Peers as treatments*

Brian Krauth
Simon Fraser University
March 24, 2011

Abstract

Models of social interactions are often estimated under the strong assumption that an individual’s choices are a linear function of the average characteristics of his or her reference group. This paper interprets social interactions in a less restrictive potential outcomes framework in which interaction with a given individual is considered a treatment with an unknown additive treatment effect. I show that when the peer characteristics in the linear regression are discrete, the results of the conventional approach can often be interpreted as measuring the average treatment effect of peers in each category.

1 Introduction

Most empirical research on social interaction effects uses a behavioral model in which an individual’s behavior responds to the observed behavior (endogenous effects) and observed characteristics (contextual effects) of peers. This formulation, generally associated with Manski (1993), has inspired a great deal of econometric research into the development of methods for modeling endogenous effects and for empirically distinguishing them from contextual effects (what Manski calls the “reflection problem”) and from endogenous peer selection.

The modeling of contextual effects has seen much less attention. Theory provides little guidance as to which background characteristics of peers should matter, so applied researchers will typically include the peer group averages of whatever variables are available and potentially interesting. The ad hoc nature of these specifications leads to difficulties in both comparing results across studies, and in interpreting results (Cooley, 2009).

This paper argues that many of these difficulties are the result of conceptualizing contextual effects as the direct effect of peer characteristics on one’s outcomes. That is, applied researchers often act as if they were trying to measure the direct effect of the race, gender, ability, parental education, or socioeconomic status of peers. This leads to several problems. First, even when individuals are randomly assigned into groups,

*This paper is very preliminary research, so comments are greatly appreciated. Revisions available at http://www.sfu.ca/~bkrauth/
characteristics are not randomly assigned to individuals. For example, an individual's socioeconomic status is systematically related to his or her race, parental education, and many other factors. In order to isolate the true effect of peer socioeconomic status in this context, one must find and estimate the correct specification that includes all relevant contextual effects. This leads researchers to estimate "kitchen-sink" regressions with many contextual effects. These specifications often have both large standard errors and many remaining candidates for omitted contextual variables. Second, the relevant policy experiment is often difficult to evaluate when individual characteristics are correlated. For example, even if it is possible to distinguish empirically between the effect of peer family income and that of peer parental education, most feasible policy changes that affect the distribution of one contextual variable across peer groups will have related effects on the distribution of the other.

This paper considers an alternative approach in which people are influenced not by the characteristics of their peers but by the peers themselves. In other words, each person has a direct, individual-specific influence (i.e., a treatment effect) on the choices of his or her peers. This influence may be statistically related to background characteristics, but is distinct from them. In this framework, conventional peer effect regressions can be interpreted as measuring how the average treatment effect of peers varies across identifiable groups.

As the paper shows, this simple framework has clear implications for the modeling and interpretation of contextual effects. First, there is no unique true model for contextual effects: because the estimand of interest is now how peer effects vary across identifiable groups a researcher can construct these groups as data availability and interests dictate. Second, simple specifications based on categorical variables provide a clearer and more robust interpretation than kitchen-sink specifications. Third, the framework implies some very particular adaptations to those forms of heterogeneity one is likely to observe in applied work.

1.1 Context: Peer effects in education

Because the goal of this paper is to clarify the treatment of contextual effects in applied work, it will be useful to discuss how they have been conceptualized in existing research. The specific subject area considered will be the measurement of contextual school peer effects on educational achievement. This literature has a sufficient number of papers with similar methodologies and goals that it can be usefully summarized. The review excludes those papers whose primary object of interest is endogenous effects. The primary purpose of this review is to establish two facts: that researchers usually interpret contextual effects as direct effects of characteristics, and that researchers choose specifications with a variety of contextual variables.

