

ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

*An International Society for the Advancement of Economic
Theory in its Relation to Statistics and Mathematics*

<http://www.econometricsociety.org/>

Econometrica, Vol. 76, No. 3 (May, 2008), 643–660

IDENTIFYING SOCIAL INTERACTIONS THROUGH CONDITIONAL VARIANCE RESTRICTIONS

BRYAN S. GRAHAM

*University of California–Berkeley, Berkeley, CA 94720, U.S.A. and National Bureau of Economic
Research*

The copyright to this Article is held by the Econometric Society. It may be downloaded, printed and reproduced only for educational or research purposes, including use in course packs. No downloading or copying may be done for any commercial purpose without the explicit permission of the Econometric Society. For such commercial purposes contact the Office of the Econometric Society (contact information may be found at the website <http://www.econometricsociety.org> or in the back cover of *Econometrica*). This statement must be included on all copies of this Article that are made available electronically or in any other format.

IDENTIFYING SOCIAL INTERACTIONS THROUGH CONDITIONAL VARIANCE RESTRICTIONS

BY BRYAN S. GRAHAM¹

This paper proposes a new method for identifying social interactions using conditional variance restrictions. The method provides a consistent estimate of the social multiplier when social interactions take the “linear-in-means” form (Manski (1993)). When social interactions are not of the linear-in-means form, the estimator, under certain conditions, continues to form the basis of a consistent test of the no social interactions null with correct large sample size. The methods are illustrated using data from the Tennessee class size reduction experiment Project STAR. The application suggests that differences in peer group quality were an important source of individual-level variation in the academic achievement of Project STAR kindergarten students.

KEYWORDS: Social interactions, social multiplier, peer group, effects, Project STAR, covariance models.

MANY ECONOMIC OUTCOMES—such as earnings, academic achievement, substance abuse, criminal behavior, and technology adoption—vary substantially across observationally similar groups. A long economics of education literature documents significant variation in mean academic achievement across different classrooms within the same school (e.g., Hanushek (1971), Rivkin, Hanushek, and Kain (2005)).² A straightforward explanation for this finding is the presence of classroom-level heterogeneity or student sorting (i.e., the dis-

¹This is a substantially revised version of material that initially appeared in Chapter 1 of my doctoral dissertation. That chapter previously circulated under the title “Identifying Social Interactions Through Excess Variance Contrasts.” I would like to thank Gary Chamberlain, Larry Katz, Michael Kremer, and Caroline Hoxby for their encouragement as well as many helpful comments, corrections, and suggestions. This revision has benefited enormously from the comments of a co-editor, three anonymous referees, Guido Imbens, Michael Jansson, and Jim Powell. Yan-nis Ioannides and Giulio Zanella both read an earlier draft quite carefully: I am very thankful for their editorial as well as substantive feedback. I am thankful to Jayne Zaharias for providing background documents and to Alan Krueger for discussions about design features of Project STAR. Useful discussions with William Brock, David Card, Steven Durlauf, Jim Heckman, Bo Honoré, Richard Murnane, Mark Watson, John Willett, Jeff Zabel, and participants in the Banff Workshop on the Microeconometrics of Spatial and Grouped Data, as well as seminars at Harvard, Harvard–MIT, UCSD, Dartmouth, Princeton, Stanford, Stanford GSB, UT–Austin, Berkeley, Chicago, Wisconsin, LSE, Yale, IFS, Toulouse–GREMAQ, Munich, Northwestern, Wharton, PIC, and USC, is also gratefully acknowledged. Financial support provided by a National Science Foundation Graduate Fellowship, the Program on Justice, Welfare and Economics at Harvard University, and by the Program of Fellowships for Junior Scholars, the MacArthur Research Network on Social Interactions and Economic Inequality. All the usual disclaimers apply.

²Another example is crime, which is endemic to some neighborhoods and negligible in other seemingly similar ones (cf. Glaeser, Sacerdote, and Scheinkman (1996)). Piketty (2000) and Becker and Murphy (2000) surveyed theoretical models that generate excess between-group variance.

tribution of unobserved teacher and student characteristics might vary across classrooms).

An alternative explanation for “excess” outcome variance across groups is that it mirrors the relative salience of social interactions or peer group effects. Social interactions are present if individual behavior is affected by reference (peer) group behavior, characteristics, or both (Manski (1993), Brock and Durlauf (2001)). If students within the same classroom learn from one another, then individual achievement levels will covary positively within a classroom and display excess variation across classrooms.³

The two rival explanations for excess variance—group-level heterogeneity and social interactions—are straightforward to understand, but difficult to distinguish empirically. Manski (1993, p. 532) highlighted the “difficulty of the identification problem” in this area, a problem that has affected the persuasiveness of empirical work (cf. Durlauf (2006, p. 164)).

This paper presents new methods for adducing the presence and magnitude of social interactions based on conditional variance restrictions. The proposed methods separately identify the social interactions component of any excess variance from that due to group-level heterogeneity and/or sorting.

The main idea is straightforward. For a certain class of social interaction models the unconditional between-group variance of outcomes is the sum of three terms. The first term equals the variance of any group-level heterogeneity (e.g., of teacher quality). The second term equals the between-group variance of any individual-level heterogeneity (e.g., the variance of average student ability across classrooms). These two terms are identical to those appearing in the textbook one-way error component model (e.g., Arellano (2003, Chap. 3)). The third term reflects the strength of any social interactions. When social interactions are present, between-group variation in outcomes should mirror between-group variation in “peer quality.” The final term therefore depends on the variance of peer quality across groups.

In the standard one-way error component model the variances of the individual- and group-level heterogeneity (or effects) are separately identified by the unconditional within- and between-group sample variances. The variance of individual heterogeneity is estimated from the within-group sample variance. This estimate and the between-group sample variance are then used to estimate the variance of any group-level heterogeneity. The variance of group-level heterogeneity is excess variance: it is identified by the residual between-group variance in outcomes that remains after accounting for the contribution of individual heterogeneity.

Difficulties arise in the presence of social interactions because there are two sources of excess variance (cf. Glaeser, Sacerdote, and Scheinkman (1996)):

³One could conceivably construct social interactions processes that would generate negative covariance in outcomes across members of the same group and hence too little between-group variance in outcomes. I use the excess variance language simply because it is intuitive and consistent with most applications. Formally, however, I consider two-sided alternatives.

the standard source due to group-level heterogeneity and that arising from variation in peer quality across groups. Any excess between-group variance in, say, student achievement, may reflect heterogeneity in teacher quality as well as, or instead of, heterogeneity in peer quality. The unconditional within- and between-group sample variances cannot form the basis of a valid test for social interactions. Fortunately comparisons of within- and between-group variances across two or more subpopulations of groups can be used to construct such a test. If the distribution of group-level heterogeneity is the same across such subpopulations, while the distribution of peer quality differs, then separate identification of the two sources of excess variance is possible.

