# A Roy Model of Social Interactions

Cicala, Steve and Fryer, Roland G. and Spenkuch, Jörg L.

Harvard University, Harvard University and NBER, University of Chicago

February 2011

# A Roy Model of Social Interactions*

Steve Cicala
Harvard University

Roland G. Fryer, Jr.
Harvard University and NBER

Jörg L. Spenkuch
University of Chicago

February 2011

**Abstract**

We develop a Roy model of social interactions in which individuals sort into peer groups based on comparative advantage. Two key results emerge: First, when comparative advantage is the guiding principle of peer group organization, the effect of moving a student into an environment with higher-achieving peers depends on where in the ability distribution she falls and the effective wages that clear the social market. In this sense our model may rationalize the widely varying estimates of peer effects found in the literature without casting group behavior as an externality in agents' objective functions. Second, since a student's comparative advantage is typically unobserved, the theory implies that important determinants of individual choice operate through the error term and may, even under random assignment, be correlated with the regressor of interest. As a result, linear in means estimates of peer effects are not identified. We show that the model's testable prediction in the presence of this confounding issue—an individual's ordinal rank predicts her behavior, *ceteris paribus*—is borne out in two data sets.

*We are grateful to Gary Becker, Edward Glaeser, Bryan Graham, Richard Holden, Lawrence Katz, Steven Levitt, Franziska Michor, Bruce Sacerdote, Chris Shannon, Jesse Shapiro, Andrei Shleifer, Glen Weyl, as well as seminar participants at Harvard and Chicago for many helpful comments and suggestions. We thank Lisa Sanbonmatsu for her assistance obtaining data from the Moving to Opportunity Experiment, which were used in an earlier draft of this paper. Vilsa E. Curto, Ryan Fagan, and Wonhee Park provided excellent research assistance. Financial support from the Weatherhead Center for International Affairs and Institute for Humane Studies [Cicala], the Education Innovation Lab at Harvard University [Fryer], and the German National Academic Foundation [Spenkuch] is gratefully acknowledged. Correspondence can be addressed to the authors at Department of Economics, Harvard University, 1805 Cambridge Street, Cambridge MA 02138 [Cicala or Fryer]; Department of Economics, University of Chicago, 1126 E 59th Street, Chicago IL 60637 [Spenkuch]; or by e-mail: scicala@fas.harvard.edu [Cicala], rfryer@fas.harvard.edu [Fryer], or jspenkuch@uchicago.edu [Spenkuch]. The usual caveat applies.

# 1 Introduction

For centuries, social scientists have recognized the importance of social interactions (see , for instance, Rousseau 1762, pp.789).[1] Mill (1859), for instance, laments not only the tyranny of government, but also the tyranny of social norms. More recently, Gans (1962) describes an insidious form of social interactions in which Italian immigrant communities in Boston's West End impose costs on individuals who "act mobile."[2] Wilson (1987) provides qualitative evidence that the development of an "underclass" of black city dwellers on Chicago's South Side was due to the emigration of working families and the resulting decrease in role models and neighborhood quality; and Borjas (1995) demonstrates that the mean skill level within one's ethnic group in the previous generation is correlated with own educational achievement.[3]

To better understand these phenomena, economists have developed models of social interactions by putting environmental variables, such as the mean behavior in one's social group or the mean educational attainment in one's neighborhood, into agents' utility functions.[4] In this class of models, peers are a source of monotonic externalities; unruly peers cause more trouble and smarter peers encourage higher academic achievement (Akerlof 1997, Becker 1974, 1996, Benabou 1993, Bernheim 1994). Importantly, these externalities are simply assumed. Becker and Murphy (2000) call this sort of complementarity between individual actions and those of one's peers "the fundamental assumption in analyzing the influence of social capital...on closely related behavior" (p.9).

Empirical evidence in favor of models which predict that favorable social interactions lead to positive outcomes has been ambivalent. While many authors confirm the hypothesis using data sets ranging from primary and secondary students in Texas to freshmen in the US Air Force Academy (Carrell et al. 2009, Hanushek et al. 2003, Hoxby 2000, Hoxby and Weingarth 2005, Duflo et al. 2008, Imberman et al. 2009, Goux and Maurin 2007), others find no evidence of peer effects or occasional negative effects from increases in the peer group mean (Sanbonmatsu et al. 2006, Kang 2007, Angrist and Lang 2004, Carrell et al. 2010).

In this paper we take a primitive view of social interactions using insights from neoclassical economics. Modelling the endogneity of contacts within narrowly defined social settings we posit the existence of a 'market for peers' analogous to a traditional labor market. In the spirit of Becker (1965), agents derive utility from final goods produced by combining time and market inputs.

[1] Manski (1993) partitions the space of social interactions into three categories: endogenous effects, exogenous effects, and correlated effects. Endogenous effects occur when an individual's behavior is directly influenced by that of her peers. Exogenous (or contextual) effects occur when individual behavior is influenced by group composition or neighborhood characteristics. Correlated effects are present when individual and group behavior are related because peers share similar traits. We will not attempt to distinguish these important channels. In what follows, all of these categories interact to determine behavior through an equilibrium mechanism.

[2] Many ethnographers describe similar phenomena around the globe: the Buraku Outcastes of Japan (Devos and Wagasutma 1966); Blacks in America (Fordham and Ogbu 1986), the Maori of New Zealand (Chapple, Jefferies, and Walker 1997), Blacks on Chicago's south side circa 1930 (Drake and Cayton 1945), the working class in Britain (Willis 1977), among others.

[3] See Durlauf (2004) for a careful review of the literature on peer effects in education, crime, welfare participation, and health.

[4] Austen-Smith and Fryer (2005), Berman (2000), and Iannaccone (1992) are notable exceptions.

Production of final goods, however, occurs in peer groups, which we cast as firms in a two sector Roy model (Roy 1951). Our focus is on equilibria in which the effect of peers is mediated through an implicit price mechanism akin to Becker (1973); an agent's contribution to group production determines the share of output she receives.

In equilibrium, heterogeneity in ability leads to individuals selecting into sectors based on comparative advantage. The direction and magnitude of the impact of social interactions depends on the shape of the production functions as well as the ability distribution. We characterize equilibria under two different assumptions: (i) industry production functions are concave in labor inputs, (ii) production functions are convex. Since the number of complexities to generalize the model is only bounded by one's ability to reinterpret variables from classical price theory, we limit the model to these basic characteristics. In Appendix B, we extend the basic model to allow for many sectors and $n$-dimensional skill (Heckman and Scheinkman 1987), hierarchies and endogenous group size (Rosen 1982), and show that the basic results of our model hold when the sectoral choice problem is cast in a general social multiplier model (Becker and Murphy 2000, Glaeser, Sacerdote, and Scheinkman 2003).[5]

It is important to emphasize at the outset that there are no intrinsic externalities or complementarities built into the model. When comparative advantage is the guiding principle of peer group organization, the effect of moving a student to an environment with higher-achieving peers is an equilibrium outcome. An individual's behavior depends on where in the new distribution she lands, and on the effective 'wages' that clear the social market. Put differently, selection into peer group roles is determined by the intersection of supply and demand for various skills; and peer effects are obtained through market prices. Thus, peer quality might be a source of linear or non-linear, positive, or even negative externalities. If the demand function is approximately linear, for instance, the model predicts peer quality to exhibit positive and linear effects. Conversely, if production is subject to strong diminishing marginal returns, the model predicts estimates of peer effects to differ in size and magnitude depending on the location of the initial equilibrium and the nature of the underlying ability distribution. Effects that are non-linear and negative can arise if the 'supply' and 'demand' curves shift in opposite directions. In this sense, the model may rationalize the widely varying results on peer effects found in the empirical literature without appealing to externalities.

Moreover, our Roy model of social interactions has implications for the identification of peer effects. Social wages and a student's comparative advantage are typically unobserved, which implies that key determinants of individual choice operate through the error term. Moreover, these unobservables may vary only on the school or neighborhood level and will, even under random assignment, be correlated with the regressor of interest (e.g., poverty rates, or the mean behavior of other students). Therefore, identification of neighborhood and peer effects may be more difficult

[5]Other possible extensions would allow individuals to invest in human capital (see Becker 1964, Ben-Porath 1967, Rosen 1972), or to search for a peer group within and across sectors (Jovanovic 1979a, 1979b, Miller 1984, Neal 1999). To understand the implications of the sorting of individuals into groups when membership is costly and benefits depend monotonically on group composition see Cutler and Glaeser (1997).

than Manski (1993) describes—one not only has to solve the reflection problem and deal with the systematic sorting of individuals into peer groups, but one must also account for the presence of group level unobservables that determine behavior. Even if individual behavior were to depend directly on the group mean, we show that when contacts in the social market are endogenous the parameters of interest are typically not identified. We leave the development of general conditions for identification in the presence of these unobservables for another occasion, and pursue the more modest goal of providing evidence consistent with our model's key prediction: Since similar individuals facing the same social wages have a common comparative advantage, all else equal, a student's proclivity to 'act out' should be correlated with her ordinal rank in the ability distribution.

Using two data sources, New York City Public Schools administrative records (NYCPS) and the National Educational Longitudinal Study (NELS), we demonstrate that, *ceteris paribus*, individuals' academic rank significantly affects the probability of exhibiting problem behaviors. In the NYCPS data, which contain information on the same students gathered at multiple points in time, we exploit transitions from elementary school (5th grade) to middle school (6th grade) to estimate that a fifty percentile decrease as opposed to a fifty percentile increase in rank among schoolmates (presumably from moving to a different school with more academically able peers) is associated with roughly a five percentage point increase in the probability of a serious behavioral incident (on a base of eight percent). In doing so, we are able to control for student specific determinants of behavior, but not for systematic choice of school. To account for systematic sorting into schools, we use a student's hypothetical change in rank if they attended their zoned school (i.e. the default school based on their physical address) as an instrument for actual change in rank based on the school they choose to attend. This approach supports our findings.

Our second dataset is the National Educational Longitudinal Study (NELS). NELS allows us to relate a student's behavior in different classrooms to a proxy for her course specific rank. We show that a fifty percentile decline in rank across classes is associated with a ten percentage points higher probability that the teacher reports behavioral problems in the course for which she has the lower rank (on a base of forty percent).

While these data are inconsistent with models that predict peer quality to exhibit positive monotonic externalities (e.g., Becker 1996), we urge the reader to interpret our findings with two important caveats. First, as mentioned above, the estimates are not well identified peer effects. Second, models which predict negative peer externalities (e.g., smarter peers may decrease achievement) might also be consistent with our results.

The remainder of the paper is organized as follows. Section 2 develops our formal model, provides conditions that can explain the widely varying estimates found in the literature, and explores the implications of the comparative advantage approach for predicting the efficacy of social programs ex ante. Section 3 discusses identification of peer effects in the presence of comparative advantage, and shows evidence from two large data sets that is consistent with our model. Section 4 concludes. There are two appendices. Appendix A describes the data used in our analysis and how we construct our samples. Appendix B contains additional formal results omitted from the body of the paper.

# 2 The Market for Peers

## 2.1 The Basic Model

The economic model we propose in this section is a simplified version of the well-known multi-sector choice problem, and builds upon impressive literatures designed to understand the evolution of earnings, the (hedonic) pricing of skills, and the assignment of workers to firms (e.g., Ben-Porath 1967, Heckman and Sedlacek 1985, Heckman and Scheinkman 1987, Murphy 1986, Rosen 1974, 1982, 1983, Roy 1951, Sattinger 1979, 1993, Tinbergen 1956, and Willis and Rosen 1979). The novelty in our approach lies in the application of these classic methods to develop a theory of social interactions where contacts within a social market are endogenous and peer quality does not act as a direct externality. The paper most similar to ours is Heckman and Sedlacek (1985), who develop an equilibrium model of self-selection in the labor market. We extend Heckman and Sedlacek (1985) by considering the case of increasing returns to scale on the industry level, but we cannot implement their empirical exercise because we observe neither 'social wages' nor individuals' choices of sectors directly.

### A. Building Blocks

Let there be a continuum of agents with unit mass. Every agent is endowed with one unit of (non-transferable) time. There are two activities in which agents can engage with their peers: studying or mischief. These activities are exclusive and undertaken by separate social groups in one of two sectors: 'nerds' and 'troublemakers'.[6] Each sector, indexed by $j \in \{N, T\}$, combines the effective units of its members' time with another input we label 'capital' according to a general, twice continuously differentiable industry production function $F_j(L_j, K_j, A_j)$. Let $A_j$ denote a technology shifter, such as school or neighborhood quality, or the quantity of policing. $L_j$ represents the total supply of effective labor units to sector $j$, and $K_j$ denotes the supply of effective capital. We allow capital to broadly represent any non-human input into groups' production (e.g., textbooks, sharp scissors, but also labor market conditions, or expectations thereof).

Agents are heterogenous along two dimensions. Their varying size and strength yield differences in the ability to cause trouble, whereas heterogeneity in cognitive ability implies differences in their ability to be a true nerd. Let the continuous function $\sigma_N(i) : [0,1] \to \mathbb{R}_+$ denote the effective units of 'nerdiness' that agent $i$ is capable of contributing to the group (e.g., expertise in differential geometry). Analogously, agent $i$'s troublemaking ability is given by $\sigma_T(i) : [0,1] \to \mathbb{R}_+$. We assume that agents are solely interested in maximizing their social income

$$U(i) = \max\{\sigma_N(i) w_N, \sigma_T(i) w_T\},$$

[6] Appendix B details an $n$-sector analog to our basic model.

where $w_N$ and $w_T$ are the market clearing wages for effective units of nerd and troublemaking labor, respectively. As in Welch (1969), Heckman and Sedlacek (1985), and Heckman and Scheinkman (1987), earnings follow the linear in characteristics approach developed by Gorman (1980) and Lancaster (1966). Note that there are no explicit peer externalities built into agent's utility. That is, given $w_N$ and $w_T$, the behavior of one's peers has no influence on own decisions. In section 2.4, we relax this assumption and allow for more general utility functions (e.g. individuals care about more than just social income in school). Here, however, we present a very simple and parsimonious model to demonstrate that sorting into peer groups alone can produce "peer effects" and may help reconcile much of the conflicting empirical literature.

Individuals maximize their social income by choosing either the nerd or troublemaking sector according to a simple cut-off rule (Roy 1951). Let $\sigma(i) \equiv \frac{\sigma_N(i)}{\sigma_T(i)}$ denote agent $i$'s skill as a nerd relative to that as a troublemaker, and order agents such that $\sigma'(i) \geq 0$. The agent indifferent between the two sectors, $i^*$, has a skill ratio of

$$\sigma(i^*) = \frac{w_T}{w_N}. \tag{1}$$

All individuals with index $i \geq i^*$ join forces with the nerds, and individuals with $i < i^*$ become troublemakers. In our price theory of social interactions, comparative (rather than absolute) advantage determines an individual's choice of sector.

By individual optimization and market clearing, labor supply to both sectors is given by:

$$L_N^* = \int_{i^*}^{1} \sigma_N(s)\, ds \tag{2}$$

$$L_T^* = \int_{0}^{i^*} \sigma_T(s)\, ds. \tag{3}$$

Equations (2) and (3) characterize the supply side in the market for peers. Equilibrium, however, also depends on the demand side, and therefore on the shape of the industry production functions. Below, we consider cases in which production is concave or convex in labor inputs.

### B. Concavity in Labor Inputs

In the theory of the firm, it is assumed that labor exhibits diminishing returns to scale in labor on the firm as well as the industry level. If, for instance, increasing the number of trouble makers does more to increase the probability of getting caught than of winning a fight, then the production function in the troublemaking sector will be concave in $L_T$. In equilibrium, competition in the market for peers ensures that nerd and troublemaking labor are paid their marginal products:

$$w_j = \frac{\partial}{\partial L} F_j\left(L_j^*, K_j, A_j\right). \tag{4}$$

With free entry into both the nerd and troublemaking sectors, individual peer groups earn zero profits. That is, all output is divided among group members.[7]

Substituting (4) into (1), market clearing yields the equilibrium condition

$$\delta\left(i^{*}\right) \equiv \frac{\frac{\partial}{\partial L} F_{T}(L_{T}(i^{*}), K_{T}, A_{T})}{\frac{\partial}{\partial L} F_{N}(L_{N}\left(i^{*}\right), K_{N}, A_{N})} = \sigma\left(i^{*}\right), \tag{5}$$

where $\delta\left(i\right) : [0,1] \rightarrow \mathbb{R}_{+}$ denotes the ratio of marginal products in both sectors when the threshold index separating sectors is $i$. Since $\delta'\left(i\right) \leq 0$ for all $i$, the 'demand curve' in the market for peers is (weakly) downward sloping.[8] To see this, note that equations (2) and (3) respectively imply $\frac{dL_N}{di} < 0$ and $\frac{dL_T}{di} > 0$; and concavity of the industry production functions causes the ratio of marginal products to decrease as labor shifts from the nerd into the troublemaking sector. We can now describe equilibrium graphically.

Figure 1 shows equilibrium in the basic model when production is concave in labor inputs. As described above, it features downward sloping demand and upward sloping supply. There is a unique equilibrium at $i^{*}$ with market clearing relative prices, $\frac{w^T}{w^N}^{*}$, determined by the intersection of supply and demand. All individuals with $i < i^{*}$ select into the troublemaker sector and individuals with $i \geq i^{*}$ choose to be nerds.

Suppose there is a shift in troublemaking technology—less police surveillance or an increase in the availability of weapons—holding everything else constant. An increase in troublemaking technology is represented by an outward shift of the $\delta$-schedule in Figure 1, which results in higher relative wages for troublemakers and fewer nerds. A decrease in troublemaking technology has the opposite effect: an inward shift of the $\delta$-schedule, a decrease in the relative wages of troublemakers, and an increase in the number of agents who choose to become nerds.

Comparative statics with respect to the skill disribution, however, can be more counterintuitive. Consider, for instance, an increase in nerd skill among the population holding troublemaking skills fixed. First, an increase in agents' nerdiness shifts the supply curve inward (from $\sigma(i)$ to $\sigma(i)'$). Second, the demand curve shifts outwards (from $\delta\left(i\right)$ to $\delta\left(i\right)'$) due to the fact that with more academically able peers there will be more efficiency units of nerd skill supplied at any $i$. While both shifts lead to an unambigous rise in the relative wage of troublemakers, the effect on quantities in indeterminate.

Figure 2 illustrates the situation. In Panel A, the wage from being a nerd drops more rapidly with effective labor in the sector, leading to a larger outward shift of $\delta(i)'$, which is one factor contributing to the overall contraction in the nerd sector. The other contributing factor is that the increase in nerd skill comes disproportionately from the upper end of the distribution, so that the expansion in $\delta$ is larger at the upper end of the distribution, and the intersection of the two

[7] Assuming identical peer groups within sectors, the number of groups adjusts such that $f_j\left(l_j, k_j, a_j\right) = c_j$, where $c_j$ is the fixed cost of operating in sector $j$, and lower case symbols denote the group level analog of their upper case counterparts. Following Rosen (1982), Appendix B endogenizes group size and hierarchies within groups.

[8] Technically a demand curve gives the quantity demanded at a particular price, all else equal. $\delta\left(i\right)$, however, denotes the marginal individual consistent with a certain wage ratio. Therefore, $\delta$ should more appropriately be thought of as demand side equilibrium schedule.

curves is therefore further to the right.[9] In Panel B, social wages are less responsive to sector labor supply, and the increase in skill is more concentrated to the left of the initial equilibrium. While counterintuitive, an overall increase in the nerd skills actually yields an expansion in the size of the nerd sector.

### C. Convexity in Labor Inputs

The assumption that the marginal utility of a social activity is increasing in overall participation has been the focus of much of the work on social interactions. Study groups allow students to benefit from division of labor on a lab project, and ensure individual students do not waste time stuck on a question to which someone else in the group knows the answer (Lazear 2001).

In the traditional theory of the firm, increasing returns to scale on the industry level imply that there exists only one firm per sector because one large firm will produce more efficiently than multiple smaller ones. One way to reconcile increasing returns to scale on the industry level with the existence of many firms in a competitive market is the concept of *external increasing returns to scale* pioneered by Marshall (1890) and formalized by Ethier (1982a, 1982b). The key idea is that individual firms are price takers and produce subject to diminishing marginal returns, but their activities exhibit positive externalities strong enough to cause marginal returns on the industry level to increase. Suppose, for instance, that the size of the market or the cost of production depend on the number of firms—say, because more firms invest in a new, more efficient technology—and that firms fail to internalize this spill-over effect. Then, production might exhibit increasing returns to scale at the industry level, but decreasing marginal returns at the firm level (Murphy et al. 1989). Another example of external economies, due to Marshall (1890), are advances in "trade-knowledge" which might be hard to keep secret from firms within the same industry.[10]

In the market for peers we assume that production is concave on the peer group level, but convex on that of the industry. The activities of one large troublemaker group might be harder to coordinate than those of a smaller one leading to a higher risk of getting caught for any individual member. However, more troublemakers at a school can divert teachers' attention and thereby reduce others' probability of punishment. Yet, the inability to coordinate efficiently precludes individual groups from becoming too large and reaping the full benefits of increasing returns to scale at the industry level. Therefore, there exist multiple independent peer groups within a sector, each of which takes the size of the industry as well as the prevailing wages as given (see Becker

[9] Note that effective labor in the nerd sector is integrated on $[i^*, 1]$, so that the shift in $\delta$ will be larger for any $i^*$ the more concentrated a given increase in the distribution of $\sigma_N$ is in the upper end of the distribution.