Hanushek et al. (2002) measure the effect of % disabled, controlling for % black, % Hispanic, and % reduced-price lunch. They find:

The estimated parameters indicate that a 10-percentage-point increase in the percentage of students classified as dis-
abled increases achievement roughly 0.016 standard deviations.
(Hanushek et al., 2002, p. 597)

Figlio (2007) measures the effect of % of peers who are disruptive, controlling for % black, % male, % black male, % low-income (reduced-price lunch), % low-income male, % immigrant, and average lagged test scores. He describes his specification as follows:

The coefficient of interest is $\theta$, the coefficient on the fraction of a student’s classmates who are disruptive (measured by the fraction who get suspended at least once for five or more days). I control for other observed peer characteristics (the vector $P$): the fraction of classmates who are black, immigrants, low-income (as proxied by free lunch eligibility), male, black males, or low-income males, and the average third-grade test score – or, alternatively, the first observed test score – of the student’s classmates. (Figlio, 2007, p. 380)

Ammermueller and Pischke (2009) measure the effect of the peer average of an index for the number of books at home, controlling (in robustness checks) for % foreign born, % speaking a foreign language at home, % economically disadvantaged, and % leaving before end of academic year. They find:

On average across countries, we find that a one-standard-deviation change in our measure of peer composition leads to a 0.17-standard deviation change in reading test scores, and this estimate is marginally significant. (Ammermueller and Pischke, 2009, p. 317)

Carrell et al. (2009) measure the effect of peer average SAT, an academic composite, a fitness score, and a leadership composite. They find:

All 15 variables are jointly significant at the .01 level. This result implies that the characteristics of upperclassmen, as a whole, play an important role in freshman academic performance. (Carrell et al., 2009, p. 451)

Hanushek et al. (2009) measure the effect of % disabled, controlling for % black, % Hispanic, and % low-income. They find:

The pattern of estimates provides strong evidence that school proportion black negatively affects mathematics achievement of blacks. These effects are much larger and more precisely estimated for blacks than the corresponding estimated impacts on whites, which are generally not significantly different from zero at conventional levels. By comparison, Hispanic enrollment share appears to exert a far smaller effect, indicating that it is proportion black rather than proportion minority that is the key aspect of peer race/ethnic composition in terms of achievement for blacks and whites. (Hanushek et al., 2009, p. 375)

Gould et al. (2009) measure the effect of % immigrant, and do not include other control variables. They find:
The point estimates imply that an increase of 10 percentage points in the immigrant concentration in the 5th grade raises a native student’s dropout rate by (a statistically insignificant) 0.3-0.4 percentage points, relative to an average of 5.4% in our sample; and lowers a native student’s matriculation rate by 1.5-1.8 percentage points, relative to an average of about 61% in our sample. (Gould et al., 2009, p. 1258)

Fletcher (2010) measures the effect of peers with emotional problems, controlling for % female and % minority. He finds:

Cross-sectional results suggest that having a classmate with an emotional problem decreases reading and math scores at the end of kindergarten and first grade by over 10 percent of a standard deviation, which is one-third to one-half of the minority test score gap. (Fletcher, 2010, abstract)

Bifulco et al. (2011) measure the effect of parental education (% college-educated mother) and ethnicity (% black or Hispanic) of same-grade schoolmates. They find:

This paper uses a within-school/across-cohort design to present new evidence of the effects of high school classmate characteristics on a wide range of post-secondary outcomes. We find that increases in the percent of classmates with college-educated mothers decreases the likelihood of dropping out and increases the likelihood of attending college, despite showing no impact on a range of in-school achievement, attitudes, and behaviors. The percent of students from disadvantaged minority groups does not show any effects on postsecondary outcomes, but is associated with students reporting less caring student-teacher relationships and increased prevalence of some undesirable student behaviors during high school. (Bifulco et al., 2011, abstract)

Of the eight articles described here, all interpret their contextual effect estimates as measuring the effect of peer characteristics. Seven of the eight use specifications that reflect this interpretation, and control for the peer average of whatever potentially relevant characteristics are available in the data.

1.2 Related literature

The contemporary economics literature on measuring social effects has been primarily aimed at addressing the challenges described by Manski (1993). Manski’s analysis emphasizes two related but distinct identification problems in previous research: distinguishing true social effects from spurious social effects due to nonrandom peer selection or common shocks, and distinguishing endogenous effects from contextual effects. Manski shows that when social effects take a “linear-in-means” form and social groups are very large, contextual and endogenous effects cannot be distinguished. The selection problem has generally been addressed by looking
for natural experiments in which peer group assignment is plausibly exogenous conditional on observable factors. Methodological research on distinguishing endogenous from contextual effects has generally worked by emphasizing nonlinear models (Brock and Durlauf, 2000), imposing exclusion restrictions such as assuming the nonexistence of contextual effects Gaviria and Raphael (2001), or exploiting the implications of finite and variable group size (Graham, 2008; Bramoullé et al., 2009).