The empirical application presented below provides intuition for the identification result. The application examines the effect of peer quality on kindergarten achievement using data from the Tennessee class size reduction experiment Project STAR. Classrooms of two sizes are observed in the data set: small and large. Students and teachers within participating schools were randomly assigned to one of the two types of classrooms.⁴ In large classrooms, clusters of talented students are typically offset by corresponding clusters of below average students, resulting in little variation in mean student ability. In small classrooms, however, groups composed of mostly above or below average students are more frequently observed, generating greater variation in mean ability. As a result, the variance of peer quality (as measured by average student ability in a classroom) is greater across the set of small than it is across the set of large classrooms. In contrast to peer quality, random assignment ensures that the distribution of teacher characteristics is similar across the two types of classrooms. Under some additional restrictions on the educational production function, discussed in detail below, class type provides a plausible source of exogenous variation in the variance of peer quality.

While other authors have emphasized the covariance implications of social interactions (e.g., Glaeser, Sacerdote, and Scheinkman (1996), Glaeser and Scheinkman (2001, 2003), Solon, Page, and Duncan (2000)), their work has not provided the basis for point identification, reflecting the inherent difficulties of discriminating outcome covariance due to peer spillovers from that due to group-level heterogeneity. The present paper's main contribution is to provide transparent conditions for point identification. The proposed conditions can be credible in a variety of settings.

Section 1 formalizes the heuristic discussion given above. It discusses identification and estimation in the context of the "linear-in-means" model of social interactions.⁵ The linear-in-means model is a member of a family of social in-

⁴The actual protocol was more involved. A detailed description of the Project STAR data and experimental protocol can be found in Word et al. (1990). All details relevant to the application can be found in Section 3 below. Further particulars of the extraction used in this paper can be found in the Supplemental appendix (Graham (2008)).

⁵Recent examples of economics research using the linear-in-means model include Sacerdote (2001), Duflo and Saez (2003), and Angrist and Lang (2004). A long history of sociological re-

teraction models. While the main identification result breaks down when the model does not take the linear-in-means form, Section 2 shows that it can serve as the basis of a valid test for social interactions as long as the true model is also a member of this family.⁶ Section 3 illustrates the proposed methods using the STAR data set. Section 4 ends with some thoughts about future research.

1. IDENTIFICATION AND ESTIMATION

We observe a random sample of N classrooms with the c th classroom consisting of M_c students. Academic achievement for the i th student in the c th classroom is given by

$$(1) \quad Y_{ci} = \alpha_c + (\gamma_0 - 1)\bar{\varepsilon}_c + \varepsilon_{ci},$$

where α_c represents classroom-level heterogeneity in, say, teacher effectiveness, ε_{ci} represents student-level heterogeneity arising from, say, variation in income, family background, and ability, and $\bar{\varepsilon}_c = \underline{\varepsilon}'_c \iota_{M_c} / M_c$ is the classroom mean of ε_{ci} (where $\underline{\varepsilon}_c = (\varepsilon_{c1}, \dots, \varepsilon_{cM_c})'$ and ι_{M_c} is a $M_c \times 1$ vector of ones). It is the dependence of own achievement on peer characteristics, indexed by the parameter γ_0 , that I define as social interactions. In addition to the $M_c \times 1$ vector of individual student test scores, \underline{Y}_c , we also observe a classroom's type: small ($W_c = 1$) or large ($W_c = 0$); α_c and ε_{ci} are latent variables unobserved by the econometrician.

Consider the effect on classroom average achievement of replacing a low- ε student with a high- ε student. In the presence of positive social interactions ($\gamma_0 > 1$), mean achievement will increase for two distinct reasons: for purely compositional reasons and because the presence of a high- ε student raises peer quality. The ratio of the full to the compositional effect equals γ_0 , a social multiplier (cf. Manski (1993), Glaeser and Scheinkman 2003).⁷

Below I impose restrictions on the conditional distribution of \underline{Y}_c given W_c which identify γ_0 . To assess the economic content of these restrictions, it is helpful to make any dependence of Y_{ci} on W_c explicit at the outset. A convenient way to do so is to posit the existence of teacher- and student-specific

search is surveyed by Duncan and Raudenbush (2001). Manski (1993), Brock and Durlauf (2001), Moffitt (2001), and Glaeser and Scheinkman (2001, 2003) all surveyed research in several disciplines that use variants of the linear-in-means framework.

⁶In the Supplemental appendix (Graham (2008)) I also show that the general approach of Section 1 can be adapted to identify other members of this family. A simple example, inspired by Lazear's (2001) disruption model of peer effects, illustrates how this might be done in practice.

⁷Note that γ_0 may be a composite function of multiple "structural" parameters. In Manski (1993) it depends on the strength of what he terms "exogenous" and "endogenous" social effects. Distinguishing between these two types of social interactions is not a goal of this paper.

potential effectiveness or productivity variables. Let $A_c(w)$ be a teacher’s *potential* effectiveness when assigned to classroom of type $W_c = w$. This setup allows the effectiveness of a teacher to change with class type in a heterogeneous way. For example, “law and order” oriented teachers may be relatively more effective in large classrooms than nurturing teachers. Similarly, I let $E_{ci}(w)$ denote a student’s *potential* productivity in a class of type $W_c = w$. Some students may be better suited, in relative terms, to small versus large classrooms and vice versa. For example kinesthetic learners may benefit from assignment to a small classroom more than visual learners.

Realized teacher effectiveness and student productivity are given by

$$(2) \quad \alpha_c = (1 - W_c)A_c(0) + W_cA_c(1),$$

$$\varepsilon_{ci} = (1 - W_c)E_{ci}(0) + W_cE_{ci}(1).$$

Since W_c is binary, (2) is without loss of generality.⁸

Substituting (2) into (1) gives a correlated random coefficients model (e.g., Heckman and Vytlacil (1998), Florens et al. (2007)) of

$$(4) \quad Y_{ci} = B_{0ci} + B_{1ci}W_c,$$

where

$$B_{0ci} = A_c(0) + (\gamma_0 - 1)\bar{E}_c(0) + E_{ci}(0),$$

$$B_{1ci} = (A_c(1) - A_c(0)) + (\gamma_0 - 1)(\bar{E}_c(1) - \bar{E}_c(0))$$

$$+ (E_{ci}(1) - E_{ci}(0)).$$

Equation (4) allows for the causal effect of class type (W_c) on achievement (Y_{ci}) to be heterogeneous, possibly depending on teacher ($A_c(0)$, $A_c(1)$), own ($E_{ci}(0)$, $E_{ci}(1)$), and/or peer ($\bar{E}_c(0)$, $\bar{E}_c(1)$) characteristics.

