On the Identification of Peer Effect Models of Cognitive Achievement

Pilar Alcalde

January 30, 2014

JEL Codes: C31, I21

Keywords: peer effects, student achievement, linear model, identification, reflection problem.

Abstract

The identification problems of general peer effect models have been studied in depth. But cognitive achievement studies suffer two additional sources of bias that arise because the behavioral variable - student effort - is unobserved. This paper studies both sources of bias and analyzes the relative

*I am very grateful to Petra Todd and Kenneth Wolpin for their comments and encouragement, and to Jere Behrman, Flavio Cunha, Aureo de Paula, Claudio Sapelli and the participants at UPenn’s Empirical Micro Lunch seminar and Universidad Catolica de Chile’s seminar for their useful comments. All errors are mine.

†Universidad de los Andes, Chile, Email: palcalde@uandes.cl
success of the identification strategies most commonly used. In a quadratic model, the equilibrium parameters suffer endogeneity bias except in three particular cases, which are fairly restrictive. Even in these three cases, the structural peer effects are not identified. The most common strategies achieve partial identification at best, even after assuming a unique mechanism for peer effort, and additional assumptions are needed on the production and cost function. This implies that the current empirical studies overestimate the true effect of peers on achievement.

I Introduction

The literature of social interactions has a pending question of whether classmates have an effect on student achievement, either direct or indirect. Unfortunately, the current estimates in the literature cannot answer this question. The techniques used to determine the relation between student achievement and peer characteristics cannot inform on their true underlying relation because they suffer from endogeneity bias.

Peer effects have been studied for different situations as crime, obesity, drug use, sexual initiation, and several others (see Blume, Brock, Durlauf and Ioannides 2010, Durlauf 2004 for some surveys). In these general contexts, the estimation of peer effects is challenged by identification problems, like the Reflection and Cor-
related Effects problems. These problems have been studied and solved using different identification strategies.

But in cognitive achievement studies there are two additional sources of bias. They arise because the behavioral variable - student effort - is unobserved and correlated with observable characteristics included in the estimation. In a model where students decide their effort, there are at least two equations: an achievement production function that relates achievement and effort, and the effort equation that describes the optimal effort chosen. The parameters of these two equations are the structural parameters that measure the true effect of peers on achievement and behavior.

When effort is unobserved, one can only estimate an equilibrium relationship between achievement and observed characteristics, the achievement equation. Because effort is a decision variable this implies a distinction between the structural parameters and the equilibrium parameters - the ones that can be directly estimated from the data. This creates two identification problems. First, the equilibrium parameters will be a composite of the structural parameters. Second, the equilibrium parameters will suffer endogeneity bias - observed characteristics will be correlated with the error term in the estimation equation.

This paper studies both problems and analyzes the relative success of the identification strategies most commonly used in the achievement peer effects literature. Using a quadratic model, I show that
when effort is unobserved the equilibrium parameters suffer endogeneity bias, except in three particular cases. In the first case, effort does not vary with unobservable variables. In the second case, effort does not vary with observable characteristics. These two cases are opposite to each other and both implausible: they imply that the observable characteristics included in the model either completely describe the effort decision or have no relation with it at all. In the third case, the production function is linear in own and peer effort. This assumption is challenged by the growing empirical evidence of nonlinear peer effects (Burke and Sass, 2008; Carrell, Sacerdote and West, 2013; Cooley, 2008; Ding and Lehrer, 2007; Gibbons and Telhaj, 2008; Pinto, 2010; Sacerdote, 2001). These studies suggest that the size of peer effects depends on own and peer characteristics, although these estimates suffer endogeneity bias if the true relation is nonlinear.

These three conditions avoid the endogeneity bias, but the structural parameters of the achievement and effort equations are not identified. To solve the Reflection problem, the studies of achievement peer effects generally use four identification strategies. The paper reviews the identification power of these strategies when the production function is linear in own and peer effort. Unfortunately, none of these strategies identifies the structural parameters by itself and additional assumptions are needed.

The paper explores the identification power of two additional as-
sumptions, regarding the mechanism by which the endogenous peer effect operates. First, that peer effort does not affect directly the production function, only through the student’s effort decision. Second, that peer effort does not affect the effort decision but it is an input of the production function. These additional assumptions provide partial identification of structural peer effects at best, and some strategies are more successful than others. For example, if peer effort does not affect directly the production function, the most successful strategy is using peer achievement as a proxy for peer effort (Boucher, Bramoulle, Djebbari and Fortin [2010], Fruehwirth [2013], Kang [2007], Sacerdote [2001]). If peer effort does not affect the effort decision, the most successful strategy is forcing peer characteristics to have an effect proportional to own characteristics (Arcidiacono, Foster, Goodpaster and Kinsler [2012], Burke and Sass [2008]). Further identification requires more assumptions on the production and cost function.

My paper is not the first one to analyze the bias due to unobservables in social interactions. Fruehwirth (2013) analyzes the effect on estimation of including an unobserved contextual effect, as peer ability. This paper broadens the issue by analyzing the effect of an unobserved endogenous effect. Several authors point out that both unobserved variables, ability and behavior, might be important determinants of achievement (Bishop, 2006, Hanushek, Kain and Rivkin, 2009). To keep the arguments simple, I abstract from abil-
ity issues, but they can be easily incorporated by using ideas from Arcidiacono et al. (2012) and Burke and Sass (2008).

Fruehwirth (2013) also reviews the use of peer achievement as a proxy for peer ability. This paper revisits her argument and shows that her strategy successfully identifies all the parameters of her model when classes vary by size, which she does not allow. This is consistent with Lee (2007), who shows that the class size contains valuable information to identify peer effects. This result implies that when unobservable variables are present, their nature is not trivial for estimation: identifying the structural parameters is easier when the variable is contextual, as peer ability, than when it is endogenous, as peer effort.

In conclusion, the unbiased estimation of peer effects in cognitive achievement requires adequate measures of own and peer effort. Unfortunately, these measures are missing of studies that use standard measures of achievement as test scores and graduation rates. The biases present in the current estimates -even after the Reflection problem is solved correctly- may help understand the large diversity of results found.

The paper proceeds as follows. Section II reviews the estimation problems in models of social interactions and the most common identification strategies. Section III analyzes the endogeneity bias problem in a quadratic model, and describes the three cases with no endogeneity bias. Section IV analyzes in detail the power of the
identification strategies most commonly used to solve the Reflection Problem. Finally, Section V reviews the use of peer achievement as a proxy for peer ability and Section VI concludes.

II The Estimation of Achievement Peer Effects

Social interactions take place when individual behavior depends on the characteristics or behavior of a reference group. There are two types of mechanisms. The endogenous effect is the influence on individual behavior of current group behavior. The contextual effect is the influence on individual behavior of exogenous group characteristics. Manski (1993) also defines the correlated effect as the influence on individual behavior of common group variables, observed or not.

The distinction between peer effects is important because they have differing policy implications (Fruehwirth, 2013; Manski, 1993). A positive endogenous effect can generate a “social multiplier” effect: an increase in the behavior of individual $j$ improves directly the behavior of individual $i$ and further increases the behavior of individual $j$. Exogenous and correlated effects do not generate this “social multiplier”. This consideration is relevant, as stressed by the experiment by Carrell et al. (2013). They estimated a model of nonlinear peer effects - without distinguishing different channels for social interactions- and used their estimates to assign peer groups. The assignment was supposed to increase achievement of students.
in the lower third of the distribution. But those students lowered their achievement, and the students in the middle of the distribution improved, contrary to their predictions.

Define for individual $i$ in group $g$ her behavior of interest $Y_{ig}$ and a vector of observable individual characteristics $X_{ig}$. Peer effects are estimated using specifications of the form

$$Y_{ig} = F(X_{ig}, \bar{X}_{-ig}, \bar{Y}_{-ig}, \alpha_g, \varepsilon_{ig})$$ (1)

where $\bar{X}_{-ig}$ and $\bar{Y}_{-ig}$ are average peer characteristics and behavior - the subscript $-i$ indicates that individual $i$ is not a member of her peer group - and $\alpha_g$ is the correlated effect that measures variables common to the group. The term $\varepsilon_{ig}$ measures an individual error term, uncorrelated with the other variables.

The identification of equation (1) is challenging. It has been extensively discussed for the binary choice model (Blume et al., 2010; Brock and Durlauf, 2001a,b; Davezies, D’Haultfoeuille and Fougere, 2009; Durlauf and Ioannides, 2010), for the lineal model (Lee, 2007; Manski, 1993; Moffitt, 2001) and for social networks (Blume, Brock, Durlauf and Jayaraman, 2013; Bramouille, Djebbari and Fortin, 2009). There are three main identification problems: the Reflection Problem, the presence of unobserved group characteristics, and the non-randomness of group formation. The last two problems imply that the error terms are correlated in a group and
this effect is captured by the correlated effect (Lee, 2007). This biases the estimation if the correlated effect is not independent from own and peer observable characteristics. The Reflection problem is the separate identification of the contextual and endogenous effect (Manski, 1993). It arises because behavior comes from a simultaneous equation system - the decision variable of one agent becomes an argument in the decisions of others. Then it only presents an identification problem when the system is linear (Brock and Durlauf, 2001b; Durlauf and Ioannides, 2010). Both Reflection and Correlated Effects problems can be solved using different identification strategies discussed in the papers above.

Usually, the outcome variable $Y_{ig}$ is also a behavioral variable, e.g. drug use. But in cognitive achievement studies the outcome variable -test scores or graduation rates- cannot be chosen at discretion and the behavioral variable -effort- is unobserved. If effort depends partially on observed characteristics, this endogeneity biases the estimation of peer effects, as section III shows.

The theoretical models of achievement peer effects recognize this feature and focus on unobserved behavior. But these models are not concerned with the identification of said peer effects and the models are not directly estimable. Lazear (2001) presents a model of class size determination with peer externalities. Akerlof and Kran- ton (2002) and Austen-Smith and Fryer (2005) present identity models, where students choose their social category and school effort to
match the ideal of the category. Cicala, Fryer and Spenkuch (2011) develop a Roy model of social interactions in which the production of study or mischief occurs within peer groups and individuals choose their sector to maximize “social income”.