More recently, the literature has begun to address new issues outside of Manski’s original framework. Cooley (2009), Burke and Sass (2009), and Arcidiacono et al. (2011) model peer influence as essentially unobserved, and observed peer variables as proxies for the unobserved interaction that is ultimately of interest. Another developing branch of research (Graham et al., 2010; Carrell et al., 2010) is concerned with the use of estimated peer effect models to consider the consequences of counterfactual allocations of individuals to peer groups.

2 The basic model

I start by describing a simple model that is similar in structure to the standard linear-in means model of peer effects. The approach described here can be usefully applied to less restrictive models, but the intuitions are clearer in a simple model. Section 3 considers some less restrictive models. Individuals are indexed by \( i = 1, 2, \ldots, n \). Each individual is assigned to a peer group \( g_i \) and is characterized by a set of predetermined characteristics \( (o_i, p_i, x_i) \) and an endogenous outcome \( y_i \) such that:

\[
y_i = o_i + \bar{p}_i
\]

where:

\[
\bar{p}_i = \frac{\sum_{j \neq i} I(g_i = g_j)p_j}{\sum_{j \neq i} I(g_i = g_j)}
\]

The peer effect \( p_j \) can be interpreted as the effect person \( j \) has on the outcome of his or her peers, relative to an arbitrary normalization. This effect is specific to individual \( j \) and can reflect behavior, personality attributes, or anything else about individual \( j \). The scalar \( \bar{p}_i \) is just the average value of \( p \) among individual \( i \)’s peers. Equation (1) rules out match-specific effects, where person \( j \) has a different influence on different people, and it rules out interactive effects, where the effect of person \( j \) depends on the other members of the peer group. These cases are considered in Section 3. The own effect \( o_i \) is the counterfactual outcome person \( i \) would have had if matched with “normal” peers (i.e., a peer group in which \( \bar{p}_i = 0 \)). The \( K \)-vector of background characteristics \( x_i \) does not enter directly into the structural model. Both \( o \) and \( p \) are unobserved, while \( y \) and \( x \) are observed.

Equation (1) is similar in structure to the reduced form of the standard linear in means model of peer effects (Manski, 1993):

\[
y_i = \psi_0 + \psi_1 x_i + \psi_2 \bar{x}_i + \epsilon
\]
in that outcomes are assumed to be a linear function of individual factors and group-level averages of peer factors. However, there are also important differences. Equation (2) implicitly assumes that the only thing that matters about peers is their observed characteristics, or that any other factor that matters is uncorrelated with those observed characteristics. The plausibility of this assumption requires sufficiently rich data on characteristics and sufficiently large samples to estimate models with many explanatory variables. Unfortunately, these criteria are often not met by the available data. The administrative data sets used in research on school peer effects, for example, have very limited information on family background and resources. Equation (1) avoids this strong assumption, and so is more plausible when working with limited data.

The model in equation (1) also raises an issue that is often faced in interpreting peer effect estimates: what is the relevant counterfactual? For example, suppose we are interested in the effect of racial segregation \((x_i = 1\) for black, \(x_i = 0\) for white) on academic outcomes \((y_i)\). Because race is a predetermined characteristic, changing the racial composition of a classroom necessarily involves moving individual students and not just their characteristics. The effect of increased or decreased segregation may very much depend on which white students or black students are induced to attend different schools. The linear-in-means model treats all reallocations that yield a given racial composition as equivalent, while the model in equation (1) does not.