In this paper I focus on the identifying content of $\mathbb{V}(\underline{Y}_c|W_c)$.⁹ Manski (1993) and Brock and Durlauf (2001) discussed conditional mean restrictions which identify γ_0 . The conditional variance of teacher effectiveness and student pro-

⁸If W_c is continuous, (2) can be replaced with the assumption that $A_c(w)$ and $E_{ci}(w)$ are low-order polynomial functions of w with random coefficients (cf. Florens et al. (2007)):

$$(3) \quad \alpha_c = A_c(W_c), \quad A_c(w) = U_{0c} + U_{1c}w + \dots + U_{Pc}w^P,$$

$$\varepsilon_{ci} = E_{ci}(W_c), \quad E_{ci}(w) = V_{0ci} + V_{1ci}w + \dots + V_{Pci}w^P.$$

Observe that while (2) is unrestrictive for W_c binary, the above specification is restrictive for W_c continuous. This extra restrictiveness, however, does allow one to weaken Assumption 1.2 given below. This extension is straightforward and not developed here.

⁹Uppercase letters denote random variables and lowercase letters denote their specific realizations. I also use the conventions $\mu_Y(x) = \mathbb{E}[Y|x] = \mathbb{E}[Y|X = x]$, $\sigma_Y^2(x) = \mathbb{V}(Y|x)$, and $\sigma_{YZ}^2(x) = \mathbb{C}(Y, Z|x)$.

ductivity given class size and type equals

$$(5) \quad \mathbb{V}(\alpha_c, \underline{\varepsilon}'_c | m, w) = \begin{pmatrix} \sigma_\alpha^2(m, w) & \sigma_{\alpha\varepsilon}(m, w)\iota_m \\ \sigma_{\alpha\varepsilon}(m, w)\iota_m & (\sigma_\varepsilon^2(m, w) - \sigma_{\varepsilon\varepsilon}(m, w))I_m + \sigma_{\varepsilon\varepsilon}(m, w)\iota_m\iota'_m \end{pmatrix},$$

where I_m is an $m \times m$ identity matrix, $\sigma_{\varepsilon\varepsilon}(m, w) = \mathbb{C}(\varepsilon_{ci}, \varepsilon_{cj} | m, w) = (1 - w)\sigma_{E(0)E(0)}^2(m, w) + \sigma_{E(1)E(1)}^2(m, w)w$, and $\sigma_\alpha^2(m, w) = (1 - w)\sigma_{A(0)}^2(m, w) + \sigma_{A(1)}^2(m, w)w$, with $\sigma_\varepsilon^2(m, w)$ and $\sigma_{\alpha\varepsilon}(m, w)$ defined similarly.

In the context of the STAR application, $\sigma_\alpha^2(m, w)$ captures heterogeneity in teacher effectiveness among those teachers assigned to classrooms of a given type ($W_c = w$) and size ($M_c = m$), $\sigma_\varepsilon^2(m, w)$ captures heterogeneity in individual productivity across students, $\sigma_{\alpha\varepsilon}(m, w)$ any covariance between teacher effectiveness and learning productivity induced by the assignment process of teachers and students to classrooms, and $\sigma_{\varepsilon\varepsilon}(m, w)$ captures the within-classroom covariance of student productivity. These latter two terms might differ from zero if, for example, teachers and students systematically match on the basis of characteristics and if students sort into homogeneous groups (e.g., due to ability tracking).

From (1) and (5), the conditional variance of observed outcomes is

$$(6) \quad \mathbb{V}(\underline{Y}_c | m, w) = \lambda^2(m, w)I_m + \left[\tau_0^2(m, w) + (\gamma_0^2 - 1) \frac{\lambda^2(m, w)}{m} \right] \iota_m \iota'_m,$$

where $\lambda^2(m, w) = \sigma_\varepsilon^2(m, w) - \sigma_{\varepsilon\varepsilon}(m, w)$ and $\tau_0^2(m, w) = \sigma_\alpha^2(m, w) + 2\gamma_0 \times \sigma_{\alpha\varepsilon}(m, w) + \gamma_0^2 \sigma_{\varepsilon\varepsilon}(m, w)$.

Under random sampling, $\mathbb{V}(\underline{Y}_c | m, w)$ is nonparametrically identified. The structure of $\mathbb{V}(\underline{Y}_c | m, w)$ is potentially informative for γ_0 . Observe that it is composed of just two unique elements: a common diagonal term of $\lambda^2(m, w) + \tau_0^2(m, w) + (\gamma_0^2 - 1)\lambda^2(m, w)/m$ and a common off-diagonal term, which measures outcome covariance across peers, of $\tau_0^2(m, w) + (\gamma_0^2 - 1)\lambda^2(m, w)/m$.¹⁰

The difference between these two terms identifies $\lambda^2(m, w)$. Unfortunately $\tau_0^2(m, w)$ and γ_0^2 are not separately identified without further assumptions. This lack of identification reflects the problem discussed in the introduction: outcome covariance across members of the same social group may reflect the presence of social interactions ($\gamma_0 \neq 1$) or simply some combination of group-level heterogeneity, sorting, and/or matching ($\tau_0^2(m, w) \neq 0$).¹¹

¹⁰Equicorrelation in outcomes across students within the same classrooms is a consequence of within-classroom exchangeability of students.

¹¹This nonidentification result is a conditional variance analog of the corollary to Manski's (1993, p. 535) Proposition 1.

Due to the equicorrelated structure of $\mathbb{V}(\underline{Y}_c|m, w)$, it is convenient to work with the within- and between-group squares of the data. Define

$$(7) \quad G_c^w = \frac{1}{M_c} \frac{1}{M_c - 1} \sum_{i=1}^{M_c} (Y_{ci} - \bar{Y}_c)^2, \quad G_c^b = (\bar{Y}_c - \mu_Y(W_c))^2,$$

where $\mu_Y(w)$ is mean achievement in classes of type $W_c = w$. The conditional expectations of G_c^w and G_c^b given $W_c = w$ are

$$(8) \quad \mathbb{E}[G_c^w|w] = \mathbb{E}\left[\frac{\lambda^2(M_c, W_c)}{M_c} \mid w\right],$$

$$\mathbb{E}[G_c^b|w] = \tau_0^2(w) + \gamma_0^2 \mathbb{E}\left[\frac{\lambda^2(M_c, W_c)}{M_c} \mid w\right],$$

where $\tau_0^2(w) = \sigma_\alpha^2(w) + 2\gamma_0\sigma_{\alpha\epsilon}(w) + \gamma_0^2\sigma_{\epsilon\epsilon}(w)$.