Most empirical papers note that peer effects are probably generated by unobserved ability, norms, and behavior rather than observed characteristics (Bishop 2006). But in practice, they give little attention to the identification of said mechanisms (Blume et al. 2010; Richards 2012), and do not analyze the consequences the unobservability has for estimation. Only a few studies analyze with more detail the existence of different mechanisms generating achievement effects, in particular Hoxby and Weingarth (2006) and Duflo, Dupas and Kremer (2011).

Instead, most empirical papers concentrate on the issue of the Reflection and Correlated Effects problems. The Reflection problem is solved using mainly four strategies. Some use instrumental variables to control the endogeneity of peer achievement (Angrist and Lang 2004; Kang 2007). Others force peer characteristics to have an effect proportional to own characteristics (Arcidiacono et al. 2012; Burke and Sass 2008). Fruehwirth (2013) proposes to use observed peer achievement as a proxy for peer unobservables. But the most popular method is to either assume or ignore this problem, by estimating an equation with only contextual effects (Ammermueller and Pischke 2009; Burke and Sass 2008; Carrell, ...

The Correlated Effects problem is solved using data with double random assignment (Foster 2006; Graham 2008; Sacerdote 2001), or controlling for fixed effects at the individual/class/teacher/school level (Burke and Sass 2008; Carrell et al. 2009; Hanushek et al. 2009; Hoxby 2000). Kang (2007) includes explicitly both random assignment and fixed effects. Finally, Lee (2007) proposes a within-transformation of the data that solves both identification problems under class size variation. This approach is used by Boucher et al. (2010).

If the assumptions required by each strategy hold, they identify the achievement model only when students do not decide their effort. If effort is endogenous and unobserved, section III shows that the estimates are biased, and section IV shows that none of these strategies have enough power to identify the structural parameters by itself.

Researchers usually think of two mechanisms by which the endogenous peer effect operates (Bishop 2006). One mechanism is production externalities: peer behavior influences how much classroom time is devoted to maintain discipline and how much is learned from classroom discussions and group projects, among others. If production externalities exist, then peer effort should enter the achievement production function. Another mechanism is utility externali-
ties: the effect of peer norms and teasing and harassment behavior on the individual’s own effort level. If utility externalities exist, then peer effort should enter the utility function.

The paper explores the identification power of two additional assumptions, regarding said mechanisms. First, only utility externalities exist. Second, only production externalities exist. Section IV shows these additional assumptions help in partial identification of some structural parameters, but full identification requires additional assumptions on the production and cost functions.

III The Student Effort Decision

Consider an effort decision framework. Each student $i$ in class $g$ of size $m_g$ has a vector of observable characteristics $X_{ig}$ of size $K$, and must decide her effort level $E_{ig}$. Effort is costly, and is necessary to produce cognitive achievement $A_{ig}$.

Both the production function of cognitive achievement $F$ and the cost function $C$ depend on own and peer effort, $E_{ig}$ and $E_{-ig}$, on own and peer observable characteristics, $X_{ig}$ and $X_{-ig}$, on common class characteristics $\alpha_g$, and on a productivity shock $u_{ig}$ and cost shock $v_{ig}$ respectively, that are uncorrelated with observed characteristics.
The student chooses her effort to maximize her utility:

\[
\max_e U_{ig} = A_{ig} - C_{ig} \tag{2}
\]

s. to \( A_{ig} = F(E_{ig}, \bar{E}_{-ig}, X_{ig}, \bar{X}_{-ig}, \alpha_g, u_{ig}) \)

\[
C_{ig} = C(E_{ig}, \bar{E}_{-ig}, X_{ig}, \bar{X}_{-ig}, v_{ig})
\]

The bias problem can be analyzed in this general model. But to highlight it, suppose the production and cost functions are quadratic, this is, both contain the quadratic and interaction terms for all inputs:

\[
A_{ig} = a + X_{ig}b + \bar{X}_{-ig}c + d_1 E_{ig} + d_2 E^2_{ig} + d_3 \bar{E}_{-ig}^2 + d_4 E^2_{-ig} + d_5 E_{ig} \bar{E}_{-ig} + E_{ig}X_{ig}f + E_{ig} \bar{X}_{-ig}g + \bar{E}_{-ig}X_{ig}h + \bar{E}_{-ig}\bar{X}_{-ig}n + pE_{ig}u_{ig} + \alpha_g + u_{ig} \tag{3}
\]

\[
C_{ig} = \tilde{a} + X_{ig} \tilde{b} + \bar{X}_{-ig} \tilde{c} + \tilde{d}_1 E_{ig} + \tilde{d}_2 E^2_{ig} + \tilde{d}_3 \bar{E}_{-ig}^2 + \tilde{d}_4 E^2_{-ig} + \tilde{d}_5 E_{ig} \bar{E}_{-ig} + E_{ig}X_{ig} \tilde{f} + E_{ig} \bar{X}_{-ig} \tilde{g} + \tilde{E}_{-ig}X_{ig} \tilde{h} + \tilde{E}_{-ig}\bar{X}_{-ig} \tilde{n} + \tilde{p}E_{ig}v_{ig} + v_{ig}
\]

The quadratic terms for observed characteristics have been omitted for simplification, without loss of generality. They can be added without changing the insights from the analysis.

This quadratic specification is a particular case, but it is still
fairly general. It can include many of the models used in the peer effects literature. Naturally, the linear-in-means model is a particular case. But it can also accommodate models that include the mean and standard deviation of peer characteristics \cite{burke2008, vigdor2004, zabel2008} and semiparametric partial linear models \cite{ding2007, rodriquez2010} by using polynomial approximations.

It can also include “share” specifications, popular in the non-linear literature. In these models achievement depends on the share of peers in a specific segment of the distribution of a contextual variable \cite{lavy2009, meowan2006}, or on interactions between the relative position of the student and the share of peers in the segment \cite{burke2008, carrell2013, gibbons2008, hoxby2006, sacerdote2001, zimmerman1999}. To include these variables, just define $X_{igk} = 1$ if the student belongs to segment $k$, and $ar{X}_{-igk}$ becomes the share of classmates in segment $k$.

The best response function for each student is given by the first order condition

\[
\begin{align*}
  &d_1 + 2d_2 E_{ig} + d_5 \bar{E}_{-ig} + X_{ig} f + \bar{X}_{-ig} g + p u_{ig} \\
  &- \left( \tilde{d}_1 + 2\tilde{d}_2 E_{ig} + \tilde{d}_5 \bar{E}_{-ig} + X_{ig} \tilde{f} + \bar{X}_{-ig} \tilde{g} + \tilde{p} u_{ig} \right) = 0
\end{align*}
\]

The second order condition requires that $2d_2 - 2\tilde{d}_2 < 0$. Define for
all parameters \( \{ \theta, \tilde{\theta} \} \) the composite parameter \( \Delta \theta = \theta - \tilde{\theta} \). Then the best response function \( G \) for each student is given by

\[
E_{ig} = \frac{\Delta d_1 + \Delta d_5 \tilde{E}_{-ig} + X_{ig} \Delta f + \tilde{X}_{-ig} \Delta g + p_{ig} - \tilde{p}_{ig}}{-2\Delta d_2} \tag{4}
\]

The production and cost functions in equation (3) correspond to the structural equations of the model. If the outcome of interest is achievement, it is possible to define the structural contextual effect on achievement \( \delta_A \) and the structural endogenous effect on achievement \( \lambda_A \) as the marginal effects of peer characteristics and peer behavior on the production function \( F \):

\[
\begin{align*}
\delta_{Aig} &= c + gE_{ig} + n\tilde{E}_{-ig} \\
\lambda_{Aig} &= d_3 + 2d_4\tilde{E}_{-ig} + d_5E_{ig} + X_{ig}h + \tilde{X}_{-ig}n
\end{align*} \tag{5}
\]

Similarly, if the outcome of interest is effort, it is possible to define the structural contextual effect on effort \( \delta_E \) and the structural endogenous effect on effort \( \lambda_E \) as the marginal effects of peer characteristics and peer behavior on the best response function \( G \):

\[
\begin{align*}
\delta_E &= \frac{\Delta g}{-2\Delta d_2} \\
\lambda_E &= \frac{\Delta d_5}{-2\Delta d_2} \tag{6}
\end{align*}
\]

As in Bisin, Moro and Topa (2011), a Nash equilibrium in class
$g$ is defined as a vector of effort levels $\{E_{ig}^*\}_{i \in g}$—not necessarily unique—such that the best response function (4) is satisfied jointly for all $i \in g$, and the equilibrium aggregate $E_g^*$ satisfies the natural consistency condition $E_g^* = \frac{1}{m_g} \sum_{i \in g} E_{ig}^*$. Appendix A shows there is a unique equilibrium if $\Delta d_5 \neq -2\Delta d_2$, with no equilibrium otherwise.

The unique equilibrium is given by the function $G^*$:

$$E_{ig}^* = G_0^* + X_{ig}G_1^* + \bar{X}_{-ig}G_2^* + u_{ig}G_3^* + \bar{u}_{-ig}G_4^* + v_{ig}G_5^* + \bar{v}_{-ig}G_6^*$$

(7)

where the parameters $\{G_k^*\}_{k=0}^6$ depend on class size $m_g$ and the subset of structural parameters that enter the best response function (4): $\{\Delta d_1, \Delta d_2, \Delta d_5, \Delta f, \Delta g, p, \bar{p}\}$. The specific functions of the parameters can be found in Appendix B.

Equilibrium effort depends on peer characteristics, even if there are no structural contextual effects on effort ($\Delta g = 0$). As long as own characteristics and peer effort enter the student decision, the equilibrium behavior of the student is to incorporate her peer variables in her own decision. When effort is unobserved, the correlation between observed peer characteristics and the error term will provide biased estimates of the parameters.