To introduce some additional notation, let \(X = (x_1, x_2, \ldots, x_n)\), \(P = (p_1, p_2, \ldots, p_n)\), \(O = (o_1, o_2, \ldots, o_n)\), and \(G = (g_1, g_2, \ldots, g_n)\). Because the focus here is on modeling contextual effects, I make some simplifying assumptions about group selection. Indexing is arbitrary so I assume that \((o_i, p_i, x_i)\) is independent of \((o_j, p_j, x_j)\). In addition I assume purely random assignment of individuals to groups:

\[
Pr(g_i = g|O, P, X) = Pr(g_i = g)
\]

Variations on random assignment are discussed in Section 3.

2.1 What can we measure?

Now, suppose that we estimate equation (2) by OLS but the true causal model is given by equation (1). How do the coefficients of the OLS regression relate to causal parameters of interest?

First, suppose that \(x\) is a vector of binary variables indicating membership in one of \(K + 1\) mutually exclusive and exhaustive categories:

\[
x_{ik} = \begin{cases} 
1 & \text{if individual } i \text{ is in category } k \\
0 & \text{otherwise}
\end{cases}
\]

Then without further loss of generality:

\[
E(o_i|x_i) = \phi_0 + \phi_1 x_i
\]

\[
E(p_i|x_i) = \theta_0 + \theta_1 x_i
\]

That is, \(\theta_{1k}\) (the \(k\)th element of the vector \(\theta_1\)) can be interpreted as the average peer effect from members of category \(k\) relative to the base
category (category 0):

\[ \theta_{1k} = E(p_i | i \text{ is in category } k) - E(p_i | i \text{ is in category } 0) \]

For example, if \( x_i \) is a single binary variable indicating gender, then \( \theta_i \) is the difference in average peer effect between males and females. Proposition 1 shows that \( \theta_1 \) can be estimated by OLS.

**Proposition 1** Given (1), (3), and (5), the OLS regression of \( y_i \) on \((x_i, \bar{x}_i)\) consistently estimates \( \theta_1 \).

Proof: First note that:

\[
E(y_i | \mathbf{X}, \mathbf{G}) = E\left( \alpha_0 + \frac{\sum_{j \neq i} I(g_i = g_j)p_j}{\sum_{j \neq i} I(g_i = g_j)} \right) \mathbf{X}, \mathbf{G}) \\
= E(\alpha_0 | \mathbf{X}, \mathbf{G}) + \sum_{j \neq i} I(g_i = g_j)E(\frac{p_j}{\sum_{j \neq i} I(g_i = g_j)}) (\mathbf{X}, \mathbf{G}) \\
= E(\alpha_0 | \mathbf{x}_i) + \sum_{j \neq i} I(g_i = g_j)(\theta_0 + \theta_1 \mathbf{x}_i) \\
= (\phi_0 + \phi_1 \mathbf{x}_i + \theta_1 \mathbf{x}_i) \\
= (\phi_0 + \theta_0) + \phi_1 \mathbf{x}_i + \theta_1 \bar{x}_i
\]

Since \((\bar{x}_i, \mathbf{x}_i)\) can be calculated from \((\mathbf{X}, \mathbf{G})\) the law of iterated expectations gives:

\[
E(y_i | \bar{x}_i, \mathbf{x}_i) = E(E(y_i | \mathbf{X}, \mathbf{G})|\bar{x}_i, \mathbf{x}_i) \\
= E((\phi_0 + \theta_0) + \phi_1 \mathbf{x}_i + \theta_1 \bar{x}_i | \bar{x}_i, \mathbf{x}_i) \\
= (\phi_0 + \theta_0) + \phi_1 \mathbf{x}_i + \theta_1 \bar{x}_i
\]

so we can estimate \( \theta_1 \) by OLS regression of \( y_i \) on \((\bar{x}_i, \mathbf{x}_i)\).

The key feature of Proposition 1 is that it applies to any predetermined vector of categorical variables \( \mathbf{x} \). There is no requirement that \( \mathbf{x} \) itself enters into the causal model, or that other peer characteristics are excluded from the causal model. In this framework, for example, a linear regression of \( y \) on own and peer gender\(^1\) will recover the difference in average peer effects between males and females, while a regression of \( y \) on own and peer low-income status will recover the difference in average peer effects between students who are low-income and those who are not.

