Syracuse University
SURFACE

Dissertations - ALL | SURFACE

June 2020

The Makings of a Peer: Evidence on Within-Classroom Heterogeneities in Peer Effects

Jonathan L. Presler
Syracuse University

Follow this and additional works at: https://surface.syr.edu/etd

Part of the Social and Behavioral Sciences Commons

Recommended Citation

Presler, Jonathan L., "The Makings of a Peer: Evidence on Within-Classroom Heterogeneities in Peer Effects" (2020). *Dissertations - ALL*. 1257.
https://surface.syr.edu/etd/1257

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

Abstract

This dissertation consists of three essays on peer effects in elementary school classrooms using data from New York City (NYC) public schools. Each chapter explores a different component of within-classroom interactions in order to build towards an understanding of what makes a peer relevant.

The first chapter proxies for the social network using a set of shared characteristics (homophily). In this coauthored work, we are not interested in measuring the classroom peer effect, but rather we estimate and then rank each network's effect on academic spillovers. This answers the question: which characteristics are socially important, and by how much?

The second chapter uses a novel method to measure the social importance of classmates based on student proximity in the lunch line over the course of the school year. The result is a revealed friendship network which I use to estimate peer effects. To my knowledge, this is the first paper to measure the social importance of classmates and use this to estimate classroom peer effects.

In the third chapter, I use reduced form models to test for the existence of obesity spillovers in elementary school classrooms. I find evidence of significant causal social effects in both BMI and exposure to overweight and obese students.

The Makings of a Peer:

Evidence on Within-Classroom Heterogeneities in Peer Effects

By

Jonathan Presler
B.A., Washington University in Saint Louis, 2009

Dissertation
Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Economics

Syracuse University
June 2020

© Copyright by Jonathan Presler, 2020

All Rights Reserved

Acknowledgements

This dissertation would not have been possible without the help and support of my advisor, Amy Ellen Schwartz. Her guidance, experience, and feedback are greatly appreciated and have been instrumental in shaping me during this process. I am also indebted to William C. Horrace for his willingness to invest in me, provide feedback, and talk through identification issues and research ideas. I also thank Alfonso Flores-Lagunes for always being available to provide feedback, thoughtful questions, and advice. I am grateful to the other members of my dissertation defense committee: Maria Zhu, Hyunseok Jung, and chair Michah Rothbart. I would also like to thank the faculty and students who participate in the Education Social Policy Workshops. I learned much from you all during my times as both presenter and participant.

In addition to support from members of the Syracuse University community, I would like to express my appreciation for my close friends and family who have encouraged me and helped to keep me sane throughout this process. I thank my parents for their support and encouragement throughout my life. My biggest debt is to my wife Christy, for her love, sacrifice, and support that has made this whole journey possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One

What Makes a Classmate a Peer?

Examining which peers matter in NYC elementary schools

William C. Horrace

Syracuse University

Hyunseok Jung

University of Arkansas

Jonathan Presler

Syracuse University

Amy Ellen Schwartz

Syracuse University

1.1 Abstract

A growing literature explores peer effects in educational outcomes (e.g., standardized test scores), but constructing student peer networks through friendship surveys or similar means may be prohibitively costly. A reasonable alternative to surveys is using shared student characteristics as a proxy for the network. In addition to being relatively inexpensive to implement, peer networks constructed from demographic data like ethnicity or gender are often strictly exogenous and constant over time, two traits that greatly simplify peer effect estimation. While researchers believe certain demographic characteristics are important for understanding peer effects, there is little evidence on the relative importance of these factors. This paper provides empirical evidence on the relative importance of peer effects by gender, ethnicity, neighborhood, bus stop, bus route, language, and country of birth for New York City elementary school classrooms. Conditional on being in the same classroom, we find that the most important student peer effects are shared bus, gender, and country of birth. In doing so, we consider identification of peer effects within the classroom setting, and tackle issues related to the 'big data' associated with large urban school systems.

Acknowledgements

We gratefully acknowledge funding from the Institute of Education Sciences (award number: R305A170270). We would like to thank the NYC Department of Education for providing data and for their support, especially the Office of Pupil Transportation, Alexandra Robinson, and Tim Calabrese. For data support, advice, and suggestions, we also thank Meryle Weinstein, Sarah Cordes, and Joanna Bailey. We thank seminar participants at the the Maxwell School at Syracuse University, the AEFP conference, and Daniel Patrick Moynihan Summer Workshop in Education and Social Policy for useful comments on previous drafts. The opinions expressed are those of the authors and do not represent views of the NYC Department of Education.

1.2 Introduction

Peer effects are thought to be an important factor in the education production function, but peer effects remain poorly understood. The promise of peer effects is twofold: a near costless improvement to educational outcomes through optimal assignment of students to classrooms and a "social multiplier effect" for interventions and investments.[1] Despite their importance, econometric estimation of peer effects in education remains difficult for a number of reasons addressed below. Previous work has produced varied results, and there is little consensus on the magnitude of peer effects or which estimates are reasonable.[2] The primary challenge in any peer effect analysis is that the 'true' network is almost never observed, and can only be approximated.[3] Empirical specification of networks involves determining the shape of the network (who is connected to whom?) and the strength of the network (the magnitude of the individual connections within the network).[4] Neither of these empirical choices is trivial, which may explain the varied results in the literature. However, one reasonable approach to determining the shape and strength of any network is to use the concept of homophily. That is, students with similar demographic characteristics in a classroom are connected, and the strengths of connections are based on the fraction of characteristics that any two students share.[5] Peer networks constructed from demographic data, like ethnicity and gender, are strictly exogenous and likely to be constant over time, two features that greatly simplify peer effect estimation. Moreover, school administrators are usually in possession of demographic data, avoiding the need to conduct student friendship surveys like

1 Bennett and Bergman (2018) provide a nice example of social multipliers in action using an attendance intervention. A good survey is Epple and Romano (2001).

2 The survey of the empirical effects literature in B. Sacerdote (2001) highlights this variety.

3 An exception may be production networks where worker interactions may be observed. See W. C. Horrace, Liu, and Patacchini (2016) for example.

4 The direction of the network is also important, but here if A affects B, then B also affects A with equal strength.

5 De Giorgi, Pellizzari, and Redaelli (2010) estimate peer effects with random assignment of students to different sections of nine college courses. Students are connected if they share at least one section, and the strength of the connection is the fraction of shared sections.

the “AddHealth” data (Patacchini, Rainone, and Zenou, 2017), or to mine social media data like Facebook to understand how students sort into friendships (Mayer and Puller, 2008).[6]

If understanding peer effects based on demographic networks is important, then this paper attempts to answer the question, ‘which demographic characteristics produce the largest effects?’ That is, if there are K demographic characteristics used to define strength in a student network, which of these characteristics are most important? One way to do this would be to estimate K peer effect models, each based on a single network formed from a different demographic characteristic (e.g., one model for ethnic peers, one model for gender peers, etc.). Ranking these estimated effects and applying a suitable multivariate inference procedure may give a sense of which effects are most important to the given outcome variable (e.g., test scores). However, this approach is akin to data mining. An alternative approach is to specify a single model with K networks each representing a different characteristic and each producing a unique peer effect estimate. With such a model, multivariate inference techniques or likelihood ratio tests may be used to determine the most important (largest) effects at a pre-specified error rate. Detailing this empirical approach in an education production function is a contribution of this paper.

Although no longer a data mining exercise, this approach is not without its complications. With many networks, identification of the model may be difficult, so another contribution of this paper is a discussion of identification in the context of multiple exogenous networks in a classroom setting. Since the difficulties of the linear-in-means model were first discussed in Manski (1993), much has been done to understand identification of peer effects from exogenous networks. Essentially, peer effects estimators from exogenous networks are identified (under certain regularity conditions), if polynomial functions of the the matrix that defines the network are linearly independent (Bramoullé, Djebbari, and Fortin, 2009). This is a very general result. Alternatively, peer effects may be identified

6 In both these instances networks are arguably endogenous. See Goldsmith-Pinkham and Imbens (2013), Kelejian and Piras (2014), Qu and L.-f. Lee (2015), Hsieh and L. F. Lee (2016), **2016** and Patacchini, Rainone, and Zenou (2017) for remedies. Endogenous networks are beyond the scope of this paper.

if either there is sufficient variability in peer group sizes (L.-F. Lee, 2007) or if peer groups are partially overlapping (De Giorgi, Pellizzari, and Redaelli, 2010).[7] We discuss identification issues in the classroom setting for the case when networks are defined on exogenous demographic characteristics. All our discussions are for the case of 'exclusive averaging' where individual students are not their own peers.

The question of who matters in the network is still open, and previous work spans a spectrum of network approximations. College roommate studies, such as B. Sacerdote (2001), examine small networks of two student peers, who are almost surely part of a larger (unobserved) network. Because the observed network is small relative to the larger network, Sacerdote argues that his estimates are a lower bound of the 'true' peer effects. Other work examines larger networks, such as the school-grade (Hoxby, 2000 and many others) or squadrons in the US Air Force Academy (Carrell, Fullerton, and West, 2009 and Carrell, B. I. Sacerdote, and West, 2013). However, these approximations likely include students peers that are not part of the 'true' peers of any individual under study. We can think of this as an issue of connection density - networks with low connection density may overstate the true network and networks with high connection density may understate it. Thus, it is imperative that we understand the potential determinants of connections within any network in order to better understand which peers matter (network shape). Moreover, it is likely that even within the set of individuals who do matter, some will matter more than others (network strength).

It is important to distinguish our study of elementary school peers from papers like B. Sacerdote (2001) and Carrell, B. I. Sacerdote, and West (2013), which study college students. Our demographic characteristic approach is more likely to be useful in educational settings where students have less autonomy in their peer choices outside the classroom. In fact, our identification discussion assumes that the primary grouping mechanism is the elementary school classroom, and that students in different classrooms are not peers by assumption. This assumption would clearly be violated

7 This is an oversimplification of identification, but we discuss more details in the sequel.

for students whose peer networks extend beyond the classroom, in which case our networks would understate the scope of true network. That said, we are also potentially overstating the classroom network with our homophily assumption. However, *any* presumed network structure suffers from this ambiguity.

This paper examines the effects of classmates on student achievement and may provide evidence regarding how peer groups form in New York City elementary schools. We approximate the peer network within the classroom and evaluate which groups of peers are most important for student achievement. The paper borrows from the spatial econometrics literature (e.g. L.-F. Lee, 2007) to introduce a model specification that incorporates multiple peer effects. We have student-level data on shared classroom, bus route, bus stop, residential census tract, and exogenous demographic characteristics of the students (e.g., gender, race, birth country, etc.), which may induce peer effects through homophily. We assume that academic peer effects occur *only* in the classroom, but within the classroom each student has multiple sub-reference groups determined by the aforementioned data, and he/she interacts with peers within the groups simultaneously but with different intensities.[8] In addition to adding sparsity to the school-wide network matrices, this assumption allows us to consider identification from within classroom variation rather than from variation at the school level, avoiding 'big data' issues associated with samples of students from the large urban school districts.

Then, instead of estimating an overall (aggregate) peer effect for a single network with varying connection strengths, we disaggregate the total peer effect into components based on multiple networks with similar connection strengths. Even though this is not a decomposition *per se*, the variability in multiple peer effect estimates provides insight into network formation, as we determine along which demographic strata students may form their peer groups and the relative

8 In our context, it may be socially beneficial for students to evaluate or compare their performances to the subset of students with same gender or census tract. The heterogeneity in classroom composition produces variability in network structure.

importance of these groups. This may be interpreted as a network approximation: the set of reference groups may approximate the unknown true peer network if the reference group set is carefully chosen.[9] This decomposition (or disaggregation) of the peer effects allows us to relax the implicit assumption found in De Giorgi, Pellizzari, and Redaelli (2010) that networks (course sections) have equal weight. There are many empirical settings in which it is not reasonable to assume that the strata along which individuals sort themselves are of equal importance. In our context, it is theoretically ambiguous whether being the same gender in the classroom is more important than being on the same bus in peer formation, but it is unlikely that they are of equal importance.

Aside from not requiring survey data or experiments, our approximation has two advantages over existing approximations. First, if students are randomly assigned to classrooms, and we can control for sub-reference group-specific effects, then we don't need to worry about network endogeneity problems. We are using a set of exogenous proxies for the network, which helps to circumvent the endogeneity problem. In this regard, our approach can be seen as an instrumental variable estimator.[10] Second, our approach can solve the reflection problem of Manski (1993). The reflection problem is the identification issue that arises in the linear-in-mean model due to lower variability in network linkages. However, in our model, the multiple and overlapping reference group structure produces variability in the network to help identification of peer effects.

The contribution of this paper is clear. It allows us to identify peer effects without full knowledge of network structure, while avoiding the problem of network endogeneity and the reflection problem. We apply this model to student-level elementary school data from the New York City Department of Education (NYCDOE) to approximate the peer network and estimate the effect of peers on student achievement in reading and mathematics. We find the strongest networks are shared gender, country of birth, and bus route, and these are of similar relative importance across mathematics and reading scores. Of lesser importance are shared bus stop, ethnicity, and language spoken

9 This network approximation is conceptually similar to spline or polynomial approximation methods. Alternatively, it also can be interpreted as a variant of the synthetic control approach.

10 Kelejian and Prucha (2004) consider instrumental variable estimation of spatial models with multiple networks

at home. Peer effects appear to be stronger in reading than mathematics scores, but the relative ranking of the effects are similar.

The next section discusses the data, the model, homophily and identification of educational peer effects. Section 3 presents the main estimation results and performs multivariate inference to determine the largest demographic peer effects, using the theory of ranking and selection. Section 4 contains some robustness checks. Section 5 concludes and discusses areas for future research.

1.3 Data and Empirical Model

1.3.1 Data

Data come from the New York City Department of Education (NYCDOE) administrative database. The data are student level, and include demographic characteristics such as gender, race, age, grade, neighborhood in which the student lives, country of birth, language spoken at home, and a poverty indicator. We also have information on which school students attend and the method of transportation they use to get to school (bus, metro, other). Information regarding buses is detailed and includes bus stop location, time of pickup, and bus route number. We also have some academic information, such as assigned homeroom class, attendance rate, and mathematics and reading test scores (grades 3-5 only). Current-year outcomes are from years 2013-2015, although we have lagged outcomes from 2012 as well.

Following are details regarding how the sample is constructed. Our data encompasses the universe of students in the NYC public schools for the period 2012-2015 (2012 is used to construct lag variables). We limit our sample to K-5 students. Focusing on elementary students has several benefits for estimating classroom peer effects. The first is data driven: we have information on each student's homeroom classroom. In middle school and beyond, students tend to switch classrooms as they move between subjects, but for elementary students the homeroom class is their main physical location. Second, younger students are more likely to use buses to get to school. This is by design, as younger students are not expected to walk as far as older students. Students below grade three are eligible for subsidized (or free) transportation, if they live more than a half mile from school. This moves up to one mile for students below grade six,[11] and the distance older students are expected to travel without assistance continues to grow with age. Thus the bus as a network is most prominent among elementary school students. Finally, elementary school students do not choose their classes, and so while the choice of which students are placed in a class may or may not be random (decided by the principal), it is orthogonal to the student's choice of peer group.