Because effort is unobserved, the structural equations in (3) cannot be estimated directly. But it is possible to replace unobserved
effort for their equilibrium (7), and obtain the equilibrium achievement equation:

\[ A_{ig} = P_0 + X_{ig}P_1 + \bar{X}_{-ig}P_2 + X_{ig}P_3X_{ig}' + \bar{X}_{-ig}P_4X_{ig}' + \]

\[ X_{ig}P_5\bar{X}_{-ig}' + \bar{\varepsilon}_{ig} \]  

(8)

\[ \tilde{\varepsilon}_{ig} = u_{ig}X_{ig}Q_1 + \bar{\mu}_{-ig}X_{ig}Q_2 + v_{ig}X_{ig}Q_3 + \bar{\nu}_{-ig}X_{ig}Q_4 + \]

\[ u_{ig}\bar{X}_{-ig}Q_5 + \bar{\mu}_{-ig}\bar{X}_{-ig}Q_6 + v_{ig}\bar{X}_{-ig}Q_7 + \]

\[ \bar{\nu}_{-ig}\bar{X}_{-ig}Q_8 + \alpha_g + \varepsilon_{ig} \]  

(9)

\[ \varepsilon_{ig} = u_{ig}R_1 + \bar{\mu}_{-ig}R_2 + v_{ig}R_3 + \bar{\nu}_{-ig}R_4 + u_{ig}^2R_5 + \bar{\mu}_{-ig}^2R_6 + \]

\[ v_{ig}R_7 + \bar{\nu}_{-ig}^2R_8 + u_{ig}\bar{\mu}_{-ig}R_9 + u_{ig}v_{ig}R_{10} + \]

\[ u_{ig}\bar{\nu}_{-ig}R_{11} + \bar{\mu}_{-ig}v_{ig}R_{12} + \bar{\nu}_{-ig}\bar{\nu}_{-ig}R_{13} + v_{ig}\bar{\nu}_{-ig}R_{14} \]

where the parameters \( \{P_k\}_{k=0}^5, \{Q_k\}_{k=1}^8 \) and \( \{R_k\}_{k=1}^{14} \) depend on class size \( m_g \), the parameters of the production function, and the parameters of equilibrium effort \( \{G_k^*\}_{k=0}^6 \). The specific functions can be found in Appendix B.

One can define the equilibrium contextual effect \( \tilde{\delta} \) as the effect of peer characteristics on equilibrium achievement:

\[ \tilde{\delta}_{ig} = P_2 + P_4\bar{X}_{-ig}' + P_4'\bar{X}_{-ig}' + P_5X_{ig}' + \]

\[ u_{ig}Q_5 + \bar{\mu}_{-ig}Q_6 + v_{ig}Q_7 + \bar{\nu}_{-ig}Q_8 \]  

(9)

which is a random variable depending on own and peer shocks.
Equilibrium parameters in equation (8) are a complicated function of the structural parameters. Without further assumptions, the structural parameters and structural peer effects cannot be identified from the estimated equilibrium parameters. Also note there is no effect of peer achievement on own achievement: any empirical correlation is spurious and its interpretation depends on the functional form of the production and cost functions.

Additionally, the structure of the error term \( \tilde{\varepsilon}_{ig} \) implies an endogeneity problem. Even if the shocks are uncorrelated with observed characteristics in the structural equations, behavior creates a correlation between them in the equilibrium achievement equation, and the estimation of the equilibrium parameters \( \{ P_k \}_{k=0}^5 \) is biased. This endogeneity bias exists unless the parameters \( Q_k = 0, k = 1, \ldots, 8 \). This happens in only three cases:

1. \( p = \bar{p} = 0 \), this is, there is no variation in effort based on unobservables.

2. \( f = g = h = n = 0 \) and \( \Delta f = \Delta g = 0 \), this is, there is no interaction between own and peer effort and observables in the production function, and there is no variation in effort based on observables.

3. \( d_2 = d_4 = d_5 = 0, f = g = h = n = 0 \) and \( p = 0 \), this is, the production function of achievement is linear in own and peer effort, with no assumption on the cost function.
Only in these three cases there is no endogeneity bias. Cases 1 and 2 are implausible, because they imply that the observable characteristics included in the model either explain all the variation in effort or do not explain effort at all. Case 3 is the most plausible, but it is still restrictive, because peer effects are constant for the whole population. This assumption is challenged by the growing empirical evidence of nonlinear peer effects, although these estimates suffer endogeneity bias if the true relation is nonlinear.

These three assumptions also change the relation between equilibrium and structural parameters. This may help in the identification of structural parameters and peer effects. The following sections analyze, for each case, this change in the relation, its identification power, and what happens under the additional assumption that there is only one mechanism for the endogenous effect - peer effort affects either the effort of the student or the production function but not both.

**Case 1: No variation in effort based on unobservables**

When $p = \tilde{p} = 0$, Appendix B.1 shows how the equations for equilibrium effort and achievement change. There is no variation in equilibrium effort based on unobservables, and the observable characteristics of the students explain all the variation in effort. This implies that the error structure in equation (8) is greatly simplified, but the equilibrium parameters $\{P_k\}_{k=0}^5$ are the same as in the general case.
In this case the structural peer effects are the same as in equations (5) and (6), and they cannot be identified. The equilibrium contextual effect is now given by

\[ \tilde{\delta}_{ig} = P_2 + P_4 \bar{X}_{ig} + P_4' \bar{X}'_{ig} + P_5 X'_{ig} \]  

(10)

and can be correctly estimated from the data.

The assumption of only one mechanism for the endogenous effect does not help identification. If there are no production externalities (\( \lambda_A = 0 \)), then \( d_3 = d_4 = d_5 = 0 \) and \( h = n = 0 \). This simplifies the relation between equilibrium and structural parameters, but not enough to provide identification. If there are no utility externalities (\( \lambda_E = 0 \)), then \( \Delta d_5 = 0 \). Equilibrium effort becomes a simple dominant strategy but the achievement equation does not change. These results are shown in Appendix B.1.

**Case 2: No variation in effort based on observables**

When \( f = g = h = n = 0 \) and \( \Delta f = \Delta g = 0 \), Appendix B.2 shows how the equations for equilibrium effort and achievement change. There is no relation between equilibrium effort and the observable variables included in the specification. This simplifies the equilibrium achievement equation: it becomes linear in observed characteristics (\( P_3 = P_4 = P_5 = 0 \)) and the equilibrium parameters \( P_1 \) and \( P_2 \) directly estimate the structural parameters \( b \) and \( c \) respectively.
The structural peer effects are now

$$\delta_A = \delta = c$$

$$\lambda_{Aig} = d_3 + 2d_4 \bar{E}_{-ig} + d_5 E_{ig}$$

$$\delta_E = 0$$

$$\lambda_E = \frac{\Delta d_5}{-2\Delta d_2}$$

and therefore this case provides partial identification: if the Correlated Effects problem from $\alpha_g$ is solved correctly, the structural contextual effect on achievement can be identified from the equilibrium contextual effect.

The assumption of only one mechanism for the endogenous effect does not help identification of other structural parameters. If there are no production externalities, then $d_3 = d_4 = d_5 = 0$. If there are no utility externalities, then $\Delta d_5 = 0$; this is a restrictive case, because it assumes there are no peer effects on effort of any kind. Both cases simplify the structure of the error $\varepsilon_{ig}$, but not enough to provide identification of the structural endogenous effect. These results are shown in Appendix B.2.

**Case 3: Linear production function**

The final case is when $d_2 = d_4 = d_5 = 0$, $f = g = h = n = 0$ and $p = 0$: the production function is linear in own and peer effort. But there is no assumption on the cost function - besides it is quadratic.
- and thus equilibrium effort can depend both on observable and unobservable factors. Appendix [B.3] shows how the equations for equilibrium effort and achievement change.

The parameters \( \{ P_k \}_{k=0}^5 \) and \( \{ R_k \}_{k=1}^{14} \) from the equilibrium achievement equation are greatly simplified, but not equal to the structural parameters. The structural peer effects simplify to

\[
\begin{align*}
\delta_A &= c \\
\lambda_A &= d_3 \\
\delta_E &= \frac{\Delta g}{-2\Delta d_2} \\
\lambda_E &= \frac{\Delta d_5}{-2\Delta d_2}
\end{align*}
\]

(11)

The equilibrium contextual effect simplifies to

\[
\tilde{\delta} = P_2 = c + d_1 G^*_2 + d_3 \left( G^*_1 + G^*_2 \frac{m_g - 2}{m_g - 1} \right)
\]

which can be directly estimated from the data, but it does not help identify the structural peer effects.

Now, the assumption of only one mechanism for the endogenous effect does help partial identification in one case: with no utility externalities \( (\Delta d_5 = 0) \). In this case, equilibrium effort becomes a dominant strategy, and using variation in class size ([Lee 2007]) it is possible to identify the composite peer effect \( \lambda_A \delta_E \). The assumption of no production externalities \( (d_3 = 0) \) simplifies the structure of the
error $\varepsilon_{ig}$, but not enough to provide identification. This results are shown in Appendix B.3.

IV Identification of the Structural Peer Effects in the Linear Case

The three cases with no endogeneity bias are restrictive, but the case of a linear production function is the most plausible of the three. Under this assumption, we analyze the identification strategies most commonly used to solve the Reflection Problem, and explore their identification power. These strategies are:

1. Assume there is no endogenous effect.

2. Use a within transformation and class size variation.

3. Force peer characteristics to have an effect proportional to own characteristics.

4. Use peer achievement as a proxy for peer effort.

These four strategies are not successful at identifying the structural parameters when the effort decision is unobserved. Some of them achieve partial identification when more assumptions are made, for example, when only one mechanism is allowed for the endogenous effect. If there are no production externalities ($d_3 = 0$), using peer achievement as a proxy for peer effort identifies the structural
contextual effect on achievement and the structural contextual effect on effort. If there are no utility externalities ($\Delta d_5 = 0$), forcing peer characteristics to have an effect proportional to own characteristics identifies the proportionality factor, the structural contextual effect on achievement, and the composite parameter $\delta E\lambda_A$. Further identification requires more assumptions on the production and cost function.

Under linearity, the equilibrium contextual effect can be correctly estimated after solving the Correlated Effects problem. This can be done with any of the three strategies described in Section II. Therefore, assume there is no Correlated Effects problem, or that it is possible to solve it.

### IV.1 No Endogenous Effect

This assumption implies that $\Delta d_5 = 0$ and $d_3 = 0$, so there is no effect of peer effort on achievement. In general models of social interactions, this assumption is sufficient to identify the contextual effects, but not here.