If \( \mathbf{x} \) is not a categorical vector, random assignment still implies the regression coefficients can be usefully interpreted. Without loss of generality, let the best linear predictors of \( \alpha_i \) and \( p_i \) be defined by:

\[
o_i = \tilde{\phi}_0 + \tilde{\phi}_1 \mathbf{x}_i + u_i \quad \text{where } E(\mathbf{x}, u_i) = 0 \\
p_i = \tilde{\theta}_0 + \tilde{\theta}_1 \mathbf{x}_i + v_i \quad \text{where } E(\mathbf{x}, v_i) = 0
\]

\(^1\)Since peers are randomly assigned here, controlling for own characteristics is not strictly required. However, since it is a common and sensible practice the presentation here does so.
Then $\tilde{\theta}_1$ can be interpreted as the slope of the best linear predictor of an individual’s effect on his or her peers given his or her characteristics.

**Proposition 2** Given (1), (3), and (6), the OLS regression of $y_i$ on $(x_i, \bar{x}_i)$ consistently estimates $\tilde{\theta}_1$.

Proof: Substituting (6) into (1):

$$y_i = \tilde{\phi}_0 + \tilde{\phi}_1 x_i + u_i + \sum_{j \neq i} I(g_j = g_i)(\tilde{\theta}_0 + \tilde{\theta}_1 x_j + v_j)$$

$$= (\tilde{\phi}_0 + \tilde{\theta}_0) + \tilde{\phi}_1 x_i + \tilde{\theta}_1 \bar{x}_i + \left( u_i + \sum_{j \neq i} I(g_j = g_i)v_j \right)$$

Independence and random assignment (3) imply that $E(x_i v_j) = 0$ and $E(\bar{x}_i u_i) = 0$, and equation (6) implies $E(x_i u_i) = 0$ and $E(\bar{x}_i v_j) = 0$, so

$$E\left(x_i \left( u_i + \sum_{j \neq i} I(g_j = g_i)v_j \right) \right) = E\left(\bar{x}_i \left( u_i + \sum_{j \neq i} I(g_j = g_i)v_j \right) \right) = 0$$

and so the equation above can be consistently estimated by the OLS regression of $y_i$ on $(x_i, \bar{x}_i)$.

These results imply that random assignment, additivity, and the absence of match-specificity are sufficient for the linear-in-means model to consistently estimate quantities we may be interested in. There is no requirement that the explanatory variables include every relevant peer characteristic, or that the coefficients on those peer characteristics be interpreted as measuring the direct effect of peer characteristics as is commonly done in the literature.

### 3 Extensions

#### 3.1 Match-specific effects

Now suppose that peer effects are match specific:

$$y_i = a_i + \frac{\sum_{j \neq i} I(g_j = g_i)p_{ij}}{\sum_{j \neq i} I(g_j = g_i)}$$

where $p_{ij}$ is now the effect of person $j$ on person $i$. Match-specific effects allow a statistical relationship between $p_{ij}$ and $(a_i, x_i)$, so the necessary independence assumptions are less straightforward. Specifically, I assume that $(a_i, x_i)$ is independent of $(a_j, x_j)$ and $(p_{jk}, p_{kj})$ for all $i \neq j \neq k$, and that assignment is purely random:

$$\Pr(g_i = g | O, X, P) = \Pr(g_i = g)$$

As before, assume that $x$ is a $k$-vector indicating membership in one of $k + 1$ mutually exclusive and exhaustive categories. Without loss of generality, let:

$$E(p_{ij} | x, x_j) = \gamma_0 + \gamma_1 x_i + \gamma_2 x_j + \gamma_3 x_i x_j$$

$$E(a_i | x) = \lambda_0 + \lambda_1 x_i$$

8
where $\gamma_0$ and $\gamma_1$ are scalars, $\gamma_1$ and $\gamma_2$ are $K$-vectors whose $k$th element is denoted by $\gamma_{1k}$ and $\gamma_{2k}$ respectively, and $\gamma_3$ is a $K \times K$ matrix whose $k$th row is denoted by $\gamma_{3k}$ and whose $k, \ell$ element is denoted by $\gamma_{3k\ell}$.