[10] Starting with Arrow (1962) and Ethier (1982a) externalities have often been used to model equilibria in which firms compete against each other in the presence of increasing returns to scale. Romer (1986, 1987), Lucas (1988), and Prescott and Boyd (1987), for instance, demonstrate the importance of knowledge accumulation and specialization for economic growth. Murphy et al. (1989) argue that increasing retruns in the form of aggregate demand spill-overs might explain why some countries appear to be stuck in a unindustrialized equilibrium. Krugman (1991) shows how increasing returns to scale shape economic geography. In the literature on international trade external increasing returns to scale have, for instance, been used to explain the pattern of trade between developed countries (see Helpman 1984 and Krugman 1995 for reviews). For a model of trade in which returns to scale are internal to firms see Krugman (1979).

and Murphy 1992 on the trade-off between returns to scale due to specialization and coordination costs).

Assuming that the industry production function is convex in labor adds one additional wrinkle. Since peer groups do not take the positive externalities of their hiring decisions into account, the marginal product of labor is smaller on the group than on the industry level. Thus, in equilibrium, wages cannot equal the marginal product of labor for the whole industry. Instead, under free entry into the market for peers, a zero-profit condition determines equilibrium wages. As in the case of concave production, we assume that all industry output goes to the agents in the particular sector.[11]

Let profits in sector $j$ be given by

$$\pi_j = F_j\left(L_j, K_j, A_j\right) - w_j L_j. \tag{6}$$

Assuming $F_j\left(0, K_j, A_j\right) = 0$ for all $K_j$ and $A_j$, equation (6) directly implies a wage per efficiency unit equal to

$$w_j = \frac{\int_0^{L_j} \frac{\partial}{\partial L} F_j\left(s, K_j, A_j\right) ds}{L_j} \equiv \overline{\frac{\partial}{\partial L} F_j\left(L_j, K_j, A_j\right)}$$

where $\overline{\frac{\partial}{\partial L_j} F_j\left(L_j, K_j, A_j\right)}$ denotes the the average marginal product of labor on the industry level.[12]

Market clearing implies

$$\sigma\left(i^*\right) = \frac{w_T}{w_N} = \frac{\overline{\frac{\partial}{\partial L} F_T(L_T(i^*), K_T, A_T)}}{\overline{\frac{\partial}{\partial L} F_N\left(L_N\left(i^*\right), K_N, A_N\right)}} \equiv \delta\left(i^*\right).$$

Note that $\delta'(i) \geq 0$ for all $i$. In words, the convexity assumption implies that the ratio of marginal products increases as labor flows from the nerd to the troublemaking sector. Thus, under increasing returns to scale in labor, the 'demand' schedule is upward sloping.

Figure 3 depicts equilibrium and illustrates the comparative statics under increasing returns to scale. The key difference, when production is convex, is that there may exist multiple equilibria. Sufficient conditions are readily derived.

**Proposition 1** *There exist at least two equilibria with a positive mass of both nerds and troublemakers if:*

$$(i)\ \delta(0) > \sigma(0), \delta(1) > \sigma(1),\ \text{and } \delta(i) < \sigma(i)\ \text{ for any } i \in (0,1),$$

*or if*

$$(ii)\ \delta(0) < \sigma(0), \delta(1) < \sigma(1), \text{and } \delta(i) > \sigma(i)\ \text{ for any } i \in (0,1).$$

*There exists at least one equilibrium with a positive mass of both nerds and troublemakers, and*

[11] The following derivations do not depend on the absence of profits in equilibrium. An alternative assumption leaving our conclusions essentially unaffected would be that labor is compensated with a fixed share of revenues.

[12] Since firms within a sector take wages as given and compete for labor, wages are also equal to the marginal product of labor on the firm level. In equilibrium the number of firms must adjust such that each of them earns zero profits.

*another one in which all agents become either nerds or troublemakers if*

$$(iii)\ \delta(1) > \sigma(1) \text{ and } \delta(i) < \sigma(i) \text{ for any } i \in (0,1),$$

*or if*

$$(iv)\ \delta(0) < \sigma(0) \text{ and } \delta(i) > \sigma(i) \text{ for any } i \in (0,1).$$

**Proof.** By continuity of $\delta$ and $\sigma$ the proof follows immediately from the Intermediate Value Theorem and by recognizing that if $\delta(1) \geq \sigma(1)$ or $\delta(0) \leq \sigma(0)$ a corner solution may obtain. ■

Figure 3 depicts a scenario in which increasing marginal product yields multiple equlibria – at the origin, $i^*$, and $i^{*\prime}$. Only equilibria in which the 'demand curve' intersects the 'supply curve' from above are locally stable. To see this, consider the adjustment process following a small shock to wages. From the initial equilibrium at $i^*$, a small decrease in relative wages (along the $\delta$-schedule) will lead to labor flowing out of the troublemaking and into the nerd sector, which will cause relative wages to decline further and lead to even more agents switching sectors. The process continues until the market reaches a new equilibrium at the origin. Conversely, a small increase in wages (along the $\delta$-schedule) will lead to labor flowing into the troublemaking sector. This causes relative wages to increase even more, thereby inducing more nerds to become troublemakers until the market reaches equilibrium at $i^{*\prime}$. Similar reasoning shows that the equilibrium at $i^{*\prime}$ is stable. Given the existence of multiple equilibria, our model may rationalize starkly different behaviors of agents in observationally similar markets.[13]

The case of convex production is closely related to models of a 'social multiplier' (Becker and Murphy 2000, Glaeser et al. 2003). In these models, social spillover effects arise because an individual's marginal utility from taking a particular action is assumed to increase in the number of agents in her reference group who behave in the same way. In our model, an agent's productivity increases as others join the same sector, which raises her wage and thereby the net utility gain from choosing this sector. In symbols:

$$\frac{d}{dL_j}\left(w_j\sigma_j(i) - w_{j'}\sigma_{j'}(i)\right) =$$

$$\left[\frac{\partial}{\partial L_j}\left(\overline{\frac{\partial}{\partial L_j}F_j(L_j, K_j, A_j)}\right)\sigma_j(i) - \frac{\partial}{\partial L_{j'}}\left(\overline{\frac{\partial}{\partial L_{j'}}F_{j'}(L_{j'}, K_{j'}, A_{j'})}\right)\frac{dL_{j'}}{dL_j}\sigma_{j'}(i)\right] > 0$$

for $j \neq j'$. That is, the net utility from choosing sector $j$ over $j'$ increases in the amount of labor employed in $j$. Our theory could thus be interpreted as providing alternative micro-foundations for the assumption of increasing marginal utility.[14] Here, however, the ability to sort into social sectors has the potential to obfuscate the complementarity between individual and group behavior.

[13] When the marginal product for one social sector is increasing while the other is decreasing, the relative demand schedule may alternate between sloping upwards and downwards.

[14] In Appendix B, we show our core results hold when introducing a second social activity to a general social multiplier model.

## 2.2 Reinterpreting the Peer Effects Literature Through the Lens of a Roy Model

There is a large literature on peer effects in schools, neighborhoods, and other venues in which individuals interact. Surprisingly, research designs which exploit experimental and quasi-experimental variation often point in conflicting directions with comparable samples. In this section we show that our Roy model of social interactions is flexible enough to reconcile the seemingly disparate evidence. Put differently, we show that a model in which 'peer effects' are due to the systematic sorting of individuals within social markets can produce the same empirical patterns that have traditionally been attributed to models in which peer effects take the form of direct externalities.

We divide the empirical literature on peer effects into four mutually exclusive categories: analyses that report no significant peer effects, effects which are linear and positive (i.e. smarter peers increase achievement), effects that are non-linear and positive, and analyses that find negative peer effects.

In what follows we assume that production functions are concave in labor inputs and provide sufficient conditions for our model to reconcile the findings of various studies.[15] We do not attempt to explain every nuance in the empirical literature on peer effects. For sure, there exist several competing models all of which can explain some aspect in isolation. Our goal is to develop a tractable model which can reconcile broad but seemingly disparate findings.

### A. No Peer Effects

One strand of the literature argues that peer effects are negligible (Angrist and Lang 2004, Cullen et al. 2006, Evans et al. 1992, Lefgren 2004, Lyle 2007, Stinebrickner and Stinebrickner 2006). Angrist and Lang (2004) evaluate Boston's Metco program, which buses minority students from high poverty neighborhoods in Boston to wealthier suburban schools. Their results indicate that, although the new Metco students are on average lower achieving, the change in peer group induced by these students does not affect test scores of elementary and middle school students in the suburban schools. Cullen et al. (2005) analyze roughly fifteen thousand students who applied to nineteen schools through the Chicago Public Schools choice program. Using data from lotteries, their results imply that the academic impact of attending a new school with higher-performing peers is negligible.

Using solely the lens and language of our model, this implies that the marginal student, $i^*$, in these settings remains the same notwithstanding a change in peer group composition. Hence, the choice of sector for all other individuals does not change as well. To see why this might be the case in practice, consider Boston's Metco program (cf. Angrist and Lang 2004). A potential explanation for the lack of peer effects in this study is that Metco students are few relative to non-Metco ones. Despite the fact that average ability declines and both the 'supply' and 'demand'

[15] It is straightforward to conduct an identical analysis under the assumption of convex production functions, which we leave to the reader.

curves shift downward, the impact of Metco students on relative wages is likely small. As the relative ability of their non-Metco peers remains the same, our model predicts that almost none of them change sectors, leading to negligible peer effects.

Figure 4 illustrates this point. Imagine an increase in the number of Metco students, which shifts the supply curve downward (from $\sigma$ to $\sigma'$) and the demand side equilibrium schedule from $\delta$ inward to $\delta'$. Notice, large shifts only occur to the left of the initial equilibrium. If this is indeed what happened, then it is not surprising that the marginal student would remain virtually the same.

### B. Linear in Means

Another portion of the peer effects literature shows that peer effects operate linearly, typically based on mean group characteristics. Hanushek et al. (2003) show, in a large matched panel data set of third through sixth graders in Texas public schools, that a one standard deviation increase in mean peer test score results in a .20 standard deviation increase in own test scores. Hoxby (2000) uses year to year variation in class-level gender and race composition; finding effects that range from .15 to .40 points for every one point increase in the class mean reading score.[16]

Through the lens of our model, this implies that a constant fraction of individuals shift sectors for every one unit increase in peers' mean test score. Figure 5 shows an example in which peer effects would operate linearly. In this example, the supply schedule (i.e. the distribution of relative ability) shifts almost parallel close to the initial equilibrium. Furthermore, the demand curve has constant negative slope around the initial equilibrium and is relatively unresponsive to shifts labor supply. Therefore, changes in peer ability lead to constant changes in relative wages, and a constant fraction of individuals switch sectors—resulting in linear peer effects.[17]

An explanation along these lines may partially explain the results of Hoxby (2000). As Hoxby (2000) identifies peer effects through plausibly random variation in gender and race composition in classrooms, it may be reasonable to assume that, over the relevant range, the ability distribution shifts one-to-one with its mean. Moreover, given the limited variation in cohorts' gender and racial composition, the demand schedule might be approximately linear in a neighborhood around the 'initial equilibrium'. Whether these or equivalent conditions do indeed hold in Hoxby (2000), or in any other analysis which reports linear peer effects, is unknown.

### C. Heterogeneous Positive Effects

A third category of the literature describes positive, but non-linear peer effects. Hoxby and Weingarth (2005), for instance, exploit a desegregation program in Wake County, NC, which produces exogenous changes in classroom peer groups. They find that based on a linear-in-means model

[16] Other contributions in this vein include Boozer and Cacciola (2001), Gaviria and Raphael (2001), Kang (2007), and Goux and Maurin (2007).

[17] We emphasize that the conditions we provide in the text are sufficient, but in no way necessary.

a student's test score is expected to increase .25 standard deviations given a 1 standard deviation increase in peers' mean score. However, when they allow their results to differ depending on the decile of peer performance, they find that students benefit more from peers with an achievement level similar to theirs. For example, students in the bottom decile benefit most from the addition of students in the second and third deciles (a 10% increase in peers at the 15th percentile increases their performance by .19 standard deviations more than an additional 10% of students in the 8th decile). Carrell et al. (2009) investigate peer effects among freshmen at the Air Force Academy who are randomly assigned to squadrons. They show that a one standard deviation increase in peers' average verbal SAT score results in a .565 standard deviation increase in freshman fall GPA for students in the bottom third of the expected achievement distribution compared to .361 and .312 for those in the middle and top third. Using Census data, Crane (1991) shows that the fraction of high-status workers in a neighborhood is negatively related to the likelihood of teen pregnancy and dropping out of school. These effects become much stronger at the lowest levels of high-status workers.[18]

Figure 6 considers a scenario consistent with the results of Crane (1991) under the auspices of our model. The left panel depicts the situation in a neighborhood with a large number of high achievers (e.g., nerds in the language of the previous section). An increase in the presence of highly skilled individuals (shifting $\sigma$ to $\sigma'$) marginally decreases $i^*$. The right panel features an identical inward shift of the supply curve, but a substantially larger increase in the nerd sector. The heterogenous effects for a given change in supply are due to the confluence of a concave relative demand curve, so that demand is more 'elastic' in Panel B, and that the initial size of the nerd sector is smaller. Thus the shift in relative supply is actually larger in the neighborhood of the marginal $i^*$. Taken together, these structural differences in market for peers yields markedly different behavioral responses to an identical change in skill composition.

### D. Heterogeneous Negative Effects

In stark contrast to the previously mentioned studies, a nascent literature provides credible evidence that own achievement might decline in peer quality. Lavy et al. (2009) use a sample of over a million students taking British age-14 tests to examine peer effects in English high schools. Exploiting the fact that students in their sample enter high school, and thus encounter a peer group that is 87% new on average, they demonstrate that peer effects are different for boys and girls. Girls are positively affected by peers in the top 5% (.07 standard deviation for a 10% increase) while boys are negatively affected (-.05 standard deviations). For boys, the negative effects are strongest among those at the top of the achievement distribution.

Based on the non-linear results in Carrell et al. (2009), Carrell et al. (2010) implemented an experiment at the US Air Force Academy aimed at increasing the GPA of incoming freshmen who were predicted to fall in the bottom tercile of the achievement distribution. To achieve this

[18] Similar non-linear peer effects are found in Burke and Sass (2008), Cooley (2010), Ding and Lehrer (2007), Duflo et al. (2008), Figlio (2007), Imberman et al. (2009), Zimmer and Toma (2000), and Zimmerman (2003).

goal squadrons in the treatment group were negatively sorted, while the composition of those in the control group continued to be random. Yet, the experiment did not have the intended effect. Students in the treatment group projected to fall in the bottom tercile of the distribution, i.e. those students the experiment was designed to help, experienced a .054 point *decline* in GPA compared to their counterparts in the control group.

One possible explanation for these perplexing results is illustrated in Figure 7. In Panel A, we consider a candidate relative supply distribution $\sigma'$ when a squadron has a high fraction of high achieving cadets due to random assignment. Compared to the marginal distribution of the cohort in bold lines, the higher fraction of cadets with high scores is due random assignment comes from a reduction in the fraction of students from both the middle *and bottom* of the distribution. The resulting equilibrium consists of a higher fraction of nerds in the squadron. The scores of those at the bottom of the distribution improves due to either some actually associating with nerds, or due to the fact that that their fellow troublemakers are more academically inclined. When squadrons are arranged according to negative sorting, on the other hand, there is an increase in the fraction of students from *both* the top and bottom tercile, as the distribution becomes $S$-shaped when the middle tercile is removed. This mean-preserving spread of the distribution has relative demands at the boundaries equal to the marginal distribution, but is much flatter in the interior. Instead of exposing the more academically disadvantaged peers to talented study partners, the intra-squadron social dynamics result in greater isolation from them.

Another striking example of peer effects in this category comes from the Moving to Opportunity (MTO) experiment which provided housing vouchers for families in high poverty neighborhoods in Baltimore, Chicago, Los Angeles, and New York City to relocate to lower poverty neighborhoods (Kling et al. 2005, Kling et al. 2007). Evaluations of MTO show that female youth were affected positively by living in an ostensibly better neighborhood. Relative to the control group, female youth are 6.9% less likely to have ever had anxiety symptoms, are 9.1% less likely to have consumed alcohol during the past month, and have .08 fewer lifetime arrests for violent crimes. In contrast, male youth are affected negatively. Relative to the control group, male youth are 8.7% more likely to have had serious nonsports accidents, are 10.3% more likely to have smoked during the past month, and have .15 more lifetime arrests for property crimes.

Our model would predict the findings in MTO if *relative* wages increased (compared to the old neighborhood) for boys, but decreased for girls. Figure 8 depicts shifts of the supply and demand curves which could produce such a result. The top two panels refer to boys and the bottom two to girls. The panels on the left illustrate the conditions in the pre-treatment neighborhood and the panels on the right demonstrate an equilibrium in the treatment neighborhood. The set of students who switch sectors is bounded by $i^*$, the marginal individual in the old neighborhood, and by $\widetilde{i}$, the counterfactual marginal individual given the relative wages in the new environment. In moving to the new neighborhood there are two opposing effects on the relative demand schedule. The nerdier population tends to shift the $\delta$-schedule outwards, while at the same time greater educational resources and supervision in the new neighborhood increase the relative returns to studying, which would shift the $\delta$-schedule inwards. In contrast to models that predict a positive influence from the new environment, here the direction of the treatment effect is determined by the net effect on social wages.

In Figure 8, the inward shift in supply by going from one environment to the other is the same for boys and girls. That is, boys and girls in the experimental group face the same set of smarter peers after they move. If, however, the relative demand curve for girls shifts sufficiently far inward to overcome the outward shift in relative supply, then relative equilibrium wages move in opposite directions. A smaller shift in the demand curve for boys could be caused by multiple factors. If the nerd production function for girls is less concave with respect to labor, they will be less likely to be competed out of the group in an environment with greater effective supply. Conversely, if the troublemaking production function for boys is more concave with respect to labor, the very force that kept them from making trouble in their old neighborhood will raise the relative return to doing so in the new neighborhood where such skills are scarce.

### 2.3 Predicting the Efficacy of Social Interventions in the Presence of Comparative Advantage

The fundamental problem our model highlights is that the treatment effect of a social environment is an endogenous process that is determined *after* any potentially random assignment to a neighborhood, classroom, etc. Without observing social wages *ex ante*, there is no sure way of determining which of the conditions outlined above will hold before a program commences. Ultimately, an intervention's effect depends on production technologies, group specific capital, skill distributions, and the resulting market clearing prices, all of which are generally unobserved. Thus, it may seem that our theory has no ex ante predictions. This is only partially correct. Our model suggests a heuristic that can potentially help policy makers predict outcomes of small scale interventions which do not change equilibrium prices.

Using the lens of comparative advantage, if a policy maker is interested in predicting the behavior of a child after moving to a new neighborhood, a new school, or new classroom, then the relevant statistic is the behavior of children with the same characteristics in the new environment. The reason is simple: children with similar characteristics who face the same social wages will likely have a common comparative advantage and can be expected to behave similarly.

The challenge is to find a way to compare agents across markets. Let $\Theta_j$ denote the set of individual characteristics which determine sectoral choice with intervention $j$. This may include, for example, test scores or innate ability in a school intervention, or height, weight, and motivation in a neighborhood intervention. If one can identify $\Theta_j$ before an intervention commences, then students can be matched across social markets and the heuristic is straightforward.

Consider a few thought experiments. If $\Theta_j$ is test scores, then one can compare individuals across cities on their scores. If $\Theta_j$ is innate ability, methods developed in Hansen et al. (2004) to extract measures of ability can be used to match individuals with the same ability across markets. If $\Theta_j$ involves non-cognitive skills such as those psychologists often refer to as "The Big Five"—

Openness, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability (e.g., Digman 1990)—one can develop pre-intervention surveys along these dimensions and match students on these five measures. Difficulties, however, arise when we have no theory or empirical evidence to inform $\Theta_j$. In this case, one might use administrative or survey data to match on as many variables as possible, recognizing that the prediction will have more noise.

Assuming $\Theta_j$ is in hand, and that we can calculate relative skill, $\sigma(i)$, the prediction from our heuristic is directly related to traditional program evaluations. Let $Y(i)$ be an indicator variable equal to one if individual $i$ chooses to be a nerd (and zero otherwise) in the old environment, and let $Y'(i)$ denote $i$'s choice in the new environment. Then the average treatment effect from manipulating the environment for all $i$ is equal to

$$ATE = \mathbb{E}\left[Y'(i) - Y(i)\right] = i^* - i^{*\prime},$$

where $i^*$ and $i^{*\prime}$ denote the marginal individual in the old and new environment, respectively. In words, the average treatment effect is simply the fraction of individuals who switch sectors.