Some elementary schools in NYC serve students through grade six, but we do not include these students. Students typically begin switching classrooms in middle school, and this gives students some choice over their classes (thereby giving them some agency over their peer group) and causes our homeroom class variable to be unreliable. For these reasons, we do not include all sixth graders in our sample. The group of sixth graders that remains ends up being a small and selected group of students which is observably different from the rest of the sample. To avoid these issues, we restrict our analysis to K-5 students. We also have concerns that there may be structural differences in schools that serve more than our target grades (such as K-12 or K-8 schools), so we omit schools serving students above grade five.[12]

11 Transportation takes the form of a yellow school bus or a metro card with the full fare for transport to and from school. Schools may opt to give students bus access or metro passes - this choice is left to the discretion of school principles, but the decision is school-wide (within a school we do not see some students given a bus and others not given that option). Students who do not receive full transportation support may receive a half-fare metro card, depending upon their distance from the school and student age.

12 For example, there may be older students who ride the bus with the students in our sample. While our approach ignores these out-of-class peer effects, having students on the bus whom we do not observe dilutes the chance for students who ride the same bus to be in the same class. This can be systematically different for portions of the sample, and if the group of students choosing to attend K-8 schools is different from those attending K-5 schools, we introduce bias into our peer effect estimates by making the weighing matrix a function of school type (and the differences in student characteristics that comes with that). We avoid this issue by ignoring schools serving higher than grade five.

We limit our analysis to the general education population, removing special education students from our sample. Because we are concerned that some classrooms which appear small in our sample may be Integrated Co-Teaching (ICT) classrooms, we remove students in classrooms of less than 20 students. Thus, 5.8% of students are removed from our sample because they may be in ICT classrooms.

We construct networks that may be important for student's choice of peers and estimate the (relative) importance of these networks. We include several networks related to students who ride the bus, but only 9.7% of elementary students ride the bus and about 34.8% of school-years offer the bus to any students.[13] In order to estimate the effects of these networks we further limit our sample to classrooms in which at least two students ride a bus (this may or may not be the same bus).

Demographic information includes the language spoken at home and the country of birth. Most students speak English at home (58.9%), and were born in the US (89%). We construct a top 10 list of languages spoken at home (besides English) and a top 10 list of countries of birth (including the US) and utilize this subset of languages in constructing our networks. In particular, students in 'other' categories (not in the top 10 categories) are NOT connected in these networks.

1.3.2 Homophily

This paper constructs peer networks based on the idea that students are drawn to other students with similar observable characteristics. This concept is called homophily. Ideally we need data with many such characteristics. A number of papers have studied homophily, but typically only a few characteristics are considered. Most frequently used are gender and race. For example, Mayer and Puller (2008) construct a structural model for network formation and use Facebook

13 We do not directly observe which schools offer the bus, but this is inferred. We assume that if a school has at least 5 students and at least 1% of students ride the bus, then the school offers the bus. We cannot use a strict cutoff when we observe a single student ride the bus because some students are offered exceptions (ex: for medical reasons).

data from universities when Facebook was limited to college students. They show that race and gender are associated with an increased probability that two students are Facebook friends. Other work has looked primarily within a demographic sub-group to see whether their connections tend to be stronger. For example, Xu and Fan (2018) provide evidence for homophily by ethnicity and immigration status. We wish to include networks that have empirical justification, such as gender and race. However, since our goal is to identify characteristics associated with the largest peer effect, we include a number of other networks which are easily observed by students. Examples might include neighborhood of residence, bus route, or native language.

We couch our discussion in terms of a single grade consisting of n students, although it is easy to generalize this to a school or a school system. Each student i is endowed with K exogenous characteristics indexed by $k = 1,...,K$. For each characteristic k we specify a symmetric $n \times n$ adjacency matrix, G_k, with typical element g_{ijk}, $i, j = 1,...,n$, where $g_{ijk} = 1$ if students $i \neq j$ share characteristic k, $g_{ijk} = 0$ otherwise. The students in the grade are randomly partitioned into classrooms $c = 1,...,C$. The number of students in each classroom c is n_c, so that $n_1 + n_2 + n_C = n$. We assume that peer effects only occur within a classroom, and that students do not interact with students in other classrooms in any appreciable way in the generation of academic achievement.[14] Therefore, we specify a symmetric $n \times n$ classroom adjacency matrix, S, with typical element s_{ij}, $i, j = 1,...,n$, where $s_{ij} = 1$ if students i and j are in the same classroom, $s_{ij} = 0$ otherwise. Hence, S is block diagonal with blocks $\iota_{n_{ct}} \iota'_{n_{ct}}$, where $\iota_{n_{ct}}$ is an $n_{ct} \times 1$ vector of ones. It is fine that the diagonal of S is populated with ones, because we ultimately interact the two adjacency matrices, so our final network matrix for characteristic k (before row normalization) is $W_k^* = G_k \circ S$, where $\circ$ is the Hadamard product for element-by-element multiplication.[15] In the

14 Obviously there are times in the school day when students interact with students in other classrooms (e.g., gym or lunch periods). However, we assume that these interactions do not create academic spillovers. Therefore, if the outcome variable is not directly related to classroom interactions (e.g., absenteeism), our model may not be applicable for estimating peer effects.

15 This approach may not be limited to school achievement. We can also imagine a similar exercise examining

sequel W_k is the row-normalized version of W_k^*. That is, each row of W_k sums to 1, so that WY is a vector of peer group averages.

As an example, consider a grade of eight students $\{A,B,C,D,E,F,G,H\}$, partitioned into two classrooms: $\{A,B,C,D\}$ and $\{E,F,G,H\}$. Then:

$$S = \begin{array}{c c} \begin{matrix} A & B & C & D & E & F & G & H \end{matrix} & \\ \begin{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{bmatrix} & \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \\ G \\ H \end{matrix} \end{array}$$

Now partition students into a female group, $\{ADFG\}$, and a male group, $\{BCEH\}$. Then

$$G_{gender} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$W_{gender}^* = G_{gender} \circ S = \begin{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{bmatrix}$$

A few comments are in order. First, the characteristic matrices, W_k^*, may not always partition the students (as *gender* does in the example above). For example, not all students ride the bus, so defining peers as bus-mates will ignore those who are not bus riders. Second, n has the potential to be very large. In our example, it is more than 50,000 students. As identification may hinge on the linear independence of polynomial functions of W_k^*, empirically checking for identification of

productivity within a firm where workers are partitioned into teams to produce output, and there are productivity peer effects within teams, but not across teams.

any model that incorporates this matrix can be numerically costly. Fortunately, the block diagonal nature of the matrix simplifies these calculations. This is an added benefit of our assumption that academic peer effects only exist within the classroom.

Given these K characteristic network matrices, a reasonable approach is to sum them to a single, aggregate network, $W^* = W_1^* + ... + W_K^*$.[16] Let the row-normalized, aggregate network be W, and a typical Spatial Auto-Regressive (SAR) social network model for the $n \times 1$ outcome vector Y and $n \times k$ matrix of exogenous variables X is:

$$Y = \alpha \iota_n + \lambda WY + X\beta + \theta WX + U, \tag{1.1}$$

where λ is the (endogenous) *peer effect*, θ is the (exogenous) *contextual effect*, and the network, W, assumes that each characteristic, k, has equal weight in the peer network. Our goal here is to disaggregate the network into its individual characteristic components, W_k:

$$Y = \alpha \iota_n + \sum_{k=1}^{K} \lambda_k W_k Y + X\beta + \sum_{k=1}^{K} \theta_k W_k X + U, \tag{1.2}$$

where W_k is the row-normalized version of W_k^*, the λ_k are peer effects for each characteristic, and the θ_k are contextual effects. We discuss identification of these models in the next section.

1.3.3 Identification

Identification in network models is a complicated issue that is a function of many factors related to selection of peers into groups, unobserved group effects, and the fact that Y appears on the right-hand side of equation above. Fortunately, 'nonidentification of parameters in the reduced form is rare' (Blume et al., 2011, p.893). We begin with a discussion of identification of the parameters in the aggregate network model of equation (2.1), and then discuss the disaggregate network model in equation (2.2). Following Bramoullé, Djebbari, and Fortin (2009), we always

16 This is the approach of De Giorgi, Pellizzari, and Redaelli (2010) but for k indexing different sections of nine courses to which college students are randomly assigned.

assume that the outer product of $(\iota_n X)$ is full rank and $E(U|X,W) = 0$.[17] We also require the regularity conditions that $|\lambda| \cdot ||W_k|| < 1$ in equation (2.1) and $\sum_k |\lambda_k| \cdot ||W_k|| < 1$ in equation (2.2).

Given that our networks are exogenous by design, there are several ways to determine that the structural parameters in equation (2.1) are identified. Bramoullé, Djebbari, and Fortin (2009) show that for $\lambda\beta + \theta \neq 0$, a sufficient condition for identification is that the matrices I_n, W and W^2 are linearly independent.[18] A way to check linear independence is to vectorize I_n, W and W^2 into three $n^2 \times 1$ vectors, horizontally concatenate these into an $n^2 \times 3$ matrix, and check that the rank of the resulting matrix is equal to 3. If n is large, checking this rank condition could be difficult. However, in our case it is simplified by the fact that W is block diagonal.[19] Therefore, we only have to check the rank of the $n^2 \times 3$ matrix with the $\sum_c n_c^2$ rows of zeros (associated with the zero blocks of W) removed. Alternatively, we can check if the condition holds for each diagonal block in W. If at least one block (classroom) satisfies the sufficient condition, then the model is identified. If there are few blocks (classrooms) that satisfy the condition, then the model is weakly identified.[20] Note that the Bramoullé, Djebbari, and Fortin (2009) sufficient condition says nothing specific about the structure of the network; in this sense the result is quite general. However, there are two commonly employed sufficient conditions that arise directly from specific structure of the network.

First, L.-F. Lee (2007) shows that a sufficient condition for identification of the structural parameters in equation (2.1) is that there is *any* variability in group sizes in the network, where a group is a subset of connected peers in W with identical characteristics, $k = 1,..,K$. As an extreme example,

17 The moment condition implies no correlated effects. Per Manski (1993), correlated effects are when 'individuals in the same group tend to behave similarly because they... face similar institutional environments.' The usual solution to the correlated effects problem is to include network- or group-level fixed effects in the specification, which we do in the sequel.

18 If there are group fixed effects in the model we also require that W^3 be linearly independent.

19 Another benefit of assuming that only in class peers matter.

20 Weak identification often manifests itself as slow convergence of numerical estimators, large standard errors and highly unstable results.

if there is only one class ($C = 1$) and all students are identical, then the network is 'complete,' and there is only one group of size n, so the model is not identified. Otherwise, in our context W is block diagonal, and the network W is not complete. Since there is likely variability in classroom sizes, n_c, it would take fortuitous circumstances for the model in equation (2.1) to be not identified by group variability. Obviously, if there is little variability in group sizes the model may only be weakly identified. Second, De Giorgi, Pellizzari, and Redaelli (2010) show that a sufficient condition for identification is that groups are partially overlapping. The block diagonal structure of W actually hinders this form of identification, but we can check if this condition is satisfied within the diagonal blocks. The model is identified as long as the students are not all identical or if student groups do not completely partition the characteristics types. An example of the latter condition would be if the only two characteristics were race and gender, and all classes consisted of black females and white males. We can think of partially overlapping groups creating identification through exclusion restriction of peers at the student level. Given these sufficient conditions and the structure of W, for reasonably large C and K it is likely that the model in equation (2.1) will be identified.

Is the disaggregate model in equation (2.2) identified? Since the W_k are block diagonal, it suffices to check the Bramoullé, Djebbari, and Fortin (2009) sufficient condition in each block (classroom) for each characteristic k. Therefore, if there is any block (classroom) that satisfies the sufficient condition for every characteristic k the model in equation (2.2) is identified. If there are few blocks that satisfy the condition for every k, then the model is weakly identified. Barring this, if there is any variability in class and group size in each characteristics network, W_k, then the sufficient condition of L.-F. Lee (2007) holds. Even if the fraction of student groups within demographic networks is fixed across classes (which they are not in our application), then any variability in class sizes would produce variability in group sizes and identification of equation (2.2). Notice that none of the W_k will individually satisfy the De Giorgi, Pellizzari, and Redaelli (2010) sufficient condition. For example, if all students are either boys or girls (the usual case), then these two groups would not partially overlap. However, across the different W_k the condition is satisfied at the student level, as long as it is satisfied for the matrix W in equation (2.1). Therefore, as long as K is not too large and C is reasonably large, the model in equation (2.2) is likely identified. Ultimately we have a panel of data, so our weighting matrices will vary over time, reinforcing identification. Finally, if K is very large relative to C, then it may be difficult to satisfy the regularity condition $\sum_k |\lambda_k| \cdot ||W_k|| < 1$. If this sum is close to 1 in practice, numerical estimation will be difficult as the solution will be near the border of the parameter space of the λ_k's.

1.3.4 Empirical model

There are a few differences between equation (2.2) and the model that we ultimately estimate. First, we have a panel of data, so our model can accommodate time-varying networks. This should assist with identification of the model's time-invariant parameters. This also allows us to add a one period lag of the dependent variable (Y_{t-1}) on the right-hand side of the model to control for student heterogeneity in test scores. To help control for the correlated effects of Manski (1993) we include classroom dummy variables. Therefore, our empirical model is:

$$Y_{ct} = \sum_{k=1}^{K} \lambda_k W_{k,ct} Y_{ct} + \sum_{k=1}^{K} \theta_k W_{k,ct} Z_{ct} + X_{ct}\beta + Y_{c,t-1}\gamma + \delta_{ct} \cdot \iota_{n_{ct}} + U_{ct} \tag{1.3}$$

where δ_{ct} is a classroom (c) fixed effect at time (t). We include seven weighting matrices (W_k, $k = 1,...,7$): shared *Gender*, shared *Bus Route*, shared *Country* of Birth, shared *Bus Stop*, shared *Ethnicity*, shared *Language* Spoken at Home, and shared Residential *Census Tract*. In the covariate vector, X, we include: a gender indicator (*Female*), an indicator for 'free or reduced price lunch' (*Poverty*), *Age* in months, Age^2, an indicator for a bus rider (*Bus*), and three indicator variables for race (*Asian*, *Black* and *White* with Hispanic excluded). For the contextual effects (Z in equation 2.3) we exclude exogenous covariates that are used to construct the weighting matrices (W_k). That is, Z contains a subset of the covariates, including *FRPL*, *Age*, Age^2, and Y_{t-1}. Since there are seven demographic weighting matrices and four covariates in Z, there are twenty eight contextual effect (θ_k) in our model. Therefore, we do not report contextual effects in our results to save space. We estimate the model using quasi-MLE (L.-F. Lee, 2007). Estimation details are in Appendix A.