We might be interested in estimating the average effect of peers from category $k$ (relative to peers in the base category) on individuals from category $\ell$:

$$MSPE_k = E \left( p_{i,j} \right| i \text{ is in category } \ell, \ j \text{ is in category } k \bigg) - E \left( p_{i,j} \right| i \text{ is in category } \ell, \ j \text{ is in category } 0 \bigg)$$

or the overall average effect of peers from category $k$ relative to peers in the base category:

$$APE_k = E(p_{i,j}| j \text{ is in category } k) - E(p_{i,j}|j \text{ is in category } 0)$$

$$= \gamma_{2k} + E(x_i')\gamma_{3k}$$

Proposition 3 shows that both of these quantities can be estimated by OLS regressions.

**Proposition 3** Given (7), (8), and (9), an OLS regression of $y_i$ on $(x_i, x_{\bar{i}})$ and their interaction consistently estimates MSPE, and an OLS regression of $y_i$ on $(x_i, x_{\bar{i}})$ consistently estimates APE.

Proof: For the first result, note that:

$$E(y_i|X, G) = E(o_i|X, G) + \frac{\sum_{j \neq i} I(g_j = g_i)E(p_{i,j}|X, G)}{\sum_{j \neq i} I(g_j = g_i)}$$

$$= E(o_i|x_i) + \frac{\sum_{j \neq i} I(g_j = g_i)E(p_{i,j}|x_i, x_j)}{\sum_{j \neq i} I(g_j = g_i)}$$

$$= \lambda_0 + \lambda_1 x_i + \sum_{j \neq i} I(g_j = g_i)(\gamma_0 + \gamma_1 x_i + \gamma_2 x_j + x_i'\gamma_3 x_j)$$

$$= \lambda_0 + \lambda_1 x_i + \gamma_0 + \gamma_1 x_i + \gamma_2 x_i + x_i'\gamma_3 x_i$$

Again, we can apply the law of iterated expectations to get:

$$E(y_i|x_i, x_{\bar{i}}) = E(E(y_i|X, G)|x_i, x_{\bar{i}})$$

$$= E((\lambda_0 + \gamma_0) + (\lambda_1 + \gamma_1)x_i + \gamma_2 x_i + x_i'\gamma_3 x_i)$$

$$= (\lambda_0 + \gamma_0) + (\lambda_1 + \gamma_1)x_i + \gamma_2 x_i + x_i'\gamma_3 x_i$$

We can estimate this equation by OLS regression to estimate $(\gamma_2, \gamma_3)$, which can then be used to construct estimates of the MSPE. To get the second result, apply the law of iterated expectations again to get:

$$E(y_i|x_{\bar{i}}) = E(E(y_i|x_i, x_{\bar{i}})|x_{\bar{i}})$$

$$= E((\lambda_0 + \gamma_0) + (\lambda_1 + \gamma_1)x_i + \gamma_2 x_i + x_i'\gamma_3 x_i)$$

$$= (\lambda_0 + \gamma_0) + (\lambda_1 + \gamma_1)E(x_i|x_{\bar{i}}) + \gamma_2 x_i + E(x_i|x_{\bar{i}})'\gamma_3 x_i$$

$$= (\lambda_0 + \gamma_0 + (\lambda_1 + \gamma_1)E(x_i) + (\gamma_2 + E(x_i)'\gamma_3) x_i$$

(10)
Since this CEF is linear in $\bar{x}_i$, the OLS regression of $y_i$ on $\bar{x}_i$ will consistently estimate $(\gamma_2 + E(x_i')\gamma_3)$, the $k$th element of which is $APE_k$. Since $\bar{x}_i$ and $\bar{x}_i$ are independent and therefore orthogonal, the coefficient on $\bar{x}_i$ in the OLS regression of $y_i$ on $(x_i, \bar{x}_i)$ will have the exact same probability limit, and so will also consistently estimate $APE$.