Interpreting our model more loosely, one could also think of outcomes $(Y_0, Y_1)$ which are different from an agent's actual choice of sector, but nevertheless depend on it. For instance, let $Y_1(i|\Xi)$ denote $i$'s test score (conditional on environmental variables $\Xi$) if $i$ chooses to be a nerd, whereas $Y_0(i|\Xi)$ is her potential outcome as a troublemaker. In this case, the average treatment effect from transplanting a population of unit mass into a new environment characterized by $\Xi'$ is

$$\begin{aligned}ATE = &\int_0^{\min\{i^*(\Xi), i^*(\Xi')\}} \left(Y_0\left(s|\Xi'\right) - Y_0\left(s|\Xi\right)\right) ds \\ &+ 1_{\{i^*(\Xi) < i^*(\Xi')\}} \int_{i^*(\Xi)}^{i^*(\Xi')} \left(Y_0\left(s|\Xi'\right) - Y_1\left(s|\Xi\right)\right) ds + 1_{\{i^*(\Xi) > i^*(\Xi')\}} \int_{i^*(\Xi')}^{i^*(\Xi)} \left(Y_1\left(s|\Xi'\right) - Y_0\left(s|\Xi\right)\right) ds \\ &+ \int_{\max\{i^*(\Xi), i^*(\Xi')\}}^{1} \left(Y_1\left(i|\Xi'\right) - Y_1\left(i|\Xi\right)\right) ds,\end{aligned}$$

where $i^*(\cdot)$ denotes the marginal individual in a given environment, and $1_{\{.\}}$ is an indicator function equal to one if the condition in braces is satisfied. The first and last row in the equation above give the change in test scores for those individuals who do not switch sectors, whereas the middle row denotes the change in the outcome for those agents who do switch sectors (e.g., for nerds who become troublemakers or vice versa). Even if changing the environment from $\Xi$ to $\Xi'$ is beneficial in the sense that it raises both $Y_0$ and $Y_1$ for every individual, the average treatment effect could still be negative if the difference between $Y_0$ and $Y_1$ (conditional on the environment) is sufficiently large compared to the effect of environmental variables.

## 2.4 An Extension to the Basic Model

In this subsection, we outline an extension to the basic model presented above that recognizes the tradeoff that may exist between social activities and the ancillary costs or benefits of associating with a social group. In particular, students may value good grades for their own sake, or for the higher future standard of living that accompanies graduation. This raises the opportunity cost of associating with troublemakers, and encourages association with nerds even in the presence of low social wages.

To fix ideas, one can think of $\theta_j$ as the utility-scaled effect that association with sector $j \in \{N, T\}$ has on educational achievement or expected future income. To make the analysis tractable, we assume linearity as well as additive separability in the utility derived from social income and that incurred from $\theta_j$. More specifically, agent $i$'s utility is given by

$$U(i) = \max_{j \in \{N,T\}} \{w_j \sigma_j(i) + \theta_j\}. \tag{7}$$

It follows straightforwardly that agent $i$ will choose to become a troublemaker if and only if

$$w_T \sigma_T(i) - w_N \sigma_N(i) \geq \theta_N - \theta_T.$$

In words, for $i$ to join forces with the troublemakers it must be the case that the psychic benefit from joining, $(w_T \sigma_T(i) - w_N \sigma_N(i))$, outweighs the utility loss due to lower achievement, $(\theta_N - \theta_T)$.[19]

More generally, however, equilibrium is determined by the condition:

$$w_T(i^*) \sigma_T(i^*) - w_N(i^*) \sigma_N(i^*) = \theta_N - \theta_T, \tag{8}$$

where $w_j(i^*)$ denotes the equilibrium wage in sector $j$. From our basic model we know that, given social wages, the net psychic benefit from being a troublemaker, i.e. the left hand side of equation (8), is decreasing in an individual's index, whereas the right handside is constant from individual agents' point of view. Hence, as before, individuals with index $i \leq i^*$ become troublemakers, and those for whom $i > i^*$ choose the nerd sector. That is, equilibrium labor supply continues to be given by equations (2) and (3), and equilibrium wages will be determined as described above.

The bottom line is that the intuition of our basic model – individuals choose sectors based on their social income and this type of sorting can explain much of the evidence on peer effects without direct externalities – holds whenever $\theta_N(i^*) - \theta_T(i^*)$ is sufficiently small, or, more loosely, the complimentarties are not "too strong." However, if individuals are sufficiently forward looking and take into account important outcomes such as their future income in their current decision making, our model of sorting cannot explain peer effects. Thus, even in a world with important complementarities, our model may be applicable to subgroups known to have high discount rates (e.g., poor minorities, ) or age-groups that are sufficiently myopic.

[19] Note that if $\theta_N(i^*) \approx \theta_T(i^*)$, then agents will base their choice of sector (almost) exclusively on differences in social income, resembling our basic model. If, for instance, achievement is primarily determined by ability, or playing video games and bullying equally distract from studying, then it might very well be the case that $\theta_N$ and $\theta_T$ are of similar maginitude.

# 3 Empirical Implications

## 3.1 The Empirical Content of a Roy Model of Social Interactions

In the previous section we have outlined a new way of thinking about social interactions using sorting and comparative advantage as the guiding principle of peer group organization. In this section, we consider the implications of the comparative advantage approach for the identification of peer effects. In doing so, we follow the literature and assume that there might exist other factors besides *social* payoffs that determine the utility of being a troublemaker or a nerd, e.g., personal and neighborhood characteristics, or the behavior of one's peers.

To fix ideas, consider a student's choice of becoming a troublemaker, $T$, or a nerd, $N$. Let $\mathbf{X}_i$ be a set of individual level covariates, and let $\mathbf{Z}_m$ denote factors varying only at the school, neighborhood, or market level. Mean behavior in market $m$ is given by $\overline{y}_m$, and $\nu_{im}$ represents an error term known to the individual, but not the econometrician. Intuitively, $\nu_{im}$ captures all unobserved factors influencing the difference in utility between $T$ and $N$. Student $i$ chooses to become a troublemaker if and only if

$$u(T;\mathbf{X}_i,\mathbf{Z}_m) - u(N;\mathbf{X}_i,\mathbf{Z}_m) = \kappa + \mathbf{X}_i'\boldsymbol{\beta}_0 + \mathbf{Z}_m'\boldsymbol{\gamma}_0 + \alpha_0\overline{y}_m + \nu_{im} \geq 0. \tag{9}$$

Social wages and individual ability are typically not directly observable. Thus, the comparative advantage approach can be viewed as providing a more explicit theory of the error term. Following our theoretical model, decompose $\nu_{im}$ into: the net market payoff from being a troublemaker and some other random variable:

$$\nu_{im} = (w_{Tm}\sigma_{Ti} - w_{Nm}\sigma_{Ni}) + \epsilon_i. \tag{10}$$

Note that only $\epsilon_i$ and $\sigma_{ji}$, $j \in \{N, T\}$, are possibly independent and identically distributed across individuals, whereas $w_{Tm}$ and $w_{Nm}$ (both of which are measured in utility units) vary only at the market or group level. Therefore, our theory stipulates the existence of group level unobservables.

Recall, in the presence of group level unobservables not all parameters in the binary choice model are identified from cross-sectional data (Blume et al. 2010, Brock and Durlauf 2007). While $\boldsymbol{\beta}_0$ can be consistently estimated without imposing parametric assumptions (using methods outlined in Heckman 1990), $\boldsymbol{\gamma}_0$ and $\alpha_0$—the coefficients of interest in the majority of applied work—cannot. Nonidentification is due to the fact that $\nu_{im}$ depends on $\mathbf{Z}_m$ and $\overline{y}_m$ in an unknown way. Therefore, only the linear combination of market level observables and unobservables is identified (see Brock and Durlauf 2007 for a formal argument).

It is important to note that nonidentification as a result of group level unobservables is quite distinct from endogeneity due to systematic sorting of individuals into social markets (such as neighborhoods and classrooms), or the reflection problem, which poses that $\alpha_0$ cannot be identified if $\mathbf{Z}_m$ and $\overline{y}_m$ are linearly dependent (Manski 1993).[20] While applied researchers have often found

[20] If $\mathbf{Z}_m$ exhibits sufficient variation and $\gamma \neq 0$, then the binary choice model of social interactions does not suffer

clever strategies to deal with these two problems, group level unobservables have received much less attention. However, there are several notable exceptions. Cooley (2010) motivates her instrument in the presence of unobserved differences in teacher quality. Hoxby (2000) uses panel data to remove the effect of group level unobservables which do not vary over time; and Graham (2008) shows how conditional variance restrictions can be used to identify endogenous peer effects when individual and group level unobservables are uncorrelated.

To appreciate the consequences of group level unobservables, denote $i$'s observed behavior by

$$y_i = \begin{cases} 1 & \text{if } u(T; \mathbf{X}_i, \mathbf{Z}_m) - u(N; \mathbf{X}_i, \mathbf{Z}_m) \geq 0 \\ 0 & \text{otherwise} \end{cases},$$

and consider the case in which there are no endogenous peer effects, i.e. $\alpha_0 = 0$. By the Frisch-Waugh Theorem and assuming that $\mathrm{Cov}(\mathbf{X}_i^*, \nu_{im}) = 0$, the probability limit of the ordinary least squares estimator of $\alpha_0$ from regressing $y_i$ on $\mathbf{X}_i$, $\mathbf{Z}_m$, and $\overline{y}_m$ equals

$$\mathrm{plim}\, \hat{\alpha}_{OLS} = \alpha_0 + \frac{\mathrm{Cov}(\overline{y}_m^*, \nu_{im})}{\mathrm{Var}(\overline{y}_m^*)} = \frac{\mathrm{Cov}(\overline{y}_m^*, \nu_{im})}{\mathrm{Var}(\overline{y}_m^*)},$$

where $\overline{y}_m^*$ denotes the residual from projecting $\overline{y}_m$ onto $\mathbf{X}_i$ and $\mathbf{Z}_m$. Only if $\overline{y}_m^*$ and $\nu_{im}$ are uncorrelated, will $\hat{\alpha}_{OLS}$ be consistent.

To see that under the assumptions of our model $\mathrm{Cov}(\overline{y}_m^*, \nu_{im}) > 0$, condition on $\mathbf{X}_i$ and $\mathbf{Z}_m$ and note that, according to (9), individual $i$ in social market $m$ chooses to become a troublemaker if and only if

$$\nu_{im} \geq \xi(\mathbf{X}_i, \mathbf{Z}_m) \tag{11}$$

where $\xi(\mathbf{X}_i, \mathbf{Z}_m) \equiv -(\kappa + \mathbf{X}_i' \boldsymbol{\beta}_0 + \mathbf{Z}_m' \boldsymbol{\gamma}_0)$. Now, decompose $\nu_{im}$ into the market specific mean social payoff, $\overline{v}_m$, and deviations around the mean, $\widetilde{v}_{im}$, which are distributed according to some cumulative distribution function $\Phi_m(\cdot)$. That is, let $v_{im} = \overline{v}_m + \widetilde{v}_{im}$. With this notation in hand, $y_i = 1$ if and only if

$$\widetilde{v}_{im} \geq \xi(\mathbf{X}_i, \mathbf{Z}_m) - \overline{v}_m,$$

and the fraction of individuals who are troublemakers in market $m$ is equal to

$$\overline{y}_m = \mathbb{E}_{\mathbf{X}_i}\left[1 - \Phi_m\left(\xi(\mathbf{X}_i, \mathbf{Z}_m) - \overline{v}_m\right)\right].$$

Unless $\mathbf{X}_i$ and $\mathbf{Z}_m$ fully determine $\overline{v}_m$ (in which case there is no role for social) it will be the case that $\frac{d\overline{y}_m^*}{d\overline{v}_m} > 0$, as $\frac{d\overline{y}_m}{d\overline{v}_m} > 0$. With this caveat in mind, it follows that $\mathrm{Cov}(\overline{y}_m^*, \nu_{im}) > 0$.

The intuition for this result is straightforward. Under the assumptions of our model, a particular behavior will be more prevalent in markets in which the social net payoff to it is higher. It follows that although endogenous social interactions might not be a driver of behavior, i.e. $\alpha_0 = 0$, linear-in-means estimates will be biased toward finding this form of peer effects—even under random

from the reflection problem, as the limited range of the outcome rules out perfect linear dependence (Brock and Durlauf 2007).

assignment to social markets and if one resolves the reflection problem.

It is not the case, however, that all estimates of peer effects are upward biased. Suppose, for instance, that $\overline{y}_m$ denotes mean test scores in social market $m$. As shown in Section 2, one cannot sign the change in relative wages as the distribution of ability changes. Even if mean test scores were a sufficient statistics for the whole ability distribution, we would not be able to determine the covariance between $\overline{y}_m^*$ and $\nu_{im}$. It is still the case, however, that (due to the presence of group level unobservables) $\alpha_0$ and $\boldsymbol{\gamma}_0$ are not identified.

## 3.2 Evidence Consistent with a Roy Model Approach to Social Interactions

Lack of identification does not imply that the comparative advantage approach has no empirical content. As the net market payoff to troublemakers is a declining function of nerd ability, even purely ordinal information, such as a ranking of individuals, may be useful. For *any* set of equilibrium prices, individuals with low cognitive ability are more likely to become troublemakers than ones with high nerd ability. If comparative advantage shapes social interactions, then individual behavior should depend on rank relative to others *within* the same market, and changes in rank induced by moves across markets should be systematically related to changes in behavior. Hence, our model predicts that, all else equal, an individual's behavior should be related to her rank within the relevant social market. For instance, children whose rank declines in transitioning from elementary to middle school should be more likely to develop behavioral problems than those whose rank increases.

The ideal data to test our theory would span multiple markets—say, schools or classrooms—and contain information on social wages, agents' choices of sector as well as all of their skills. With such data in hand we could test *directly* whether comparative advantage determines behavior by comparing potential 'social earnings' across sectors and relating them to agents' choices. Alternatively, data on only a subset of skills, social wages and individuals' choices of sector would allow us to follow Heckman and Sedlacek (1985), who combine information on individual characteristics, wages, sectoral choices, and aggregate wage bills to estimate a structural model of self-selection in the labor market, as well as the demand for observed and unobserved skill. We are unaware of such data.

In the absence of any information on social wages, and in lieu of imposing restrictive assumptions to ensure identification, we pursue the more modest goal of providing reduced form evidence which suggests that, within a market, individual behavior depends on one's rank. We leave the important question surrounding identification in the presence of group level unobservables for another occasion. The evidence we present below is purely suggestive, as there exist several competing models that share the same prediction. A trivial example is a model in which the intrinsic utility from being a nerd depends on one's class rank. Therefore, we urge caution when evaluating the empirical evidence in favor of the comparative advantage approach.

In what follows, we investigate the relationship between a student's relative academic ranking and behavioral outcomes in two large data sets: New York City Public Schools (NYCPS) administrative data from 2003/04 through 2008/09, and the National Education Longitudinal Study of 1988 (NELS). Since our theory predicts that this relationship can be non-linear we estimate semi-parametric specifications, as described in Yatchew (1998).

### A. Evidence from New York City Public Schools

The New York City Public Schools (NYCPS) data contain student-level administrative information on approximately 1.1 million students across the five boroughs of the NYC metropolitan area. The data include student race, gender, free and reduced-price lunch eligibility, behavior, attendance, and matriculation with course grades for all students, as well as state math and English/Language Arts (ELA) test scores for students in grades three through eight. We have NYCPS data spanning the 2003/04 to 2008/09 school years. Summary statistics for the variables we use in our core specifications are displayed in Appendix Table 1.

Using the NYCPS data, we estimate models of the form

$$\Delta y_i = f(\Delta x_i) + \mathbf{X}_i'\beta + School_i + Year_i + \epsilon_i, \tag{12}$$

restricting our attention to the set of students who change schools in the transition from elementary to middle school. Our behavioral measure, $y_i$, in each year is an indicator equal to one if a student has at least one reported behavioral incident from that year and zero otherwise; $\Delta y_i \in \{-1, 0, 1\}$. The three most common behavioral incidents in our data are "engaging in an altercation or physically aggressive behavior with other student(s)," "behaving in a manner that disrupts the educational process (horseplay)," or "engaging in verbally rude or disrespectful behavior / insubordination."

A student's rank in fifth grade is the student's percentile ranking based on achievement on the New York State exam relative to other students who are in the same school in fifth grade.[21] We also compute each student's position relative to peers in her sixth grade school using the *fifth* grade test scores. This captures the student's ranking in the new school at the beginning of the school year; $\Delta x_i$ denotes the difference between these two rankings. We report results using both math and ELA scores to compute the change in percentile. Finally, we include school fixed effects (for both a student's elementary and middle school), year fixed effects, and a standard set of covariates that includes the test score in the same subject from the previous year, an exhaustive set of race dummies, sex, free lunch eligibility, English Language Learner (ELL) status, and special educa-

[21] The state math and ELA tests, developed by McGraw-Hill, are high-stakes exams conducted in the winters of third through eighth grade. Students in third, fifth, and seventh grades must score level 2 or above (out of 4) on both tests to advance to the next grade without attending summer school. The math test includes questions on number sense and operations, algebra, geometry, measurement, and statistics. Tests in the earlier grades emphasize more basic content such as number sense and operations, while later tests focus on advanced topics such as algebra and geometry. The ELA test is designed to assess students on three learning standards—information and understanding, literary response and expression, critical analysis and evaluation—and includes multiple-choice and short-response sections based on a reading and listening section, along with a brief editing task. Content breakdown by grade and additional exam information is currently available at <http://www.emsc.nysed.gov/osa/pub/reports.shtml>.

tion designation. By including these covariates we attempt to control for factors which plausibly influence changes in behavior and might be correlated with rank.

Our estimates of the link between changes in rank and changes in behavior are displayed in Figure 9. Independent of whether we calculate rank based on ELA or math scores, the behavior of students whose rank decreases in going from elementary to middle school worsens significantly compared to students whose relative standing improves. A student experiencing a 50 percentile decline in rank is approximately five percentage points more likely to have a behavioral incident on record than a student whose rank improves by 50 percentiles—with the estimated effect being slightly larger if we calculate rank based on math scores than if we do so based on ELA scores. Given sample means (and standard deviations) of .087 (.282) for sixth grade and .049 (.215) in fifth grade, our estimates are non-trivial in size.

Although the NYCPS data allow us control for a students' natural proclivities to cause trouble by relating changes in behavior to changes in rank induced by the transition to middle school, there exists the possibility that our results are driven by systematic school choice. That is, students who chose an academically less challenging environment might have experienced less of an increase in behavioral problems, even if their rank had not improved.

To address the concern of systematic sorting into schools, we instrument for a student's change in rank with their predicted change in rank based on the school they were zoned to attend (given their residential address).[22] More specifically, we estimate two-stage least squares (2SLS) specifications corresponding to:

$$\Delta y_i = \alpha \Delta x_i + \mathbf{X}_i'\beta + School_i + Year_i + \epsilon_i \,, \tag{13}$$

where the first stage is given by

$$\Delta x_i = \delta \widehat{\Delta x_i} + \mathbf{X}_i'\gamma + School_i + Year_i + \nu_i,$$

and $\widehat{\Delta x_i}$ denotes student $i$'s counterfactual change in rank at the beginning of sixth grade had all students attended the schools for which they were zoned. More specifically, let $x_{i,t-1}$ denote student $i$'s test score in fifth grade and let $\text{rank}_I\left(x_{i,t-1}\right)$ be the percentile ranking of a student with score $x_{i,t-1}$ among the set of students $I$, given their respective test scores at $t-1$. Then,

$$\widehat{\Delta x_i} \equiv \underset{S_{i,t}}{\text{rank}}\left(x_{i,t-1}\right) - \underset{S_{i,t-1}}{\text{rank}}\left(x_{i,t-1}\right),$$

where $S_{i,t-1}$ and $S_{i,t}$ are the sets of students who are zoned for the same elementary and middle school as $i$, respectively.

Table 1 presents 2SLS estimates of the partial correlation between school rank and behavior, as well as the corresponding OLS estimates for comparison. In panel A we use ELA scores to construct rank, whereas math scores are used in panel B. Based on the OLS point estimates one

[22] For more information on New York's school choice plan, see <http://schools.nyc.gov/ChoicesEnrollment/default.htm>.

would expect a student experiencing a 50 percentile decline in rank to be circa 3.5 (panel A) or 5–6 (panel B) percentage points more likely to have a behavioral incident on record than a student whose rank improves by 50 percentiles—consistent with our previous semi-parametric results.

Due to the large number of observations, our OLS estimates are very precise. Unfortunately, this is not the case when we estimate equation (13) by 2SLS. Although the first stage F-statistic is well above conventional critical values (Stock and Yogo 2005), our instrument explains very little residual variation in the excluded variable, as evidenced by small values of Shea's $R^2$ (Shea 1997). One potential explanation for this is that only 46.0% (53.6%) of students attend the middle (elementary) school for which they are zoned.