Consider the ratio of the *difference* in between-group squares across small and large classrooms to the corresponding difference in within-group squares:

$$(9) \quad \frac{\mathbb{E}[G_c^b|W_c = 1] - \mathbb{E}[G_c^b|W_c = 0]}{\mathbb{E}[G_c^w|W_c = 1] - \mathbb{E}[G_c^w|W_c = 0]}$$

$$= \gamma_0^2 + \frac{\tau_0^2(1) - \tau_0^2(0)}{\mathbb{E}[G_c^w|W_c = 1] - \mathbb{E}[G_c^w|W_c = 0]}.$$

The left-hand side of (9) is the population analog of a Wald estimator. Its numerator is a contrast of observed or actual between-classroom variance in student achievement across small and large type classrooms. If $\tau_0^2(1) = \tau_0^2(0)$, this contrast is purged of the influence of teacher heterogeneity, matching, and sorting. In that case it solely reflects differences in the variance of peer quality across the two types of classrooms, *as amplified by social interactions*. The denominator also equals the difference in the variance of peer quality across the two types of classrooms, but unamplified by social interactions.

Equation (9) identifies γ_0^2 if $\tau_0^2(1) = \tau_0^2(0)$. There are several sets of primitive assumptions which will guarantee this condition. Here I work with the following three assumptions.

ASSUMPTION 1.1 —(Independent Random Assignment):

$$F(A_c(w), \underline{E}_c(w)|W_c) = F(A_c(w)) \times \prod_{i=1}^{M_c} F(E_{ci}(w)|W_c),$$

$$w \in \mathcal{W} = \{0, 1\}.$$

ASSUMPTION 1.2 —(Stochastic Separability):

$$A_c(1) \sim A_c(0) + \kappa_0.$$

ASSUMPTION 1.3 —(Peer Quality Variation):

$$\mathbb{E}[G_c^w | W_c = 1] \neq \mathbb{E}[G_c^w | W_c = 0].$$

To develop some intuition regarding when Assumption 1.1 is likely to be reasonable it is helpful to think about the classroom formation process. Classroom formation consists of two, potentially interrelated steps (cf. Becker and Murphy (2000, Chap. 5)). First students sort into groups on the basis of peer characteristics. Second they sort into groups on the basis of teacher characteristics. Assumption 1.1 places restrictions on these processes. In particular, it follows when teacher and student assignment to classrooms follows a “double randomization” procedure.¹² In the first step of the procedure the planner chooses a feasible distribution of classroom types (W_c) and sizes (M_c). In the second step students and teachers are independently and randomly assigned to classrooms. This latter step ensures that (i) $A_c(w)$ and $E_{ci}(w)$ are independent of W_c , (ii) $E_{ci}(w)$ is independently distributed within classrooms (i.e., no sorting), and (iii) that $A_c(w)$ is independent of $E_{ci}(w)$ (i.e., no matching). The absence of matching and sorting implies that

$$(10) \quad \sigma_{\alpha\epsilon}(w) = \sigma_{\epsilon\epsilon}(w) = 0,$$

in which case $\tau_0^2(w) = \sigma_\alpha^2(w)$.

Assumption 1.2 states that the marginal distributions of $A_c(1)$ and $A_c(0)$ are equal up to an additive shift. It follows from assuming, for example, that the effect of class type on student outcomes does not depend on teacher characteristics (i.e., $A_c(1) = A_c(0) + \kappa_0$). It also holds under much weaker restrictions.¹³ For example, it holds if

$$\begin{pmatrix} A_c(1) \\ A_c(0) \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \mu_A + \kappa_0 \\ \mu_A \end{pmatrix}, \sigma_A^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right),$$

which in turn implies that $\mathbb{E}[A_c(1)|A_c(0)] = \kappa_0 + \mu_A + \rho[A_c(0) - \mu_A]$. Thus Assumption 1.2 does allow for individual teachers to have a comparative advantage ($\rho \neq 0$) in either small or large classroom instruction.¹⁴ However, it does restrict the population variance of small and large classroom teacher effectiveness to be the same (i.e., $\sigma_{A(0)}^2 = \sigma_{A(1)}^2$). This is a technological restriction

¹²I am thankful to a co-editor for suggesting the double randomization terminology.

¹³I am grateful to Michael Jansson for help in formulating and understanding this assumption.

¹⁴If $\rho < 0$, teachers with above average effectiveness in large classrooms tend to be relatively less effective in small classrooms; if $\rho > 0$, the opposite pattern holds.

on the educational production function and, as such, no feasible assignment scheme can ensure its satisfaction. If it is satisfied $\sigma_\alpha^2(w) = \sigma_\alpha^2$, which when combined with (10), implies that $\tau_0^2(w)$ is constant in w .

It is important to emphasize that Assumptions 1.1 and 1.2 do not imply that the effect of class type on student outcomes is homogeneous. Indeed it may vary with both teacher and student, either own or peer, characteristics. This is important as Project STAR has produced evidence that the effect of class size depends on student race, gender, and socioeconomic background (e.g., Krueger and Whitmore (2001)).

Assumption 1.3 is a rank restriction. It requires that the variance of peer quality *differs* across the two types of classrooms. Together Assumptions 1.1, 1.2, and 1.3 formalize the requirement that W_c generates exogenous variation in the variance of peer quality across small and large classrooms. Consequently, the population least squares regression of G_c^b onto a constant and $\mathbb{E}[G_c^w|W_c]$ identifies γ_0^2 . While $\mathbb{E}[G_c^w|W_c]$ is not directly observed, (8) shows that it is identified, giving the following result.

PROPOSITION 1.1: *Under Assumptions 1.1, 1.2, and 1.3, γ_0^2 is identified by the left-hand side of (9).*

PROOF: The result follows directly from inspection of (9). Q.E.D.

Proposition 1.1 can be derived under weaker conditions than those imposed by Assumptions 1.1–1.3.¹⁵ However, I prefer the formulation given here for three reasons. First, Assumption 1.1 is constructive; it can be automatically satisfied by a well-defined assignment mechanism. Second, Assumption 1.2 is simple to interpret and sensitivity to deviations from it can be easily assessed.^{16,17} Third, I show in the next section that the Wald estimate, under Assumptions 1.1 and 1.2, can form the basis of a valid test for social interactions even if the true model does not take the linear-in-means form. Assumption 1.3 is, of course, testable.