The contextual effects become

$$
\delta_A = c
$$

$$
\delta_E = \frac{\Delta g}{-2\Delta d_2}
$$

$$
\tilde{\delta} = c + d_1 \frac{\Delta g}{-2\Delta d_2}
$$
The equilibrium contextual effect depends on the structural contextual effect on achievement, the effect of own effort on achievement, and the structural contextual effect on effort. Thus, it is not possible to separate the direct effect of peer characteristics on effort or on achievement.

This implies that linear peer effect models that assume no endogenous effect do not estimate the structural contextual effect on achievement, but a composition of the total effect of peer characteristics on achievement (Ammermueller and Pischke, 2009; Burke and Sass, 2008; Carrell et al., 2009; Foster, 2006; Hanushek et al., 2009; McEwan, 2003). If $d_1$ and $\delta_E$ are positive, the structural contextual effect on achievement will be smaller than the results from these studies.

**IV.2 Within transformation**

In general models of social interactions, this strategy allows to identify the structural peer effects (Bramoulle et al., 2009; Davezies et al., 2009; Lee, 2007). But when effort is unobserved, this strategy is not enough - as shown in Appendix C.1 - and more assumptions are required. For example, one can identify the structural contextual effect on achievement only if observable characteristics have no effect on effort, or if the contextual structural effect on achievement is the only mechanism for social interactions.
The intuition is that this strategy allows to estimate correctly the equilibrium parameters $P_1$ and $P_2$. But these equilibrium parameters already depend on class size, thus the within transformation does not provide additional variation to help identification. This argument is similar to Davezies et al. (2009), who note that the within transformation strategy implicitly requires that the parameters do not depend on class size for identification.

This implies that linear peer effect models that use the within transformation do not estimate the structural peer effects on achievement, but a composition (Bramoulle et al., 2009). If all structural parameters are positive, the structural peer effects on achievement will be smaller than the results from these studies, but the size of the bias depends on sample class size.

Appendix C.1 shows that assuming one mechanism for the endogenous effect only provides partial identification of the composite parameter $\lambda A \delta_E$ when there are no utility externalities.

### IV.3 Peer Characteristics are Proportional

Arcidiacono et al. (2012) and Burke and Sass (2008) estimate a model that only considers student observed and unobserved fixed effects, with no endogenous effects. Their identifying assumption is that the effect of peer characteristics is proportional to that of own characteristics. For the model presented here, this implies that there
exists $\gamma > 0$ such that

\[
c = \gamma b \quad \tilde{d}_3 = \gamma \tilde{d}_1 \\
d_3 = \gamma d_1 \quad \tilde{d}_4 = \gamma \tilde{d}_2 \\
g = \gamma f \quad \tilde{g} = \gamma \tilde{f}
\]

both in the production and cost function. Note there is no assumption on $\tilde{d}_5$ and the structural peer effects in equation (11) can still have different size.

The equations for equilibrium effort and achievement change accordingly, but there is no further simplification and the structural parameters are not identified. Partial identification is achieved if one assumes a unique mechanism for the endogenous effect, as Appendix C.2 shows. Under no utility externalities the strategy is relatively successful, because the proportionality factor $\gamma$, structural contextual effect on achievement $\delta_A$, and the composite parameter $\delta_E \lambda_A$ are identified. Assuming no production externalities is very restrictive, because it implies there is no effect of own and peer effort in the production function ($d_1 = d_3 = 0$).

IV.4 Peer Achievement as a Proxy for Peer Effort

Some authors think of peer achievement as a proxy for peer unobservables, effort in this case (Boucher et al., 2010; Fruehwirth, 2013; Kang, 2007; Sacerdote, 2001). But this proxy depends on both effort
and observed characteristics. This can confound even more the relation between equilibrium and structural parameters, and thus needs to be analyzed with care. Fruehwirth (2013) reviews the use of peer achievement as a proxy for peer ability. Here, we discuss its use as a proxy for peer effort.

This strategy is also relatively successful. The intuition is that it adds another moment to the estimation - the correlation between own and peer achievement - and one can use the relation between this correlation and the structural parameters for identification.

As Appendix C.3 shows, this strategy uses the moment given by the effort best response (4). Because effort is unobserved, it can be written in terms of observed characteristics and achievement, using the production function of achievement. Replacing this new relation in the effort best response yields a different equilibrium achievement equation, in which individual achievement depends on peer achievement.

Without further assumptions the structural parameters are again not identified. Partial identification is achieved by assuming only one mechanism for peer effort. If there are no utility externalities one can identify the composite parameter \( d_3/d_1 \), which is related to the structural endogenous effect on achievement. If there are no production externalities one can identify the structural peer effects \( \lambda_E \) and \( c \), and thus it is the most successful of the identification strategies studied.
V  A Note on Identification using Peer Achievement as a Proxy for Peer Ability

Fruehwirth (2013) also reviews the use of peer achievement as a proxy for peer ability, but she concludes that the structural parameters are not identified. This section revisits her argument and shows that her strategy does successfully identify all the parameters of her model using variation in class size, which she does not allow.\footnote{Even without using class size variation, under her assumptions she could identify the contextual effect for both observed and unobserved inputs by estimating in two stages. Her equation (3.1) can be directly estimated from the data to identify $\alpha_X$ and $\tilde{\alpha}_X$. Then, one can use this information to estimate equation (3.4) that includes peer achievement, and identify $\tilde{\alpha}_u$.}

This result implies that when unobservable variables are present, their nature is not only relevant because of their implications, but also for their estimation. An unobserved contextual variable like peer ability only has a first round of direct effects and requires less assumptions to identify the structural parameters. Instead, an unobserved endogenous variable like peer effort can generate multiplier effects and requires more assumptions to identify the structural parameters. Therefore, the implications for applying empirical estimates to the questions of regrouping students can be quite different depending on the source of unobservables.

Suppose the following model. There is no effort decision, and there is an unobserved variable that generates contextual effects, namely peer ability. The following equation corresponds to equa-
tion (3.1) from Fruehwirth (2013), including the notation, with the difference that classes have size $m_g$ that varies in the data.

$$Y_{ig} = X_{ig} \alpha_X + \bar{X}_{-ig} \tilde{\alpha}_X + u_{ig} + \bar{u}_{-ig} \tilde{\alpha}_u + \mu + \epsilon_{ig} \quad (12)$$

The notation is equivalent to the linear production function reviewed in this paper, where $b = [\alpha_X, 1]$, $c = [\tilde{\alpha}_X, \tilde{\alpha}_u]$ and $\alpha_g = \mu$. The paper presents assumptions sufficient so the parameters $\tilde{\alpha}_X$ can be consistently estimated.

Similar as the previous section, Fruehwirth (2013) proposes to write peer ability as a function of observed characteristics and achievement, using the achievement production function (12) for $j \neq i$. Replacing this new relation in student’s $i$ production function yields an equilibrium achievement equation where all variables are observed, and individual achievement depends on peer achievement. This is shown in Appendix D. The equilibrium parameters, which can be directly estimated from the data, uniquely determine the structural parameters in equation (12). The unobserved contextual effect $\tilde{\alpha}_u$ comes directly from the equilibrium parameter of peer achievement. The individual effect $\alpha_X$ and observed contextual effect $\tilde{\alpha}_X$ can be identified from the equilibrium parameters of observed characteristics, using class size variation.

This is consistent with Lee (2007), who shows that the class size contains valuable information to identify peer effects. The intuition
is that class size variation mechanically introduces variation in the effect that each particular student has on her peers (smaller effect in larger classes) and this contains valuable information to identify peer effects. The use of peer achievement as a proxy for peer unobserved variables includes the additional moment necessary for identification of an unobserved contextual input. Therefore the strategy proposed by Fruehwirth (2013) is successful.

VI Conclusion

This paper studies two additional sources of bias that arise because the behavioral variable - student effort - is unobserved. To avoid the endogeneity bias in the equilibrium parameter, one must make relatively strong assumptions on the production and cost functions of the student problem. But these three sets of assumptions can be subject to critique. The first two - equilibrium effort does not depend on observable characteristics, or does not depend on unobserved shocks - are implausible. The third case - the production function is linear in own and peer effort - is challenged by the empirical evidence of nonlinear peer effects.

The second estimation problem is the identification of the structural parameters from the estimated equilibrium parameters. The distinction between peer effects is important because they have differing policy implications (Fruehwirth, 2013; Manski, 1993). The
most common strategies achieve partial identification at best, even after assuming a unique mechanism for peer effort, and additional assumptions are needed on the production and cost function. This implies that the current empirical studies overestimate the true effect of peers on achievement.

In conclusion, the unbiased estimation of peer effects in cognitive achievement requires adequate measures of own and peer effort. Unfortunately, these measures are missing from administrative databases that contain standard measures of achievement as test scores and graduation rates. An advance in this area will be necessary to determine the true effect that classmates have on student achievement.

References


Angrist, Joshua and Kevin Lang (2004), ‘Does school integration
generate peer effects? evidence from Boston’s METCO program’, 


Bisin, Alberto, Andrea Moro and Giorgio Topa (2011), The empirical content of models with multiple equilibria in economies with social interactions, NBER working paper no. 17196, National Bureau of Economic Research.

Blume, Lawrence, William Brock, Steven Durlauf and Rajshri Jayaraman (2013), Linear social interactions models, NBER working paper no. 19212, National Bureau of Economic Research.

Blume, Lawrence, William Brock, Steven Durlauf and Yannis Ioan-


Burke, Mary and Tim Sass (2008), Classroom peer effects and student achievement, Working Paper 18, National Center for Analysis of Longitudinal Data in Education Research.


Gibbons, Stephen and Shqiponja Telhaj (2008), Peers and achievement in England’s Secondary Schools, Discussion paper no.1, SERC.


Hoxby, Caroline and Gretchen Weingarth (2006), Taking race out


Lavy, Victor, Olmo Silva and Felix Weinhardt (2009), The Good, the Bad and the Average: Evidence on the scale and nature of ability peer effects in schools, NBER working paper no. 15600, National Bureau of Economic Research.


Rodriguez, Jorge (2010), La no linealidad del efecto par educacional: Evidencia para Chile, Working paper, Ministerio de Hacienda, Chile.