We also apply the theory of 'ranking and selection' to determine a subset of largest (best) $\lambda_1, ..., \lambda_K$ at a pre-specified error rate, $\alpha \in (0, 0.5)$. The procedures account for the inherent multiplicity and uncertainty in the ranked $\hat{\lambda}_k$'s. See W. C. Horrace and Parmeter (2017), who recently apply ranking and selection to economics journal citation counts to determine a subset of the 'best' journals. Let the ranked population peer effects be $\lambda_{[K]} \geq \lambda_{[K-1]} \geq ... \geq \lambda_{[1]}$. Ranking and selection procedures recognize the uncertainly in the ranked estimates, $\hat{\lambda}_{(K)} \geq \hat{\lambda}_{(K-1)} \geq ... \geq \hat{\lambda}_{(1)}$, so the $(j)^{th}$ demographic characteristic in the sample may not correspond to the $[j]^{th}$ demographic characteristic in the population, in general. In particular, the $(K)^{th}$ demographic characteristic in the sample may not correspond to the largest demographic characteristic in the population, $[K]$. Let $\kappa = \{1, 2, ..., K\}$ be the set of population indices for our K demographic characteristics. Then assuming (asymptotic) normality of the estimates $\hat{\lambda}_1, ..., \hat{\lambda}_K$ with general variance covariance structure, the procedures identify the 'subset of the best' (largest) peer effect indices, $\zeta \subset \kappa$, such that:

$$Pr\{[K] \in \zeta\} \geq 1 - \alpha, \tag{1.4}$$

where $\alpha = 0.05$, typically. In other words, the population indices in ζ cannot be statistically distinguished from the largest unknown population index, $[K]$, with probability at least $1 - \alpha$. If the inference is very sharp, then ζ may a singleton. If the inference is very weak, then ζ may equal κ. The cardinality of ζ is increasing in K. That is, as the inference needs to make more pairwise comparisons, it is hard to distinguish $[K]$ at a fixed error rate, α. This is the concept of 'multiplicity.' To conduct the inference we need critical values drawn fro a k-dimensional multivariate normal distribution with covariance structure determined from the Hessian of the quasi-MLE estimation procedure. These critical values were all around 2.5, larger that the usual critical of 1.96 from a univariate normal and, hence, accounting for the multiplicity.

1.4 Main Results

We begin with our results for mathematics and reading scores for fourth and fifth graders in Table 1.1. To save space we do not report the contextual effect results (θ_k, $k = 1,...,28$), only the peer effects (λ_k, $k = 1,...,7$) and the marginal effects (β) of the covariates (X) in equation 2.3. The test scores are normalized Z-scores. The top panel of the table contains the estimated peer effects (λ_k) for each demographic network (W_k) in rank order for the mathematics test scores. For example, for 'Mathematics Test Scores' the *Gender* peer effect is a significant 0.051. That is, as the average mathematics score of a student's gender peers increase by one unit, the student's own test score is improved by approximately 0.051 units. Because this is a multiplier effect, it is more precise to note that own test scores are multiplied by $\frac{1}{1-\lambda_k}$,or an increase of .0527 units. With small values of λ, the approximation is reasonable for quick reading of the results.

All the mathematics score peer effects are significant except for *Census Tract*, with *Gender* (0.051), *Bus Route* (0.049), *Country* of birth (0.043), *Bus Stop* (0.041), *Ethnicity* (0.041), and *Language* (0.022) in rank order.[21] The ranking and selection procedure determined that the subset of the largest peer effects contains all demographic characteristics except Residential *Census Tract* at the 95% level. In other words, the inference is not very sharp, and all the peer effects are equally large (in a statistical sense).

Table 1.1 also contains the results for Reading Test Z-Scores. Even though the sample is identical to that for the mathematics scores, there are some interesting differences in the demographic effects of peers across the two sets of results. First, the mathematics peer effects are always larger than the reading peer effects except for *Ethnicity* (compare 0.041 for mathematics to 0.030 for reading). Second, the peer effects rankings are about the same for reading and mathematics scores. The only differences are that *Bus Route* and *Country* of birth are switched in the reading score

21 It is important to note that we have not included fixed effects for specific bus stops and bus routes in this draft due to computational constraints, although we have included an indicator for whether students ride the bus. As a result, these estimates may be biased by correlated effects related to the bus.

ranking, as are *Ethnicity* and *Language* spoken at home. Finally, the subsets of the best (largest) peer effects, based on ranking and selection inference, are different for mathematics and reading. While the mathematics score subset contained all demographic characteristics other that *Census Tract* at the 95% level, the reading score subset only includes *Gender* (0.089), *Bus Route* (0.079), *Country* of birth (0.083) and *Bus Stop* (0.053). In particular, *Ethnicity* (0.030), and *Language* spoken at home (0.032) are not in contention for the largest demographic peer effect at the 95% level. That is, while *Ethnicity* and *Language* spoken at home are individually statistically significant, they are not as important as the other sources of homophily. Hence, if an aggregate weighting scheme is desired for mathematics test scores, then perhaps these demographic variables should receive less weight than the others.

It is interesting that bus route and bus stop peers have a relatively strong effect on academic performance in NYC elementary schools. In fact these peers are as important as gender, country of birth and ethnicity peers. One explanation for this is that sharing a bus with classmates increases student face time, strengthening peer bonds. Another explanation is anecdotal: parents often accompany their elementary school children at bus stops. Therefore, the peer connections may be happening through the parents. Perhaps parents share information at the bus stop that help their children perform better in school. It could also be the case that elementary students do homework together on the bus before arriving to school. This could also occur on the way home from school. This is all anecdotal, but it does provide food for thought for policy makers.

As an experiment we re-estimate the model seven times, including only *one* of our seven demographic peer networks in each estimation. The goal is to see how the rankings of the peer effects change when we estimate them in isolation. The results are in Table 1.2 . These models include the same contextual effects and marginal effects as in Table 1.1, but we not report them to focus on the peer effects. The peer effects in Table 1.2 appear in the same order as in Table 1.1 for the sake of comparison, and stark differences in the rankings emerge. In Table 1.2 , *Ethnicity* is now the largest peer effect for mathematics test scores, while *Country* of birth is the largest for reading test scores. *Gender* still has a large peer effect, but it has dropped to second in the ranking for both tests. The difference in the rank statistics highlights the value of the proposed model in equation (2.3). It is important to simultaneously control for all networks if one is interested in ranking the relative importance of the demographic peers.

1.5 Robustness Checks

1.5.1 Randomness of classroom assignment

One key assumptions in educational peer effect models is that classroom assignment is random (e.g., De Giorgi, Pellizzari, and Redaelli, 2010). While we cannot prove this random assignment, we provide some evidence to support this claim for these data. For each of the demographic peer networks we would like to show that classroom assignment is not a function of the associated demographic variable. To do this, we consider a series of multinomial logit models of the form:

$$Class_i = \alpha + X_i\beta_{gst} + \varepsilon_i$$

We limit the sample to a single grade (g) in a single school (s) in a single year (t), so there may only be 3 or 4 classrooms for each regression. *Class* is a categorical variable for assignment of student i to one of these classrooms, and X_i is a binary indicator for some demographic characteristic of student i. For example, X may be an indicator of female, white, black, Hispanic, Asian/other or 'bus rider'. We convert demographic indicators with multiple categories (i.e. 'language spoken at home' and 'country of birth') to binary indicators (i.e., 'English speaker' and 'foreign born,' respectively). The result is a series of estimates ($\tilde{\beta}_{gst}$) and t-statistics, capturing the marginal effect and significance of the relevant demographic variable (X) on class assignment. Then, for each school-grade-year we randomly assign the students to the same classrooms and re-run the series of multinomial logit for each X. We then compare the two sets of regression results for each X with a plot of the non-randomized t-statistics on the *ranked* randomized t-statistics. Each dot in the panels of Figures 1.1 and 1.2 represents a pair of t-statistics (one non-randomized, one randomized) for each grade-school-year. Each panel represent a single demographic indicator and contains a line with a slope of one (a 45 degree line), so deviations in the dot patterns from this line represent a difference in the significance of the non-randomized and randomized regressions, and (perhaps) a departure from random assignment to classrooms. Figure 1.1 contains the plots for female, ‘English speaker,’ ‘foreign born’ and ‘bus rider,’ while figure 1.2 contains plots for Hispanic, black, white and Asian. All plots suggest random assignment of students to classrooms. Therefore, we need not be concerned about selection and may treat our demographic peer networks as exogenous.

1.5.2 Other robustness checks

Tables 3-5 modify the specification in Table 1.1 to see if the main peer effect rankings and magnitudes are sensitive to different specifications. The general finding is that the relative rankings of the peer effects tend to be robust, while the magnitudes tend to change. This may suggest that the complete model is the preferred specification if one is only concerned with understanding the relative importance of different demographic peers in elementary school education test scores.

Table 3 contains the model in 2.3, but removes the bus networks. The bus networks tend to be sparser then the other demographic networks, and we want to be sure that this sparseness is not corrupting the other estimated peer effects. Again, for comparison purposes the peer effects appear in the same order as Table 1. For mathematics test scores the ranking of the peer effects are relatively unchanged without the bus networks. Only *Ethnicity* and *Country* of birth are switched in the ranked effects in Table 3. All the peer effects are smaller for the mathematics score compared to Table 1, but the relative rankings are about the same. The same holds for the reading test scores with one clear exception: the *Ethnicity* peer effect is insignificant without the bus networks. Interestingly, now the indicator for riding the bus is positive and significant for both mathematics (0.700) and reading (0.672) test scores. It is negative and insignificant for both scores in the complete model of Table 1. There appears to be a linkage between bus ridership and ethnicity in our results.

Like most large cities, New York is geographically segregated by ethnicity and income, so a geographical exploration of school bus routes and the ethnic enclaves that they serve may shed light on this linkage and how it affects educational performance.

Next, we remove network connections between English speakers and also between US born students in the *Language* spoken at home and the *Country* of birth networks, respectively, the two largest groups in these demographic networks. The idea is to make these two peer networks more sparse and re-estimate the models. The model did not converge for reading test scores, so it is not reported in Table 4. For the mathematics test scores in Table 4, we see the peer effect rankings are fairly robust to this change. The magnitudes of the effects also tend to be larger.

Because gender and ethnicity are the most commonly investigated demographic determinants of educational outcomes in the literature, we run the model with just these two networks for comparison. Table 5 contains the results with only *Gender* and *Ethnicity* peer networks. In these models ethnicity peers are important for mathematics scores (0.037), but not so for reading scores (0.019).

1.6 Conclusions and Future Research

In this paper we disaggregate the peer effect network into demographic strata and discuss identification when peer effects are limited to the classroom. Applying the method to elementary school students in NYC and appealing to ranking and selection methods we determine that with 95% confidence, shared gender, bus route, country of birth, bus stop, ethnicity and language spoken at home are simultaneously important for improving mathematics test scores. However, for reading test scores only shared gender, bus route, country of birth and bus stop are important. The technical implications are clear, some demographic characteristics are more important that others in educational spillovers. In particular, it seems that assigning students, who share the bus, to the same class may benefit those students in term of test scores, *ceteris paribus*. Prior to this study, the importance of bus ridership had never been explored

The implication for empiricists is that certain demographic characteristics should receive more weight than others in construction of aggregate weighting matrices when estimating academic peer effects for school children. Of course the resulting aggregate matrix will be estimated with uncertainty. Precisely how this uncertainty would effect the standard errors in a second stage regression is unclear and is left for future research. There are a couple additional avenues for future research. First, it would be interesting to experiment with different row-normalization schemes for the disaggregate networks. If there are K weighting matrices, perhaps it makes sense to normalize each row to K or $1/K$ (rather than to the usual value of 1). How this would effect the estimated peer effect is not clear and may be worth exploring. Second, there appears to be a linkage between bus ridership and ethnicity in our results that we did not explore. Like most large cities, New York is geographically segregated by ethnicity and income, so a geographical exploration of school bus routes and the ethnic enclaves that they serve may shed light on this linkage and how it affects educational performance. This is currently being explored by one of the authors.

1.7 Tables

Table 1.1 Main Demographic Peer Effect Results

	Mathematics		Reading	
Peer Effects (λ_k):				
Gender	0.051**	(0.011)	0.089**	(0.011)
Bus Route	0.049**	(0.008)	0.079**	(0.009)
Country	0.043**	(0.011)	0.083**	(0.012)
Bus Stop	0.041**	(0.009)	0.053**	(0.010)
Ethnicity	0.041**	(0.007)	0.030**	(0.008)
Language	0.022**	(0.007)	0.032**	(0.008)
Census Tract	0.003	(0.005)	-0.004	(0.005)
Marginal Effects (β):				
Lag Test Score	0.701**	(0.003)	0.674**	(0.004)
Age	0.014	(0.008)	0.040**	(0.009)
Age Squared	-0.00008*	(0.00003)	-0.00016**	(0.00003)
Female	-0.008	(0.005)	0.049**	(0.005)
Poverty	-0.070**	(0.006)	-0.061**	(0.007)
Bus	-0.018	(0.010)	-0.019	(0.012)
Asian	0.136**	(0.010)	0.101**	(0.011)
Black	0.012	(0.010)	-0.007	(0.011)
White	0.040**	(0.009)	0.025*	(0.010)
Observations (n):	53,118		53,493	

Models include classroom fixed effects, 28 contextual effects, fixed for census tract, country of birth, language spoken at home. Hispanic is the omitted reference group in the marginal effects. * Significant at the 5% level and ** at the 1% level. Standard errors in parentheses.

Table 1.2 Peer Effects (λ) from Individual Models

	Mathematics		Reading	
Gender	0.029**	(0.011)	0.043**	(0.012)
Bus Route	0.012	(0.007)	0.010	(0.008)
Country	0.028*	(0.012)	0.046**	(0.013)
Bus Stop	0.019**	(0.007)	0.011	(0.008)
Ethnicity	0.037**	(0.007)	0.019*	(0.008)
Language	0.024**	(0.007)	0.026**	(0.008)
Census Tract	0.003	(0.005)	-0.007	(0.005)
Observations	53,118		53,493	

Models include classroom fixed effects, 28 exogenous effects, lagged test score, age, age squared, and fixed effects for gender, poverty, ethnicity, census tract, country of birth, language spoken at home, and whether student rides the bus. Ethnicity is defined with Hispanic as the reference group. * Significant at the 5% level and ** at the 1% level. Standard errors in parentheses.