Under Assumptions 1.1 and 1.2 we have, for $\theta_0 = (\tau_0^2, \gamma_0^2)'$ and $\rho(Z_c, \theta) = G_c^b - \tau^2 - \gamma^2 G_c^w$, the conditional moment restriction

$$(11) \quad \mathbb{E}[\rho(Z_c, \theta_0)|W_c] = 0,$$

¹⁵All that is required is that $\tau_0^2(w)$ is constant in w . This necessarily requires stochastic separability (Assumption 1.2), but need not exclude the possibility of sorting and matching across classrooms of the same type (Assumption 1.1). For example, if we were willing to assume that $E_{ci}(1) \sim E_{ci}(0) + \xi_0$, then it would be reasonable to assume that $\sigma_{\alpha_{se}}(w)$ and $\sigma_{\alpha_{pe}}(w)$, while nonzero, are constant in w .

¹⁶A formal sensitivity analysis of this assumption is provided in the Supplemental appendix (Graham (2008)).

¹⁷If W_c is continuous, Assumption 1.2 can be weakened under a strengthening of Assumption 1.3. This extension is not pursued here.

which can be used to construct a generalized method-of-moments (GMM) estimator.

In many applications Assumptions 1.1 and 1.2 will only be plausible after first conditioning on a set of additional classroom-level covariates. Let $W_c = (W_{1c}, W'_{2c})'$ with W_{2c} now denoting class type and W_{1c} denoting additional classroom-level covariates (e.g., observed teacher and school characteristics). Replacing Assumptions 1.1, 1.2, and 1.3 with analogous statements that apply only to subpopulations homogeneous in W_{1c} leads to a straightforward generalization of Proposition 1.1. Defining $\rho(Z_c, \tau^2(W_{1c}), \gamma^2) = G_c^b - \tau^2(W_{1c}) - \gamma^2 G_c^w$, where $\tau^2(W_{1c})$ is some smooth function of W_{1c} , then yields the conditional moment restriction

$$(12) \quad \mathbb{E}[\rho(Z_c, \tau_0^2(W_{1c}), \gamma_0^2) | W_c] = 0.$$

Under standard regularity conditions GMM applied to either (11) or (12) will consistently estimate γ_0^2 .

2. ALTERNATIVE MODELS OF SOCIAL INTERACTIONS

Proposition 1.1 is specific to the linear-in-means model. While this model is used in the majority of empirical social interactions research, it does have a number of unattractive features. It implies that average outcomes are invariant to reallocations of individuals to different groups (although they do affect their marginal variance). The model is thus ill suited for evaluating equity and efficiency trade-offs that might arise in, say, the context of school choice programs (cf. Graham, Imbens, and Ridder (2006, 2007)).

Furthermore, although the dependence of own outcomes on mean peer characteristics can be given a choice-theoretic motivation, as in Brock and Durlauf (2001), different economic theories naturally suggests different statistical models. Lazear's (2001) work on educational production, for example, posits that peer spillovers operate through a disruption externality. This theory suggests that a more appropriate statistical model of student achievement might be

$$(13) \quad Y_{ci} = \alpha_c + (\gamma_0 - 1)h(\varepsilon_{ci}, \underline{\varepsilon}_c) + \varepsilon_{ci},$$

where $h(\varepsilon_{ci}, \underline{\varepsilon}_c) = \min\{\varepsilon_c\}$. Equation (13) is consistent with the intuition that low productivity students slow the rate of their classmate's learning; perhaps by asking questions of the teacher to which their classmates already know the answer. In this model average outcomes are maximized by perfectly stratifying groups by ability. In contrast, such stratification has no outcome benefit and a substantial equity cost when spillovers take the linear-in-means form.

The "bad apple" model, like the linear-in-means one, has the implication of a homogeneous peer effect (i.e., the effect of peers on own outcomes is the same for all group members). A model which allows for heterogeneous

peer effects is (13) with $h(\varepsilon_{ci}, \underline{\varepsilon}_c) = (\bar{\varepsilon}_c - \varepsilon_{ci})^2$. In this model the peer effect is individual-specific, depending on the distance between own and average classroom ability.¹⁸

In this section I show that the estimator introduced above can form the basis of a consistent test for social interactions when the true model is not of the linear-in-means form. This requires that the true model belongs to the family given by (13), with $h(\cdot, \cdot)$ invariant to permutations of the elements of its second argument, $\underline{\varepsilon}_c$. This restriction arises naturally out of exchangeability considerations. This family includes several focal models as special cases.¹⁹

The intuition for this robustness result is straightforward. Under the null of no social interactions $\gamma_0^2 = 1$, all members of (13) simplify to $Y_{ci} = \alpha_c + \varepsilon_{ci}$. The estimator of Section 1 will be consistent for $\gamma_0^2 = 1$ in such cases. This ensures the correct large sample size of tests based on $\hat{\gamma}^2$.

PROPOSITION 2.1 —(Testing in the Presence of Misspecification): *Let γ_*^2 denote the right-hand side of (9) when the true data generating process is given by (13). Under Assumptions 1.1, 1.2, and 1.3 a test of the null that $\gamma_*^2 = 1$ is a test of the null that $\gamma_0 = 1$ with correct large sample size.*

PROOF: We have, using Assumptions 1.1 and 1.2,

$$\begin{aligned} \mathbb{E}[G_c^b|w] &= \sigma_\alpha^2 + (\gamma_0 - 1)^2 \text{Var}(\bar{h}(\underline{\varepsilon}_c)|w) + \mathbb{E}\left[\frac{\lambda^2(M_c, W_c)}{M_c} \middle| w\right] \\ &\quad + 2(\gamma_0 - 1) \text{Cov}(\bar{h}(\underline{\varepsilon}_c), \bar{\varepsilon}_c|w), \end{aligned}$$

where $\bar{h}(\underline{\varepsilon}_c) = \sum_{i=1}^{M_c} h(\varepsilon_{ci}, \underline{\varepsilon}_c)/M_c$. The probability limit in question is therefore, using Assumption 1.3, given by

$$\begin{aligned} \gamma_*^2 &= 1 + (\gamma_0 - 1)^2 \frac{\text{Var}(\bar{h}(\underline{\varepsilon}_c)|W_c = 1) - \text{Var}(\bar{h}(\underline{\varepsilon}_c)|W_c = 0)}{\mathbb{E}[G_c^w|W_c = 1] - \mathbb{E}[G_c^w|W_c = 0]} \\ &\quad + 2(\gamma_0 - 1) \frac{\text{Cov}(\bar{h}(\underline{\varepsilon}_c), \bar{\varepsilon}_c|W_c = 1) - \text{Cov}(\bar{h}(\underline{\varepsilon}_c), \bar{\varepsilon}_c|W_c = 0)}{\mathbb{E}[G_c^w|W_c = 1] - \mathbb{E}[G_c^w|W_c = 0]}, \end{aligned}$$

which will generally differ from 1 only if $\gamma_0 \neq 1$. Correct asymptotic size follows from the fact that under the null $\gamma_*^2 = 1$ for the entire family of models. Q.E.D.