Nevertheless, not including school fixed effects, the 2SLS estimates are remarkably similar to their OLS counterparts; and statistically significant at 1%-level. If, however, we include school fixed effects, the point estimates shrink by more than half and are statistically indistinguishable from zero, though still negative. Given that our instrument, $\widehat{\Delta x_i}$, explains less than one percent of the residual variation in $\Delta x_i$ when school fixed effects are included, we strongly urge caution when interpreting the results presented in Table 1. In order to better account for the possibility of bias due systematic sorting of students into schools, we turn the National Education Longitudinal Study.

### B. Evidence From the National Educational Longitudinal Study

The National Education Longitudinal Study of 1988 (NELS) was initiated in 1988 with a cohort of 24,599 eighth graders, who were then resurveyed through four follow-ups in 1990, 1992, 1994, and 2000. The available information on these students covers a wide range of topics including: school, work, and home experiences; educational resources and support; the role in education of their parents and peers; neighborhood characteristics; educational and occupational aspirations; as well as other student perceptions. For the first three waves, students also completed achievement tests in reading, social studies, mathematics and science. In addition to collecting information on students' course work and grades in high school as well as postsecondary transcripts, their teachers, parents, and school administrators were also surveyed. Appendix Table 2 displays summary statistics for all variables we use in our analysis.

We examine NELS data from 1988 and 1990, when students were in eighth and tenth grade. An important limitation of the NELS data is that only 25 students per school were surveyed, yielding a noisy measure of rank. To lessen the impact of measurement error, we limit our sample to students in classrooms with at least five observations.[23] Yet, NELS allows us take advantage of the fact that the data include teacher reports on behavior and student self-reported grades from exactly two subjects in the same year. By using a model that relates changes in a student's behavior across classrooms to changes in her rank we can implicitly account for students' natural tendencies to cause trouble and rule out that systematic sorting into schools drives our results.[24] More specifically, we

[23] We obtain qualitatively identical results for alternative threshold levels of ten and zero.

[24] Students strategically sorting into classrooms based on other unobservables is still a potential concern.

estimate a model of the form

$$\Delta y_i = f(\Delta x_i) + \mathbf{X}_i'\beta + Grade_i + \epsilon_i, \tag{14}$$

where $y_i$ is an indicator for whether the teacher reported that the student had any behavioral problems, and $\Delta y_i$ refers to the difference in this indicator across subjects within the same year. Teachers were asked whether the student had a problem in any of six different categories: the student performed below his ability, the student did not complete homework, the student was frequently absent, the student was frequently tardy, the student was inattentive, or the student was disruptive. Our indicator variable is equal to one if the teacher reported that the student had at least one of these behavioral problems. Note: the NELS measure of behavioral problems encompasses a far more benign set of 'offenses' than those typically reported in NYCPS. We use student self-reported grades to compute subject-specific rank $x_i$, and let $\Delta x_i$ denote the difference in these ranks across subjects within the same year. Moreover, $\mathbf{X}_i$ includes: the mean score across subjects from the same year and its square, race, sex, English Language Learner status, indicator variables for parents' marital status, indicator variables for parents' education, indicator variables for school type (public, Catholic, or other private), indicator variables for school location (urban, suburban, or rural), indicator variables for socioeconomic status quartiles, birth year indicators, and birth month indicators; $Grade_i$ marks a grade level fixed effect.

As was the case in the NYCPS data, we find that changes in a student's rank within a social market are related to changes in her behavior (see Figure 10). For instance, students whose rank is 50 percentiles lower in English class than in Math class are estimated to be approximately ten percentage points more likely to act out in the former than the latter. Taken at face value rank appears to have a substantial influence on behavior.[25]

Broadly summarizing, the results presented in this section suggest that students' behavior deteriorates as their academic rank declines. The empirical evidence is thus consistent with a comparative advantage approach to social interactions.

# 4 Concluding Remarks

While social scientists have long been concerned with the influence of social interactions, empirical estimates of peer effects vary widely in the literature. Using different sources of (plausibly) exogenous variation some studies find negligible effects, others find effects that are positive and linear in mean peer characteristics, and occassionally peer effects have been found to be negative. We develop a Roy model of social interactions which, through comparative advantage, has the potential to

[25] In NELS the sample mean (and standard deviation) of our indicator for having at least on behavioral incident is .427 (.495) in math, .402 (.490) in history, .426 (.495) in science, and .429 (.495) in English class. Every student has information on at most two of these subjects. See also Table A.2 and the description of NELS in the Data Appendix. Instead of using an indicator variable for whether the teacher reports any behavioral incidents, we have also constructed a summary index of children's behavior by factor analyzing different teacher reported behaviors. For both outcomes our results are qualitatively identical.

provide a parsimonous explanation for the disparate empirical evidence. In our model 'peer effects' arise endogenously due to the sorting of individuals *within* narrowly defined social settings, such as neighborhoods or classrooms. Consequently, the comparative advantage approach has important implications for the (non)identification of peer effects—even if individuals are randomly assigned to social markets. In addition, since an unobserved wage is the critical determinant of selection, assigning students to an environment with higher academic abilities is no guarantee that they will face higher wages for academic effort. This places severe constraints on the external validity of any randomized trial.

Our data exercise provides suggestive evidence that the key prediction distinguishing our model from traditional approaches is borne out in two datasets. However, it is important to note that other models which predict both positive *and* negative peer effects, might also be able to reconcile our empirical evidence as well as the existing literature.

At its core, our theory builds upon impressive literatures designed to understand the evolution of earnings, the (hedonic) pricing of skills, and the assignment of workers to firms. The novelty in our approach lies in the application of these classic methods to develop a theory of social interactions to characterize equilibria in different social markets. The insights emerging from this approach may also be useful in understanding a variety of other social phenomena, one particular example being identity choice.

## References


[1] Akerlof, George A. "Social Distance and Social Decisions." *Econometrica* 65, no. 5 (1997): 1005-1028.

[2] Angrist, Joshua D., and Kevin Lang. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review* 94, no. 5 (2004): 1613-1634.

[3] Arrow, Kenneth J. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 39 (1962): 155-173.

[4] Austen-Smith, David, and Roland G. Fryer Jr. "An Economic Analysis of 'Acting White'." *Quarterly Journal of Economics* 120, no. 2 (2005): 551-583.

[5] Becker, Gary S. *Human Capital*, 1st ed. New York: Columbia University Press, 1964.

[6] Becker, Gary S. "A Theory of the Allocation of Time." *Economic Journal* 75, (1965): 493-517.

[7] Becker, Gary S. "A Theory of Marriage, Part I." *Journal of Political Economy* 81, no. 4 (1973): 813-846.

[8] Becker, Gary S. "A Theory of Social Interactions." *Journal of Political Economy* 82, no. 6 (1974): 1063-1093.

[9] Becker, Gary S. *Accounting for Tastes*. Cambridge: Harvard University Press, 1996.

[10] Becker, Gary S., and Kevin M. Murphy. "The Division of Labor, Coordination Costs, and Knowledge." *Quarterly Journal of Economics* 107, no. 4 (1992): 1137-1160.

[11] Becker, Gary S., and Kevin M. Murphy. *Social Economics: Market Behavior in a Social Environment*. Cambridge: Belknap Press of Harvard University, 2000.

[12] Benabou, Roland. "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics* 108, no. 3 (1993): 619-652.

[13] Ben-Porath, Yoram. "The Production of Human Capital and the Lifecycle of Earnings." *Journal of Political Economy* 75 (1967): 352-365.

[14] Berman, Eli. "Sect, Subsidy, and Sacrifice: An Economist's View of Ultra-Orthodox Jews." *Quarterly Journal of Economics* 115, no. 3 (2000): 905-953.

[15] Bernheim, B. Douglas. "A Theory of Conformity." *Journal of Political Economy* 102, no. 5 (1994): 841-877.

[16] Blume, Lawrence E., William A. Brock, Steven N. Durlauf, and Yannis M. Ioannides. "Identification of Social Interactions," forthcoming in *Handbook of Social Economics*, edited by J. Benhabib, A. Bisin, and M. Jackson. Amsterdam: Elsevier, 2010.

[17] Boozer, Michael A., and Stephen E. Cacciola. "Inside the 'Black Box' of Project Star: Estimation of Peer Effects Using Experimental Data." Economic Growth Center Discussion Paper No. 832. Yale University. 2001.

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

[19] Brock, William A., and Steven N. Durlauf. "Identification of Binary Choice Models with Social Interactions." *Journal of Econometrics* 140 (2007): 52-75.

[20] Burke, Mary A., and Tim R. Sass "Classroom Peer Effects and Student Achievement." Federal Reserve Bank of Boston Working Paper. 2008.

[21] Carrell, Scott E., Richard L. Fullerton, and James E. West. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement." *Journal of Labor Economics* 27, no. 3 (2009): 439-464.

[22] Carrell, Scott E., Bruce I. Sacerdote, and James E. West. "From Natural Variation to Optimal Policy: A Cautionary Tale in How Not to Improve Student Outcomes." Unpublished Manuscript. 2010.

[23] Chapple, Simon, Richard Jefferies, and Rita Walker. "Maori Education Participation and Performance: A Literature Review and Research Programme." Paper for the New Zealand Ministry of Education. 1997.

[24] Cooley, Jane C. "Desegregation and the Achievement Gap: Do Diverse Peers Help?" Unpublished manuscript. Department of Economics, University of Wisconsin at Madison. 2010.

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

[26] Cullen, Julie Berry, Brian A. Jacob, and Steven D. Levitt. "The Effect of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools." *Journal of Public Economics* 89 (2005): 729-760.

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

[28] Devos, George and Hiroshi Wagasutma. *Japan's Invisible Race: Caste in Culture and Personality.* Berkeley: University of California Press, 1966.

[29] Digman, J. M. "Personality structure: Emergence of the five-factor model." *Annual Review of Psychology* 41 (1990): 417-440.

[30] Ding, Weili, and Steven F. Lehrer. "Do Peers Affect Student Achievement in China's Secondary Schools?" *Review of Economics & Statistics* 89, no. 2 (2007): 300-312.

[31] DiPasquale, D., and E. L. Glaeser. "The Los Angeles Riot and the Economics of Urban Unrest." *Journal of Urban Economics* 43, no. 1 (1998): 52-78.

[32] Drake, St. Clair, and Horace Cayton. *Black Metropolis: A Study of Negro Life in a Northern City.* Chicago: University of Chicago Press, 1945.

[33] Duflo, Esther, Pascaline Dupas, and Michael Kremer. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." NBER Working Paper No. 14475. 2008.

[34] Durlauf, Steven N. "Neighborhood Effects." In *Handbook of Urban and Regional Economics, Vol. IV, Geography and the Cities*, edited by J. Vernon Handerson and Jacques-Francois Thisse. North-Holland, Amsterdam, 2004.

[35] Ethier, Wilfred J. "National and International Returns to Scale in the Modern Theory of International Trade." *American Economic Review* 72, no. 3 (1982a): 389-405.

[36] Ethier, Wilfred J. "Decreasing Costs in International Trade and Frank Graham's Argument for Protection." *Econometrica* 50, no. 5 (1982b): 1243-1268.

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

[38] Figlio, David N. "Boys Named Sue: Disruptive Children and their Peers." *Education Finance and Policy* 2, no. 4 (2007): 376-394.

[39] Fordham, Signithia, and John Ogbu. "Black Students' School Successes: Coping with the Burden of 'Acting White'." *The Urban Review* 18, no. 3 (1986): 176–206.

[40] Gans, Herbert. *The Urban Villagers: Group and Class in the Life of Italian-Americans.* New York: The Free Press, 1962.

[41] Gaviria, Alejandro, and Steven Raphael. "School-Based Peer Effects and Juvenile Behavior." *Review of Economics & Statistics* 83, no. 2 (2001): 257-268.

[42] Glaeser, Edward L., Bruce Sacerdote, and Jose A. Sheinkman. "The Social Multiplier." *Journal of the European Economic Association* 1 (2003): 345-353.

[43] Gorman, W. "A Possible Procedure for Analysing Quality Differentials in the Egg Market." *Review of Economic Studies* 47, no. 5 (1980): 843-856.

[44] Goux, Dominique, and Eric Maurin. "Close Neighbours Matter: Neighbourhood Effects on Early Performance at School." *The Economic Journal* 117 (2007): 1193-1215.

[45] Graham, B. S. "Identifying Social Interactions Through Conditional Variance Restrictions." *Econometrica* 76, no. 3 (2008): 643-660.

[46] Hansen, Karsten T., James J. Heckman, Kathleen M. Mullen. "The Effect of Schooling and Ability on Achievement Test Scores." *Journal of Econometrics* 121 (2004): 39-98.

[47] Hanushek, Eric. A, et al. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18 (2003): 527-544.

[48] Heckman, James J. "Varieties of Selection Bias." *American Economic Review Papers and Proceedings* 80 (1990): 313-318.

[49] Heckman, James J., and Jose A. Scheinkman. "The Importance of Bundling in a Gorman-Lancaster Model of Earnings." *Review of Economic Studies* 54, no. 2 (1987): 243-255.

[50] Heckman, James J., and Guilherme Sedlacek. "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-selection in the Labor Market." *Journal of Political Economy* 93, no. 6 (1985): 1077-1125.

[51] Helpman, Elhanan. "Increasing Returns, Imperfect Markets, and Trade Theory." In *Handbook of International Economics, Vol. 1*, edited by R. W. Jones and P. B. Kenen. Amsterdam: Elsevier, 1984.

[52] Hoxby, Caroline. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper No. 7867. 2000.

[53] Hoxby, Caroline, and Gretchen Weingarth. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." Unpublished manuscript. 2005.

[54] Iannaccone, Laurence R. "Sacrifice and Stigma: Reducing Free-riding in Cults, Communes, and Other Collectives." *Journal of Political Economy* 100, no. 2 (1992): 271-291.

[55] Imberman, Scott, Adriana D. Kugler, and Bruce Sacerdote. "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees." NBER Working Paper No. 15291. 2009.

[56] Jovanovic, Boyan. "Firm-specific Capital and Turnover." *Journal of Political Economy* 87, no. 6 (1979a): 1246-1260.

[57] Jovanovic, Boyan. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87, no. 5 (1979b): 972-990.

[58] Kang, Changhui. "Classroom peer effects and academic achievement: Quasi-randoization evidence from South Korea." *Journal of Urban Economics* 61, no. 3 (2007): 458-495.

[59] Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics* 120, no. 1 (2005): 87-130.

[60] Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75, no. 1 (2007): 83-119.

[61] Krugman, Paul. "Increasing Returns, Monopolistic Competition, and International Trade." *Journal of International Economics* 9 (1979): 469-479.

[62] Krugman, Paul (1991). "Increasing Returns and Economic Geography." *Journal of Political Economy* 99, no. 3 (1991): 483-499.

[63] Krugman, Paul. "Increasing Returns, Imperfect Competition and the Positive Theory of International Trade." In *Handbook of International Economics* Vol. 3, edited by G. Grossman and K. Rogoff. Amsterdam: Elsevier, 1995.

[64] Lancaster, Kelvin J. "A New Approach to Consumer Theory." *Journal of Political Economy* 74 (1966): 132-157.

[65] Lavy, Victor, Olmo Silva, and Felix Weinhardt. "The Good, the Bad, and the Average: Evidence on the Scale and Nature of Ability Peer Effects in Schools." NBER Working Paper No. 15600. 2009.

[66] Lazear, Edward P. "Education Production." *Quarterly Journal of Economics* 116, no.3 (2001): 777-803.

[67] Lefgren, Lars. "Educational peer effects and the Chicago public schools." *Journal of Urban Economics* 56 (2004): 169-191.

[68] Lucas, Robert E. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1988): 3-42.

[69] Lyle, David S. "Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point." *Review of Economics & Statistics* 89, no. 2 (2007): 289-299.

[70] Manski, Charles F. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60, no. 3 (1993): 531-542.

[71] Marshall, Alfred. *Principles of Economics*, Vol. 1. London: Macmillan, 1890.

[72] Mill, John Stuart. *On Liberty*. London: John W. Parker & Son, West Strand, 1859.

[73] Miller, Robert. "Job Matching and Occupational Choice." *Journal of Political Economy* 92, no. 6 (1984): 1086-1120.

[74] Murphy, Kevin J. "Incentives, Learning, and Compensation: A Theoretical and Empirical Investigation of Managerial Labor Contracts." *RAND Journal of Economics* 17, no. 1 (1986): 59-76.

[75] Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. "Industrialization and the Big Push." *Journal of Political Economy* 97, no. 5 (1989): 1003-10026.

[76] Neal, Derek. "The Complexity of Job Mobility among Young Men." *Journal of Labor Economics* 17, no. 2 (1999): 237-261.

[77] Prescott, Edward C., and John Boyd. "Dynamic Coalitions, Growth, and the Firm." In *Contractual Arrangements for Intertemporal Trade*, edited by E.C. Prescott and N. Wallace. Minneapolis, University of Minnesota Press: 1987.

[78] Romer, Paul M. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94, no. 5 (1986): 1002-1037.

[79] Romer, Paul M. "Growth Based on Increasing Returns Due to Specialization." *American Economic Review Papers and Proceedings* 77, no. 2 (1987): 56-62.

[80] Rosen, Harvey S. "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics* 13, no. 2 (1982): 311-323.

[81] Rosen, Sherwin. "Learning by Experience as Joint Production." *Quarterly Journal of Economics* 86, no. 3 (1972): 366-382.

[82] Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82, no. 1 (1974): 34-55.

[83] Rosen, Sherwin. "Specialization and Human Capital." *Journal of Labor Economics* 1, no. 1 (1983): 43-49.

[84] Rousseau, Jean-Jacques. *Emile*, or *On Education*. Translated by Allan Bloom. New York: Basic Books, 1762.

[85] Roy, A.D. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers (New Series)* 3 (1951): 135-146.

[86] Sanbonmatsu, Lisa, Jeffrey Kling, Greg Duncan, and Jeanne Brooks-Gunn. "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment." *Journal of Human Resources* 41, no. 4 (2006): 649-691.

[87] Sattinger, Michael. "Differential Rents and the Distribution of Earnings." *Oxford Economic Papers* 31, no. 1 (1979): 60-71.

[88] Sattinger, Michael. "Assignment Models of the Distribution of Earnings." *Journal of Economic Literature* 31, no. 2 (1993): 831-880.

[89] Shea, John. "Instrument Relevance in Multivariate Linear Models: A Simple Measure." *Review of Economics & Statistics* 79, no. 2 (1997): 348–352.

[90] Stinebrickner, Todd R., and Ralph Stinebrickner "What Can Be Learned about Peer Effects Using College Roommates? Evidence from New Survey Data and Students from Disadvantaged Backgrounds." *Journal of Public Economics* 90 (2006): 1435-1454.

[91] Stock, James H., and Motohiro Yogo. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by D.W.K. Andrews and J.H. Stock. Cambridge: Cambridge University Press, 2005.

[92] Tinbergen, Jan. "On the Theory of Income Distribution." *Weltwirtschaftliches Archiv* 77 (1956): 155-173.

[93] Welch, Finis. "Linear Synthesis of Skill Distribution." *Journal of Human Resources* 4, no.3 (1969): 311-327.

[94] Willis, Paul. *Learning to Labor: How Working Class Kids Get Working Class Jobs.* New York: Columbia University Press, 1977.

[95] Willis, Robert J., and Sherwin Rosen. "Education and Self-Selection." *Journal of Political Economy* 87, no. 5 (1979): S7-S36.

[96] Wilson, W. J. *The truly disadvantaged: the inner city, the underclass, and public policy.* Chicago: University of Chicago Press, 1987.

[97] Yatchew, Adonis. "Nonparametric Regression Techniques in Economics." *Journal of Economic Literature* 36, no. 2 (1998): 669-721.

[98] Zimmer, Ron W., and Eugenia F. Toma. "Peer Effects in Private and Public Schools across Countries." *Journal of Policy Analysis and Management* 19, no. 1 (2000): 75-92.

[99] Zimmerman, David J. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics & Statistics* 85, no. 1 (2003): 9-23.

# A Appendix A: Data Appendix

## A.1 New York City Public Schools

### Demographic Variables

Demographic variables that should not vary from year to year (race, gender) were pulled from New York City enrollment files from 2003/04 through 2008/09, with precedence given to the most recent files. Race consisted of the following categories: black, Hispanic, white, Asian, and other race. These categories were considered mutually exclusive. The "other race" category consisted of students who were coded as "American Indian." Gender was coded as male, female, or missing.

Demographic variables that may vary from year to year (free lunch status, English Language Learner status, and special education designation) were only pulled from the enrollment file corresponding to the same year as the observation. A student was considered eligible for free lunch if he was coded as "A" or "1" in the raw data, which corresponds to free lunch, or "2", which corresponds to reduced-price lunch. A student was considered non-free lunch if the student was coded as a "3" in the NYC enrollment file, which corresponds to full price lunch. All other values, including blanks, were coded as missing. For English Language Learner status, a student was given a value of one if he was coded as "Y" for the limited English proficiency variable. All other students in the NYC data were coded as zero for English Language Learner status. Special education was coded similarly.

### New York State Test Scores

NYC state test scores were constructed from the NYC test score files for 2003/04 through 2008/09 for English/Language Arts (ELA) and math. School-wide rankings were constructed based on these test scores.