Table 1.3 No Bus Networks

	Mathematics		Reading	
Peer Effects (λ_k):				
Gender	0.031**	(0.011)	0.043**	(0.012)
Country	0.025*	(0.011)	0.043**	(0.013)
Ethnicity	0.034**	(0.007)	0.015	(0.008)
Language	0.017*	(0.007)	0.021**	(0.008)
Census Tract	0.004	(0.005)	-0.006	(0.005)
Marginal Effects (β):				
Lag Test Score	0.700**	(0.003)	0.672**	(0.004)
Age	0.013	(0.008)	0.040**	(0.009)
Age Squared	-0.00007*	(0.00003)	-0.00017**	(0.00003)
Female	-0.009*	(0.005)	0.057**	(0.005)
Poverty	-0.071**	(0.006)	-0.064**	(0.007)
Asian	0.140**	(0.010)	0.106**	(0.011)
Black	0.000003	(0.000025)	0.000044	(0.000028)
White	0.010	(0.010)	-0.008	(0.011)
Observations (n):	53,118		53,493	

We run the baseline model without the bus networks, as we may be concerned about sparse networks. Models include classroom fixed effects, 28 contextual effects, fixed for census tract, country of birth, language spoken at home. Hispanic is the omitted reference group in the marginal effects. * Significant at the 5% level and ** at the 1% level. Standard errors in parentheses.

Table 1.4 Alternate Language and Country Networks

	Mathematics	
Peer Effects (λ_k):		
Gender	0.062**	(0.011)
Bus Route	0.052**	(0.008)
Country	0.055**	(0.011)
Bus Stop	0.043**	(0.009)
Ethnicity	0.046**	(0.007)
Language	0.031**	(0.004)
Census Tract	0.005	(0.005)
Marginal Effects (β):		
Lag Test Score	0.700**	(0.003)
Age	0.014	(0.008)
Age Squared	-0.00008*	(0.00003)
Female	-0.007	(0.005)
Poverty	-0.069**	(0.006)
Bus	-0.018	(0.010)
Asian	0.135**	(0.010)
Black	0.011	(0.010)
White	0.040**	(0.009)
Observations (n):	53,118	

Models include classroom fixed effects, 28 contextual effects, fixed for census tract, country of birth, language spoken at home. Hispanic is the omitted reference group in the marginal effects. * Significant at the 5% level and ** at the 1% level. Standard errors in parentheses.

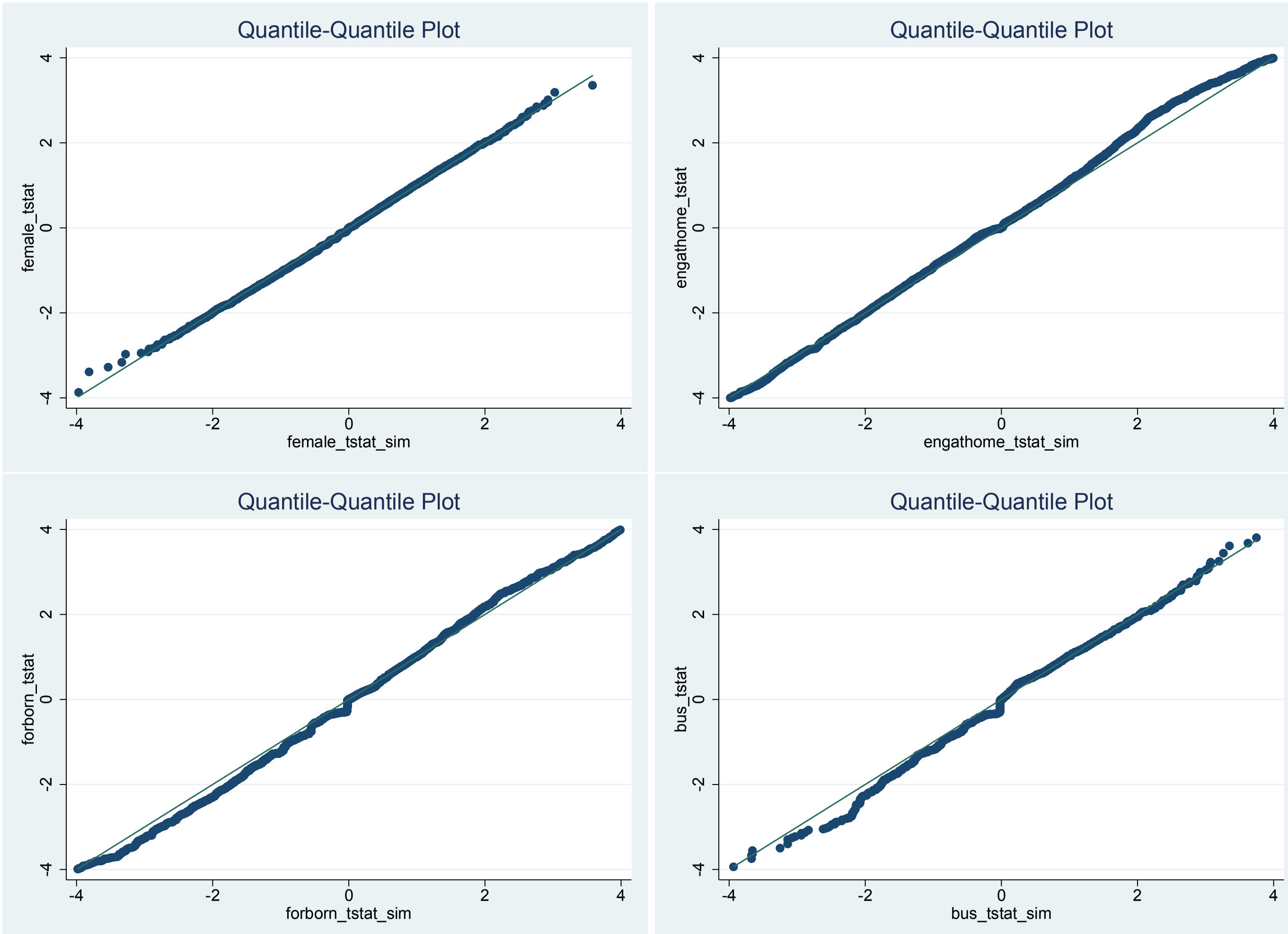
Table 1.5 Only Gender and Ethnicity Peers

	Mathematics		Reading	
Peer Effects (λ_k):				
Gender	0.030**	(0.011)	0.043**	(0.012)
Ethnicity	0.037**	(0.007)	0.019*	(0.008)
Marginal Effects (β):				
Lag Test Score	0.699**	(0.003)	0.671**	(0.003)
Age	0.014	(0.008)	0.040**	(0.009)
Age Squared	-0.00008	(0.00003)	-0.00017	(0.00003)
Female	-0.009*	(0.004)	0.057**	(0.005)
Poverty	-0.069**	(0.006)	-0.062**	(0.007)
Bus	-0.018**	(0.007)	-0.018*	(0.007)
Asian	0.139**	(0.010)	0.105**	(0.011)
Black	0.010	(0.010)	-0.008	(0.011)
White	0.041**	(0.009)	0.026**	(0.010)
Observations (n):	53,118		53,493	

Models include classroom fixed effects, 28 contextual effects, fixed for census tract, country of birth, language spoken at home. Hispanic is the omitted reference group in the marginal effects. * Significant at the 5% level and ** at the 1% level. Standard errors in parentheses.

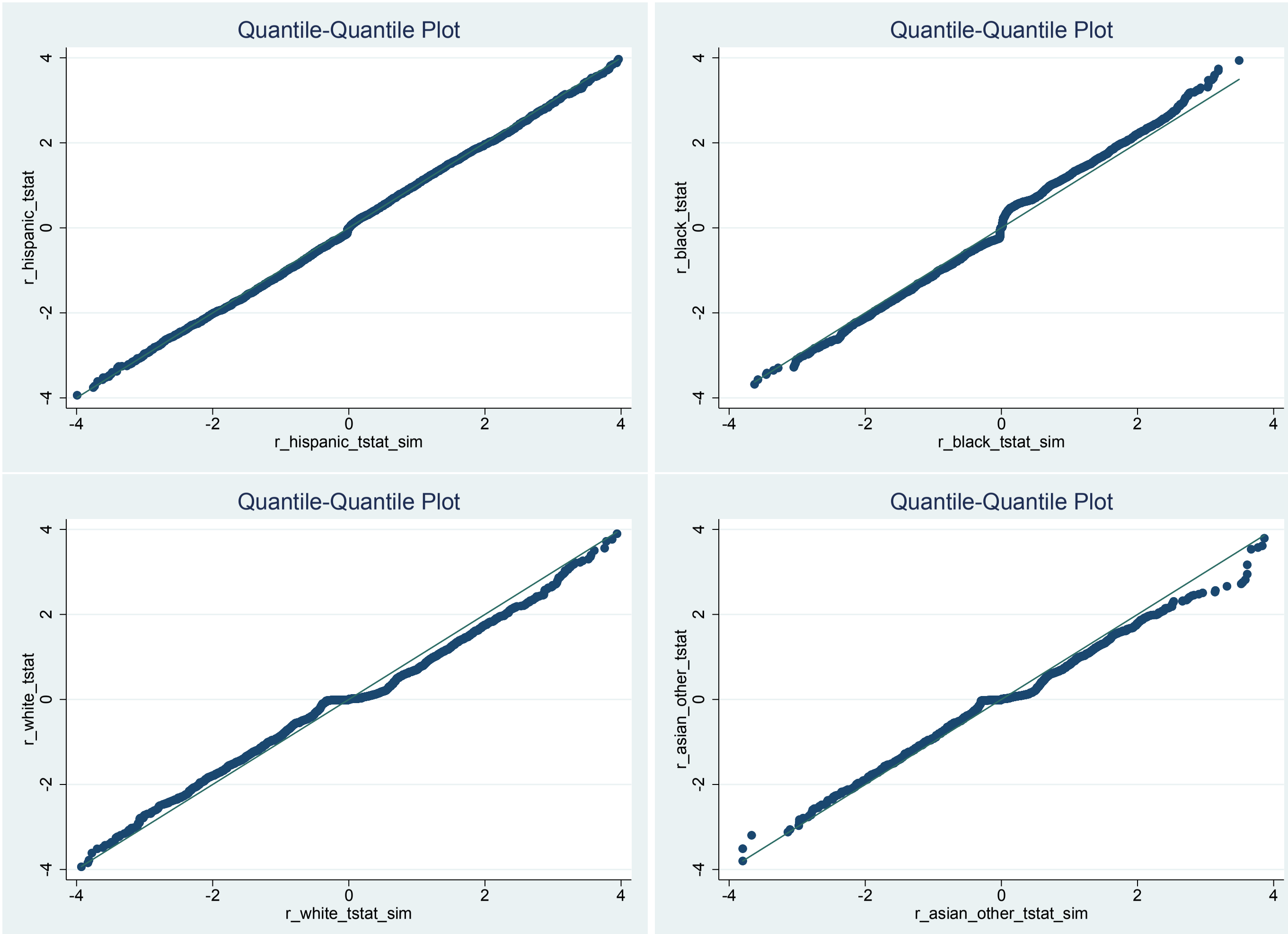
1.8 Figures

Figure 1.1 Quantile-Quantile Plots for Females



A series of multinomial logits are run to estimate the importance of gender, whether english is spoken at home, whether a student is foreign born, or whether the student rides the bus in class assignment. In each pair of graphs, the left plots the t-statistics from these against the t-statisitics from a similar exercise in which we randomly assign students to classrooms. Thus we are plotting these two distributions against one another, and if the distributions are the same, we should expect a straight line of slope one. We argue that these provide evidence that class assignment is consistent with a random process. This test performs best when the networks are large (ex: females are about half the student population) and is more noisy when the network is small (ex: bus, which is less than 10 percent of students).

Figure 1.2 Quantile-Quantile Plots for Ethnicity



A series of multinomial logits are run to estimate the importance of gender, whether english is spoken at home, whether a student is foreign born, or whether the student rides the bus in class assignment. In each pair of graphs, the left plots the t-statistics from these against the t-statisitics from a similar exercise in which we randomly assign students to classrooms. Thus we are plotting these two distributions against one another, and if the distributions are the same, we should expect a straight line of slope one. We argue that these provide evidence that class assignment is consistent with a random process. This test performs best when the networks are large (ex: hispanics) and is more noisy when the network is small (ex: only 16.5 percent of students in the sample are white).

1.9 Appendix: Estimation procedures

The empirical model is given by

$$Y_{ct} = \sum_{k=1}^{K} \lambda_k W_{k,ct} Y_{ct} + \sum_{k=1}^{K} \theta_k W_{k,ct} X_{ct} + X\beta + Y_{c,t-1}\gamma + \delta_{ct} \cdot \iota_{ct} + U_{ct} \tag{1.5}$$

where the term $\sum_{k=1}^{K} \theta_k W_{k,ct} X_{ct}$ is to measure the exogenous social effect, $Y_{c,t-1}$ is the lag of outcome which is included to account for individual effect, δ_{ct} is classroom fixed effect and $\iota_{n_{ct}}$ is a $n_{ct} \times 1$ vector with ones where n_{ct} is the total number of students in classroom c at time t.

In order to apply MLE, we assume that each element in U_{ct} is $iid(0, \sigma^2)$.[22]. To remove the time varying classroom fixed effect, following L.-f. Lee and Yu (2010), we consider an transformation which can eliminate the classroom fixed effects while maintaining interdependency between the disturbances. Let the orthonormal matrix of the with transformation matrix $Q_{ct} = I_{n_{ct}} - \frac{1}{n_{ct}} \iota_{n_{ct}} \iota_{n_{ct}}^{'}$ be $[P_{ct}, \iota_{n_{ct}}/\sqrt{n_{ct}}]$. The columns in P_{ct} are eigenvectors of Q_{ct} corresponding to the eigenvalue one, such that $P_{ct}^{'} \iota_{n_{ct}} = 0$, $P_{ct}^{'} P_{ct} = I_{n_{ct}-1}$ and $P_{ct} P_{ct}^{'} = Q_{ct}$. Then, premultiplying (A.1) by $P_{ct}^{'}$ leads to

$$P_{ct}^{'} Y_{ct} = P_{ct}^{'} \sum_{k=1}^{K} \lambda_k W_{k,ct} Y_{ct} + P_{ct}^{'} \sum_{k=1}^{K} \theta_k W_{k,ct} X_{ct} + P_{ct}^{'} X_{ct} \beta + P_{ct}^{'} Y_{c,t-1}\gamma + P_{ct}^{'} U_{ct} \tag{1.6}$$

Let $\bar{Y}_{ct} = P_{ct}^{'} Y_{ct}$, $\bar{X}_{ct} = P_{ct}^{'} X_{ct}$, $\bar{Y}_{c,t-1} = P_{ct}^{'} Y_{c,t-1}$, $\bar{U}_{ct} = P_{ct}^{'} U_{ct}$, $\bar{W}_{k,ct} = P_{ct}^{'} W_{k,ct} P_{ct}$ for $k = 1, ..., K$. Due to $P_{ct}^{'} W_{k,ct} = \bar{W}_{k,ct} P_{ct}^{'}$, (A.2) can be written as

$$\bar{Y}_{ct} = \sum_{k=1}^{K} \lambda_k \bar{W}_{k,ct} \bar{Y}_{ct} + \sum_{k=1}^{K} \theta_k \bar{W}_{k,ct} \bar{X}_{ct} + \bar{X}_{ct}\beta + \bar{Y}_{c,t-1}\gamma + \bar{U}_{ct} \tag{1.7}$$