Zimmerman, David (1999), Peer effects in academic outcomes: Ev-
idence from a natural experiment, Technical Report 52, Williams College.
Appendix

A Equilibrium of the Student Effort Game

Redefine the best response function $G$ for each student by adding and subtracting the terms $\Delta d_5 E_{ig} / -2\Delta d_2 (m_g - 1)$ and $X_{ig} \Delta g / -2\Delta d_2 (m_g - 1)$. Reordering the terms, the best response function becomes:

$$E_{ig} = \frac{\Delta d_1 (m_g - 1)}{-2\Delta d_2 (m_g - 1) + \Delta d_5} + \frac{\Delta d_5 m_g E_g}{-2\Delta d_2 (m_g - 1) + \Delta d_5} + \frac{X_{ig} ((m_g - 1)\Delta f + \Delta g)}{-2\Delta d_2 (m_g - 1) + \Delta d_5} + \frac{m_g X_g \Delta g}{-2\Delta d_2 (m_g - 1) + \Delta d_5} + \frac{\tilde{p}_u_{ig} (m_g - 1)}{-2\Delta d_2 (m_g - 1) + \Delta d_5} - \frac{\tilde{p}_v_{ig} (m_g - 1)}{-2\Delta d_2 (m_g - 1) + \Delta d_5}$$

Now, take average over all $i \in g$ to obtain:

$$\bar{E}_g = \frac{\Delta d_1 (m_g - 1)}{-2\Delta d_2 (m_g - 1) + \Delta d_5} + \frac{\Delta d_5 m_g \bar{E}_g}{-2\Delta d_2 (m_g - 1) + \Delta d_5} + \frac{(m_g - 1)\bar{X}_g (\Delta f + \Delta g)}{-2\Delta d_2 (m_g - 1) + \Delta d_5} + \frac{\bar{p}_u_g (m_g - 1)}{-2\Delta d_2 (m_g - 1) + \Delta d_5} - \frac{\tilde{p}_v_g (m_g - 1)}{-2\Delta d_2 (m_g - 1) + \Delta d_5}$$

This function $\bar{E}_g = H(\bar{E}_g)$ describes the average behavior of the class as a function of average effort.

Because the function $H$ is linear, a unique Nash equilibrium ex-
exists if $H' \neq 1$, this is

$$\frac{\Delta d_5 m_g}{-2\Delta d_2(m_g - 1) + \Delta d_5} \neq 1$$

$$\Delta d_5 \neq -2\Delta d_2$$

To prove existence, if $H' < 1$, there exists $\tilde{E}_L$ and $\tilde{E}_H$ with $\tilde{E}_L < \tilde{E}_H$, such that $H(\tilde{E}_g) > \tilde{E}_g$ for all $\tilde{E}_g < \tilde{E}_L$ and $H(\tilde{E}_g) < \tilde{E}_g$ for all $\tilde{E}_g > \tilde{E}_H$ respectively. Similarly, if $H' > 1$, there exists $\tilde{E}_L$ and $\tilde{E}_H$ with $\tilde{E}_L < \tilde{E}_H$, such that $H(\tilde{E}_g) < \tilde{E}_g$ for all $\tilde{E}_g < \tilde{E}_L$ and $H(\tilde{E}_g) > \tilde{E}_g$ for all $\tilde{E}_g > \tilde{E}_H$ respectively. Then, it is possible to bound the domain of $H$ to $[\tilde{E}_L, \tilde{E}_H]$, and by Brouwer’s Fixed Point Theorem\footnote{Every continuous function from a closed ball of an Euclidean space to itself has a fixed point.} a Nash equilibrium $\tilde{E}_g^* = H(\tilde{E}_g^*)$ exists.

To prove uniqueness, suppose that $\tilde{E}_0$ and $\tilde{E}_1$ are two fixed points of $H$, with $\tilde{E}_1 > \tilde{E}_0$ without loss of generality. Then, there exists $a > 0$ such that $\tilde{E}_1 = \tilde{E}_0 + a$. Then $\tilde{E}_1 = H(\tilde{E}_1) = H(\tilde{E}_0 + a)$. If $H' < 1$, then

$$H(\tilde{E}_0 + a) = H(\tilde{E}_0) + \frac{\Delta d_5 m_g a}{-2\Delta d_2(m_g - 1) + \Delta d_5} < \tilde{E}_0 + a$$

which is a contradiction. And if $H' > 1$, then

$$H(\tilde{E}_0 + a) = H(\tilde{E}_0) + \frac{\Delta d_5 m_g a}{-2\Delta d_2(m_g - 1) + \Delta d_5} > \tilde{E}_0 + a$$
which is also a contradiction.

Finally, if $H' = 1$ there is no equilibrium, except for the case where $H(0) = 0$ in which there are infinite equilibria.

**B Equilibrium Functions and Parameters**

The equilibrium equations for effort (7) and achievement (8) have parameters that are functions of the structural parameters of the production and cost functions. Equilibrium effort is given by the function $G^*$:

$$
E_{ig}^* = G_0^* + X_{ig}G_1^* + \bar{X}_{-ig}G_2^* + u_{ig}G_3^* + \bar{u}_{-ig}G_4^* + v_{ig}G_5^* + \bar{v}_{-ig}G_6^*
$$

where

$$
\begin{align*}
D_m &= 4\Delta d_2^2 - 2 \frac{\Delta d_2 \Delta d_5 m_g}{m_g - 1} + \frac{\Delta d_5^2}{m_g - 1} \\
G_0^* &= \frac{\Delta d_1}{D_m} \left( -2 \Delta d_2 + \Delta d_5 \frac{2m_g - 1}{m_g - 1} \right) \\
G_1^* &= \frac{\Delta f}{D_m} \left( -2 \Delta d_2 + \Delta d_5 \frac{m_g}{m_g - 1} \right) + \frac{\Delta g}{D_m} \frac{\Delta d_5}{m_g - 1} \\
G_2^* &= \frac{\Delta f}{D_m} \Delta d_5 + \frac{\Delta g}{D_m} \left( -2 \Delta d_2 + 2 \Delta d_5 \right) \\
G_3^* &= \frac{p}{D_m} \left( -2 \Delta d_2 + \Delta d_5 \frac{m_g}{m_g - 1} \right) \\
G_4^* &= \frac{p}{D_m} \Delta d_5 \\
G_5^* &= \frac{-p}{D_m} \left( -2 \Delta d_2 + \Delta d_5 \frac{m_g}{m_g - 1} \right) \\
G_6^* &= \frac{-p}{D_m} \Delta d_5
\end{align*}
$$

(13)
Define some auxiliary parameters, composed of parameters from the production function and the equilibrium effort:

\[
\begin{align*}
B_1 &= d_1 + 2d_2G_0^* + d_5G_0^* \\
B_2 &= d_3 + 2d_4G_0^* + d_5G_0^* \\
B_3 &= G_1^* + G_2^* \frac{m_g - 2}{m_g - 1} \\
B_4 &= G_3^* + G_4^* \frac{m_g - 2}{m_g - 1} \\
B_5 &= G_5^* + G_6^* \frac{m_g - 2}{m_g - 1} \\
B_6 &= 2d_2G_3^* + d_5 \frac{G_4^*}{m_g - 1} \\
B_7 &= 2d_2G_5^* + d_5 \frac{G_6^*}{m_g - 1} \\
B_8 &= 2d_4 \frac{G_4^*}{m_g - 1} + d_5G_3^* \\
B_9 &= 2d_4 \frac{G_6^*}{m_g - 1} + d_5G_5^* \\
B_{10} &= 2d_4G_4^* + d_5B_4 \\
B_{11} &= 2d_4G_6^* + d_5B_5 \\
B_{12} &= 2d_4B_4 + d_5G_4^* \\
B_{13} &= 2d_4B_5 + d_5G_6^*
\end{align*}
\]
With some algebra, the equilibrium achievement equation is:

\[ A_{ig} = P_0 + X_{ig}P_1 + \bar{X}_{-ig}P_2 + X_{ig}P_3X_{ig}' + \bar{X}_{-ig}P_4\bar{X}_{-ig}' + X_{ig}P_5\bar{X}_{-ig}' + \bar{\varepsilon}_{ig} \]  \hspace{1cm} (14)

where

\[
\begin{align*}
P_0 &= a + (d_1 + d_3)G_0^* + (d_2 + d_4 + d_5)G_0^{*2} \\
P_1 &= b + G_0^*(f + h) + B_1G_1^* + B_2\frac{G_2^*}{m_g - 1} \\
P_2 &= c + G_0^*(g + n) + B_1G_2^* + B_2B_3 \\
P_3 &= G_1^*\left(d_2G_1^* + \frac{d_5^*}{G_2}m_g - 1 + f\right)' + G_2^*\left(\frac{d_4G_2^*}{(m_g - 1)^2} + h\right)' \\
P_4 &= G_2^*(d_2G_2^* + d_5B_3 + g)' + B_3(d_4B_3 + n)' \\
P_5 &= G_1^*(d_2G_2^* + d_5B_3 + g)' + \\
&\frac{1}{m_g - 1}G_2^*(d_5G_2^* + 2d_4B_3 + n)' + hB_3' + fG_2'' \\
\end{align*}
\]

and the error term \( \bar{\varepsilon}_{ig} \) is given by

\[
\bar{\varepsilon}_{ig} = u_{ig}X_{ig}Q_1 + \bar{u}_{-ig}X_{ig}Q_2 + v_{ig}X_{ig}Q_3 + \bar{v}_{-ig}X_{ig}Q_4 + \\
u_{ig}\bar{X}_{-ig}Q_5 + \bar{u}_{-ig}\bar{X}_{-ig}Q_6 + v_{ig}\bar{X}_{-ig}Q_7 + \bar{v}_{-ig}\bar{X}_{-ig}Q_8 + \\
\alpha_g + \bar{\varepsilon}_{ig} \]  \hspace{1cm} (15)
where

\[
Q_1 = G_3^* f + \frac{G_4^*}{m_g - 1} h + (B_6 + p) G_1^* + B_8 G_2^*
\]

\[
Q_2 = G_4^* f + B_4 h + B_{10} G_1^* + \frac{B_{12}}{m_g - 1} G_2^*
\]