Proposition 2.1 implies that even if the true data generating process is not of the linear-in-means form, we can still, under certain conditions, test for social

¹⁸Levin (2001), citing social cognitive learning research suggesting that students benefit from homogeneous grouping, worked with a peer effect model in this spirit.

¹⁹It is also possible to establish identification of γ_0 for specific members of the family defined by (13). An example of how to do this in practice is provided in the Supplemental appendix.

interactions by proceeding as if it were. Of course, this test has no particular optimality properties and may have low power for some alternatives.²⁰

3. APPLICATION: PEER EFFECTS AND LEARNING AMONGST PROJECT STAR KINDERGARTEN STUDENTS

This section presents an empirical application of the identification strategy outlined in Section 1. The application uses data from the class size reduction experiment Project STAR to assess the role of peer spillovers in learning. While the STAR experiment was originally designed to assess the effect of class size on student achievement, its design is fortuitously well suited to studying peer group effects.

Full details on Project STAR are provided by Word et al. (1990), from which the following information is drawn. In the fall of 1985, entering kindergarten students in each of 79 project schools, located throughout the state of Tennessee, were randomly assigned to one of three class types within their school: small, with 13 to 17 students; regular, with 22 to 25 students; and regular with a full time teacher's aide, also with 22 to 25 students. Teachers were randomly assigned to one of the three class types in a second step. Schools participating in the project were required to be large enough to accommodate at least three kindergarten classes. Legislation also specified stratification across inner city, urban, suburban, and rural schools.

Random assignment, for both students and teachers, was to class type *not* to specific classrooms. For 31 of the 79 participating schools this distinction is without content, since in those schools there were only three classrooms (one of each type). For the 48 schools with more than three classrooms, teachers and students could have been, in principle, nonrandomly allocated across classrooms of the *same type*.

During the first year, 6,325 students across 325 different classrooms participated in the project. At the end of the year *Stanford Achievement Tests in Mathematics and Reading* were administered (no pre-intervention test scores are available). I have normalized the total scaled math and reading test scores by their sample mean and standard deviation.²¹ The extraction used below includes data on 6,172 students across 317 classrooms, the maximum usable sample.²² The analysis reported below uses only the kindergarten data.

I base estimation on the conditional moment restriction given by equation (12). I include in W_{1c} a vector of school dummy variables indicating in

²⁰Asymptotic power greater than size will typically follow if either $\text{Var}(\bar{h}(\underline{\varepsilon}_c)|W_c = 1) \neq \text{Var}(\bar{h}(\underline{\varepsilon}_c)|W_c = 0)$ or $\text{Cov}(\bar{h}(\underline{\varepsilon}_c), \bar{\varepsilon}_c|W_c = 1) \neq \text{Cov}(\bar{h}(\underline{\varepsilon}_c), \bar{\varepsilon}_c|W_c = 0)$ hold. When, for example, $h(\varepsilon_{ci}, \underline{\varepsilon}_c) = \bar{\varepsilon}_c/\sqrt{M_c}$ and the variance of ε_{ci} is constant in W_c , both these conditions fail and the linear-in-means test will have power equal to size.

²¹Scaled scores are constructed by the test publisher such that a 1-point increase in test scores represents the same amount of incremental knowledge at all points along the test score scale.

²²Full details of the extraction used can be found in the Supplemental appendix.

which of the 79 Project STAR schools the c th classroom was located as well as a binary variable indicating whether the c th classroom is of the regular-with-aide type and let W_{2c} be a binary indicator for whether the c th classroom is of the small type. Regular classrooms are the excluded group.

Validity of (12) requires that Assumptions 1.1 and 1.2 hold within subpopulations of classrooms homogeneous in W_{1c} . Identification requires that Assumption 1.3 hold.

In the 31 schools with just three kindergarten classrooms, Assumption 1.1 follows directly from the experiment's protocol. In the remaining 48 schools, within-class-type sorting and matching could, in principle, invalidate restriction (12). This seems unlikely as even in larger schools STAR's protocol left little scope for systematic classroom formation. Furthermore, as I observe students during their very first year of formal schooling, it seems likely that administrators would have had little a priori information upon which to systematically group students or match them with teachers. Nevertheless, the absence of explicit double randomization across the set of larger schools is a concern. Therefore I use the 31 smaller schools, where the experiment's protocol explicitly induced double randomization, to directly test my maintained assumption of "as if" double randomization across all schools.

Assumption 1.2 is a restriction on the educational production function. By virtue of random assignment, the distribution of teacher *characteristics* should be similar across small, regular, and regular-with-aide classrooms within the same school. Constancy of the distribution of teaching effectiveness, α_c , across small and regular classrooms follows if teacher characteristics and class type are stochastically separable in the educational production function.²³ Although this allows for heterogeneity across teachers in their effectiveness across small and large classrooms, it restricts the population variance of teacher potential effectiveness in the two types of classrooms to be the same. This is a strong assumption. It seems a priori plausible that the salience of specific teacher characteristics for student achievement varies with class size.²⁴ For example, the importance of a teacher's ability to maintain order and discipline may grow with class size. Similarly, a teacher's ability to customize his pedagogy to the needs of individual students may be more important in smaller classrooms.

I address possible bias due to nonseparability of teacher attributes and class type in educational production in two ways. First, in the Supplemental appendix (Graham (2008)) I undertake a formal sensitivity analysis. There I conclude that the degree of nonseparability would have to be implausibly large to fully explain the magnitude of my estimates of γ_0^2 . Second, also in the Supplemental appendix, I construct a formal test for nonseparability bias and fail to reject Assumption 1.2.

²³Note that the variance of teaching quality/effectiveness is allowed to differ across schools as well as in those classrooms which include a teacher's aide.

²⁴I thank Gary Chamberlain (and many subsequent seminar participants) for clearly articulating this concern to me.

The results of both the sensitivity analysis and the formal test are consistent with qualitative evidence collected in conjunction with Project STAR. Classroom observations were taken for a subsample of teachers. In the final Project STAR report Word et al. (1990, p. 123) concluded that “a reader given an unmarked narrative would be hard pressed to decide whether it was [from] a regular or small class. Teachers apparently made few changes in curriculum, lesson format, or methods based on class size.” Complementary statistical analyses concluded that “teachers appeared to maintain the same pattern of instruction regardless of class size” (Word et al. (1990, p. 121)). Since Project STAR teachers appear to have used the same instructional technology regardless of class size, the assumption that the mapping from teacher characteristics to outcomes is stochastically separable from class size seems reasonable.²⁵

Finally, as discussed in the introduction, Assumption 1.3 is intuitive in the STAR setting: under random assignment the distribution of average ability across small classrooms should be more dispersed than the corresponding distribution across large classrooms. This assumption is testable.