Note that $\bar{U}_{ct} \sim (0, \sigma^2 \iota_{n_{ct}-1})$. Then the likelihood function for (A.3) is given by

$$\ln L_{ct}(\Lambda, B, \sigma^2) = -\frac{n_{ct}-1}{2} \ln(2\pi\sigma^2) + \ln|\bar{S}_{ct}(\Lambda)| - \frac{\bar{\varepsilon}_{ct}(\psi)^{'} \bar{\varepsilon}_{ct}(\psi)}{2\sigma^2} \tag{1.8}$$

where $\Lambda = (\lambda_1, ..., \lambda_K)^{'}$, $B = (\Theta, \beta, \gamma)$ with $\Theta = (\theta_1, ..., \theta_K)^{'}$, $\bar{S}_{ct}(\Lambda) = I_{n_{ct}} - \sum_{k=1}^{K} \lambda_k \bar{W}_{k,ct}$, and $\bar{\varepsilon}_{ct}(\psi) = \bar{Y}_{ct} - \sum_{k=1}^{K} \lambda_k \bar{W}_{k,ct} \bar{Y}_{ct} - \sum_{k=1}^{K} \theta_k \bar{W}_{k,ct} \bar{X}_{ct} - \bar{X}_{ct}\beta - \bar{Y}_{c,t-1}\gamma$ where $\psi = (\Lambda, B)$. With some

22 If we assume non-normality for the error terms, then MLE is Quasi-MLE

algebra, we can show that (A.4) can be written without P_{st} as

$$\ln L_{ct}(\Lambda, B, \sigma^2) = -\frac{n_{ct}-1}{2}\ln(2\pi\sigma^2) - \ln\left(1-\sum_{k=1}^{K}\lambda_k\right) + \ln|S_{ct}(\Lambda)| - \frac{\varepsilon_{ct}(\psi)' \cdot Q_{ct} \cdot \varepsilon_{ct}(\psi)}{2\sigma^2} \tag{1.9}$$

where $S_{ct}(\Lambda) = I_{n_{ct}} - \sum_{k=1}^{K}\lambda_k W_{k,ct}$, and $\varepsilon_{ct}(\psi) = Y_{ct} - \sum_{k=1}^{K}\lambda_k W_{k,ct}Y_{ct} - \sum_{k=1}^{K}\theta_k W_{k,ct}X_{ct} - X_{ct}\beta - Y_{c,t-1}\gamma$. We need to restrict the parameter space for Λ to guarantee that $|S(\Lambda)|$ and $\ln\left(1-\sum_{k=1}^{K}\lambda_k\right)$ are strictly positive so that the likelihood is well defined. $|S(\Lambda)|$ will be strictly positive when $\sum_{k=1}^{K}|\lambda_k| < 1$ as our network matrices are row-normalized.

The log-likelihood function for entire classes and times is simply $\ln L(\Lambda, B, \sigma^2) = \sum_{c=1}^{C}\sum_{t=1}^{T} L_{st}(\Lambda, B, \sigma^2)$. To simplify estimation, we concentrate out B and σ in (A.5). The QMLE of B and σ^2 given Λ is $\hat{B}(\Lambda) = \left(\sum_{c=1}^{C}\sum_{t=1}^{T}\chi_{ct}' \cdot Q_{ct} \cdot \chi_{ct}\right)^{-1}\left(\sum_{c=1}^{C}\sum_{t=1}^{T}\chi_{ct}' \cdot Q_{ct} \cdot \mu_{ct}(\Lambda)\right) = \Phi(\Lambda)^{-1}\Psi(\Lambda)$ where $\chi_{ct} = \sum_{k=1}^{K}\theta_k W_{k,ct}X_{ct} + X_{ct}\beta + Y_{c,t-1}\gamma$, $\mu_{ct}(\Lambda) = Y_{ct} - \sum_{k=1}^{K}\lambda_k W_{k,ct}Y_{ct}$, $\Phi(\Lambda) = \sum_{c=1}^{C}\sum_{t=1}^{T}\chi_{ct}' \cdot Q_{ct} \cdot \chi_{ct}$ and $\Psi(\Lambda) = \sum_{c=1}^{C}\sum_{t=1}^{T}\chi_{ct}' \cdot Q_{ct} \cdot \mu_{ct}(\Lambda)$, and

$$\hat{\sigma}^2(\Lambda) = \frac{\sum_{c=1}^{C}\sum_{t=1}^{T}\varepsilon_{ct}'(\psi) \cdot Q_{ct} \cdot \varepsilon_{ct}(\psi)}{\sum_{c=1}^{C}\sum_{t=1}^{T}(n_{ct}-1)} = \frac{\Upsilon(\Lambda) - \Psi(\Lambda)' \cdot \Phi(\Lambda) \cdot \Psi(\Lambda)}{\sum_{c=1}^{C}\sum_{t=1}^{T}(n_{ct}-1)}. \tag{1.10}$$

where $\Upsilon(\Lambda) = \sum_{c=1}^{C}\sum_{t=1}^{T}\mu_{ct}(\Lambda)' \cdot Q_{ct}\mu_{ct}(\Lambda)$. Then, the concentrated log-likelihood function in Λ is

$$\ln L^c(\Lambda) = -\sum_{c=1}^{C}\sum_{t=1}^{T}\frac{n_{ct}-1}{2}[\ln(2\pi)+1] - \ln\left(1-\sum_{k=1}^{K}\lambda_k\right) + \ln|S_{ct}(\Lambda)| - \frac{n_{ct}-1}{2}\ln\hat{\sigma}^2(\Lambda) \tag{1.11}$$

Then the QMLE, $\hat{\Lambda}$, is the maximizer of (A.7), and the QMLE of B and σ are $\hat{B}(\hat{\Lambda})$ and $\hat{\sigma}^2(\hat{\Lambda})$, respectively. Finally, we apply adaptive lasso technique to select a set of networks (reference groups) that is most important to students.[23] For this, we are adding a penalty term to (A.7) such that

$$Q(\Lambda) = \ln L^c(\Lambda) - \sum_{k=1}^{K} p_{\tau,k}(|\lambda_k|) \tag{1.12}$$

where $p_{\tau,k}(|\lambda_k|) = \frac{\tau}{\hat{\omega}_k}|\lambda_k|$, τ is a tuning parameter, and $\hat{\omega}_k$ is some $\sqrt{n}$ consistent estimator. We select the optimal tuning parameter using the ERIC (Hui et al., 2015)[24] and use the estimates

23 The shrinkage technique is also helpful to avoid a overfitting problem potentially embedded in our problem as we do not have a prior which network or reference groups are relevant.

from (A.7) for $\hat{\omega}_k$. For computation, we use the local quadratic approximation algorithm (Fan and Li, 2001) and stop iterations if the difference between subsequent estimates are below 10^{-12}, and any estimate whose absolute value is less than 10^{-4} is shrunk to 0.

24 We also can consider half sample cross-validation.

Chapter Two

You Are Who You Eat With

Evidence on academic peer effects from school lunch lines

Jonathan Presler

Syracuse University

2.1 Abstract

There is a sizable literature estimating the magnitude and importance of peer effects in education on academic outcomes. Most studies define who is in the peer group, but little work has been done to measure the intensity of connection - how important is each peer? Equally weighting all peers in a reference group assumes that all peers are equally important and may bias estimates towards zero by underweighting important peers and overweighting unimportant peers. I examine classmates using a novel approach to measure the intensity of connection between these students. I use administrative transaction data from the New York City Department of Education to observe the lunch line on a daily basis and use lunch line proximity as a measure of connection strength. The result is a revealed friendship network which I use to identify peer effects. I find that students who eat together are important influencers of one another's academic performance, with stronger effects in math than in reading. Further exploration of the mechanisms supports the claim that these are friendship networks. I also compare the strength of connections from different portions of the school year and find that connections formed at the beginning of the school year are most important, consistent with the story that long-term friendships are more important than short-term friendships.

Acknowledgements

Support for this project comes from the National Institute of Health (R01 grant: 5R01DK097347-02). I would like to thank the NYC Department of Education's Policy Research and Support Group and Office of Food and Nutrition Services for providing data and Stephen O'Brien, and Armando Taddei for their support. For data support, advice, and suggestions, I also thank Meryle Weinstein, Amy Ellen Schwartz, William C. Horrace, Hyunseok Jung, Alfonso Flores-Lagunes, and Joanna Bailey. I thank seminar participants at the Maxwell School at Syracuse University and at the APPAM conference for useful comments on previous drafts. The opinions expressed are those of the author and do not represent the views of the NYC Department of Education or NIH.

2.2 Introduction

Researchers and policymakers have long thought peer effects to be an important component of education.[1] Current policy discussions around tracking, school choice, and school integration are often based on the assumption that peers play an influential role in education production. A primary focus of the literature is the effect peers have on test scores, but estimates vary widely[2] due to differences in context, methodology, and how the peer group is defined. Social networks are by their nature hierarchical and complex, and each individual is uniquely impacted by a different set of peers. Understanding which peers are relevant is critical to identifying meaningful estimates of peer effects - both for research relevance and tautologically as many models include average group characteristics or outcomes. If the reference group has little impact on a student, or is too broadly defined, we may understate the importance of peers. However, if we define the set of peers too specifically, we may miss other influencers and misstate their importance.

A central question when identifying peer effects is how to define peers. Much of the literature has focused on the scope of the peer group - defining (or assuming) who is and who is not important, but little has been done to measure connection strength - how important each relevant peer is. Heterogeneity of effects may be explored based on race or sex (for example), but intensity of relationship is typically relegated to binary indicators centered around homophily. Bramoullé, Djebbari, and Fortin (2009) points out that if individuals act outside of a group framework,[3] instead having individualized peer groups, then the endogenous and exogenous effects discussed

1 For example, the Coleman Report (Coleman et al., 1966) looked at the achievement gap between whites and blacks and found differences in peer effects to be a contributor.

2 See B. Sacerdote (2011) for an excellent review of the education peer effect estimates. These estimates range from slight negative in Vigdor and Nechyba (2007) near -0.1 to slight positives in Burke and Sass (2007) near 0.05 to Hoxby (2000) with large estimates of 0.3 to over 6. These papers show the range of estimates, but most in the literature fall near the middle range.

3 Bramoullé, Djebbari, and Fortin (2009) uses the term “group interactions” to refer to the idea that people interact and behave in groups, and that all members in the group are equally important. The group may be grade-cohorts or classrooms, for example.

in Manski (1993) can be separately identified. Separate identification of these effects is policy relevant and may imply different prescriptions. For example, Hoxby (2000) shows that elementary students benefit from the presence of a higher percentage of girls in the elementary school classroom. What this does not tell us is whether interaction with girls is important (the endogenous effect), or if it is related to, for example, behavioral differences where young boys demand more teacher time (an exogenous effect). If the former is true, it may be desirable for classrooms to be structured such that student interaction in mixed gender groups is increased. If it is the latter it may be beneficial for teachers to have an aide help manage student behavior.

This paper measures the intensity of connection between students in the same classroom and uses this network structure to determine social spillovers on math and reading test scores. We use administrative point of service (POS) data from the New York City Department of Education (NYCDOE) to observe daily lunch transactions. We use these daily observations to determine which students frequently stand near one another in the lunch line as a measure of friendship. This friendship network has several important features. First, the measure of contact is based on administrative data rather than surveys. This means that friendships are revealed rather than stated, and their revelation means that the friendships need to be reciprocated.[4] Additionally, friendships change and we observe the result of daily decisions students make, rather than a single survey snapshot. Second, our measure of contact between students occurs during lunch, which is an important social space. Lunch is a relatively unstructured environment within school, allowing students more freedom to interact outside direct supervision from teachers and without direct academic consequences. This makes connections observed during the lunch period socially meaningful. Third, connections can have varying strengths,[5] and we allow them to vary on a

4 In survey data like Add Health, students nominate their friends. Lin and Weinberg (2014) use Add Health to show that reciprocated friendships are stronger than unreciprocated friendships on a variety of outcomes, including academics.

5 For example: student A considers student B to be her best friend, student C is a friend, and student D is simply a classmate.

continuous scale of importance. Next, these networks are constructed fresh during the school year as students do not have control over their classroom assignment, and we can observe how friendship significance evolves over time. Finally, because this network is individual-specific, we do not observe the perfect collinearity between group mean characteristics and mean expected group outcome (the reflection problem).[6]

We find evidence of significant social spillovers on academic outcomes. Effects are stronger in math than in reading, with a one standard deviation improvement in classroom peer performance resulting in an increase in own math performance between 7.5% and 11.1% of a standard deviation. This is on par with our estimates of the black and white achievement gap. Improvement in classroom peer performance in reading by a standard deviation improves own performance by between 4.1% and 6.3% of a standard deviation.

We measure friendship between students based on daily observations of contact over the entire school year. This allows us to treat friendships in different portions of the year as separate networks and observe the evolution of these connections over time. Our results suggest that connections formed at the start of the school year are most important, a finding that is consistent with Patacchini, Rainone, and Zenou (2017) in which long-term friends are found to be more influential than short-term friends.

There are two relevant literatures to which this paper contributes. The first is the literature around education peer effects on academic outcomes. We see a large variation in the reference group used by researchers. Cohort, or school-grade level, is frequently used as a reference group in order to avoid concerns around sorting into classrooms. Hoxby (2000) examines elementary school students at the cohort level and finds that students perform better in classrooms with higher proportions female, regardless of gender. Burke and Sass (2013) find cohort effects near zero using a student fixed effect model incorporating average peer performance. However, they find that that classroom peers are more important than cohort peers and produce meaningful peer

6 As described in Manski (1993), and in section 2.5.1.

effects. This suggests that while choosing the cohort is a convenient way to avoid selection problems, estimates derived from this reference group may understate the overall peer effect. This is intuitive, as peer performance is often measured using mean outcome. Incorporating unimportant peers who contribute little to the outcome should bias estimates towards zero.

It is sometimes possible to zoom in further than the classroom level to examine the effect of students who almost certainly are in one anothers' social network. B. Sacerdote (2001) does this in the college setting by looking at the effect of roommates on one another. He finds that roommates are important sources of peer effects, but that these are lower bounds on the effects of peers because roommates are typically only a piece of a student's social network. Another way to look beyond the school, cohort, or classroom is to use information on the network structure itself. The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a workhorse in the peer effects literature as a result of its friendship survey of middle and high school students.[7] Lin (2010) uses the friendship network described in Add Health to identify academic peer effects among high schoolers using maximum likelihood estimation of a spatial autoregressive model (SAR). The friendship survey allows researchers to pick a set of relevant peers from the school (networks are within school, not within classroom), but the researcher does not know the relative importance of these students (ex: who is the student's closest and most influential friend).[8]

This paper also contributes to the the partially overlapping networks literature. The key insight from this literature is that identification problems such as reflection are solved when multiple reference groups partially overlap. The intuition is that collinearity issues related to the

7 A litany of papers have been written using Add Health, including Bifulco, Jason M. Fletcher, and Ross (2011) and Patacchini, Rainone, and Zenou (2017). The appeal of this data, which surveys and follows up with students who were 7-12th graders in US public schools during the 1994-1995 school year, is that it asks students who their friends are and includes survey responses on a variety of outcomes from GPA to smoking and drinking habits.