\[
Q_3 = G_5^* f + \frac{G_6^*}{m_g - 1} h + B_7 G_1^* + \frac{B_9}{m_g - 1} G_2^*
\]

\[
Q_4 = G_6^* f + B_5 h + B_{11} G_1^* + \frac{B_{13}}{m_g - 1} G_2^*
\]

\[
Q_5 = G_7^* g + \frac{G_4^*}{m_g - 1} n + (B_6 + p) G_2^* + B_8 B_3
\]

\[
Q_6 = G_8^* g + B_4 n + B_{10} G_2^* + B_{12} B_3
\]

\[
Q_7 = G_9^* g + \frac{G_6^*}{m_g - 1} n + B_7 G_2^* + B_9 B_3
\]

\[
Q_8 = C_6 g + B_5 n + B_{11} C_2 + \frac{B_{13}}{m_g - 1} B_3
\]

and finally, the error term \( \varepsilon_{ig} \) is given by

\[
\varepsilon_{ig} = u_{ig} R_1 + \bar{u}_{-ig} R_2 + v_{ig} R_3 + \bar{v}_{-ig} R_4 + u_{ig}^2 R_5 + \bar{u}_{-ig}^2 R_6 + v_{ig}^2 R_7 + \bar{v}_{-ig}^2 R_8 + u_{ig} \bar{u}_{-ig} R_9 + u_{ig} v_{ig} R_{10} + u_{ig} \bar{v}_{-ig} R_{11} + \bar{u}_{-ig} v_{ig} R_{12} + \bar{u}_{-ig} \bar{v}_{-ig} R_{13} + v_{ig} \bar{v}_{-ig} R_{14}
\]

(16)
where

\[
R_1 = B_1 G_3^* + B_2 \frac{G_4^*}{m_g - 1} + pG_0^* + 1
\]

\[
R_2 = B_1 G_4^* + B_2 B_4
\]

\[
R_3 = B_1 G_5^* + B_2 \frac{G_6^*}{m_g - 1}
\]

\[
R_4 = B_1 G_6^* + B_2 B_5
\]

\[
R_5 = d_2 G_3^{2*} + d_4 \left( \frac{G_4^*}{m_g - 1} \right)^2 + d_5 G_3^* \frac{G_4^*}{m_g - 1} + pG_3^*
\]

\[
R_6 = d_2 G_4^{2*} + d_4 B_4^2 + d_5 G_4^* B_4
\]

\[
R_7 = d_2 G_5^{2*} + d_4 \left( \frac{G_6^*}{m_g - 1} \right)^2 + d_5 G_5^* \frac{G_6^*}{m_g - 1}
\]

\[
R_8 = d_2 G_6^{2*} + d_4 B_5^2 + d_5 G_6^* B_5
\]

\[
R_9 = (B_6 + p) G_4^* + B_8 B_4
\]

\[
R_{10} = (B_6 + p) G_5^* + B_8 \frac{G_6^*}{m_g - 1}
\]

\[
R_{11} = (B_6 + p) G_6^* + B_8 B_5
\]

\[
R_{12} = B_{10} G_5^* + B_{12} \frac{G_6^*}{m_g - 1}
\]

\[
R_{13} = B_{10} G_6^* + B_{12} B_5
\]

\[
R_{14} = B_7 G_6^* + B_9 B_5
\]

**B.1 No variation in effort based on unobservables**

When \( p = \bar{p} = 0 \), equilibrium effort for student \( i \) becomes

\[
E_{ig}^* = G_0^* + X_{ig} G_1^* + \bar{X}_{ig} G_2^*
\]
with \( G_0^*, G_1^*, G_2^* \) same as in equation (13) above, and \( G_3^* = G_4^* = G_5^* = G_6^* = 0 \). This implies that the auxiliary parameters \( B_k = 0, k = 4, \ldots, 13 \). Then, all parameters \( Q_k = 0, k = 1, \ldots, 8 \) in equation (15), and all parameters \( R_k = 0, k = 2, \ldots, 14 \) in equation (16).

The error term becomes \( \tilde{\varepsilon}_{ig} = \alpha_g + u_{ig} \), and the equilibrium achievement equation is

\[
A_{ig} = P_0 + X_{ig}P_1 + \bar{X}_{-ig}P_2 + X_{ig}P_3X_{ig} + \bar{X}_{-ig}P_4\bar{X}_{-ig} + X_{ig}P_5\bar{X}_{-ig} + \alpha_g + u_{ig}
\]

with equilibrium parameters \( \{P_k\}_{k=0}^5 \) same as equation (14).

The assumption of only one mechanism for the endogenous effect does not help identification. If \( \lambda_A = 0 \), then \( d_3 = d_4 = d_5 = 0 \) and \( h = n = 0 \). Equilibrium effort does not change, and because \( B_2 = 0 \) the equilibrium achievement equation is simplified to

\[
A_{ig} = P_0 + X_{ig}P_1 + \bar{X}_{-ig}P_2 + X_{ig}P_3X_{ig} + \bar{X}_{-ig}P_4\bar{X}_{-ig} + X_{ig}P_5\bar{X}_{-ig} + \tilde{\varepsilon}_{ig}
\]  

(17)
where

\[
\begin{align*}
P_0 &= a + d_1 G_0^* + d_2 G_0^{*2} \\
P_1 &= b + G_0^* f + B_1 G_1^* \\
P_2 &= c + G_0^* g + B_1 G_2^* \\
P_3 &= G_1^* (d_2 G_1^* + f)' \\
P_4 &= G_2^* (d_2 G_2^* + g)' \\
P_5 &= G_1^* (d_2 G_2^* + g)' + f G_2^{*'}
\end{align*}
\]

If \( \lambda_E = 0 \), then \( \Delta d_5 = 0 \). Equilibrium effort becomes

\[
E_{ig}^* = \frac{\Delta d_1}{-2\Delta d_2} + X_{ig} \frac{\Delta f}{-2\Delta d_2} + \bar{X}_{-ig} \frac{\Delta g}{-2\Delta d_2}
\]

but the equilibrium parameters do not change - only with the new \( G_0^*, G_1^*, G_2^* \) - and thus the structural parameters are not identified.

\section*{B.2 No variation in effort based on observables}

When \( f = g = h = n = 0 \) and \( \Delta f = \Delta g = 0 \), equilibrium effort for student \( i \) becomes

\[
E_{ig}^* = G_0^* + u_{ig} G_3^* + \bar{u}_{-ig} G_4^* + v_{ig} G_5^* + \bar{v}_{-ig} G_6^*
\]

with \( G_0^*, G_3^*, G_4^*, G_5^*, G_6^* \) same as in equation (13) above, and \( G_1^* = G_2^* = 0 \). This implies that the auxiliary parameter \( B_3 = 0 \). Then, all
parameters \( Q_k = 0, k = 1, \ldots, 8 \) in equation (15).

The error term becomes \( \bar{\varepsilon}_{ig} = \alpha_g + \varepsilon_{ig} \), and the equilibrium achievement equation is

\[
A_{ig} = P_0 + X_{ig}b + \bar{X}_{-ig}c + \alpha_g + \varepsilon_{ig}
\]

where

\[
P_0 = a + (d_1 + d_3)G_0^* + (d_2 + d_4 + d_5)G_0^{*2}
\]

with error term \( \varepsilon_{ig} \) same as equation (16), and parameters \( P_3 = P_4 = P_5 = 0 \). The equilibrium parameters \( P_1 \) and \( P_2 \) directly estimate the structural parameters \( b \) and \( c \), which also corresponds to the structural contextual effect on achievement.

The assumption of only one mechanism for the endogenous effect does not help identification of other structural parameters. If \( \lambda_A = 0 \), then \( d_3 = d_4 = d_5 = 0 \). The auxiliary parameters \( B_k = 0, k = 2, 8, \ldots, 13 \) and this simplifies the structure of the error \( \varepsilon_{ig} \) to

\[
\varepsilon_{ig} = u_{ig}R_1 + \bar{u}_{-ig}R_2 + v_{ig}R_3 + \bar{v}_{-ig}R_4 + u_{ig}^2R_5 + \bar{u}_{-ig}^2R_6 + u_{ig}v_{ig}R_7 + \bar{u}_{-ig}v_{-ig}R_8 + u_{ig}\bar{u}_{-ig}R_9 + u_{ig}v_{ig}R_{10} + u_{ig}\bar{v}_{-ig}R_{11} + v_{ig}\bar{v}_{-ig}R_{14}
\]

(18)
where

\begin{align*}
R_1 &= B_1 G_3^* + p G_0^* + 1 \\
R_2 &= B_1 G_4^* \\
R_3 &= B_1 G_5^* \\
R_4 &= B_1 G_6^* \\
R_5 &= d_2 G_3^* + p G_3^* \\
R_6 &= d_2 G_4^* \\
R_7 &= d_2 G_5^* \\
R_8 &= d_2 G_6^* \\
R_9 &= (B_6 + p) G_4^* \\
R_{10} &= (B_6 + p) G_5^* \\
R_{11} &= (B_6 + p) G_6^* \\
R_{14} &= B_7 G_6^*
\end{align*}

with \( R_{12} = R_{13} = 0 \). Unfortunately the parameters do not simplify enough to provide identification of \( \Delta d_5 \).

If \( \lambda_E = 0 \), then \( \Delta d_5 = 0 \). Equilibrium effort becomes a dominant strategy:

\[
E_{ig}^* = \frac{\Delta d_1}{-2 \Delta d_2} + u_{ig} \frac{p}{-2 \Delta d_2} + v_{ig} \frac{\tilde{p}}{-2 \Delta d_2}
\]

with \( G_4^* = G_6^* = 0 \). The auxiliary parameters \( \{B_k\}_{k=1}^{13} \) and the pa-
rameters of the error term \( \{R_k\}_{k=1}^{14} \) are greatly simplified and do not depend on class size, but the simplification is not enough to identify the structural parameters.