### Behavior

The number of behavioral incidents for each student was determined from NYC files listing all recorded behavioral incidents from 2004/05 through 2008/09. Students not listed in this file but with a valid test score from the same year were assumed to have zero behavioral incidents. We constructed a behavioral incident indicator with a value of one if the student was listed for a behavioral incident in the file from the relevant year, zero if the student had a valid test score from the same year, and missing otherwise.

## A.2 National Educational Longitudinal Study

### Demographic Variables

Demographic variables were taken from the baseline year of the survey. These included: race, sex, English Language Learner status, parents' marital status, parents' education, school type (public, Catholic, or other private), school location (urban, suburban, rural), socioeconomic status, birth month, and birth year.

**Behavior**

Behavior variables were constructed using data from teacher reports on individual students. Teachers were asked to indicate whether the student had problems in each of the following areas: the student performs below his ability, the student does not complete homework, the student is frequently absent, the student is frequently tardy, the student is inattentive, or the student is disruptive. In the baseline year (eighth grade), each student had one teacher report from either Math or Science and another from either English or History, for a total of two teacher reports. Similarly, each student had two reports from the first follow-up year (tenth grade). Teacher reports were also administered in the second follow-up year (twelfth grade) but only in one subject, so these reports are excluded from our analysis, which takes advantage of within-year across-subject variation. For each student, we constructed an indicator that is equal to one if the student's teacher reports that the student has a problem in at least one of the six categories and zero otherwise. The outcomes used for our analysis are the within-year differences across subjects in the behavioral indicator.

**Grades**

The dataset contains self-reported grades for the baseline, first follow-up, second follow-up years. In the baseline year, students were asked to report for each subject (Math, Science, English, and History) whether their grades since sixth grade had been "mostly A's (90-100)," "mostly B's (80-89)," "mostly C's (70-79)," "mostly D's (60-69)," or "mostly below D (below 60)." Similarly, in the first follow-up year, students were asked to report for each subject whether their grades from ninth grade until now were "mostly A's," "about half A's and half B's," "mostly B's," "about half B's and half C's," "mostly C's," "about half C's and half D's," "mostly D's," or "mostly below D." These responses were converted to the average of the corresponding grade point values on a 4.0 scale, where 1.0 corresponds to D, 2.0 corresponds to C, 3.0 corresponds to B, and 4.0 corresponds to A. These grade values were used to compute a student's percentile rank within each class.

**Test Scores**

The dataset contains test scores for each student from Math, Science, English, and History for each year. We construct a test score control that is the mean of the test scores from the two subjects for which there are teacher reports in the baseline year and first follow-up year. We also construct its square and use both as controls in our estimates.

# B Appendix B: Additional Results

## B.1 Optimal Sorting in the Presence of Comparative Advantage

We conclude our modelling exercise by considering how a social planner interested in maximizing student's utility from social interactions would sort a fixed set of children into classrooms taking comparative advantage into account. Let the set of children have unit mass, and let their relative ability, $\sigma \equiv \frac{\sigma_N}{\sigma_T}$, be distributed according to the cumulative distribution function $\Phi$. For tractability, assume there are only two classrooms, and capital is fixed. Both the nerd and the troublemaking sectors within each classroom are competitive, and industry production functions are concave in labor.

With respect to the assignment of nerd and troublemaker labor to classrooms and industries the solution to the social planner's problem coincides with the allocation that would obtain in a competitive market in which children can switch classrooms and sectors costlessly. To see this note that as firms in each sector earn zero profits and the price of output remains fixed, maximizing total utility is equivalent to maximizing total output. The competitive market outcome ensures maximal output by equalizing the marginal product of labor across sectors and classrooms. From this market analog, three necessary conditions for an interior solution readily emerge:

$$\frac{\partial}{\partial L_1^N} F_1^N \left(L_1^N, K_1^N, A_1^N\right) = \frac{\partial}{\partial L_2^N} F_2^N \left(L_2^N, K_2^N, A_2^N\right) \tag{15}$$

$$\frac{\partial}{\partial L_1^T} F_1^T \left(L_1^T, K_1^T, A_1^T\right) = \frac{\partial}{\partial L_2^T} F_2^T \left(L_2^T, K_2^T, A_2^T\right) \tag{16}$$

$$\Phi^{-1}\left(i^*(L_1^T + L_2^T)\right) = \frac{\frac{\partial}{\partial L_1^T} F_1^T \left(L_1^T, K_1^T, A_1^T\right)}{\frac{\partial}{\partial L_1^N} F_1^N \left(L_1^N, K_1^N, A_1^N\right)}, \tag{17}$$

where $i^*(L_1^T + L_2^T)$ denotes the marginal individual if there are exactly $L_1^T + L_2^T$ efficiency units of troublemaking labor employed. More precisely, $i^*\left(L_1^T + L_2^T\right)$ is the $i$ which solves $L_1^T + L_2^T = \int_0^i \sigma_T(s)\,ds$. Notice, together with a market clearing condition, equations (15)–(17) are also sufficient. The first and second condition respectively state that the marginal products of nerds and troublemakers must be equal across classrooms. If not, total output, and hence the payment to group members, could be increased by reallocating labor from one classroom to the other. As there are no wage differences across classrooms within a sector, these conditions also imply that all individuals are indifferent between classrooms. The third condition ensures that the labor supply to each sector supports the prevailing wages as an equilibrium. That is, the troublemaker with the highest relative nerd skill has to be indifferent between the two sectors. The last condition must hold if the solution to the planner's problem is interior, as otherwise total utility could be increased if the marginal troublemaker became a nerd, or vice versa.

**Proposition 2** *If there exists an interior allocation with nerds and troublemakers in each classroom that satisfies equations (15)–(17) and market clearing, then it is the unique optimum to the planner's* *problem.*

**Proof.** Let $L^T \equiv L_1^T + L_2^T$, and substitute out $L_2^T$ in (16). Holding $L_T$ fixed, the left-hand side is decreasing and the right-hand side is increasing in $L_1^T$ by concavity of the production functions. Given that the solution is interior, we can thus uniquely determine the amount of troublemaker labor in classroom 1 conditional on $L^T$. Let this quantity be denoted by $l_1^T\left(L^T\right)$. By the Implicit Function Theorem $l_1^T$ is continuously differentiable in a neighborhood around $L^T$. Totally differentiating (16) with respect to $L^T$ and rearranging shows that $0 < \frac{d}{dL^T} l_1^T\left(L^T\right) < 1$. For any $L^T$ we can determine the marginal individual, $i^*\left(L^T\right)$, from equation (3).

Similarly, define $L^N \equiv \int_{i^*(L^T)}^1 \Phi_N^{-1}(s)\, ds$, where $\Phi_N$ denotes the cumulative distribution function of $\sigma_N$, and substitute out $L_2^N$ in equation (15). Again, concavity of the production functions implies that there can only exist one $L_1^N$ for every $L^T$ such that (15) is satisfied. Denoting this by $l_1^N\left(L^T\right)$ and differentiating both sides of (15) with respect to $L^T$ gives $\frac{d}{dL^T} l_1^N\left(L^T\right) < 0$. Since the left-hand side of (17) is increasing and the right-hand side is decreasing in $L^T$, we can determine a single value for $L^T$ by substituting $l_1^T\left(L^T\right)$ and $l_1^N\left(L^T\right)$ into equation (17). With $L^T$, $L_1^T$, and $L_1^N$ in hand, market clearing implies unique values for $L_2^T$, and $L_2^N$.

This shows that there exists at most one interior allocation satisfying equations (15)–(17) and market clearing. To see that this allocation, if it exists, must be the unique optimum, note that at any other point at least one of the three equations above must be violated, which means that total output can be increased by reallocating labor across sectors or classrooms. ■

Comparative statics follow from (15)–(17). For instance, suppose the troublemaking technology in classroom 1, i.e. $A_1^T$, increases and that $\frac{\partial^2}{\partial L_1^T \partial A_1^T} F_1^T\left(L_1^T, K_1^T, A_1^T\right) > 0$. For equation (17) to continue to hold, $L^T$ must increase, whereas $L_2^T$, $L_1^N$, and $L_2^N$ decrease for all marginal products to equalize and the market to clear. Not surprisingly, the social planner would assign more children to become troublemakers (reallocating some of them from classroom 2 to 1 in the process), and fewer to become nerds.

However, there could also be cases in which a corner solution obtains. Intuitively, for this to happen there must exist large productivity differences between sectors or classrooms. For instance, if for a given $L^T$ it was the case that $\frac{\partial}{\partial L_1^T} F_1^T\left(L^T, K_1^T, A_1^T\right) \geq \frac{\partial}{\partial L_2^T} F_2^T\left(0, K_2^T, A_2^T\right)$ then the planner would not assign any troublemakers, i.e. children for whom $\sigma < \Phi^{-1}\left(i^*(L^T)\right)$, to classroom 2, but all of them to classroom 1 instead. A similar inequality might define a corner solution for nerds. It is important to note that equation (17) changes if we are at a corner, but it must continue to hold in modified form if there are to be both nerds and troublemakers at all. E.g., if there are no nerds in classroom 1 and no troublemakers in classroom 2, then (17) becomes $\Phi^{-1}\left(i^*(L^T)\right) = \frac{\frac{\partial}{\partial L_1^T} F_1^T\left(L_1^T, K_1^T, A_1^T\right)}{\frac{\partial}{\partial L_2^N} F_2^N\left(L_2^N, K_2^N, A_2^N\right)}$.

The equations above suggest a very simple intuition: The nature (and shape) of group production is critical in determining optimal sorting. At an abstract level, three general insights are gleaned. First, if there are two classrooms (or neighborhoods) that have equal production technologies and capital, the quantity of efficiency units will be equalized across classrooms. Note, this does not imply that there must be the same number of nerds in each classroom—just the same number of effective units. This may imply one super nerd in one classroom and several lesser nerds in another.

Second, if there are classrooms that are quite different—one promoting nerd production (an abundance of textbooks) and another favoring the production of troublemaking (sharp scissors)—there may be interesting corner solutions to the optimal sorting problem. All nerds should be placed in their preferred classroom and troublemakers in the other.

Third, suppose there exist classrooms with heterogenous production technology and mobile capital. In this case, a social planner will combine nerds, more effective nerd technology, and nerd capital until the marginal products are equalized across classrooms. Whether or not there are corner solutions depends on how large differences in the *fixed* factors of production are.

### B.2 Bundling and $n$-dimensional Skill

In Section 2 we have developed a model of social interactions in which individuals with two distinct types of ability choose between two peer groups, each of which only values one particular skill. In this appendix we treat the general case in which individuals posses $n$ skills and choose between $m$ sectors, which potentially value all (or some subset of) skills. In doing so we rely heavily on Heckman and Scheinkman (1987), both in notation and style.

Let the vector of skill endowments of individual $i \in I = [0, 1]$ be given by $\boldsymbol{\sigma}(i) = [\sigma_1(i), \ldots, \sigma_n(i)]^T$, and assume that the functions $\sigma_j : I \to \mathbb{R}_+$ are measurable and bounded for all $j \in \{1, \ldots, n\}$. For ease of notation dispose of group specific technology as well as capital, and let output of peer group $k \in \{1, \ldots, l_s\}$ in sector $s \in \{1, \ldots, m\}$ be given by the production function $f_s(\mathbf{a}_s^k)$, which is assumed to be once continuously differentiable and admit finite first derivatives. $\mathbf{a}_s^k$ denotes the total efficiency units of different skills used by group $k$. That is, a peer group consisting of the set of members $C_s^k \subseteq I$ commands total skill

$$\mathbf{a}_s^k = \int_{C_s^k} \boldsymbol{\sigma}(i) di.$$

Given a set of $n \times 1$ vectors $\mathbf{w}_s$ denoting the compensation of skill in each sector, peer groups choose skill inputs to maximize profits

$$\pi_s^k = \max_{\mathbf{a}_s^k} f_s(\mathbf{a}_s^k) - \mathbf{a_s^k}^T \mathbf{w}_s,$$

and agents join the sector offering the highest social income

$$U_i = \max \left\{ \boldsymbol{\sigma}(i)^T \mathbf{w}_1, \ldots, \boldsymbol{\sigma}(i)^T \mathbf{w}_m \right\}.$$

Note that individuals 'sell' all of their skill endowments to a single group. This form of bundling is the distinctive aspect of Heckman and Scheinkman (1987), and may result in skill prices varying across sectors. Without such variation individuals would be indifferent between sectors, rendering their choice problem moot.

## A. Concavity in Production

Before characterizing the equilibrium we first introduce some terminology.

**Definition 3** *An* allocation *is a partition of $I$ into disjoint measurable subsets $C_s^k$ with $k \in \{1, \ldots, l_s\}$ and $s \in \{1, \ldots, m\}$.*

**Definition 4** *A* feasible state *is a set of vectors $\mathbf{a}_s^k$, $k \in \{1, \ldots, l_s\}$, $s \in \{1, \ldots, m\}$, for which there exists an allocation $C_s^k$ with $\mathbf{a}_s^k = \int_{C_s^k} \boldsymbol{\sigma}(i)di$. For any feasible state let $\mathbf{a}_s = \int_{C_s} \boldsymbol{\sigma}(i)di$ denote the total efficiency units of skill used in sector $s$.*

**Definition 5** *A* competitive equilibrium *is an allocation with an associated $n \times m$ wage matrix $\mathbf{w} = [w_{js}]$, $j \in \{1, \ldots, n\}$, $s \in \{1, \ldots, m\}$, such that:*
*(i) Peer groups maximize profits, i.e. for all $k \in \{1, \ldots, l_s\}$ and all $s \in \{1, \ldots, m\}$, $\mathbf{a}_s^k = \int_{C_s^k} \boldsymbol{\sigma}(i)di$ solves* $\max_{\mathbf{a}} f_s(\mathbf{a}) - \mathbf{a}^T \mathbf{w}_s$.
*(ii) Agents maximize their social income. That is, for almost all $i \in I$, $i \in C_s = \cup_{k=1}^{l_s} C_s^k$ implies $\boldsymbol{\sigma}(i)^T \mathbf{w}_s \geq \boldsymbol{\sigma}(i)^T \mathbf{w}_{s'}$ for all $s' \in \{1, \ldots, m\}$.*

Notice that the first and second condition together imply that a peer group in sector $s$ cannot achieve higher profits by hiring a subset of individuals $D \subseteq C_{s'}$ from sector $s'$ at wage rates $\mathbf{w}_{s'}$.

If production is concave in inputs then any optimal allocation must assign equal amounts of skill to all (identical) firms within a sector. The following two lemmas from Heckman and Scheinkman (1987) greatly simplify the search for competitive equilibria on the industry level.

**Lemma 6** *For all $k \in \{1, \ldots, l_s\}$, $s \in \{1, \ldots, m\}$, if $\mathbf{a}_s^k$ is feasible, then $\mathbf{a}_s^k = l_s^{-1}\mathbf{a}_s$ is feasible as well.*

**Proof.** See Heckman and Scheinkman (1987). ■

**Lemma 7** *The set of feasible industry states*

$$K = \left\{ (\mathbf{a}_1, \ldots, \mathbf{a}_m) \,|\mathbf{a}_s \in \mathbb{R}^n, \mathbf{a}_s = \int_{C_s} \boldsymbol{\sigma}(i)di, \cup_{s=1}^m C_s = I, C_s \cap C_{s'} = \varnothing \text{ if } s \neq s' \right\}$$

*is convex.*

**Proof.** See Heckman and Scheinkman (1987). ■

Subject to feasibility the competitive equilibrium maximizes total output. Lemma 1 allows us to write

$$\max_{\mathbf{a}} \sum_{s=1}^m l_s f_s \left( l_s^{-1} \mathbf{a}_s \right) \tag{18}$$
$$\text{s.t.} \quad \mathbf{a} = (\mathbf{a}_1, \ldots, \mathbf{a}_m) \in K.$$

As (18) is a convex optimization problem by Lemma 2 and concavity of $f$, the following necessary and sufficient condition characterizes any solution $\widetilde{\mathbf{a}}$:

$$\sum_{s=1}^{m} \nabla f_s \left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)^T \left(\widetilde{\mathbf{a}}_s - \mathbf{a}'_s\right) \geq 0 \text{ for all } \mathbf{a}' \in K. \tag{19}$$

Note that, for all $\widetilde{\mathbf{a}}, \mathbf{a}' \in K$, $\sum_{s=1}^{m} \left(\widetilde{\mathbf{a}}_s - \mathbf{a}'_s\right) = \mathbf{0}$, since feasibility requires all skill to be employed. It follows straightforwardly that whenever there exists an allocation such that

$$\nabla f_s \left(l_s^{-1} \widetilde{\mathbf{a}}_s\right) = \nabla f_{s'} \left(l_{s'}^{-1} \widetilde{\mathbf{a}}_{s'}\right) \quad \text{for all } s, s' \in \{1, \ldots, m\},$$

then (19) is satisfied and the optimum can be decentralized by the wage structure

$$\mathbf{w}_s = \frac{\partial f_1 \left(l_1^{-1} \widetilde{\mathbf{a}}_1\right)}{\partial \mathbf{a}} \quad \text{for all } s \in \{1, \ldots, m\}.$$

In such a case bundling imposes no binding constraint on the problem, i.e. $\widetilde{\mathbf{a}} \in \operatorname{Int}(K)$, and wages are equal in all sectors of the economy. Therefore, agents are indifferent between sectors.

Differences in production technologies or factor immobility, however, might very well preclude marginal productivities from being equalized across sectors—especially in the case of social interactions in which it is perceivable that some peer groups have no use for certain skills.

As an example, recall the $2 \times 2$ case of Section 2. In this model a nerd possesses $\sigma_T(i)$ units of troublemaking ability, which must be employed in the nerd sector; despite the marginal product of troublemaking skill being zero in the nerd and strictly positive in the troublemaking sector. In such cases bundling does impose binding constraints, and $\widetilde{\mathbf{a}} \in \operatorname{Bd}(K)$.[26] If skill endowments are sufficiently diverse for peer groups to employ more than one 'type' of agent, i.e. individuals with distinct combinations of skill, then wages cannot be equalized across sectors.

To see this consider two agents, $i$ and $i'$, with $\frac{\sigma_j(i)}{\sigma_{j'}(i)} \neq \frac{\sigma_j(i')}{\sigma_{j'}(i')}$ for some $j, j' \in \{1, \ldots, n\}$, both of whom are employed in sector $s$; and suppose $\frac{\partial f_s\left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)}{\partial a_j} \neq \frac{\partial f_s\left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)}{\partial a_{j'}}$. Competition among groups within $s$ results in zero profits on the margin, or

$$\left[\frac{\partial f_s \left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)}{\partial a_j} - w_{js}\right] \sigma_j(i) = \left[\frac{\partial f_s \left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)}{\partial a_{j'}} - w_{j's}\right] \sigma_{j'}(i)$$
$$\left[\frac{\partial f_s \left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)}{\partial a_j} - w_{js}\right] \sigma_j(i') = \left[\frac{\partial f_s \left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)}{\partial a_{j'}} - w_{j's}\right] \sigma_{j'}(i')$$

which implies[27]

$$\frac{\partial f_s \left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)}{\partial a_j} = w_{js} \quad \text{and} \quad \frac{\partial f_s \left(l_s^{-1} \widetilde{\mathbf{a}}_s\right)}{\partial a_{j'}} = w_{j's}. \tag{20}$$

[26] Geometrically, $-\sum_{s=1}^{m} \nabla f_s$ defines a hyperplane to the feasible set at $\widetilde{\mathbf{a}}$.

[27] The same conclusion holds if $\frac{\eta_j(i)}{\eta_{j'}(i)} = \infty$ due competiton among firms.

If peer groups in sector $s'$ also employ two distinct types of individuals, then

$$\frac{\partial f_{s'}\left(l_{s'}^{-1}\widetilde{\mathbf{a}}_{s'}\right)}{\partial a_j} = w_{js'} \quad \text{and} \quad \frac{\partial f_{s'}\left(l_{s'}^{-1}\widetilde{\mathbf{a}}_{s}'\right)}{\partial a_{j'}} = w_{j's'}. \tag{21}$$

Since bundling prevents marginal products from being equalized, (20) and (21) imply wage differences across sectors. With wages differing across sectors most agents are not indifferent. Mathematically, the set of agents in sector $s$ is then given by

$$C_s = \left\{ i \in I | \boldsymbol{\sigma}(i)^T \mathbf{w}_s \geq \boldsymbol{\sigma}(i)^T \mathbf{w}_{s'} \text{ for all } s' \in \{1, \ldots, m\} \right\},$$

and in $\boldsymbol{\sigma}$-space the set of agents who choose $s$ forms a convex cone (through the origin).

## B. Convexity in Production

As mentioned in Section 2, if production is convex in skill inputs we cannot assume marginal product pricing because peer groups would incur losses. With increasing returns to scale on the peer group level a competitive equilibrium in the usual sense does not exist. However, under the assumption that new peer groups can form costlessly within a sector (free entry), a zero-profit condition helps determine the equilibrium.