8 There are some exceptions. Patacchini, Rainone, and Zenou (2017) divides the networks into students who are friends in both waves and those who are not, finding that students who are long-term friends are more important than those who are friends in just one, both in the short and long term.

characteristics and outcome of the reference group are broken up when individuals participate in more than one network, and these networks share some but not all members. To our knowledge, an early draft of Laschever (2013) was the first paper to propose this method. Bramoullé, Djebbari, and Fortin (2009) formalize the partially overlapping reference group approach and shows explicit conditions for overcoming the reflection problem. De Giorgi, Pellizzari, and Redaelli (2010) provides an empirical example of partially overlapping networks. Students are randomly assigned to nine college courses. Strength of connection between individuals is based on the number of courses students take with one another, and the authors estimate an overall peer effect for this network. W. Horrace et al. (2019) relaxes the assumption that all networks are equally important and estimates the relative importance of networks based on ideas of homophily.

The contributions of this paper are as follows. First, we measure opportunities for contact between students in the lunch line to demonstrate a novel approach for constructing a revealed friendship network. Data in which social connections are observed are rare, and this is a large hurdle in empirical peer effects research. Second, our measure of contact allows us to not only refine the classroom environment and weight peers according to a revealed friendship network, but also overcome the binary nature typically seen in friendship networks wherein students are either friends or they are not. We weight friendships on what is essentially a continuous scale of importance, meaning that our results are not predicated on overweighting unimportant students and underweighting important peers as group averages necessarily do. Third, to our knowledge this paper presents an empirical application with the largest set of partially overlapping networks (one for each day of the school year) such that the measure of connection between individuals can be thought of as continuous. Finally, each student has a unique reference group allowing us to decompose the peer effect into its "endogenous" and "exogenous" components[9] using a linear in means model, a distinction which is important for accurate estimation of spillover effects.

The layout of this paper continues as follows. Section 2.3 describes the data we use and how

9 As described in Manski (1993), and in section 2.5.1.

we construct our sample. Section 2.4 discusses how we measure contact between students. We discuss identification issues, how we overcome them, and the model we use in Section 2.5. Section 2.6 presents our baseline estimates. Section 2.7 shows that our estimates are robust to several additional specifications and robustness checks. Section 2.8 concludes.

2.3 Data and Sample Construction

2.3.1 Data

Data used in this paper are student level and come from the New York City Department of Education (NYCDOE) administrative database for the 2018 academic year, with lag outcomes coming from the 2017 academic year. We observe student characteristics such as sex, race, grade, zip code, poverty status, and whether a student is an English language learner. We also observe homeroom classroom assignment and test score outcomes for reading and math. For elementary school students, homeroom class assignment indicates the student's primary classroom. In addition, we observe student lunch transactions at the point of sale (POS). The data indicate the exact timing of lunch purchase transactions for students (to the second) for every day during the school year. We use this transaction information to observe the order of students in the lunch line and measure social connections in the classroom by observing which students are often in proximity to one another in the lunch line.

2.3.2 The point of sale system

The New York City Department of Education (NYCDOE) has been increasing the use of point of sale (POS) systems in their school cafeterias since 2010. By academic year 2018, 88.0% of schools had the system installed at the start of the school year. These schools served 90.2% percent of the over one million students in the school district. Implementation started in large schools first, with a focus on middle and high schools where the district felt these systems would do the most good. However, by academic year 2018 the system was in 93.4% of elementary schools, and these schools served 95.5% of grade 1-5 students.

Table 2.1 shows that the makeup of the schools with POS systems is slightly different than the full school district. Students in schools with a POS system are more likely to be Hispanic or Asian/other and less likely to be black. This is likely because schools with POS systems are more likely to be located in Staten Island and Queens, and less likely to be in the Bronx or Brooklyn. These differences, while statistically significant, are not large.

The primary way students interact with the POS system is either by entering a PIN in a keypad or a cafeteria worker uses a list of names and faces to enter the transaction as students move through the line. This is not standardized over the district, can vary by school, and is not observed.

2.3.3 Sample

The sample is taken from the universe of students in the NYC public schools for the year 2018. We examine elementary school students for two main reasons. First, homeroom assignment corresponds to the student's primary classroom in elementary school. Second, elementary school students are more likely to participate in the school lunch program than middle or high school students because they have less autonomy and do not have the same outside options as older students who may be allowed to go off campus during lunch. School lunch participation for our sample is 65.8%. Additionally, we limit our analysis to students in general education classrooms.

Table 2.1 illustrates our sample selection process. We begin with all general education fourth and fifth grade students in schools with a POS system in place for the entire academic year. The POS system is important because we measure social connection through repeated observation of lunch transactions, such that students who are observed together frequently are considered friends. We restrict to fourth and fifth grade students because standardized tests begin in the third grade, and we include a lag test score in our model. We are unable to measure connection to students who never participate in school lunch, and 3.7% of students fall into this category. We lose 7.5% of students because they are missing either a current year score or a lag test score for both math and reading. Some students do not participate in standardized tests, so we lose another 2.3% of students from test non-participation in both math and reading. We lose less than half a percent of students from the following three reasons. First, we exclude lunch transactions occurring before 10am and after 2pm. Transactions occurring outside this window are rare and may be improperly coded breakfast transactions, transactions entered after the fact (such that timing is not indicative of the lunch line order), or simply an unreasonable assigned lunch time.[10] Second, we remove transactions occurring more than an hour earlier or later than the mean transaction time for a classroom. These students appear to be "out of line", and as a result are not relevant for determining who is next to whom in the lunch line. Including them would simply add noise to our estimates, so we remove them. Third, we remove transactions which occur simultaneously for the entire classroom. This is indicative of an unusual event, such as a field trip, and gives no information relevant to the lunch line order. The final exclusion we make is excluding students in classrooms with less than 20 students, resulting in the largest loss of students (15.78%). We choose to look only at classrooms that are larger than twenty students because we are concerned classrooms that appear smaller may be integrated co-teaching (ICT) classrooms, and we do not want unobserved peers.[11]

Table 2.3 gives some summary statistics regarding our sample. Our sample includes fewer black students and more Hispanic and Asian/other students. This is likely the result of where the POS systems have been implemented, as the Bronx and Brooklyn are underrepresented while Queens

10 Some lunch times are even more unreasonable than these bounds we place on lunch times, as in Brand's (2019) article "Why do some NYC school kids still eat lunch before some of us have had breakfast?" However, times like these are even more of an anomaly for elementary students than the high schoolers discussed in the article.

11 ICT classrooms combine general education students and students with disabilities together. Students learn from the general education curriculum and are taught by a team of two teachers: one general education teacher and one special education teacher. ICT classrooms typically have a ratio of 40% students with disabilities and 60% general education students.

and Staten Island are overrepresented in locations having received POS systems. Because implementation has occurred in a large proportion of schools, discrepancies are small. Test scores are normalized z scores across grade level in the school district, so our sample is slightly higher performing than average. Average class size in our sample is 25.6 students, and the lunch participation rate is 65.8%. Math and Reading scores are z-scores standardized to zero for the entire NYC public school student population.

2.4 Network Construction

2.4.1 Defining Social Distance

This paper uses a novel approach to measure contact between students and reveal the classroom friendship network. We observe the timing of every lunch transaction in the POS system every day during the school year, and this timing is precise to the second. This allows us to observe the lunch line both in terms of physical order and in the timing of movement through the line. We use this information to construct a peer network, but first discuss how to extract a meaningful social distance from this information. There are two ways we might consider using this information. The first is to use the actual timing of the transactions to measure distance between students. However, this is not our preferred method, as we believe this to be a noisy estimate of social proximity between students. We discuss this measure further in section 2.7.4.

Our preferred method is to transform the near-continuous timing data into ordinal data. This allows us to think about distance as the physical proximity of students to one another rather than temporal proximity. We argue that because lunch is a relatively unstructured and social time, students' primary concern is who they are able to socialize with in the line and then during lunch. The simplest way to transform the observed order into a social distance is to look at whether any two students i and j are within some threshold distance (number of students) of one another. Our baseline model uses a threshold distance of one - whether two students are next to one another in line. For robustness, we also look at larger threshold distances in Section 2.7.2.

It is worth discussing the implication of the observed lunch line order, as the ordering process is a black box to us as researchers, and the method of ordering likely varies by classroom. We discuss some possibilities for how students are ordered, fitting them into three categories: students have agency over their choice of line position, students are ordered by someone else (such as the teacher), and students have agency within a constraint. We then provide evidence that in the majority of classrooms, students either have at least some agency over their position in line or the order they are given changes frequently.

First we discuss situations in which students have agency over their position in line. Students must balance a choice between being in line with their friends and their preference for being towards the front or back of the line. For most students, we believe the choice of being in line with friends is more important than their line position. If this is true, then it is clear that the line order contains information relevant to the social network in the classroom. However, it is possible that many students' preference for being at the front of the line (for example) dominates their desire to be near friends. A classroom in which all students wish to be first would see a race to the front of the line. Thus line order depends upon classroom geography and where students sit in relation to the door (start of the line), with students sitting near one another tending to line up near one another. If students sitting near one another are more likely to talk to one another or work together during class time, then this gives us another reason physical proximity in the lunch line would be socially important. In both of these situations, students who are near one another in the lunch line would be expected to be more influential in one another's social network - at least as it relates to academics within the classroom - than a randomly selected classmate. The truth is likely some combination of these two situations. For students geographically near the door, they have the option to be first or wait for their friends. Those further from the door do not have this choice. Thus in a classroom in which all students wish to be first, the benefits to rushing decrease in distance from the door. A tipping point could occur at which point students switch from racing to the front of the line to waiting for their friends.

Second, it is possible that students could be ordered by their teacher according to some metric - perhaps alphabetical. We do not observe names, and so we cannot test this hypothesis directly (although we do look at how much strict ordering exists in our sample). If students are ordered based on name, we expect little reason for these students to be socially more important than other students.[12] The teacher could order students by some other method - perhaps according to student characteristics (demographic, performance, or behavioral). If we believe that students with similar characteristics are more likely to be friends with one another (homophily), then observing similar characteristics in students near one another may be indistinguishable from an external ordering placed upon the students according to this same set of criteria. These students may also be more socially important to one another than a randomly selected classmate, as W. Horrace et al. (2019) shows.

Finally, there is the possibility that students sort into the lunch line based on some combination of autonomy and rules. For example, the teacher may dismiss students from their classroom tables, so that students form a line within a subset of the classroom - they have autonomy within a constraint. Students face a similar decision whether to line up next to friends (within the constraint group) or in terms of optimal position. Notice that both physical and social positioning are constrained, as a student with preference for the front of the line may not have a choice over line position until the first half of the line is filled. Similar to when students have full autonomy, line order likely reflects some level of student importance - either through selecting friend groups or the importance of the constraint group (such as classroom geography). The result is similar to that of full autonomy, but the effect of these peers is likely smaller than under full autonomy, as this is a group of "next-best" friends.

12 Outside the notion that students of similar cultural or ethnic backgrounds might have similar names and thereby be grouped together. While some work looks at the ability to predict ethnicity based on names, such as Elliott et al. (2009) and Ryan et al. (2012), the success of these algorithms is still limited. Predicting ethnicity based on alphabetical ranking within an average group size of around 25.6 would be unsuccessful.

While the line-up process is itself unobserved, we provide evidence that students have at least some agency over this decision by considering whether students are ordered into roughly the same order each day. To do this, we construct a measure of within-classroom noise as detailed in Appendix 2.11.2. The measure M is based on the number of order inversions (swaps in the order of students i and j) observed in the order over the year, and it is normalized such that it is invariant to classroom size and participation rate. Figure 2.4 shows the distribution of that measure and that the bulk of classrooms (average measure value is 0.203) are closer to a uniformly random distribution (value of 0.25) than fully ordered (value of zero), but that there is more order than complete randomness.[13] This is consistent with the idea that students in most classrooms have agency over their position in line, and choose positions in ways that are varied but less than random (ex: in order to be with their friends). It is important to note that the distribution of this measure has a small tail with what may be considered abnormally low noise (where we may think classrooms are ordered). If we let 0.1 be the threshold below which classrooms are ordered, about 3% of classrooms are ordered.[14] In Section 2.7.3 we remove these as a robustness check.

2.4.2 Scaling from daily observations to the friendship network

We observe daily lunch transaction timings over the entire school year, which we translate into the lunch line order for each day. The next step is to zoom out to the full year, such that students observed in frequent close proximity to one another on individual days are considered friends. Because some students do not participate in school lunch every day (or are absent from school), this process is less straightforward than we might like. On a day that a student does not attend school, we miss their signal of who they would choose to stand in line next to on that day, and they also limit the choice set of the students who remain (by removing themselves from the candidate pool).

13 Appendix 2.11.2 outlines how the measure behaves under changes in class size, participation rate, and levels of randomness.

14 134 of 4,077 classrooms are below the 0.1 threshold.

We start to think about constructing the network by averaging daily observations together, akin to what De Giorgi, Pellizzari, and Redaelli (2010) do with classes. This proximity matrix gives us the percent of days each student is near each other student. We may be concerned that students with low participation will appear to have artificially low connections measured by this proximity matrix. There are two ways we can address this. First, we can simply row-normalize the average proximity matrix such that each row sums to one. Row normalization is common in the literature to transform a proximity matrix to the weighting matrix used in estimation, because it improves interpretability of results by appropriately weighting influential peers for the given student such that we have the weighted average (characteristics or outcome) of the peer group. Each row i indicates student i's relevant peer group, appropriately weighted. We plan to row normalize for the interpretation benefits, but it is important to notice that row-normalization changes the interpretation of the proximity matrix from the percent of school days both students are near one another to be the percent of days student i is present that student i was near every other student j. For student i with low participation, this moves their average connections with students from near zero to the percentage of times i participated and was near each other student j. By increasing the weight on the days a student does participate, we have addressed the issue of not observing who a student would choose to be near if they did attend. However, it is not clear that we have addressed the second problem in which student i is removed from the choice set of other students.

In order to address this second concern, we construct a weighting matrix for the percent of times we observe students near one another when both are present:

$$p_{ij} = \frac{\sum_{d=1}^{D} S_d(i,j)}{\sum_{d=1}^{D} \delta_d(i,j)} \text{ for } i \neq j \text{ ; and } p_{ij} = 0 \text{ for } i = j \tag{2.1}$$

where $S_d(i,j)$ indicates that students i and j are next to one another on day d and $\delta_d(i,j)$ indicates that both i and j are participating in lunch on day d. The weighting matrix is then $W = \{w_{ij}\}$, which is row-normalized prior to estimation (each row sums to one). While averaging the daily observations as done in De Giorgi, Pellizzari, and Redaelli (2010) and the method described in equation 2.1 lead to different weighting matrices before row-normalization, row normalization makes them identical. Figures 2.2 and 2.3 provide visual examples to illustrate how we convert daily observations into weighting matrix (according to equation 2.1). In Figure 2.2, I show a basic example using two example days with four students, and how these combine into a weighting matrix. Figure 2.3 adds a little bit more complexity, illustrating with five students over six days and including both student absences and row-normalization. In Figure 2.3, it appears that row normalization has dampened the effect, but this is not the case. Instead, it proportionally reweights the proximity matrix so that each row, when multiplied by Y or X, creates a weighted average of the relevant peer outcome or characteristics.