I make the additional parametric assumptions that $\mu_Y(W_c) = W'_{1c}\pi_1 + W'_{2c}\pi_2$, and that $\tau_0^2(W_{1c}) = W'_{1c}\beta_0$, and base estimation on the unconditional moment restriction

$$(14) \quad \mathbb{E}[W_c(G_c^b - W'_{1c}\beta_0 - G_c^w\gamma_0^2)] = 0.$$

Table I reports estimates of γ_0^2 using the Project STAR math and reading data.²⁶ Here I focus on the column 1 mathematics results (the reading results are qualitatively similar, although less well identified). The estimate of γ_0^2 in this case equals roughly 3.5, suggesting a social multiplier of approximately 1.9.²⁷ The no social interactions null is easily rejected at the 5 percent level. The γ_0^2 parameter appears to be strongly identified. The first stage F -statistic is over 40, suggesting that weak instruments are not a concern. The results for reading achievement are similar, albeit less precisely estimated.

The estimates of γ^2 reported in Table I suggest that social interactions substantively contributed to the learning process of Project STAR kindergarten students. To get a rough sense of the magnitude of the implied social multiplier, observe that under random assignment to classrooms a 1 standard deviation change in peer quality, $\bar{\varepsilon}_c$, is given by $\sigma_\varepsilon/\sqrt{M}$.²⁸ Thus the effect on student

²⁵Additional evidence from the education and economics of education literature relevant to assessing the plausibility of Assumption 1.2 is surveyed in the Supplemental appendix (Graham (2008)).

²⁶The feasible estimator replaces G_c^b with $\widehat{G}_c^b = (\widehat{Y}_c - W'_{1c}\widehat{\pi}_1 - W'_{2c}\widehat{\pi}_2)^2$, where $\widehat{\pi}_1$ and $\widehat{\pi}_2$ are least squares estimates.

²⁷In principle the sign of γ_0 is undetermined, but it is likely positive in the present setting.

²⁸This calculation makes the auxiliary assumption that the variance of ε_{ci} is homoscedastic with respect to class size.

TABLE I
GMM ESTIMATES OF γ^2 FOR NORMALIZED MATH AND READING ACHIEVEMENT TEST SCORES^a

	(1) Math	(2) Math	(3) Reading	(4) Reading
γ^2	3.47 (1.03)	2.33 (1.13)	5.28 (2.48)	2.11 (1.74)
Regular-with-aide	0.01 (0.02)	-0.00 (0.02)	0.03 (0.03)	0.02 (0.03)
Large-school \times small	—	0.04 (0.04)	—	0.07 (0.05)
Large-school \times regular-with-aide	—	0.01 (0.040)	—	-0.00 (0.04)
p -value $H_0: \gamma^2 = 1$	0.018	0.24	0.086	0.523
$F_{(df1,df2)}$ 1st-Stage	46.8 _(1,236)	19.0 _(1,234)	11.3 _(1,236)	6.0 _(1,234)
Number of classrooms	317	317	317	317
School fixed effects	Yes	Yes	Yes	Yes

^aEstimates based on the sample described in the text. Column 1 reports GMM estimates of γ_0^2 for math achievement based on (14) with the included instrument, W_{1c} , equaling a binary variable for whether a classroom is of the regular-with-aide type, and full vector school dummies, and the excluded instrument, W_{2c} , equaling a binary variable for whether a classroom is of the small type. G_c^b is formed by replacing $\mu_Y(W_c)$ in (7) with the fitted value associated with the ordinary least squares fit of Y_{ci} onto W_{1c} and W_{2c} . Replacing $\mu_Y(W_c)$ with a consistent estimate does not alter the first order asymptotic sampling properties of $\hat{\gamma}^2$. Column 2 reports results for a model where W_{1c} is augmented by two interaction variables: interactions of an indicator variable for whether a school contains more than three kindergarten classrooms with each of the two class type indicator variables. Columns 3 and 4 report parallel estimates using the reading achievement data.

achievement of a 1 standard deviation change in peer quality under this assignment regime *relative* to a 1 standard deviation change in own ability is given by $(\gamma_0 - 1)/\sqrt{M}$. Across classrooms of 15 and 22 students this corresponds to relative changes in mathematics achievement of 0.22 and 0.18, respectively. For comparison purposes, the relative effect on mean achievement of being assigned to a small classroom equals 0.20 (i.e., π_2/σ_ϵ).²⁹ The achievement benefit of assignment to a regular classroom with above average peers, was about equal to that of assignment to a small classroom of average peers.

To test for the possibility that the small class type dummy variable is an invalid instrument for G_c^w , I exploit the 31 three classroom schools where teachers and students were randomly assigned to classrooms. I reestimate γ_0^2 after augmenting W_{1c} with the interactions of a dummy for whether a classroom is in one of the 48 larger schools with the small and regular-with-aide class type dummies. If within-class-type student sorting and/or student-teaching matching in larger schools is driving my point estimates, then the between-classroom

²⁹This number equals the preliminary (not reported) estimate of π_2 in $\mu_Y(W_c) = W'_{1c}\pi_1 + W'_{2c}\pi_2$ divided by the square root of the sample average of $M_c \cdot G_c^w$, which is a consistent estimate of σ .

variance contrast across small and regular classrooms in larger (more than three classrooms) schools should differ significantly from that observed in small (exactly three classrooms) schools. As is apparent from the coefficient on the large school by small classroom interaction (column 2), there is no evidence of such a difference. The estimate of γ_0^2 reported in column 2 is identified by small school variance contrasts alone. It is insignificantly different from the column 1 estimate, albeit also insignificantly different from the no social interactions null of one.

The balance of the evidence suggests that Assumptions 1.1, 1.2, and 1.3 are reasonable in the context of the Project STAR experiment with the point estimates of γ_0^2 suggesting sizable peer spillovers.

4. SUMMARY

This paper has proposed a new method for identifying social interactions models using conditional variance restrictions. The method provides a consistent estimate of the social multiplier when social interactions take the linear-in-means form (Manski (1993)). When social interactions are not of the linear-in-means form, the estimator, under certain conditions, continues to form the basis of a consistent test of the no social interactions null. The proposed methods are straightforward to implement. The GMM representation of the identifying conditions makes inference standard.

The methods are illustrated using data from Tennessee class size reduction experiment Project STAR. The application suggests that differences in peer group quality were an important source of individual-level variation in the academic achievement for Project STAR kindergarten students. This result contributes to a longstanding and controversial debate on the salience of peer group effects in the learning process (e.g., Coleman (1966), Hoxby (2002), Angrist and Lang (2004)).