3.2 Non-additive group effects

The model so far does not allow things like thresholds or decreasing returns to scale. This can be accounted for by considering the peer group as a whole. That is, suppose that each individual can be characterized by $(o_i, x_i)$ where $o_i$ is the unobserved own-effect and $x_i$ is an observed $k$-vector of binary variables indicating membership in one of $k + 1$ mutually exclusive and exhaustive categories. Each peer group $g$ is characterized by $(p_g, z_g)$ where $p_g$ is an unobserved scalar and $z_g$ is an observed $m$-vector of indicators indicating which of $m + 1$ group types $g$ falls into. Let:

$$y_i = o_i + p_g(i)$$

Again I assume random assignment, i.e., that $(o_i, x_i)$ is independent of $(p_g(i), z_g(i))$. Without loss of generality, let:

$$E(o_i|x_i, z_g(i)) = \psi_0 + \psi_1 x_i$$
$$E(p_g(i)|x_i, z_g(i)) = \gamma_0 + \gamma_1 x_i + \gamma_2 z_g(i) + x'_i \gamma_3 z_g(i)$$

Then:

$$E(y_i|x_i, z_g(i)) = E(o_i|x_i, z_g(i)) + E(p_g(i)|x_i, z_g(i))$$
$$= E(o_i|x_i) + E(p_g(i)|x_i, z_g(i))$$
$$= \psi_0 + \psi_1 x_i + \gamma_0 + \gamma_1 x_i + \gamma_2 z_g(i) + x'_i \gamma_3 z_g(i)$$
$$= (\psi_0 + \gamma_0) + (\psi_1 + \gamma_1) x_i + \gamma_2 z_g(i) + x'_i \gamma_3 z_g(i)$$

so an OLS regression of $y_i$ on $x_i, z_g(i)$ and an interaction term will recover:

$$E(p_g(i)|x_i, z_g(i) = z) - E(p_g(i)|x_i, z_g(i) = 0) = \gamma_2 z + x'_i \gamma_3 z$$

i.e., the average effect of peer groups in category $m$ on individuals in category $k$, relative to the effect of peer groups in the base category on those same individuals.

3.3 Random assignment within groups (in progress)

Most studies of educational peer effects are based on a research design in which individuals are non-randomly assigned to large groups (e.g., schools) but then are randomly assigned to smaller groups within those large groups (e.g., grade cohorts or classrooms).

Individual $i$ is characterized by the predetermined variables $(o_i, p_i, x_i, s_i)$ and the endogenous outcome $y_i$, and is randomly matched with other individuals with the same value of $s$. The outcome is given by:

$$y_i = o_i + \sum_{j \neq i} I(g_j = g_i) p_j \frac{\sum_{j \neq i} I(g_j = g_i)}{\sum_{j \neq i} I(g_j = g_i)}$$
Again we assume that \((a_i, p_i, x_i, s_i)\) is independent of \((a_j, p_j, x_j, s_j)\), and that peer group is randomly assigned conditional on \(s_i\):

\[
\Pr(g_i = g|a_i, p_i, x_i, s_i) = \Pr(g_i = g|s_i)
\]

As before, assume \(x\) is a vector of categorical variables, so that:

\[
E(o_i|x_i, s_i = s) = \psi_{0s} + \psi_{1s}x_i
\]

\[
E(p_i|x_i, s_i = s) = \theta_{0s} + \theta_{1s}x_i
\]

Then:

\[
E(g_i|x, G, s) = E(o_i|x, G, s) + \frac{\sum_{j \neq i} I(g_j = g_i)E(p_j|x, G, s)}{\sum_{j \neq i} I(g_j = g_i)}
\]

\[
= E(o_i|x_i, s_i = s) + \frac{\sum_{j \neq i} I(g_j = g_i)E(p_j|x_j, s_j = s)}{\sum_{j \neq i} I(g_j = g_i)}
\]

\[
= \psi_{0s} + \psi_{1s}x_i + \frac{\sum_{j \neq i} I(g_j = g_i)(\theta_{0s} + \theta_{1s}x_j)}{\sum_{j \neq i} I(g_j = g_i)}
\]

\[
= (\psi_{0s} + \theta_{0s}) + \psi_{1s}x_i + \theta_{1s}x_i
\]

Applying the law of iterated expectations we again get:

\[
E(g_i|x_i, x_i, s_i) = (\psi_{0s} + \theta_{0s}) + \psi_{1s}x_i + \theta_{1s}x_i
\]