Admittedly there are other ways we could construct the proximity matrix, and there are potential concerns with the way we have constructed ours. Perhaps most concerning is that low-participation students could appear overly important for those they stand near when they do participate. We address this by looking at an alternate proximity matrices for robustness in Section 2.7.

It is also important to distinguish between absence and non-participation. The previous discussion dealt with absence from school, but non-participation adds an additional complication. The majority of non-participation in elementary school lunch is because students brought their own lunch from home. Thus if a student is present at school, but not participating in lunch, they are likely present in the lunch line - at least during travel from the classroom to the cafeteria - and importantly they are part of the decision process when students decide where to stand in line. Thus two students we observe as being next to one another may actually have another student between them (or more than one) during the decision process for who to stand near. This is an issue of truncated data, and likely a significant source of noise in the model. The result is that an observed distance of one between two students is actually a distance of at least one. This means connections we observe are weaker than actual connections in the classroom, and results obtained from this data are likely a lower bound on the peer effect from lunch-mates.

2.5 Methods

2.5.1 Identification

Identification of peer effects is notoriously difficult, and in this section we discuss some of the common issues and how we address them in this paper.

In his seminal paper, Manski (1993) discusses the different effects which may be captured in a naive model of peer effects. The first is the endogenous effect, which is often the effect of interest to researchers and policymakers. This is the effect of one individual's performance on the performance of another. For example, we observe an endogenous effect for two students working together on a group project if the performance of one student varies based on skill level (performance) of that student's partner. This is of interest to policymakers because the endogenous effect is a multiplier, causing spillovers to other students. If an intervention is applied to some students and improves their performance, all students will benefit because of the dependence of all students' performance on that of their peers. The endogenous effect is named because it directly places the outcome on the right hand side of our model. Use of the linear in means model structure (assuming the peer effect is a weighted average of peer performance) and maximum likelihood estimation allows us to solve for this endogeneity in our results. More details about the model follow in Section 2.5.2 and about the estimation procedure in Appendix 2.11.1. The second effect is the exogenous effect, sometimes called a contextual effect. Exogenous effects control for student characteristics in the peer group. For example, we might expect that wealthier students perform better on tests, all else equal. As a result classroom performance may increase with wealth, but this is due to characteristics of the student rather than interactions with them. The final effect Manski discusses is the correlated effect. This is often not a social effect at all, but is related to common exposure by students to the same treatment. For example, a lack of adequate facilities or a good teacher are felt by all students in the classroom, but they are not related to the students or any interaction with them.

In many reduced form models of peer effects, we cannot distinguish between exogenous and endogenous effects because the performance of the reference group is collinear with the characteristics of this group. This is known as the reflection problem (Manski 1993). However, when individuals have unique reference groups, this is sufficient to separately identify endogenous and exogenous effects. This is because the collinearity issue arises when individuals share a reference group, but when this does not exist, there is no collinearity issue to worry about here. The individual level reference groups arise because each student is next to a unique set of students each day (another student will likely be next to one of the students, but there cannot be more than one student next to both students). We may be concerned that averaging over the days could cause different students to have the same reference group. However, with 180 days and differing levels of participation among students this does not occur in our sample.

Correlated effects are commonly addressed using fixed effects (ex: Ajilore et al., 2014, Bifulco, Jason M. Fletcher, and Ross, 2011, W. C. Horrace, Liu, and Patacchini, 2016, Lin, 2015), and we follow this trend with the inclusion of classroom fixed effects. Intuitively, we can think of this as controlling for a teacher effect, although it also controls for other group treatments such as quality of the built environment, scheduled lunch time, and principle quality to name a few.

Selection can be a problem if students are sorted into reference groups based on shared characteristics. Because we are looking at the revealed friendship network within a classroom, we are not concerned with selection into the within-classroom network - this is in fact what we are interested in measuring. A potential concern is if students are sorted into classrooms based on shared characteristics such that the strength of the social spillovers within a classroom are correlated with these characteristics. We test whether student characteristics explain classroom assignment and show that class assignment based on observed student characteristics is consistent with randomness in Appendix 2.11.3.

2.5.2 Baseline Model

This paper uses a revealed friendship network to measure academic spillovers in the classroom. For our baseline model, we construct a within-classroom network according to equation 2.1. In the linear in means model we use, this network is multiplied by both the outcome Y and student characteristics X so that we can separately identify endogenous and exogenous effects. Below is the basic format of the linear in means model we estimate:

$$Y = \alpha + WY + WX + X\beta + \theta + U \tag{2.2}$$

where Y is the outcome of interest, α is a constant, W is the weighting matrix as defined at the end of section 2.4.2, θ is the classroom fixed effect, β is the estimate of own characteristics X, and U is the error term. Controls in X include lag test scores and indicators of sex, ethnicity, zip code, English language learning, and poverty status.[15]

Modeling exogenous effects allows us to control for the characteristics of students in the reference group, thereby isolating the endogenous effect of interest - the effect of one student's performance on another's. The endogenous effect is important to distinguish from the exogenous effect because it captures the spillover effects resulting from social interaction, whereas the exogenous component controls for student characteristics. References to estimates of the peer effect refer to this endogenous effect.

We estimate our model using Maximum Likelihood Estimation (MLE) and follow W. Horrace et al. (2019) and L.-f. Lee and Yu (2010). Details of the estimation procedure are found in Appendix 2.11.1.

15 The poverty indicator indicates whether a student has ever been eligible for free or reduced price lunch over all of the student's time in the NYC public school system.

2.5.3 Interpretation

It is important to note that estimates of the endogenous effect from model 2.2 are multiplier effects. This means that interpretation of the estimated structural parameter $\hat{\lambda}$ is done by converting the result as below:

$$\hat{\gamma} = \frac{1}{1-\hat{\lambda}} \tag{2.3}$$

Thus an estimate of $\hat{\lambda} = 0.05$ is interpreted as a multiplier of 1.053. This means that a ten percent improvement in test scores for a student's reference group results in a 0.53 percent improvement in the student's own test score. Notice that for small λ, the multiplier γ is comparable in magnitude.

2.6 Results

2.6.1 Baseline Results

Table 2.4 shows our baseline results for math and reading scores of fourth and fifth graders using a proximity measure in which students are next to one another in the lunch line. Our outcome of interest is test scores, and these are z-scores normalized citywide among students in the same grade. In addition to the controls shown, the model also includes fixed effects for zip code of residence. We also include these zip codes in the exogenous effect.The first line of Table 2.4 shows a math peer effect for students in the lunch line together of 0.089, which is statistically significant. As discussed in section 2.5.3 this is a multiplier effect, and so we interpret this as a multiplier of 1.098 or an increase of 0.098 units. The endogenous effect for reading is also significant, but smaller at 0.053, or a multiplier of 1.056. The fact that both estimates of the endogenous effect λ are positive is consistent with our intuition and the general findings of the literature, which is that improvements in the reference group should lead to improvements in own outcomes. We can interpret these results by saying that if a student's relevant peers exogenously improve their performance by one standard deviation, we expect to see improvements in own performance in math by 9.8% of a standard deviation. This is equivalent to the black-white test score gap in math. The gap is larger and the spillover effect is smaller in reading, so the equivalent improvement is equivalent in magnitude to about 40% of this gap.

It is difficult to associate meaning to a comparison of these estimates to static estimates because they are multipliers and therefore amplify all other elements of the education production function. We can think about the interpretation of these multiplier estimates when combined with additional external information and compute an average effect. The average classroom in our data has substantial variation in student ability, which we see manifest itself in student performance. If we collect the top performer in all classrooms, we find that the average classroom has a student performing 1.5 standard deviations above the mean.[16] This is mirrored in low performers.[17] We then collect the strongest connection we observe in each classroom, which we can think of as a student's best friend. The average student's largest connection is 0.357, meaning that over one third of the time we see both students present, they are next to one another in the lunch line. When we conduct our row normalization, the meaning is preserved, but the value of the matrix cell for the strongest connection reduces to 0.202. This means that the benefit to a student of connecting with the best student in the classroom, rather than an average student, is 0.030 in math and 0.018 in reading. This is equivalent to half the effect of poverty and nearly one third of the black-white test score gap in math, and it stems from only one peer connection. In reading the effects are smaller than math, being one third the effect of poverty and 14% of the black-white test score gap (which is larger in reading).

The fact that the spillover is larger in math than reading is consistent with the idea that students learn verbal and reading skills at home, but primarily learn math in school. We see stronger in-school math effects than reading effects in Nye, Konstantopoulos, and Hedges (2004) which shows that teachers have a greater impact on math scores than reading scores.

16 In math, the average top performer across all classrooms scores 1.52 standard deviations better than the mean. For reading the average top performer scores 1.55 standard deviations above the mean.

17 The average bottom performer across all classrooms scores 1.39 standard deviations worse than the mean in math, and 1.51 standard deviations worse in reading

Notice that the controls are performing as expected. Own student lag scores are highly significant and important. Male students perform slightly better in math but worse in reading than their female peers. English language learners and poor students do worse than native speakers and students who are not poor. The comparison group for ethnicity is Hispanic students, because these are the modal student in NYC public schools, and whites and Asians do better than them, while blacks do worse. Most of the exogenous effects are not statistically important, with the exception of friends in the Asian/other group, which has a large positive impact. Taking the math estimate, this means having all friends in the Asian/other group improves own math performance by 0.12 standard deviations as opposed to having Hispanic friends (the baseline reference group), all else equal. Notice that the exogenous effect is not a multiplier effect, but simply shows the effect of having friends from this group type. Surprisingly, the previous performance of friends does not appear to matter in math, but it is quite important in reading.

2.6.2 Evolution of the friendship network over time

The friendship network is constructed over repeated observations of the lunch line. As a result, we can explore different portions of the school year to see whether the strength of the friendship network varies over the year. It is ambiguous whether connections should be more important at the start or at the end of the year. Patacchini, Rainone, and Zenou (2017) provides evidence that students who are friends for longer periods of time are more important than short term friends, so we might expect that connections at the start of the year are more influential. On the other hand, students are still getting to know one another at the start of the year, so we may observe more noise as students sort into friendships. Additionally, testing occurs towards the end of the school year, so we might expect connections closer to the test date are most important.

Table 2.10 divides the year into halves, thirds, and quarters to compare friendship importance over time. When dividing the year into halves, we construct two proximity matrices according to equation 2.1. The first proximity matrix uses the set of days from the first half of the year, and the second matrix uses the days from the second half of the year. When dividing the year into three and four components, each matrix is constructed from the corresponding set of lunch line observations (days).

The first portion of the year seems to be most important, suggesting that friendships formed early and in place the longest are most important - even within the shorter time horizon of a single school year. This could be the result of carryover from who students know the previous year (so that they really are long-term friends), or they could be the result of newly formed friendships. Students do not have agency over their classroom assignment, and we show in Appendix 2.11.3 that class assignment is consistent with randomness along observable characteristics. While some connections will carry over from the previous year, there is no reason to think the strongest connections carry over, outside those that randomly get placed in the same classroom. This is also consistent with a story in which students either continue or begin long-term relationships with a small group of students and then try to branch out over the course of the year.[18]

2.6.3 Discussion

To our knowledge, Lin (2010) is the closest study to our own in terms of methodology, so we compare estimates. Lin uses the Add Health data to estimate peer effects on GPA using a similar spatial autoregressive model with maximum likelihood estimation. It is important to note that the sample in Lin (2010) is older than our own sample, because the youngest students in the Add Health survey are in the seventh grade. Students nominate up to five other students of each gender from their school as friends. Thus the network structure involves connections between each student and up to ten or twenty students in the school depending upon whether connections need to be reciprocated. Each of these students is equally weighted, because the researcher does not observe the relative importance of these friends. Lin (2010) estimates the magnitude of peer effects such that they improve GPA by 7.85% of the mean GPA. This is comparable with our own

18 Future versions of this paper will further explore selection.

results, in which math is above and reading below this estimate.

Our approach is to appropriately weight classmates based on the revealed friendship network. This should reduce the importance of students who are not friends, and increase the importance of students who are, thereby better targeting important peers. If friends are more important than a random peer, we should expect our results to be larger than those who give all friends equal weight. There are three potential reasons we do not see this stronger effect, which we will address one by one.

First, as touched upon previously, the papers look at different contexts. Social spillovers may be smaller for younger students than for older ones. Table 2.5 shows the results for our model when we separately run the model for each grade. The estimates for social spillovers in reading are statistically indistinguishable from one another when we compare fourth graders and fifth graders. However, the math spillover effect is much stronger in the fifth grade than in the fourth grade, which would be consistent with the conjecture that social spillovers increase with age. Thus differences in age may explain why we do not find stronger effects than an unweighted social network as in Lin (2010).

Second, while the models are similar, the data are different and dictate different models. Our data allow us to look within the classroom, and so include classroom (teacher) fixed effects. Add Health only allows network (school) fixed effects to be used. If friends share classes, we might expect them to do better or worse together due to teacher effects which cannot be controlled for in the Add Health data. A difficulty with Add Health is that we do not know if students are friends because they take classes together, or choose to take classes together because they are friends. Additionally, students are unrestricted in their choice of friends in Add Health. Thus students can nominate students outside their classes who are important and the choice set for important peers is less restrictive than nominations within a classroom. However, this last confounder is unlikely to be a large source of bias, as work such as Burke and Sass (2013) show that classmates are much more important than those outside the classroom.

Third, we could be incorrect in our supposition that friends have differing levels of importance. It may be that what really matters is the extensive margin - whether students are friends, not the intensive margin of how close the friends are. Using the Add Health data, Patacchini, Rainone, and Zenou (2017) find that students who are friends in both waves have stronger effects on one another than those who are friends in only one wave (although this paper does not look at academic outcomes). This suggests that there is indeed a hierarchy in friendships, even in the connections laid out in the Add Health network. In our data, we cannot distinguish between who is and who is not a friend outside our weighted framework. We can test whether our model performs better than randomly selected peers, and we discuss this in greater detail in section 2.7.1. We find no evidence of a peer effect when we randomly assign line orders, suggesting that this third option is not correct. We conclude that the reason our estimates are not larger than Lin (2010) must be due to a combination of context and model specification.

Our results provide evidence of strong spillover effects using a revealed friendship network. These results are on par with estimates found using a similar method. Our estimates are smaller than some reduced form estimates which do not attempt to disentangle endogenous and exogenous effects, such as those in Hoxby (2000). We are consistent with Hoxby (2000) in that we find stronger effects in math than in reading. Our results are larger than those Burke and Sass (2013) find for elementary school students using student and teacher fixed effects in combination with mean peer achievement. We can attribute this difference to a combination of methodology and definition of the reference group.