**B.3 Linear Production Function**

When \( d_2 = d_4 = d_5 = 0, f = g = h = n = 0 \) and \( p = 0 \), equilibrium effort for student \( i \) becomes

\[
E_{ig}^* = G_0^* + X_{ig}G_1^* + \bar{X}_{-ig}G_2^* + v_{ig}G_5^* + \bar{v}_{-ig}G_6^*
\]

with \( G_3^* = G_4^* = 0 \) and other parameters same as in equation (13). This implies that the auxiliary parameters \( B_k = 0, k = 4, 6, \ldots, 13 \) and

\[
\begin{align*}
B_1 &= d_1 \\
B_2 &= d_3 \\
B_3 &= G_1^* + G_2^* \frac{m_g}{m_g - 1} - 2 \\
B_5 &= G_5^* + G_6^* \frac{m_g}{m_g - 1} - 2
\end{align*}
\]

Then, all parameters \( Q_k = 0, k = 1, \ldots, 8 \) in equation (15) and the equilibrium achievement equation is simplified to

\[
A_{ig} = P_0 + X_{ig}P_1 + \bar{X}_{-ig}P_2 + \alpha_g + \varepsilon_{ig} \quad (19)
\]
where

\[
\begin{align*}
P_0 &= a + (d_1 + d_3)G_0^* \\
P_1 &= b + d_1G_1^* + d_3\frac{G_2^*}{m_g - 1} \\
P_2 &= c + d_1G_2^* + d_3B_3
\end{align*}
\]

with parameters \( P_3 = P_4 = P_5 = 0 \) and error term \( \varepsilon_{ig} \) given by:

\[
\varepsilon_{ig} = u_{ig} + v_{ig}R_3 + \bar{v}_{-ig}R_4
\]

where

\[
\begin{align*}
R_3 &= d_1G_5^* + d_3\frac{G_6^*}{m_g - 1} \\
R_4 &= d_1G_6^* + d_3B_5
\end{align*}
\]

Now, the assumption of only one mechanism for the endogenous effect does help partial identification in one case: with no utility externalities. In this case, \( \Delta d_5 = 0 \) and equilibrium effort becomes

\[
E_{ig}^* = \frac{\Delta d_1}{-2\Delta d_2} + X_{ig} \frac{\Delta f}{-2\Delta d_2} + \bar{X}_{-ig} \frac{\Delta g}{-2\Delta d_2} + v_{ig} \frac{-\bar{p}}{-2\Delta d_2}
\]

with \( G_6^* = 0 \). Equilibrium achievement is now

\[
A_{ig} = P_0 + X_{ig}P_1 + \bar{X}_{-ig}P_2 + \gamma + \varepsilon_{ig}
\]

52
where

\[
P_0 = a + (d_1 + d_3) \frac{\Delta d_1}{-2\Delta d_2},
\]

\[
P_1 = b + d_1 \frac{\Delta f}{-2\Delta d_2} + \frac{d_3}{m_g - 1} \frac{\Delta g}{-2\Delta d_2},
\]

\[
P_2 = c + d_1 \frac{\Delta g}{-2\Delta d_2} + d_3 \left( \frac{\Delta f}{-2\Delta d_2} + \frac{\Delta g}{-2\Delta d_2} \frac{m_g - 2}{m_g - 1} \right),
\]

\[
\varepsilon_{ig} = u_{ig} + v_{ig}d_1 \frac{-\tilde{p}}{-2\Delta d_2} + \bar{v}_{-ig}d_3 \frac{-\tilde{p}}{-2\Delta d_2}.
\]

Under these assumptions, if it is possible to solve the Correlated Effects problem, then one can estimate the equilibrium parameters \(P_1\) and \(P_2\) for different class sizes. Using Lee (2007), it is possible to identify the composite parameter \(d_3\Delta g/-2\Delta d_2\), which corresponds to \(\lambda_A \delta_E\).

On the other hand, if there are no production externalities, there is no identification of structural parameters. The assumption of \(d_3 = 0\) further simplifies the parameters \(P_k\) and \(R_k\) above

\[
A_{ig} = (a + d_1 G^*_0) + X_{ig}(b + d_1 G^*_1) + \bar{X}_{-ig}(c + d_1 G^*_2) + \alpha_g + \varepsilon_{ig}
\]

\[
\varepsilon_{ig} = u_{ig} + v_{ig}d_1 G^*_5 + \bar{v}_{-ig}d_3 G^*_6
\]

but the \(G^*\) parameters are the same as in equation (13), and the structural parameters are not identified.
C Identification Strategies in the Linear Case

C.1 Within transformation

The within transformation of equation (19) is

\[ A_{ig} - \bar{A}_g = (X_{ig} - \bar{X}_g) \left( P_1 - \frac{P_2}{m_g - 1} \right) + \varepsilon_{ig} - \bar{\varepsilon}_g \]

\[ \varepsilon_{ig} - \bar{\varepsilon}_g = u_{ig} - \bar{u}_g + (v_{ig} - \bar{v}_g) \left( R_3 - \frac{R_4}{m_g - 1} \right) \]

where

\[ P_1 - \frac{P_2}{m_g - 1} = b - \frac{c}{m_g - 1} + \frac{(d_1 - \frac{d_3}{m_g - 1})}{-2\Delta d_2 + \frac{\Delta d_5}{m_g - 1}} \left( \Delta f - \frac{\Delta g}{m_g - 1} \right) \]

\[ R_3 - \frac{R_4}{m_g - 1} = \frac{-\hat{p} \left( d_1 - \frac{d_3}{m_g - 1} \right)}{-2\Delta d_2 + \frac{\Delta d_5}{m_g - 1}} \]

and thus the structural parameters are not identified. The effect \( c \) could be identified only if \( \Delta f = \Delta g = 0 \) or if \( d_3 = \Delta d_5 = \Delta g = 0 \).

As an example, in this last case the composite parameter becomes

\[ P_1 - \frac{P_2}{m_g - 1} = b - \frac{c}{m_g - 1} + \frac{d_1}{-2\Delta d_2} \left( \frac{\Delta f}{\Delta d_2} \right) \]

Then, one can estimate the equilibrium equation in 2 samples with different class sizes, and use this variation to identify \( c \).

Assuming one mechanism for the endogenous effect achieves partial identification when \( \Delta d_5 = 0 \). The composite parameter be-
comes

\[
P_1 - \frac{P_2}{m_g - 1} = \frac{b - \frac{c}{m_g - 1} + \left(d_1 - \frac{d_3}{m_g - 1}\right) \left(\Delta f - \frac{\Delta g}{m_g - 1}\right)}{-2\Delta d_2}
\]

\[
= b + \frac{\frac{d_1 \Delta f}{-2\Delta d_2} - \frac{1}{m_g - 1}}{\left(c + \frac{\frac{d_1 \Delta g}{-2\Delta d_2} + \frac{d_3 \Delta f}{-2\Delta d_2}}{m_g - 1}\right)}
+ \frac{1}{\left(m_g - 1\right)^2 - 2\Delta d_2}
\]

One can estimate the equilibrium equation in 3 samples with different class size, and use this variation to identify the composite parameters \((b + d_1 \Delta f / -2\Delta d_2), (c + d_1 \Delta g / -2\Delta d_2 + d_3 \Delta f / -2\Delta d_2)\) and \((d_3 \Delta g / -2\Delta d_2)\). This last parameter corresponds to \(\lambda A \delta E\).

On the other hand, when \(d_3 = 0\) the composite parameter becomes

\[
P_1 - \frac{P_2}{m_g - 1} = b - \frac{c}{m_g - 1} + \frac{d_1 \left(\Delta f - \frac{\Delta g}{m_g - 1}\right)}{-2\Delta d_2 + \frac{\Delta d_5}{m_g - 1}}
\]

and the structural parameters are not identified.

### C.2 Peer Characteristics are Proportional

Partial identification is achieved assuming one mechanism for the endogenous effect. If \(\Delta d_5 = 0\), equilibrium achievement becomes

\[
A_{ig} = P_0 + X_{ig} P_1 + X_{-ig} P_2 + \alpha_g + \epsilon_{ig}
\]
where

\[ P_0 = a + d_1(1 + \gamma) \frac{\Delta d_1}{-2\Delta d_2} \]
\[ P_1 = b + \frac{d_1 \Delta f}{-2\Delta d_2} + \frac{1}{m_g - 1} \frac{\gamma^2 d_1 \Delta f}{-2\Delta d_2} \]
\[ P_2 = \gamma b + 2 \frac{\gamma d_1 \Delta f}{-2\Delta d_2} + \frac{m_g - 2 \gamma^2 d_1 \Delta f}{m_g - 1} \frac{\Delta f}{-2\Delta d_2} \]

One can estimate this equation in samples with different class size, and use this variation to identify the composite parameters \((b + \frac{d_1 \Delta f}{-2\Delta d_2})\), \((\gamma b + 2 \frac{\gamma d_1 \Delta f}{-2\Delta d_2})\) and \((\gamma^2 d_1 \Delta f / -2\Delta d_2)\). These correspond to three equations with three unknowns, and thus it is possible to identify \(b, \gamma, \text{ and } d_1 \Delta f / -2\Delta d_2\).

Under no production externalities \((d_3 = d_1 = 0)\), equilibrium achievement becomes

\[ A_{ig} = a + X_{ig} b + \bar{X}_{-ig} c + \alpha_g + u_{ig} \]

and thus the equilibrium contextual effect identifies the structural contextual effect on achievement \(c\). The other structural peer effects \(\Delta g, \Delta d_5\) are not identified.

