage estimates reported in table 6; reduced form estimates reported in table 7. Estimates include individual characteristics (age, gender, Hispanic background, Hispanic background unknown, race dummy variables, years at school, years at school missing, mother's education, mother's education missing, mother is a homemaker, mother is a homemaker missing, dad present, and dad present missing), and school-grade fixed effects variables.

Appendix Table 1. Variable Descriptions.

Variable	Description	Coding
	At the most recent grading period, what was your grade in each of the following subjects?	
Grade - English	...English / Language Arts	4 (A); 3 (B); 2 (C) 1 (D or lower)
Grade - Math	...Mathematics	As above
Grade - History, Social Studies	...History/Social Studies	As above
Grade - Science	...Science	As above
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)

Appendix Table 1. Variable Descriptions (Continued)

Variable	Description	Coding
	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
TV	Outside of school hours, about how much time do you spend watching television or video cassettes on an average school day?	0 (none); .5 (less than 1 hour); 1.5 (1 to 2 hours); 2.5 (3 to 4 hours); 4.5 (more than 4 hours)