Moreover, under the assumption of external increasing returns to scale (see Section 2) wages in a sector must be equal to the marginal product of labor at the firm level (which is smaller than that at the industry level). Hence, wages in sector $s$ are given by

$$\mathbf{w}_s = \nabla f_s(\mathbf{a}_s^k | \mathbf{a}_s) = \nabla f_s(\mathbf{a}_s^{k'} | \mathbf{a}_s). \tag{22}$$

Since we are only interested in the allocation of labor across industries, in what followes we abstract from the matching of agents to individuals groups within a sector. In Appendix B.3 we endogenize group size and groups' hierarchies.

Letting $F(\cdot)$ denote the industry production function net of fixed costs and using (22), the equilibrium allocation of labor, $\widetilde{\mathbf{a}}$, and number of peer groups, $\widetilde{\mathbf{l}}$, must satisfy

$$0 = \Pi_s\left(\widetilde{\mathbf{a}}_s, \widetilde{l}_s\right) \equiv \begin{cases} 0 & \text{if firms don't enter and } \widetilde{l}_s = 0 \\ F_s(\widetilde{\mathbf{a}}_s, \widetilde{l}_s) - \mathbf{w}_s^T \widetilde{\mathbf{a}}_\mathbf{s} & \text{otherwise} \end{cases}, \text{ for all } s,$$

or in vector notation

$$\boldsymbol{\Pi}\left(\widetilde{\mathbf{a}}, \widetilde{\mathbf{l}}\right) = \mathbf{0}. \tag{23}$$

Equation (22) implies that individual groups are price takers, which is only realistic if there are a large number of them within each sector. To sidestep technical difficulties associated with discreteness in $\mathbf{l}$, we allow $l_s$ to be a real number. One might think thinks of $l_s$ as the mass of firms within sector $s$.

**Proposition 8** *If $l_s \in \mathbb{R}_+$ and $\Pi_s(\overline{\mathbf{a}}, \underline{l}_s) > 0$ for some sector $s$ and some $\underline{l}_s$, where $\overline{\mathbf{a}} = \int_I \boldsymbol{\sigma}(i)\, di$, then there exists at least one pair $\left(\widetilde{\mathbf{a}}, \widetilde{\mathbf{l}}\right) \in K \times \mathbb{R}_{++}$ such that $\boldsymbol{\Pi}\left(\widetilde{\mathbf{a}}, \widetilde{\mathbf{l}}\right) = \mathbf{0}$. Hence, an equilibrium exists.*

**Proof.** To prove existence, we show existence in a trivial case. For all $s' \neq s$ let $\widetilde{\mathbf{a}}_{s'} = \mathbf{0}$ and set $l_s = 0$. Hence, $\Pi_{s'} = 0$ for all $s' \neq s$. For equation (23) to be satisfied and equilibrium to obtain there must exist some $\widetilde{l}_s$ such that $\Pi_s\left(\overline{\mathbf{a}}, \widetilde{l}_s\right) = 0$. Due to the fixed cost of production that all firms have to pay and the limited amount of skill in the economy, $\Pi_s\left(\overline{\mathbf{a}}, \overline{l}_s\right) < 0$ for some $\overline{l}_s$ large enough. The Intermediate Value Theorem then guarantees existence of $\widetilde{l}_s \in \left(\underline{l}_s, \overline{l}_s\right)$ such that $\Pi_s\left(\overline{\mathbf{a}}, \widetilde{l}_s\right) = 0$. ∎

As before, wages are determined by the marginal product of labor at the firm level, i.e. equation (22), and the set of agents in a sector is given by

$$C_s = \left\{ i \in I | \boldsymbol{\sigma}(i)^T \mathbf{w}_s \geq \boldsymbol{\sigma}(i)^T \mathbf{w}_{s'} \text{ for all } s' \in \{1, \ldots, m\} \right\}.$$

### B.3 Endogenous Hierarchy and Group Size

As noted earlier, the potential extensions to our basic economic model of peer effects are endless. In this section we exposit a rather poignant extension based on Rosen (1982). This allows us to investigate endogenous hierarchies within groups, and to derive the number and size of peer groups within a sector. In our previous analysis, agents could choose one of two sectors: nerds or troublemakers. In this extension, agents can choose to be with the main troublemaking group or branch out and create their own troublemaking organization. In equilibrium only the most able agents will serve as leaders because followers' productivity depends positively on leader ability. Moreover, peer group size is an increasing function of leader ability. In what follows, we abstract from the sorting of agents into sectors, but instead focus on the organization of individual peer groups within a given sector. In doing so we borrow heavily from Rosen (1982), both in notation and style.

Suppose that within each peer group there are two types of agents: one leader and followers. Let $\sigma^f(i) \in \mathbb{R}_+$ denote the productivity of an agent who is deemed a 'follower' in the group. Let $r$ denote the skill of the group's leader, which is also measured in efficiency units and varies from agent to agent. We think of this as the ability to lead others to a common purpose. $t(i)$ denotes the amount of time that a leader spends with follower $i$.

Ignoring capital, the social utility attributable to $r$ leading $\sigma^f(i)$ is:

$$x(i) = g(r)\, q\left(rt(i), \sigma^f(i)\right),$$

where $g(\cdot)$ and $q(\cdot)$ are smooth, twice continuously differentiable functions, that are increasing and concave in each of their arguments. The key assumption embodied in this choice of functional form is that a follower produces more output under a better leader and if the leader spends more time with him. In this setup leader's time is the scarce factor of production that in equilibrium limits group size and prevents the formation of one large group led by the most able agent. The total output of the group can be written as:

$$X = \int_{i \in I} x(s)ds = g(r) \int_{i \in I} q\left(rt(s), \sigma^f(s)\right) ds,$$

where both the set of followers, $I$, and $t(i)$ are endogenously determined.

The description of production possibilities is completed by a specification of factor supplies. Each agent is completely described by a tupel of endowed ability $(\sigma^f, r)$. Following Rosen (1982) we analyze a special case in which abilities follow a nonhomogeneous latent one factor structure:

$$\sigma^f(i) = a_\sigma + d_\sigma \xi(i) \tag{24}$$

$$r = a_r + d_r \xi(i) \tag{25}$$

where $a_l$ and $d_l$, $l \in \{\sigma, r\}$, are positive constants, and $\xi(i) \in \mathbb{R}_+$ represents 'general ability'. The distribution of $\xi$ in the population follows $m(\xi)\, d\xi$.

Given the quality and quantity of human inputs, the group's production function is defined by the allocation of leadership effort $t(\cdot)$ that maximizes total output:

$$X = \max_{t(\cdot)} \left[ g(r) \int_{i \in I} q\left(rt(s), \sigma^f(s)\right) ds + \lambda \left(1 - \int_{i \in I} t(s)ds\right)\right]$$

where we assume that each leader has $T = 1$ units of time, and $\lambda$ is a standard Lagrange multiplier.

For tractability, assume $q\left(rt(i), \sigma^f(i)\right)$ exhibits constant returns to scale. Then

$$x_i = g(r)\, \sigma^f(i)\, \theta\left(rt(i) / \sigma^f(i)\right)$$

with $\theta' > 0$ and $\theta'' < 0$. Leaders' optimal allocation of time is determined by the first order conditions:[28]

$$rg(r)\, \theta'\left(rt(i) / \sigma^f(i)\right) = \lambda, \text{ for all } i. \tag{26}$$

With $r$ fixed, (26) implies that $t(i) / \sigma^f(i) = k$, where $k$ is a constant. Thus, leadership effort is allocated in proportion to follower's skill: $t(i) = k\sigma^f(i)$. Aggregating across all workers gives $T = \int_{i \in I} t(s)\, ds \equiv kQ$, where $Q \equiv \int_{i \in I} \sigma(s)\, ds$ is the total amount of efficiency units among the followers in the group. With this in hand, the group production function can be re-written as:

$$X = \int_{i \in I} x(s)ds = g(r) \int_{i \in I} \sigma^f(s)\, \theta(rT/Q)\, ds = g(r)\, Q\theta(rT/Q).$$

The market assignment of individuals to positions within groups is based on comparative advan-

[28] Second order conditions are satisfied by concavity.

tage. Constant returns to scale in $f\left(rt,\sigma^f\right)$ and perefect substitutability of efficient follower labor imply that only rank order needs to be determined and that questions of sorting within ranks and groups are irrelevant. Thus, in this model, an agent must decide whether to be a follower in a more successful peer group using $\sigma^f(i)$, or to lead his own group of peers using $r$. Market equilibrium determines the price per efficiency unit of follower skill, $\omega$, and the leader reward function $\pi(r)$, i.e.

$$\pi(r) = \max_Q \{g(r)\, Q\theta(rT/Q) - \omega Q\}$$

Ignoring nonpecuniary differences between positions, an agent's choice rests on which alternative yields the larger social income, given wages and his latent ability.

Rosen (1982) demonstrates that an analytical solution is available if one assumes that endowed talents follow (24). Using this fact, it is straightforward to show that all individuals above a critical ability threshold, $\xi^*$, choose to lead their own peer groups and individuals below $\xi^*$ serve as followers. Total peer group output is

$$\int_{\xi^*}^{\infty} g(\xi)\, Q(\xi)\, \theta(\xi T/Q(\xi))\, m(\xi)\, d\xi \tag{27}$$

where $Q(r)$ is the amount of labor lead by $r$. The total supply of production labor, i.e. the amount to be allocated among all followers, is

$$Q^S = \int_{-\infty}^{\xi^*} \sigma^f(\xi) m(\xi)\, d\xi, \tag{28}$$

and the market demand for it is the sum of peer groups' demand:

$$Q^D = \int_{\xi^*}^{\infty} Q(\xi)\, m(\xi)\, d\xi. \tag{29}$$

Equilibrium is thus defined by a function $Q(r)$, representing the optimal allocation of production labor to each active $r$, and a partition $\xi^*$ that maximizes (27) subject to market clearing, i.e. $Q^S = Q^D$. In symbols, we seek $Q(r)$ and $\xi^*$ to maximize

$$\int_{\xi^*}^{\infty} g(\xi)\, Q(\xi)\, \theta(\xi T/Q(\xi))\, m(\xi)\, d\xi + \mu \left[\int_{-\infty}^{\xi^*} \sigma^f(\xi) m(\xi)\, d\xi - \int_{\xi^*}^{\infty} Q(\xi)\, m(\xi)\, d\xi\right].$$

The Euler condition for $Q(r)$ is

$$\left(g(r)\left[\theta - (rT/Q(r))\,\theta'\right] - \mu\right) m(r) = 0, \tag{30}$$

and the condition for the extensive margin, $\xi^*$, is

$$X^* - \mu Q(r^*) = \mu \sigma^{f*} \tag{31}$$

where $r^* \equiv r(\xi^*)$ marks the marginal leader, $\sigma^{f*} \equiv \sigma^f(\xi^*)$ is his skill as follower, and $Q(r^*)$ is the potential labor to be lead by him. $X^*$ denotes the total output of his peer group. Assuming $m(\xi) > 0$ over the range of $\xi$, (30) requires that the marginal product of labor is equalized for all observed $r$. Thus, $\mu$ is the economy wide marginal productof $Q$, equal to the wage, $\omega$, in equilibrium. The left hand side of (31) is the payoff accruing to the leader of the marginal peer group and the right hand side is his opportunity cost, i.e. what he would earn as a follower. The partition is determined by the absence of rents at the margin, as typical.

The preceding discussion shows that in equilibrium only the most able agent's will lead their own groups (with the most talented leaders commanding more efficient follower labor), while all others will serve as followers in some peer group and are rewarded based on their marginal product. Note that the peer group itself retains no profits. Leaders receive the whole surplus after paying all other factors of production. Allowing for multi-level hierarchies is a straighforward extension of the model above that is outlined in Rosen (1982).

### B.4 Embedding the Roy Model in a Social Multiplier Framework

In this subsection, we cast our Roy model of social interaction into a general model of social interactions. Let there be a continuum of agents with unit mass. Every agent is endowed with one unit of (non-transferable) time. In the canonical model of social multipliers (see Becker and Murphy 2000, for example), agents derive utility from two different kinds of activities, social and asocial. Our principal departure from the usual framework is that we assume that there are two activities in which agents can engage with their peers: studying or mischief. These activities are exclusive and undertaken by separate social groups in one of two sectors: 'nerds' and 'troublemakers'. The utility derived from participating in a social activity depends on the total effective labor supplied *to that sector only,* $L_j$ indexed by $j \in \{N, T\}$, as well as another input we label 'capital', $K_j$. We allow capital to broadly represent any non-human input into groups' production (e.g., school quality, diligence of adult supervision, etc.).

Agents are heterogenous along two dimensions. Their varying size and strength yield differences in the ability to cause trouble, whereas heterogeneity in cognitive ability implies differences in their ability to be a true nerd. Let the continuous function $\sigma_N(i) : [0,1] \to \mathbb{R}_+$ denote the effective units of 'nerdiness' that agent $i$ is capable of contributing to the group (e.g., expertise in differential geometry). Analogously, agent $i$'s troublemaking ability is given by $\sigma_T(i) : [0,1] \to \mathbb{R}_+$.

To emphasize the importance of social activity selection in the estimation of peer effects, we model the effective utility from time spent with nerds or troublemakers as perfect substitutes

$$s(i) = t_N(i)\sigma_N(i)w_N(L_N, K_N) + t_T(i)\sigma_T(i)w_T(L_T, K_T) \tag{32}$$

We can then nest utility from social activities in a quasiconcave utility function that depends on an individual's social 'production' $s(i)$ as defined in (32), asocial activities $t_a(i)$, and atomistic agents take aggregate labor and capital conditions $(\mathbf{L}, \mathbf{K})$ as given.

$$\begin{aligned} U(i) &= U(t_a(i), s(i); \mathbf{L}, \mathbf{K}) \qquad (33) \\ \text{s.t. } t_a(i) + t_N(i) + t_T(i) &= 1 \end{aligned}$$

With this formulation, we think of the marginal return to spending an additional effective unit of time with a social group as a 'wage' that is determined by aggregate social market conditions. In the absence of a second activity, this is the standard framework for analyzing social interactions. The 'social multiplier' is generated the assumption that $\frac{\partial w_j}{\partial L_j} > 0$.[29]

With perfect substitutability between social activities, utility maximization implies a simple cutoff rule that separates individuals between social sectors in a manner analogous to a classical Roy (1951) model. Let $\sigma(i) \equiv \frac{\sigma_N(i)}{\sigma_T(i)}$ denote agent $i$'s skill as a nerd relative to that as a troublemaker, and order agents such that $\sigma'(i) \geq 0$. Equilibrium 'labor supply' is determined by the set of agents whose comparative advantage determines the sector of their social activity, and the marginal rate of substitution between social and asocial activities. Utility maximization implies the first order condition of 33

$$\frac{U_a}{U_j} = \sigma_j(i) w_j(L_j, K_j) \qquad (34)$$

Diminishing marginal utility implies that more skilled agents are more socially active within their respective sector as their opportunity cost of asocial activity is higher. The agent indifferent between the two sectors, $i^*$(not necessarily unique, see below) has a skill ratio of

$$\sigma(i^*) = \frac{w_T(L_T, K_T)}{w_N(L_N, K_N)}. \qquad (35)$$

For any given equilibrium wage ratio, all individuals with index $i \geq i^*$ join forces with the nerds, and individuals with $i < i^*$ become troublemakers. In our price theory of social interactions, comparative (rather than absolute) advantage determines an individual's choice of sector. Total effective labor supply to each sector is determined by

$$L_N^* = \int_{i^*}^{1} \sigma_N(s)\, t_N[s, w_N(L_N^*, K_N)] ds \qquad (36)$$

[29] To see this, suppose studying were the only social activity, and we were interested in the effect of additional school resources $K_N$ on effective time spent studying $L_N = \int_0^1 \sigma_N(i) t_N[i, w(L_N, K_N)]$. Taking the total derivative,

$$\begin{aligned} dL_N &= [\frac{\partial w_N}{\partial L_N} dL_N + \frac{\partial w_N}{\partial K_N} dK_N] \int_0^1 \sigma_N(i) \frac{\partial t_N(i)}{\partial w_N} di \\ \frac{dL_N}{dK_N} &= \frac{\frac{\partial w_N}{\partial K_N} \int_0^1 \sigma_N(i) \frac{\partial t_N(i)}{\partial w_N} di}{1 - \frac{\partial w_N}{\partial L_N} \int_0^1 \sigma_N(i) \frac{\partial t_N(i)}{\partial w_N} di} \end{aligned}$$

While the numerator measures the average direct effect of the additional resources on time spent studying, when $\frac{\partial w_N}{\partial L_N} > 0$, this amount is amplified by the additional increase in the return to studying induced by the 'social multiplier': $\frac{\partial w_N}{\partial L_N} \int_0^1 \sigma_N(i) \frac{\partial t_N(i)}{\partial w_N} di$.

$$L_T^* = \int_0^{i^*} \sigma_T(s)\, t_T[s, w_T(L_T^*, K_T)] ds. \tag{37}$$

Equations (36) and (37) characterize the supply side in the market for peers. A labor demand schedule with respect to $i^*$ can be characterized as a curve that traces the marginal return per unit of effective time as the measure of the sector increases–taking the labor supply decisions of inframarginal agents into consideration. The relationship between total labor supply in the troublemaking sector and the position of the marginal agent is derived by taking the derivative of 37 with respect to $i^*$

$$\begin{aligned} \frac{\partial L_T^*}{\partial i^*} &= \sigma_T(i^*)\, t_T[i^*, w_T(L_T^*, K_T)] + \frac{\partial w_T}{\partial L_T^*} \frac{\partial L_T^*}{\partial i^*} \int_0^{i^*} \sigma_T(s) \frac{\partial t^*[s, w_T(L_T^*, K_T)]}{w_T} ds \\ &= \frac{\sigma_T(i^*)\, t_T[i^*, w_T(L_T^*, K_T)]}{1 - \frac{\partial w_T}{\partial L_T^*} \int_0^{i^*} \sigma_T(s) \frac{\partial t^*[s, w_T(L_T^*, K_T)]}{w_T} ds} \end{aligned} \tag{38}$$

The expression for the nerd sector is identical, but with a negative numerator to account for the opposite effect moving $i^*$ has on the measure of the group. The numerator is the mechanical effect of adding a new member to the sector, and the denominator accounts for the labor supply adjustment of inframarginal members in response to the change in group size. $\frac{\partial t^*[s, w_T(L_T^*, K_T)]}{w_T}$ is always positive by the first order condition (34). While effective labor supply always increases with the measure of the sector, whether the mechanical effect is dampened or amplified depends on whether the social wage increases or decreases with effective labor in the sector. We turn to each case in turn.

### B.4.1 Diminishing Marginal Social Product of Labor in Both Sectors

In contrast with the 'social multiplier' approach in which it is assumed $\frac{\partial w_j}{\partial L_j} > 0$, we begin with the assumption that additional labor supply in a social sector reduces the marginal product of participation, $\frac{\partial w_j}{\partial L_j} < 0$–a common assumption for the theory of the firm. In our setting this would correspond to a game–of math problems or bullying, for example–becoming less fun as the activity has to be shared with more participants.[30] As part of our analysis we introduce the 'relative labor demand curve', $\delta(i) : [0, 1] \rightarrow \mathbb{R}_+$, a schedule that traces the relative marginal return of a unit of effective labor in social activity when index $i$ is the marginal agent between sectors

$$\delta(i) \equiv \frac{w_T(L_T(i), K_T)}{w_N(L_N(i), K_N)} \tag{39}$$

The assumption of diminishing marginal product in both sectors guarantees that the relative labor demand curve will be (weakly) downward sloping, as increasing the measure of the trouble-

[30] This holds for effective units, regardless of the number of participants. A brilliant student answering all of the questions or a talented athlete hogging the ball is an example of many effective units concentrated in a single person diminishing the social utility of participation for others.

making sector with lower the numerator, which raising the denominator of (39).

### B.4.2 Increasing Marginal Social Product of Labor in Both Sectors

As discussed above, the assumption that $\frac{\partial w_j}{\partial L_j} > 0$ in a single social sector has been the focus of much of the work on social interactions. Study groups allow students to benefit from division of labor on a lab project, and ensure individual students do not waste time stuck on a question to which someone else in the group knows the answer. More troublemakers at a school can divert teachers' attention and thereby reduce others' probability of punishment. We now consider the nature of equilibria when there are two social sectors characterized by increasing marginal product.

Equation 38 shows that when $\frac{\partial w_j}{\partial L_j} > 0$, the mechanical increase in labor supply from increasing $i^*$ is amplified by an increased labor supply from inframarginal agents due to the increased marginal product of making trouble. Conversely, the accompanying decrease in labor in the nerd sector yields an amplified decrease in the return to being a nerd. The relative demand schedule is therefore upward sloping. This creates the potential for multiple equilibria along the lines of Becker (1991) and DiPasquale and Glaeser (1998).