An obvious concern is that the range of applications for which the identification approach outlined above is appropriate is limited. Such concerns need to be contextualized within what is a large, but very problematic, empirical literature. While this literature has produced substantial evidence of peer spillovers, little of it is credible. Durlauf (2006, p. 164), in assessment shared by many other reviewers, argued that “there is little reason why a skeptic should be persuaded to change his mind by the statistical evidence [on social interactions] currently available.”

In such a light the methods presented here are constructive: a specific mechanism of group formation, combined with clear technological assumptions, motivates a straightforward estimation strategy. Applications of the proposed method to observational data will, of course, be more challenging. Finding subpopulations of social groups where assignment to groups is as if random is nontrivial. However, the issues involved are not fundamentally different from, nor more intractable than, those encountered when, say, evaluating social programs.

Dept. of Economics, University of California–Berkeley, 549 Evans Hall 3880, Berkeley, CA 94720, U.S.A. and National Bureau of Economic Research; bgraham@econ.berkeley.edu; <http://www.econ.berkeley.edu/~bgraham/>.

Manuscript received June, 2005; final revision received October, 2007.

REFERENCES

- ANGRIST, J. D., AND K. LANG (2004): “Does School Integration Generate Peer Effects? Evidence From Boston’s Metco Program,” *American Economic Review*, 94, 1613–1634. [645,658]
- ARELLANO, M. (2003): *Panel Data Econometrics*. Oxford, U.K.: Oxford University Press. [644]
- BECKER, G. S., AND K. M. MURPHY (2000): *Social Economics: Market Behavior in a Social Environment*. Cambridge, MA: Harvard University Press. [643,650]
- BROCK, W. A., AND S. N. DURLAUF (2001): “Interactions-Based Models,” in *Handbook of Econometrics*, Vol. 5, ed. by J. Heckman and E. Leamer. Amsterdam: North-Holland, 3297–3380. [644,646,647,652]
- COLEMAN, J. S. (1966): *Equality of Educational Opportunity*. Washington, DC: U.S. Department of Health, Education and Welfare. [658]
- DUFLO, E., AND E. SAEZ (2003): “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment,” *Quarterly Journal of Economics*, 118, 815–842. [645]
- DUNCAN, G. J., AND S. RAUDENBUSH (2001): “Neighborhoods and Adolescent Development: How Can We Determine the Links,” in *Does It Take a Village? Community Effects on Children, Adolescents, and Families*, ed. by A. Booth and A. C. Crouter. State College, PA: Pennsylvania State University Press, 1005–1136. [646]
- DURLAUF, S. N. (2006): “Groups, Social Influences and Inequality: A Memberships Theory Perspective on Poverty Traps,” in *Poverty Traps*, ed. by S. Bowles. Princeton, NJ: Princeton University Press, 141–175. [644,658]
- FLORENS, J.-P., J. HECKMAN, C. MEGHIR, AND E. VYTLACIL (2007): “Instrumental Variables, Local Instrumental Variables and Control Functions,” Mimeo, Columbia University. [647]
- GLAESER, E., AND J. A. SCHEINKMAN (2001): “Measuring Social Interactions,” in *Social Dynamics*, ed. by S. Durlauf and P. Young. Cambridge, MA: MIT Press, 83–130. [645,646]
- (2003): “Nonmarket Interactions,” in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Vol. 1, ed. by M. Dewatripont et al. Cambridge: Cambridge University Press, 339–369. [645,646]
- GLAESER, E. L., B. SACERDOTE, AND J. A. SCHEINKMAN (1996): “Crime and Social Interactions,” *Quarterly Journal of Economics*, 111, 507–548. [643-645]
- GRAHAM, B. S. (2008): “Supplement to ‘Identifying Through Conditional Variance Restrictions’: Additional Empirical Results, Specification Tests, and Power Calculations,” *Econometrica Supplemental Material*, Vol. 76, <http://www.econometricsociety.org/ecta/Supmat/5915-extensions.pdf>. [645,646,651,655,656]
- GRAHAM, B. S., G. W. IMBENS, AND G. RIDDER (2006): “Complementarity and Aggregate Implications of Assortative Matching: A Nonparametric Analysis,” Mimeo, University of California, Berkeley. [652]
- (2007): “Measuring the Average Outcome and Inequality Effects of Segregation in the Presence of Social Spillovers,” Mimeo, University of California, Berkeley. [652]
- HANUSHEK, E. A. (1971): “Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data,” *American Economic Review*, 60, 280–288. [643]
- HECKMAN, J. J., AND E. J. VYTLACIL (1998): “Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return Is Correlated With Schooling,” *Journal of Human Resources*, 33, 974–987. [647]

- HOXBY, C. M. (2002): "The Power of Peers: How Does the Makeup of the Classroom Influence Achievement?" *Education Next*, 2, 57–63. [658]
- KRUEGER, A. B., AND D. WHITMORE (2001): "The Effect of Attending a Small Class in the Early grades on College-Test Taking and Middle School Test Results: Evidence From Project Star," *Economic Journal*, 111, 1–28. [651]
- LAZEAR, E. P. (2001): "Educational Production," *Quarterly Journal of Economics*, 116, 777–803. [646,652]
- LEVIN, J. (2001): "For Whom the Reduction Counts: A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement," *Empirical Economics*, 26, 221–246. [653]
- MANSKI, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531–542. [644,646-648,658]
- MOFFITT, R. A. (2001): "Policy Interventions, Low-Level Equilibria and Social Interactions," in *Social Dynamics*, ed. by S. Durlauf and P. Young. Cambridge, MA: MIT Press, 45–82. [646]
- PIKETTY, T. (2000): "Theories of Persistent Inequality and Intergenerational Mobility," in *Handbook of Income Distribution*, Vol. 1, ed. by A. Atkinson and F. Bourguignon. Amsterdam: North-Holland, 430–476. [643]
- RIVKIN, S. G., E. A. HANUSHEK, AND J. F. KAIN (2005): "Teachers, Schools and Academic Achievement," *Econometrica*, 73, 417–458. [643]
- SACERDOTE, B. (2001): "Peer Effects With Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 116, 681–704. [645]
- SOLON, G., M. E. PAGE, AND G. J. DUNCAN (2000): "Correlations Between Neighboring Children in Their Subsequent Educational Attainment," *Review of Economics and Statistics*, 82, 383–392. [645]
- WORD, E., J. JOHNSTON, H. P. BAIN, B. DEWAYNE FULTON, C. M. ACHILLES, M. N. LINTZ, J. FOLGER, AND C. BREA (1990): "The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project: Technical Report 1985–1990," Mimeo. [645,654,656]