Equation (23) is a standard heterogeneous-coefficient linear panel data model. One approach to this type of model is to estimate the full set of coefficients for each group \(s\), and then average to estimate \(E(\theta_{1s})\). However, while \(E(\theta_{1s})\) is related to the average peer effect defined earlier, the two are not identical. To see this, note that:

\[
E(p_i|x_i = x) = E(E(p_i|x_i, s_i)|x_i = x)
\]

\[
= E(\theta_{0s} + \theta_{1s}x_i|x_i = x)
\]

\[
= E(\theta_{0s}|x_i = x) + E(\theta_{1s}|x_i = x)x
\]

Substituting in we get:

\[
APE_k = E(p_i|i \text{ is in category } k) - E(p_i|i \text{ is in category } 0)
\]

\[
= E(\theta_{0s}|i \text{ is in category } k) + E(\theta_{1sk}|i \text{ is in category } k) - E(\theta_{0s}|i \text{ is in category } 0)
\]

Since \(\theta_{0s}\) is not identified (only \(\psi_{0s} + \theta_{0s}\) is), the value of \(APE_k\) is not identified either. To give some intuition consider the case of standard cohort-based research designs for measuring educational peer effects, where \(s\) denotes (nonrandomly assigned) schools and \(g\) denotes (randomly assigned) cohorts within schools. We cannot distinguish between \(\psi_{0s}\) (the average own-effect in the school) and \(\theta_{0s}\) (the average peer effect in the school) because a given difference in average outcomes across schools could be explained by either. But because students with a given set of characteristics are nonrandomly distributed across schools, we are only able to characterize within-school differences in average peer effects.
A more common approach is to estimate a traditional fixed effects model that allows heterogeneity across $s$ of intercepts but not of slopes. This implies the assumption:

$$E(o_i|x_i, s_i = s) = \psi_0 + \psi_1 x_i$$  \hspace{1cm} (27)$$
$$E(p_i|x_i, s_i = s) = \theta_0 + \theta_1 x_i$$  \hspace{1cm} (28)

That is, the average difference in both own and peer effects between individuals in each of the categories does not vary across $s$.

### 4 Conclusion

Returning to the context of Section 1.1, this paper has several relevant findings. First, contextual effects models can be interpreted as measuring how the effect of an individual on his or her peers varies with his or her characteristics. This interpretation does not require the problematic assumption that peer characteristics themselves have a direct effect. Second, the interpretation of peer effects made here argues in favor of particular specifications for contextual effects: contextual variables based on a set of mutually exclusive categories provide a significantly more straightforward interpretation than do contextual variables based on overlapping categories such as race and socioeconomic status.

### References


Carrell, Scott E., Bruce I. Sacerdote, and James E. West, “Be-
ware of Economists Bearing Reduced Forms? An Experiment in How

_ , Richard L. Fullerton, and James E. West, “Does Your Cohort
Matter? Measuring Peer Effects in College Achievement,” Journal of

Cooley, Jane, “Can Achievement Peer Effect Estimates Inform Policy?
A View from Inside the Black Box,” Working paper, University of

Figlio, David N., “Boys named Sue : Disruptive students and their

Fletcher, Jason M., “Spillover Effects of Inclusion of Classmates with
Emotional Problems on Test Scores in Early Elementary School,” Jour-

Gaviria, Alejandro and Steven Raphael, “School-based peer effects
(2), 257–268.

Gould, Eric D., Victor Lavy, and M. Daniele Pasceman, “Does im-
migration affect the long-term educational outcomes of natives? Quasi-

Graham, Bryan S., “Identifying social interactions through conditional

_ , Guido W. Imbens, and Geert Ridder, “Measuring the effects
of segregation in the presence of social spillovers: A nonparametric

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin, “In-
ferring program effects for special populations: Does special education
raise achievement for students with disabilities?,” Review of Economics

_ , _ , and _ , “New Evidence about Brown v. Board of Education:
The Complex Effects of School Racial Composition on Achievement,”

Manski, Charles F., “Identification of endogenous social effects: The