Figure 1: Equilibrium with Concavity in Labor Inputs

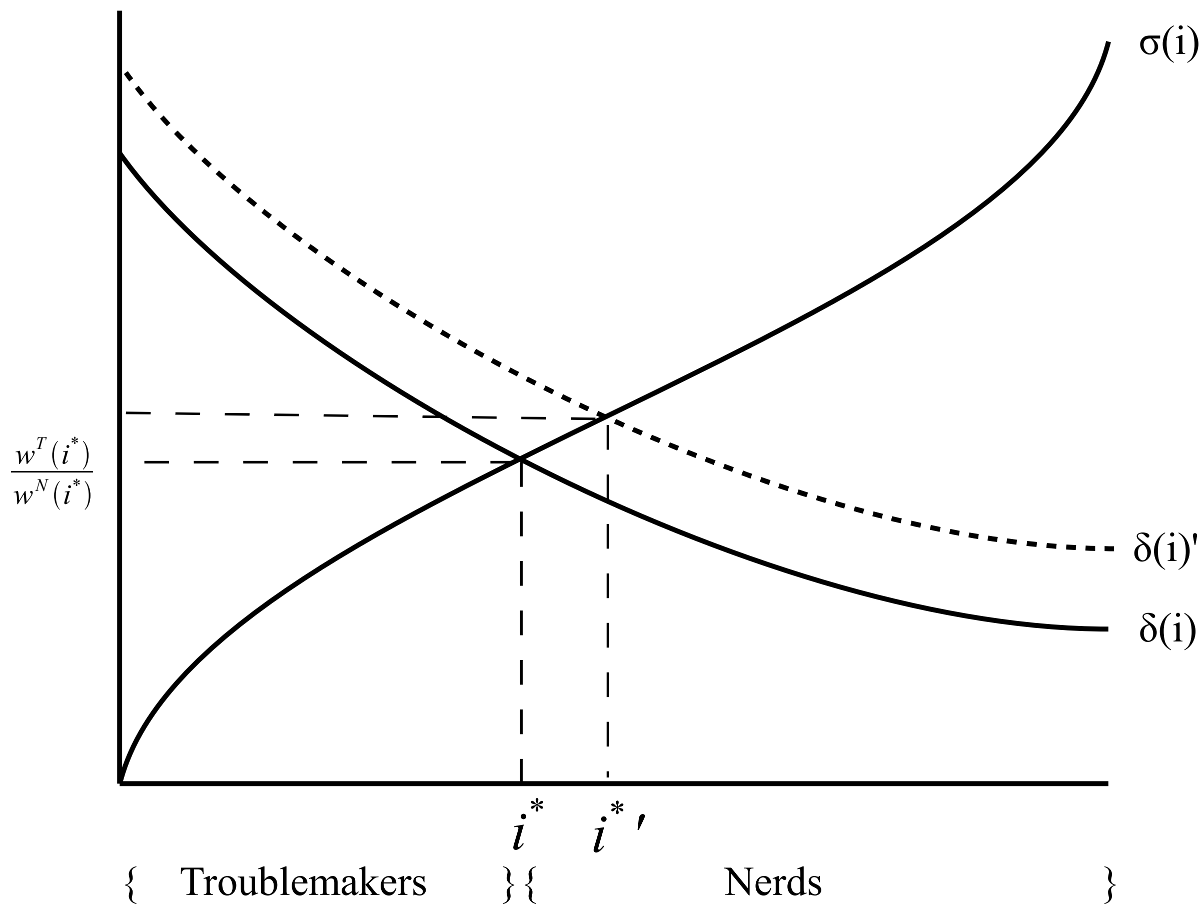


Figure 2: Comparative Statics when Production is Concave in Labor Inputs

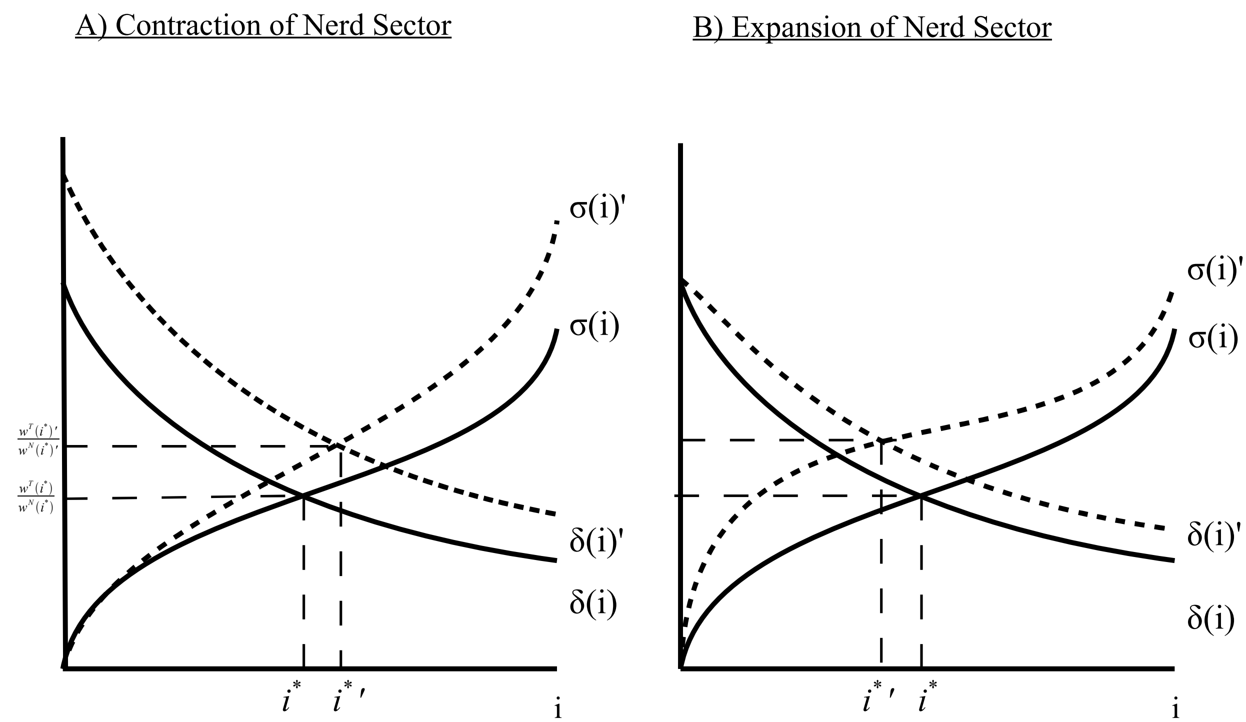


Figure 3: Multiple Equilibria when Production is Convex in Labor Inputs

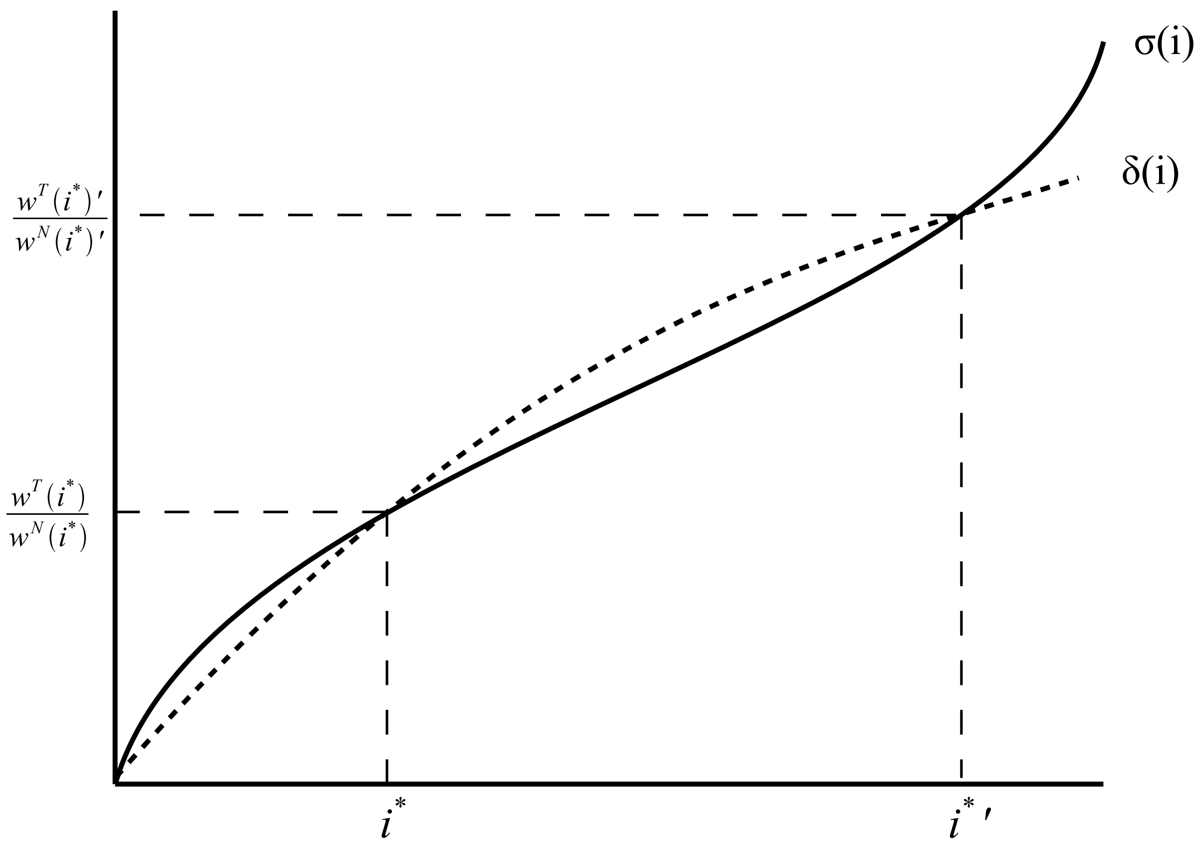


Figure 4: Reconciling No Peer Effects with the Comparative Advantage Approach

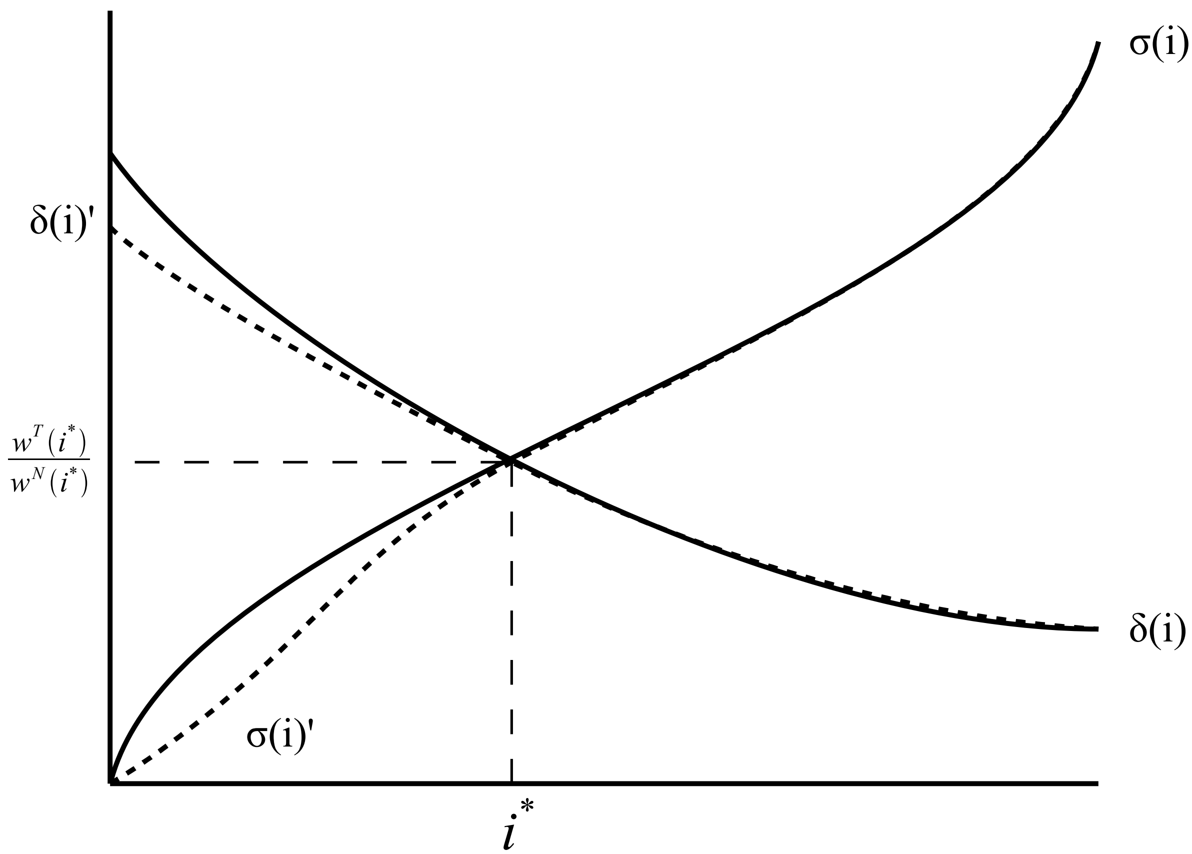


Figure 5: Reconciling Linear in Means with the Comparative Advantage Approach

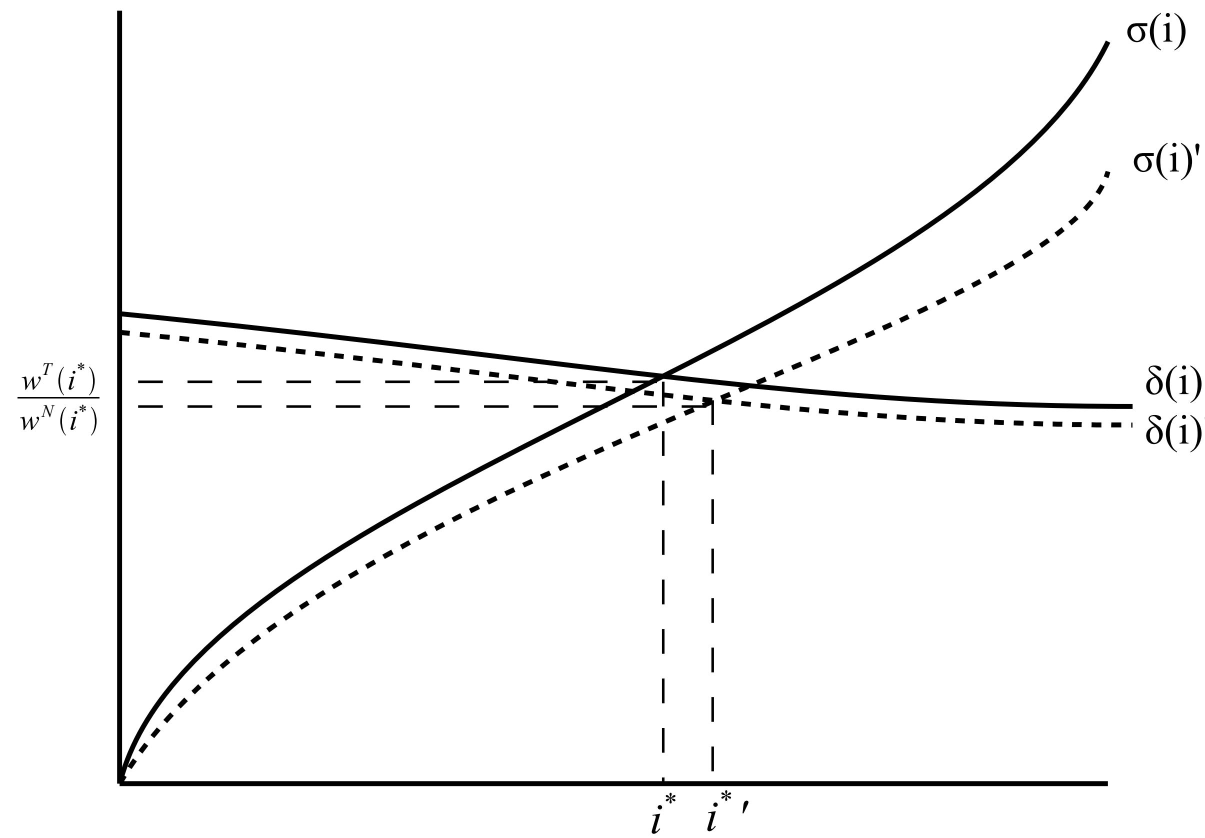


Figure 6: Reconciling Heterogenous Positive Effects with the Comparative Advantage Approach

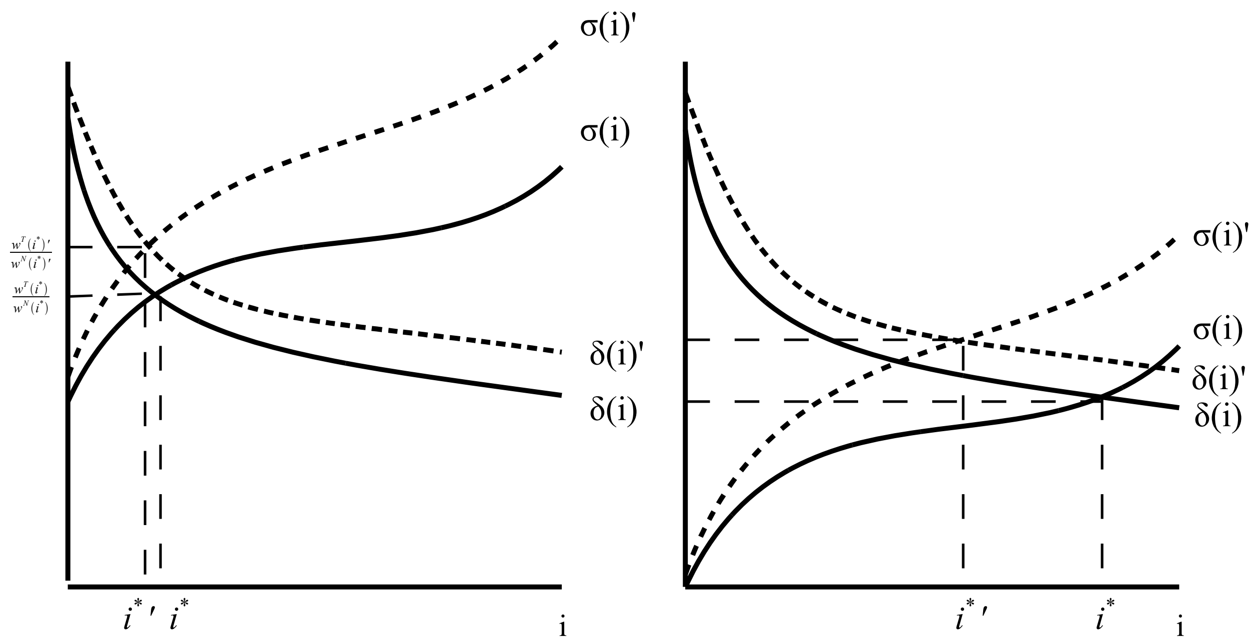


Figure 7: Reconciling Heterogenous Negative Effects in the Air Force Example with the Comparative Advantage Approach

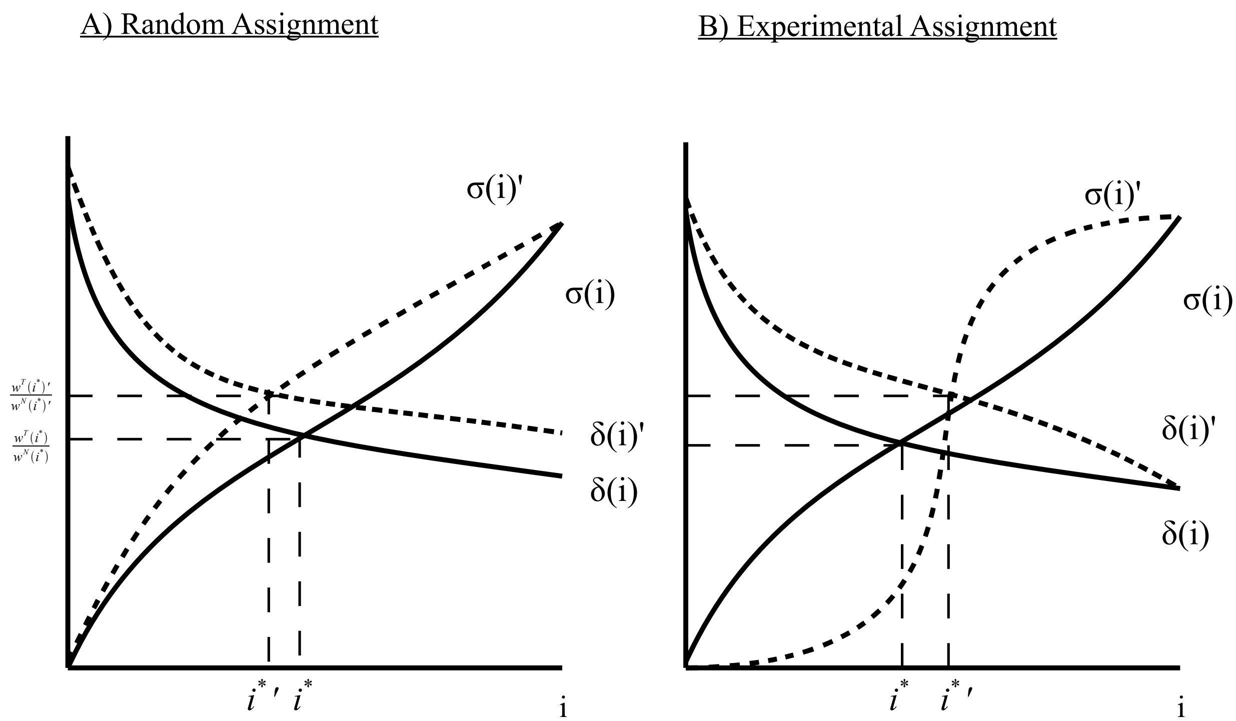


Figure 8: Reconciling Heterogenous Negative Effects in the MTO Example with the Comparative Advantage Approach