C.3 Peer Achievement as a Proxy for Peer Effort

In the linear case, the production function is given by

\[ A_{ig} = a + X_{ig} b + \bar{X}_{-ig} c + d_1 E_{ig}^* + d_3 \bar{E}_{-ig}^* + u_{ig} \]
Define $S_{ig} = A_{ig} - a - X_{ig}b - \bar{X}_{-ig}c - u_{ig}$. Then, one can write the system of equations for class $g$:

$$
\begin{bmatrix}
S_{1g} \\
S_{2g} \\
\vdots \\
S_{mg}
\end{bmatrix} =
\begin{bmatrix}
d_1 & \frac{d_3}{m_g-1} & \ldots & \frac{d_3}{m_g-1} \\
\frac{d_3}{m_g-1} & d_1 & \ldots & \frac{d_3}{m_g-1} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{d_3}{m_g-1} & \frac{d_3}{m_g-1} & \ldots & d_1
\end{bmatrix}
\begin{bmatrix}
E_{1g}^* \\
E_{2g}^* \\
\vdots \\
E_{mg}^*
\end{bmatrix}
$$

The determinant of the matrix is different than zero if $d_1, d_3 \neq 0$, and it is possible to invert this system to obtain:

$$
\begin{bmatrix}
E_{1g}^* \\
E_{2g}^* \\
\vdots \\
E_{mg}^*
\end{bmatrix} =
\begin{bmatrix}
a_1 & a_2 & \ldots & a_2 \\
a_2 & a_1 & \ldots & a_2 \\
\vdots & \vdots & \ddots & \vdots \\
a_2 & a_2 & \ldots & a_1
\end{bmatrix}
\begin{bmatrix}
S_{1g} \\
S_{2g} \\
\vdots \\
S_{mg}
\end{bmatrix}
$$

(20)

where

$$
a_1 = \frac{d_1(m_g - 1) + d_3(m_g - 2)}{d_1^2(m_g - 1) + d_1 d_3(m_g - 2) - d_3^2}
$$

$$
a_2 = \frac{-d_3}{d_1^2(m_g - 1) + d_1 d_3(m_g - 2) - d_3^2}
$$

Then, unobserved effort can be written in terms of observed characteristics and achievement:

$$
E_{ig}^* = a_1 S_{ig} + \sum_{j \neq i} a_2 S_{jg}
$$

(21)

$$
\bar{E}_{-ig}^* = a_1 \bar{S}_{-ig} + a_2 S_{ig} + a_2(m_g - 2) \bar{S}_{-ig}
$$
Replacing equations (21) in the effort best response (4), it becomes:

\[
E_{ig}^* = \frac{\Delta d_1}{-2\Delta d_2} + \lambda_E \ddot{E}_{-ig}^* + X_{ig} \frac{\Delta f}{-2\Delta d_2} + \dot{X}_{-ig} \delta_E - \frac{\tilde{p}_{ig}}{-2\Delta d_2} \\
= a_1 S_{ig} + \sum_{j \neq i} a_2 S_{ij} \\
= \frac{\Delta d_1}{-2\Delta d_2} + \lambda_E \left((a_1 + a_2(m_g - 2)) \ddot{S}_{-ig} + a_2 S_{ig}\right) + X_{ig} \frac{\Delta f}{-2\Delta d_2} + \dot{X}_{-ig} \delta_E - \frac{\tilde{p}_{ig}}{-2\Delta d_2}
\]

Replacing \( S_{ig} \) and with some algebra, this equation becomes

\[
A_{ig} = T_0 + X_{ig}T_1 + \dot{X}_{-ig}T_2 + T_3\ddot{A}_{-ig} + \epsilon_{ig} \quad (22)
\]

where

\[
T_0 = a(1 - T_3) + \frac{\Delta d_1}{-2\Delta d_2 a_1 - a_2 \lambda_E} \\
T_1 = b + \frac{1}{a_1 - a_2 \lambda_E - \frac{\Delta f}{-2\Delta d_2}} - \frac{T_3 c}{m_g - 1} \\
T_2 = c + \frac{1}{a_1 - a_2 \lambda_E} \delta_E - T_3 b - T_3 \frac{m_g - 2}{m_g - 1} c \\
T_3 = \frac{\lambda_E (a_1 + a_2(m_g - 2)) - a_2(m_g - 1)}{a_1 - a_2 \lambda_E} \\
\epsilon_{ig} = u_{ig} - T_3 \bar{u}_{-ig} + v_{ig} \frac{-\tilde{p}}{-2\Delta d_2 a_1 - a_2 \lambda_E} \frac{1}{-2\Delta d_2 a_1 - a_2 \lambda_E}
\]

Partial identification is achieved if there is only one mechanism for the endogenous effect. If \( \Delta d_5 = 0 \), then

\[
T_3 = -\frac{a_2(m_g - 1)}{a_1} = \frac{d_3(m_g - 1)}{d_1(m_g - 1) + d_3(m_g - 2)}
\]

58
and class size variation identifies the composite parameter \( d_3/d_1 \).

Other structural parameters cannot be identified.

If \( d_3 = 0 \), then \( a_1 = 1/d_1, a_2 = 0 \), and the equilibrium parameter \( T_3 \) directly identifies the structural endogenous effect on effort:

\[
T_3 = \frac{\lambda_E}{d_1} = \lambda_E
\]

The parameters of equation (22) above become

\[
T_0 = a(1 - \lambda_E) + d_1 \frac{\Delta d_1}{-2\Delta d_2} \\
T_1 = b + \frac{d_1 \Delta f}{-2\Delta d_2} - \frac{\lambda_E c}{m_g - 1} \\
T_2 = c + d_1 \delta_E - \lambda_E b - \lambda_E \frac{m_g - 2c}{m_g - 1}
\]

and thus the structural contextual effect on achievement is identified using class size variation. One can also identify the composite parameters \((b + d_1 \Delta f / -2\Delta d_2)\) and \((c + d_1 \delta_E - \lambda_E b)\).
D Peer Achievement as a Proxy for Peer Ability

For every student $j \neq i$, one can use the production function (12) to solve for $u_j$ as a function of $j$’s achievement and other inputs:

\[ u_{jg} = Y_{jg} - X_{jg} \alpha_X - X_{jg} \bar{\alpha}_X - u_{-jg} \bar{\alpha}_u - \mu - \epsilon_{jg} \quad (23) \]

Adding for all $j \neq i$ and rearranging terms yields

\[
\bar{u}_{-ig} = \frac{\bar{Y}_{-ig} - X_{ig} \alpha_X}{\phi(m_g - 1)} - \frac{X_{-ig}}{\phi} \left( \alpha_X + \bar{\alpha}_X \frac{m_g - 2}{m_g - 1} \right) \]
\[
- \frac{u_{ig} \alpha_u}{\phi(m_g - 1)} - \frac{\mu}{\phi} - \bar{\epsilon}_{-ig}
\]

where

\[ \phi = 1 + \bar{\alpha}_u \frac{m_g - 2}{m_g - 1} \]

Replacing this equation in the production function for student $i$ yields an equilibrium achievement equation, equivalent to equation (3.4) in
\[ Y_{ig} = X_{ig} \gamma_X + \bar{X}_{ig} \bar{\gamma}_X + \bar{Y}_{ig} \bar{\gamma}_Y + u_{ig} \gamma_u + \mu \gamma_\mu + \epsilon_{ig} - \bar{\epsilon}_{ig} \bar{\gamma}_Y \]

\[ \gamma_X = \alpha_X - \frac{\bar{\alpha}_u \bar{\alpha}_X}{m_g - 1 + \bar{\alpha}_u m_g - 2\bar{\alpha}_u} \]

\[ \bar{\gamma}_X = \bar{\alpha}_X - \frac{\bar{\alpha}_u (\alpha_X (m_g - 1) + \bar{\alpha}_X (m_g - 2))}{m_g - 1 + \bar{\alpha}_u m_g - 2\bar{\alpha}_u} \]

\[ \bar{\gamma}_Y = \frac{\bar{\alpha}_u (m_g - 1)}{m_g - 1 + \bar{\alpha}_u m_g - 2\bar{\alpha}_u} \]

\[ \gamma_u = 1 - \frac{\bar{\alpha}_u^2}{m_g - 1 + \bar{\alpha}_u m_g - 2\bar{\alpha}_u} \]

\[ \gamma_\mu = 1 - \frac{\bar{\alpha}_u (m_g - 1)}{m_g - 1 + \bar{\alpha}_u m_g - 2\bar{\alpha}_u} \]

When \( m_g = 2 \), this equation has the same parameters than equation (3.4) in Fruehwirth (2013).

One can identify all the structural parameters \( \alpha \) from the equilibrium parameters \( \gamma \) using class size variation. The unobserved contextual effect \( \bar{\alpha}_u \) comes directly from the correlation between own and peer achievement:

\[ \bar{\gamma}_Y = \frac{\bar{\alpha}_u (m_g - 1)}{m_g - 1 + \bar{\alpha}_u m_g - 2\bar{\alpha}_u} \]

\[ \Rightarrow \]

\[ \bar{\alpha}_u = \frac{\bar{\gamma}_Y (m_g - 1)}{m_g - 1 - \bar{\gamma}_Y (m_g - 2)} \]

To identify the individual effects \( \alpha_X \) and the observed contextual
effects $\tilde{\alpha}_X$, consider

\[
\gamma_X = \alpha_X - \frac{\tilde{\alpha}_u \tilde{\alpha}_X}{m_g - 1 + \tilde{\alpha}_u m_g - 2\tilde{\alpha}_u} = \frac{m_g (\alpha_X + \tilde{\alpha}_u \alpha_X) - (\alpha_X + 2\tilde{\alpha}_u \alpha_X + \tilde{\alpha}_u \tilde{\alpha}_X)}{m_g - 1 + \tilde{\alpha}_u m_g - 2\tilde{\alpha}_u}
\]

Take two different class sizes $m_1$ and $m_2$, and estimate $\gamma_X$ for each one separately. Define the fixed parameters $A = \alpha_X + \tilde{\alpha}_u \alpha_X$ and $B = \alpha_X + 2\tilde{\alpha}_u \alpha_X + \tilde{\alpha}_u \tilde{\alpha}_X$. Using the estimation and knowledge of $\tilde{\alpha}_u$, compute $K_m = \gamma_X (m_g - 1 + \tilde{\alpha}_u m_g - 2\tilde{\alpha}_u)$. With two class sizes one can compute the system of equations

\[
\begin{bmatrix}
    m_1 & -1 \\
    m_2 & -1
\end{bmatrix}
\begin{bmatrix}
    A \\
    B
\end{bmatrix} =
\begin{bmatrix}
    K_1 \\
    K_2
\end{bmatrix}
\]

If $m_1 \neq m_2$ the system has a unique solution $[A, B]$. Then the following system of equations

\[
A = \alpha_X + \tilde{\alpha}_u \alpha_X \\
B = \alpha_X + 2\tilde{\alpha}_u \alpha_X + \tilde{\alpha}_u \tilde{\alpha}_X
\]

provides identification of $\alpha_X$ and $\tilde{\alpha}_X$, completing the identification of the model.