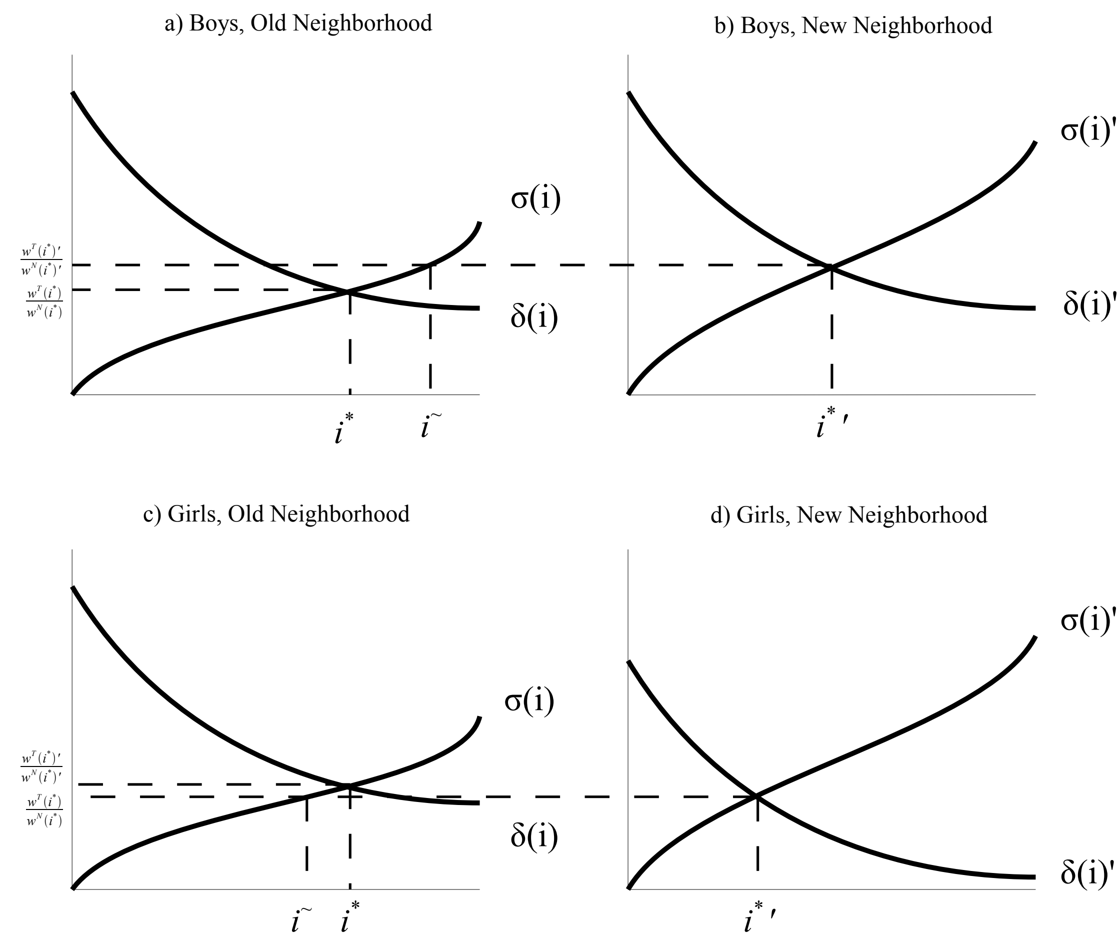
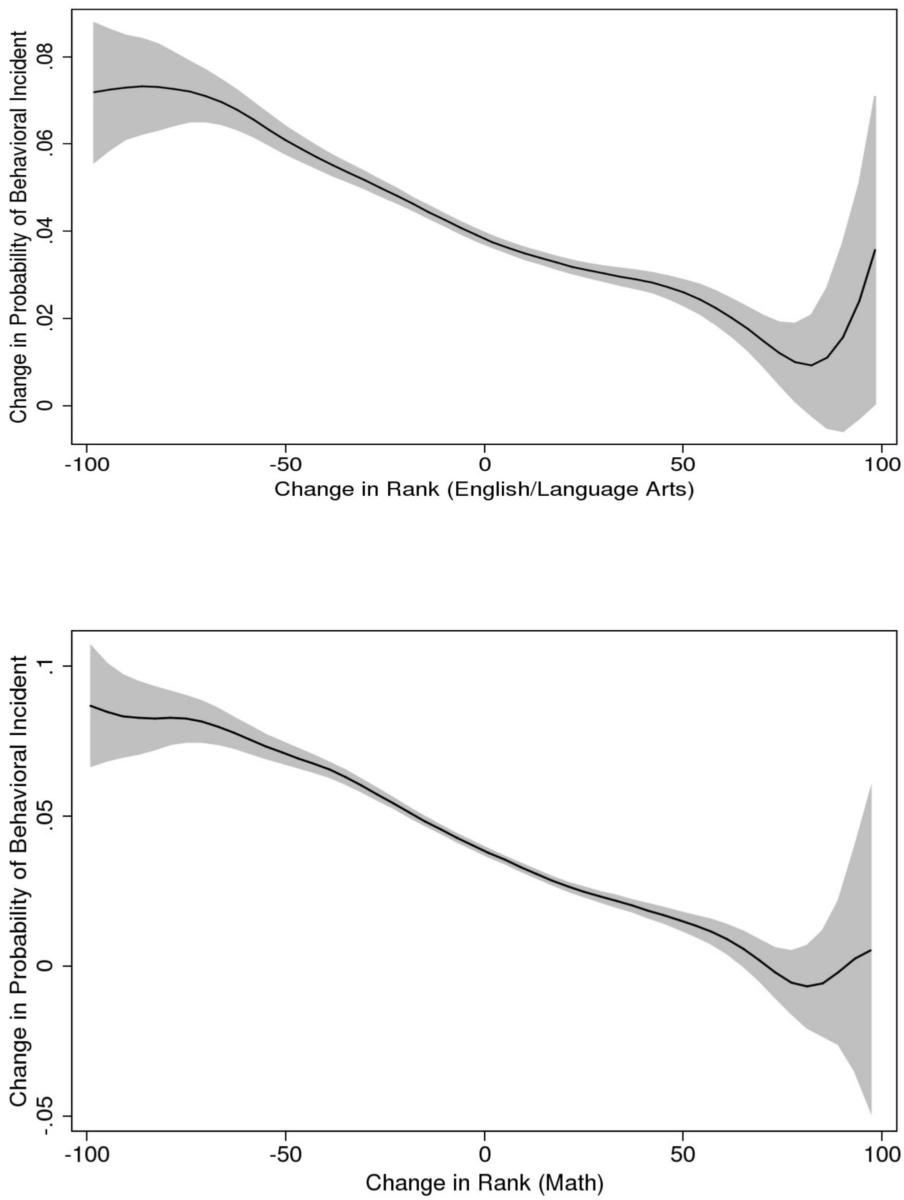
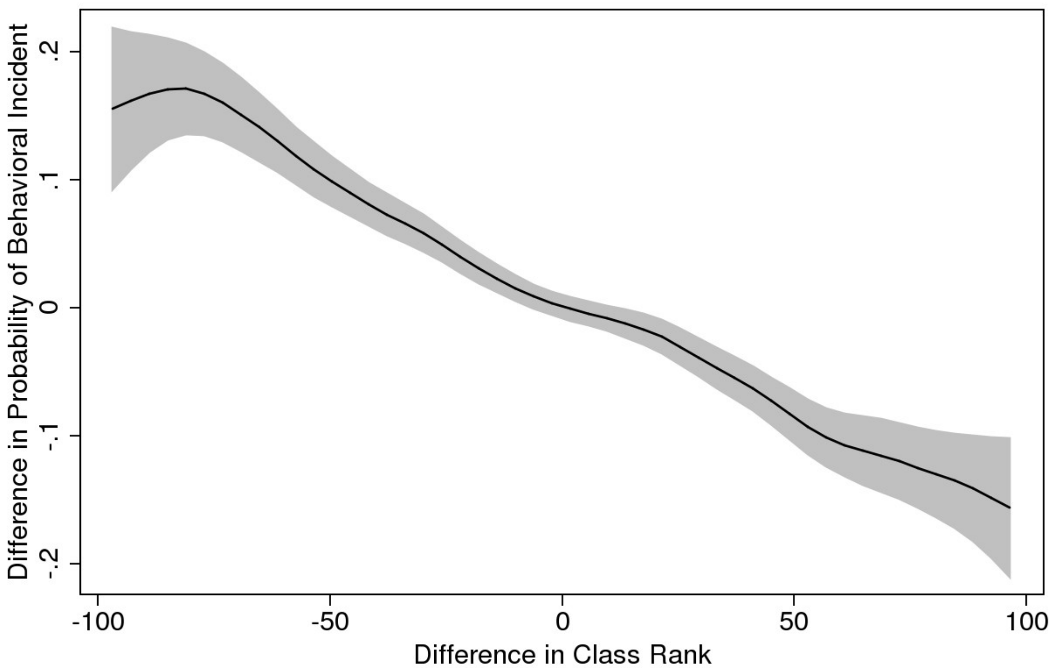


Figure 9: The Relationship between Ordinal Rank and Behavioral Incidence, New York City Public Schools




*Notes:* Panels show non-parametric estimates and the associated 95%-confidence intervals of the effect of a change in a student's class rank (in going from elementary to middle school) on the change in an indicator variable for whether she was involved in a behavioral incident, cf. equation (12). The top panel constructs rank based on English/Language Arts (ELA) test scores, whereas the lower one uses math test scores. Estimates are obtained using the differencing procedure in Yatchew (1998) and local-mean smoothing with a Gaussian kernel and a bandwidth of 10. Section 4.2 and the Data Appendix provide additional information on the exact econometric specification as well as the sample.

Figure 10: The Relationship between Class Rank and Behavioral Incidence, National Educational Longitudinal Study




*Notes:* The figure shows non-parametric estimates and the associated 95%-confidence intervals of the effect of differences in a student's course-specific rank on the difference in two course-specific behavioral outcomes, cf. equation (14). Estimates are obtained using the differencing procedure in Yatchew (1998) and local-mean smoothing with a Gaussian kernel and a bandwidth of 7.5. Section 4.2 and the Data Appendix provide additional information on the exact econometric specification as well as the sample.

Table 1: Estimates of the Relationship between Ordinal Rank and Behavior

A. Ranking Based on ELA Test Scores

| | Δ Behavioral Incident | | Δ Behavioral Incident | |
|---|---|---|---|---|
| Independent Variable | OLS | OLS | 2SLS | 2SLS |
| ΔRank (÷100) | -.033 | -.036 | -.047 | -.022 |
| | (.003) | (.003) | (.017) | (.041) |
| Test Score at Previous School (÷100) | -.030 | -.031 | -.034 | -.028 |
| | (.002) | (.002) | (.005) | (.010) |
| Male | .019 | .019 | .018 | .019 |
| | (.001) | (.001) | (.001) | (.002) |
| Black | .007 | .025 | .006 | .024 |
| | (.002) | (.003) | (.002) | (.003) |
| Hispanic | -.005 | .005 | -.006 | .005 |
| | (.002) | (.002) | (.002) | (.003) |
| Asian | -.021 | -.014 | -.021 | -.014 |
| | (.002) | (.002) | (.002) | (.003) |
| Other Race | .001 | .022 | -.004 | .018 |
| | (.012) | (.012) | (.012) | (.012) |
| Free Lunch | .012 | .011 | .011 | .011 |
| | (.002) | (.002) | (.002) | (.002) |
| English Language Learner | -.010 | -.009 | -.011 | -.009 |
| | (.003) | (.003) | (.003) | (.004) |
| Special Education | .010 | .005 | .006 | .005 |
| | (.003) | (.003) | (.004) | (.006) |
| Year Fixed Effects | Yes | Yes | Yes | Yes |
| School Fixed Effects | No | Yes | No | Yes |
| First Stage F-Stat | — | — | 5,117 | 1,224 |
| Shea's Partial R-Squared | — | — | .027 | .006 |
| R-Squared | .005 | .062 | .005 | .062 |
| Number of Observations | 241,734 | 241,734 | 233,247 | 233,247 |

B. Ranking Based on Math Test Scores

| | ΔBehavioral Incident | | ΔBehavioral Incident | |
|---|---|---|---|---|
| Independent Variable | OLS | OLS | 2SLS | 2SLS |
| ΔRank (÷100) | -.051 | -.060 | -.051 | -.019 |
| | (.003) | (.003) | (.016) | (.036) |
| Test Score at Previous School (÷100) | -.027 | -.030 | -.027 | -.023 |
| | (.002) | (.002) | (.004) | (.006) |
| Male | .022 | .021 | .021 | .021 |
| | (.001) | (.001) | (.001) | (.001) |
| Black | .007 | .024 | .006 | .024 |
| | (.002) | (.003) | (.002) | (.003) |
| Hispanic | -.005 | .004 | -.005 | .004 |
| | (.002) | (.002) | (.002) | (.002) |
| Asian | -.017 | -.010 | -.017 | -.012 |
| | (.002) | (.002) | (.002) | (.003) |
| Other Race | .003 | .024 | -.000 | .021 |
| | (.012) | (.011) | (.012) | (.012) |
| Free Lunch | .013 | .011 | .013 | .011 |
| | (.002) | (.002) | (.002) | (.002) |
| English Language Learner | -.007 | -.007 | -.006 | -.006 |
| | (.002) | (.002) | (.002) | (.002) |
| Special Education | .006 | .001 | .005 | .003 |
| | (.003) | (.003) | (.004) | (.004) |
| Year Fixed Effects | Yes | Yes | Yes | Yes |
| School Fixed Effects | No | Yes | No | Yes |
| First Stage F-Stat | — | — | 7,155 | 1,903 |
| Shea's Partial R-Squared | — | — | .036 | .009 |
| R-Squared | .006 | .061 | .006 | .060 |
| Number of Observations | 253,496 | 253,496 | 245,298 | 245,298 |

Notes: Entries are coefficients and standard errors from estimating the linear model, equation (13), by ordinary least squares and two-stage least squares, i.e. equation. The dependent variables is listed at the top of each column. The instrument for Δ Rank is the predicted change in rank based on school zoning regulations, as explained in the text and the Data Appendix. Panel A calculates a student's rank at his school based on ELA test scores, whereas panel B does so for Math test scores. Heteroskedasticity robust standard errors are reported in parentheses. In addition to the variables included in the table, indicator variables for missing values on each covariate are also included in the regressions. See the Data Appendix for the precise definition and source of each variable.

Appendix Table 1: Summary Statistics for NYCPS Data

| | Mean | SD | N |
|---|---|---|---|
| Change in behavioral incident indicator | .042 | .318 | 256,168 |
| Change in rank (English/Language Arts) | -.156 | 24.556 | 241,742 |
| Change in rank (math) | -.215 | 22.711 | 253,504 |
| Previous year test score (English/Language Arts) | 661.372 | 36.948 | 241,742 |
| Previous year test score (math) | 668.141 | 41.627 | 253,504 |
| White | .149 | .356 | 256,159 |
| Black | .314 | .464 | 256,159 |
| Hispanic | .394 | .489 | 256,159 |
| Asian | .139 | .346 | 256,159 |
| Other race | .004 | .063 | 256,159 |
| Male | .507 | .500 | 256,160 |
| Female | .493 | .500 | 256,160 |
| Free lunch | .830 | .376 | 210,290 |
| English Language Learner | .093 | .290 | 253,641 |
| Special education | .087 | .282 | 253,641 |
| Behavioral incident indicator | .089 | .285 | 256,168 |
| Previous year behavioral incidents | .083 | .531 | 256,168 |
| Observation from 2004-05 school year | .206 | .405 | 256,168 |
| Observation from 2005-06 school year | .194 | .395 | 256,168 |
| Observation from 2006-07 school year | .194 | .396 | 256,168 |
| Observation from 2007-08 school year | .201 | .401 | 256,168 |
| Observation from 2008-09 school year | .205 | .403 | 256,168 |
| Missing race | .000 | .006 | 256,168 |
| Missing sex | .000 | .006 | 256,168 |
| Missing free lunch status | .179 | .383 | 256,168 |
| Missing English Language Learner status | .010 | .099 | 256,168 |
| Missing special education status | .010 | .099 | 256,168 |
| Missing previous year behavioral incidents | .000 | .000 | 256,168 |

Notes: Entries are means and standard deviations together with the number of valid observations for each variable we use in the NYCPS data. For further details about the NYCPS data see Section A.1 in the Data Appendix.

Appendix Table 2: Summary Statistics for NELS Data

| | Mean | SD | N |
|---|---|---|---|
| Difference in behavioral incident dummy (eighth grade) | .007 | .519 | 19,782 |
| Difference in behavioral incident dummy (tenth grade) | .008 | .548 | 10,506 |
| Behavioral incident dummy (eighth grade) | .387 | .487 | 19,782 |
| Behavioral incident dummy (tenth grade) | .557 | .497 | 10,523 |
| Difference in behavioral incident factor (eighth grade) | .003 | .951 | 19,709 |
| Difference in behavioral incident factor (tenth grade) | .016 | .938 | 10,399 |
| Behavioral incident factor (eighth grade) | -.024 | .983 | 19,761 |
| Behavioral incident factor (tenth grade) | -.020 | .978 | 10,485 |
| Difference in class rank (eighth grade) | -.652 | 33.960 | 19,782 |
| Difference in class rank (tenth grade) | -.980 | 35.599 | 10,523 |
| Mean test score (eighth grade) | .060 | .868 | 19,213 |
| Mean test score squared (eighth grade) | .757 | .879 | 19,213 |
| Mean test score (tenth grade) | .093 | .870 | 10,241 |
| Mean test score squared (tenth grade) | .765 | .867 | 10,241 |
| Mean test score squared (eighth grade) | .757 | .879 | 19,213 |
| Mean test score squared (tenth grade) | .765 | .867 | 10,241 |
| Male | .497 | .500 | 21,297 |
| Female | .503 | .500 | 21,297 |
| White | .694 | .461 | 21,297 |
| Black | .111 | .314 | 21,297 |
| Hispanic | .118 | .322 | 21,297 |
| Asian | .058 | .233 | 21,297 |
| Other race | .020 | .139 | 21,297 |
| English Language Learner | .027 | .163 | 21,297 |
| Parents married | .721 | .449 | 21,297 |
| Parents divorced | .105 | .307 | 21,297 |
| Parents separated | .032 | .177 | 21,297 |
| Parents never married | .021 | .145 | 21,297 |
| Other parents' marital status | .120 | .326 | 21,297 |
| Parents' education: less than high school | .096 | .295 | 21,145 |
| Parents' education: high school graduate | .192 | .394 | 21,145 |
| Parents' education: some college | .403 | .490 | 21,145 |
| Parents' education: college graduate | .153 | .360 | 21,145 |
| Parents' education: M.A. | .096 | .295 | 21,145 |
| Parents' education: Ph.D., M.D., other | .060 | .238 | 21,145 |
| Public school (eighth grade) | .792 | .406 | 21,297 |
| Catholic school (eighth grade) | .102 | .303 | 21,297 |
| Independent/other private school (eighth grade) | .106 | .308 | 21,297 |
| School in urban area | .289 | .453 | 21,297 |
| School in suburban area | .421 | .494 | 21,297 |
| School in rural area | .290 | .454 | 21,297 |
| Lowest socioeconomic quartile | .229 | .420 | 15,897 |
| Second-lowest socioeconomic quartile | .242 | .428 | 15,897 |
| Second-highest socioeconomic quartile | .236 | .425 | 15,897 |
| Highest socioeconomic quartile | .293 | .455 | 15,897 |
| Missing sex | .000 | .000 | 21,297 |
| Missing race | .000 | .000 | 21,297 |
| Missing English Language Learner status | .000 | .000 | 21,297 |
| Missing parents' marital status | .000 | .000 | 21,297 |
| Missing parents' education | .007 | .084 | 21,297 |
| Missing school type | .000 | .000 | 21,297 |
| Missing school area | .000 | .000 | 21,297 |
| Missing socioeconomic quartile | .254 | .435 | 21,297 |
| Missing mean test score (eighth grade) | .029 | .167 | 19,782 |
| Missing mean test score (tenth grade) | .027 | .162 | 10,523 |

Notes: Entries are means and standard deviations together with the number of valid observations for each variable we use in the NELS data. For further details about the NELS data see Section A.3 in the Data Appendix, or the NELS website currently located at http://nces.ed.gov/surveys/nels88