2.7 Robustness Checks

2.7.1 Random Lunch Lines

By construction, students in each of the networks we construct share a classroom, so we might expect that they are socially important to one another regardless of proximity in the lunch line.

To test whether the spillover we estimate is simply a result of the students sharing a classroom, we randomly shuffle the lunch line for each day of the year and re-estimate the model. Results are found in Table 2.6. The placebo estimate for math is a statistically insignificant 0.004 and for reading it is also insignificant at 0.016. While the endogenous effect (as well as all the components of the exogenous effects) are insignificant, the own effects perform similarly to the baseline model. We conclude that students in close proximity to one another in the lunch line are socially more important to one another than a randomly selected classmate.

2.7.2 Alternate Distances in the Lunch Line

When constructing our baseline network, we chose to connect students who are next to one another in line. If friendship groups are larger than pairs, a larger distance may be appropriate. For example, if we observe friends A, B, and C in line together, we miss the connection between A and C if we restrict our analysis to students who are next to one another. We can increase our distance from one to test whether this is the appropriate group. Table 2.8 reports estimates when including multiple networks for each additional distance between two and six. While the inclusion of additional networks slightly dampens the effect from a distance of one, the results remain robust to the inclusion of these networks. Additionally, no higher distance is statistically significant. This indicates that students being next to one another in line is the strongest signal of connection we can measure, and while students who are further apart in line may also be friends, this signal is too noisy to be meaningful. Given the large number of observations throughout the year, we are able to observe a spectrum of connection strengths. For example, we may observe the same group of friends C, A, and B the next day. With only these two observations, we measure a stronger connection between A and B than either has with C. This is because we always observe both A and B next to one another.

It is possible to estimate the model with a single network, using higher distances (rather than putting each distance into its own separate network). Table 2.9 shows the results of this network specification for the same set of distances. The effects appear to be stronger than our baseline model. Notice that the point estimates in math increase and then decline after a distance of four. For reading the point estimates continue to rise in each specification. Standard errors are increasing for both outcomes, indicating additional noise from the inclusion of students who do not matter. We expect that the cause is scenarios like in the toy example with students A, B, and C discussed above. Students who are friends are likely near one another in line, even if they are not next to one another. What these models pick up are larger friend groups. While Table 2.8 shows that no other distance is important on its own, these estimates pick up the extent to which larger friendship groups (or possibly friends of friends) are important. We provide this as suggestive evidence that a broader set of students matter for performance in reading than in math.

2.7.3 Removing potentially ordered classrooms

When we measure variation in the line order, not all classrooms appear to give students agency over their location in line. We introduce a measure M of within-classroom noise in section 2.4.1, which is discussed in greater detail in appendix 2.11.2.

Table 2.7 shows the results of removing classrooms which exhibit small levels of variation in the observed line order. We remove classrooms with M less than 0.05, 0.1, and 0.15. These values signify low levels of variation in lunch line order throughout the year, which may be attributable to students having no agency over their position in line. Appendix 2.11.2 further discusses properties and behaviors of the measure M. In each case, the peer effect increases in M.

We provide this as evidence that the mechanism is friendship rather than the effect of time spent in the line. In classrooms with very stable line orders, students spend time in the lunch line with their neighbors more consistently than in classrooms with more variation. If the spillover mechanism was due to this time, we would expect a decrease in estimates when these classrooms are removed. We see the opposite, suggesting that connections we observe in classrooms with more autonomy are more meaningful. In addition to suggesting that friendship is the mechanism through which these spillovers are working, this suggests that peer effect estimates in my baseline model may be biased towards zero because of the additional noise added to the model by these ordered classrooms.

2.7.4 Alternate Proximity Matrix Specifications

For our main network specification, we used the order in which we observe students go through the lunch line. Another option is to use the actual timing of transactions as a measure of distance, where students further apart socially will go through the lunch line further apart from one another. In order to estimate the model, we need to construct a proximity matrix similar to the ordinal version discussed previously. To do this, we average the time between each student in the classroom on each day both students participate in the lunch line. This creates a distance matrix, which we convert into a proximity matrix by taking the reciprocal of each entry. As before, we row-normalize this proximity matrix.

Table 2.13 shows the results of using time as a measure of distance. The biggest contrast from the baseline model is that spillovers in math are statistically insignificant and smaller than those for reading (which are statistically significant). The point estimate in reading is smaller than my baseline estimate, but it is not statistically different. For both models, other controls behave similarly for the most part. The biggest exception is the exogenous effect of the lag test score in reading, which is negative and significant at the 5% level, whereas it is positive in the baseline reading model model. Additionally, racial exogenous effects are less significant.

Precisely what is being measured by a temporal distance is more difficult to pin down relative to the ordinal system discussed previously. As a result, this is not our preferred method. There are a number of ways this measure introduces unnecessary noise and uncertainty into our measure. For example, it is unclear what meaning to ascribe to a 10 second pause between adjacent transactions relative to a minute pause between them. Does the longer time signify a social distance - where students who are closer socially try to stick together temporally as well, perhaps waiting for one another and going through the POS system in quick succession? Or is it likely that the person operating the POS system had technical difficulties or was distracted by something occurring elsewhere in the lunch line or kitchen? Similarly, if there is a set of three transactions in relatively quick succession, does this indicate a friend group, or does this indicate students (or a cashier) who is relatively more adept at navigating the lunch line? We could imagine that some students are slower than others, and consistently have larger gaps in time before or after them. Does this slowness somehow make that student less likable? The temporal distance between students could just as easily be due to indecision in the lunch line, ease of distraction, chattiness with cafeteria workers, and the like. Additionally, does it make sense to ascribe the same weight to all connections with a 10 second gap between them, even if there is another transaction between them? These are some of the questions raised by using a temporal measure of social distance, and sufficient reason for us to prefer an ordinal or spatial measure of social distance.

There are other ways we could define proximity, each with positives and negatives. In this section, we examine an alternative method for defining proximity which addresses the concern that low attendance students may have outsized effects on those near them. Section 2.4.2 describes our method for constructing the proximity matrices. Recall that the two main difficulties caused by student non-participation are a missed signal of who they choose to be near and removal of the absent student from the choice set of other students. The definition of proximity proposed here builds upon our baseline measure. Our previous definition of proximity measures the percent of days two students participate in lunch during which they choose to be near one another. To address the problem posed by low-participation students, we can require students to participate a certain number of days together before their connection can be evaluated. The proximity matrix in equation (2.1) is altered such that:

$$p_{ij} = \begin{cases} \dfrac{\sum_{d=1}^{D} S_d(i,j)}{\sum_{d=1}^{D} \delta_d(i,j)}, & \text{if } i \neq j \text{ and } \displaystyle\sum_{d=1}^{D} \delta_d(i,j) \geq \eta \\ 0, & \text{otherwise} \end{cases} \tag{2.4}$$

where η is the threshold number of days students must both participate in before we count their connection. Connections between students when one or both of them are low participation students are reduced to zero unless we see enough participation from both students on the same days. The potential drawback of this method is that we lose all or most connections with low-participation classmates and may be throwing away useful signals. The benefit is that these signals we lose may be noisy.

Tables 2.11 and 2.12, we explore thresholds of between two and ten days that students must participate on the same day before we evaluate their connection. Our results remain relatively unchanged in each specification. We conclude that our baseline model is robust to the concern that low participation students may appear overly important.

2.8 Conclusion

This paper measures contact between students in the school lunch line as an indicator of connection strength between students. This allows us to define not only the scope of the reference group, but measure the intensity of the relationships within that environment. We use the revealed friendship network to separately identify endogenous and exogenous effects using a linear in means model.

Our results indicate significant social spillovers in both math and reading. These effects are stronger in math than in reading. In math, a one standard deviation improvement in peer performance results in an increase in own performance between 7.5% and 11.1% of a standard deviation. This is a significant effect, on par with our estimates of the performance gap between black and white students. Social spillovers have a multiplier effect, magnifying other inputs in the education production function. This suggests an alternative interpretation, which is that for a given intervention that improves math scores, between 7.5% and 11.1% of improvements occur through the peer effects mechanism. Spillovers are lower in reading, where an increase in peer performance of one standard deviation improves own performance by between 4.1% and 6.3% of a standard deviation.

Our measure of connection is constructed using daily observations of student contact in the lunch line. The daily nature of this data allows us to look at the evolution of these connections over time. We find evidence that connections formed at the beginning of the year are most important, which is consistent with the findings in Patacchini, Rainone, and Zenou (2017) that long-term friends are more influential than short-term friends.

There are situations in which we want to measure average peer effects within a context, such as when measuring the average effect of exposure to a specific type of student. For example, it is useful to know the spillover effects for disruptive peers, as in Carrell, Hoekstra, and Kuka (2018). However, when it is important to understand the strength of the overall peer effect within the classroom or to understand mechanisms through which policies may be working, it is important to take into account the connection complexities found in the school social network.

Peer effects are believed to be important for a number of important policy discussions such as tracking, school choice, and school integration. In order to effectively weight the costs and benefits of these policies, accurate estimates of the relevant social spillovers are imperative. Estimates that rely on the average peer effect within a reference group may understate the overall effect by not appropriately weighting the most relevant peers for each student. In this paper, we measure the strength of connection between elementary school students sharing a classroom. Understanding the relative importance of peers within the network accurately weights which peers are important for each individual and provides stronger estimates of the peer effect.

2.9 Tables

Table 2.1 Comparing students in schools with a POS system to the full sample

	NYC Student Population		Has POS System		
	Freq.	Percent	Freq.	Percent	difference:
Borough					
Manhattan	182,794	15.64%	164,258	15.59%	-0.06%
Bronx	246,967	21.14%	218,101	20.70%	-0.44%
Brooklyn	350,124	29.96%	312,236	29.63%	-0.33%
Queens	321,262	27.49%	296,443	28.13%	0.64%
Staten Island	67,307	5.76%	62,721	5.95%	0.19%
Total:	1,168,454	100%	1,053,759	100%	
Ethnicity					
hispanic	472,229	49.76%	429,074	50.41%	0.66%
black	302,744	31.90%	265,098	31.15%	-0.75%
white	174,105	18.34%	156,966	18.44%	0.10%
asian other	214,698	22.62%	198,241	23.29%	0.67%
Total:	949,078	100%	851,138	100%	

Data are from the New York City Department of Education (NYCDOE). Table depict differences between all schools and those with a point of sale (POS) system for all students (over all grades). Ethnicity information is not known for all students.

Table 2.2 Sample Selection Process

Students	Number Drop	Percent Drop	Percent Remaining	Transactions	Number Drop	Percent Drop	Percent Remaining
All 4th and 5th graders at schools using POS systems for the full year							
145,495			100.00%	16,179,728			100.00%
Participate in school lunch							
140,090	5,405	3.71%	96.29%	16,174,323	5,405	0.03%	99.97%
Have a test lag for either math or reading							
129,234	10,856	7.46%	88.82%	15,118,176	1,056,147	6.53%	93.44%
Have a test score for either math or reading							
125,890	3,344	2.30%	86.53%	14,861,855	256,321	1.58%	91.85%
Transaction time is between 10:00am and 2:00pm							
125,771	119	0.08%	86.44%	14,705,898	155,957	0.96%	90.89%
Removing transactions not occuring with student's class							
125,700	71	0.05%	86.39%	14,582,405	123,493	0.76%	90.13%
Removing transactions which are simultaneous for the entire class							
125,559	141	0.10%	86.30%	14,568,447	13,958	0.09%	90.04%
Class size is at least 20 students							
102,606	22,953	15.78%	70.52%	12,010,146	2,558,301	15.81%	74.23%

The table depicts how many students (and corresponding transactions) are lost at each point of the sample selection process.

Table 2.3 Summary Stats

variable	mean	sd	N
lunch_part_rate	0.658	0.275	102,606
lunch length	15.69	24.31	12,010,146
lunch time	12.07	0.79	12,010,146
class_size	25.57	3.21	102,606
female	0.506	0.500	102,606
grade4	0.489	0.500	102,606
grade5	0.511	0.500	102,606
ever poor	0.845	0.362	102,606
ell	0.125	0.331	102,606
ethnicity:			
hispanic	0.408	0.492	102,606
black	0.190	0.392	102,606
white	0.165	0.371	102,606
asian_other	0.237	0.425	102,606
Borough:			
manhattan	0.099	0.299	102,606
bronx	0.213	0.409	102,606
brooklyn	0.282	0.450	102,606
queens	0.335	0.472	102,606
staten_island	0.070	0.256	102,606
zmath	0.082	0.950	101,948
zread	0.088	0.953	102,244

Summary statistics for our selected sample. Lunch length calculated in minutes. Lunch time is in hours, so the mean lunch time is equivalent to 12:04.

Table 2.4 Baseline Model

	Math		Reading	
Peer Effect	0.089**	(0.009)	0.053**	(0.009)
Own Effect:				
lag test score	0.730**	(0.002)	0.660**	(0.003)
female	-0.011**	(0.004)	0.055**	(0.004)
ELL	-0.094**	(0.006)	-0.193**	(0.007)
Asian/other	0.177**	(0.005)	0.157**	(0.006)
black	-0.019**	(0.005)	-0.043**	(0.006)
white	0.076**	(0.006)	0.088**	(0.007)
ever poor	-0.059**	(0.005)	-0.056**	(0.006)
Contextual Effect:				
lag test score	-0.007	(0.010)	0.049**	(0.011)
female	0.005	(0.009)	-0.031**	(0.011)
ELL	0.020	(0.021)	0.048	(0.027)
Asian/other	0.121**	(0.019)	0.085**	(0.022)
black	-0.011	(0.020)	-0.037	(0.023)
white	0.045*	(0.021)	0.006	(0.025)
ever poor	-0.007	(0.019)	0.001	(0.023)
observations:	100,156		94,838	

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

Table 2.5 Baseline model by grade

	Fourth Grade		Fifth Grade	
	Math	Reading	Math	Reading
Peer Effect	0.070**	0.054**	0.105**	0.047**
	(0.013)	(0.013)	(0.013)	(0.013)
Own Effect:				
lag test score	0.701**	0.677**	0.761**	0.644**
	(0.003)	(0.004)	(0.003)	(0.004)
female	-0.029**	0.043**	0.007	0.066**
	(0.005)	(0.007)	(0.005)	(0.006)
ELL	-0.116**	-0.137**	-0.069**	-0.254**
	(0.008)	(0.010)	(0.008)	(0.011)
Asian/other	0.172**	0.172**	0.182**	0.144**
	(0.007)	(0.009)	(0.007)	(0.008)
black	-0.026**	-0.040**	-0.013	-0.047**
	(0.008)	(0.009)	(0.007)	(0.009)
white	0.081**	0.093**	0.071**	0.083**
	(0.008)	(0.010)	(0.008)	(0.009)
ever poor	-0.078**	-0.053**	-0.040**	-0.060**
	(0.007)	(0.009)	(0.007)	(0.008)
Contextual Effect:				
lag test score	0.010	0.031	-0.025	0.067**
	(0.015)	(0.017)	(0.015)	(0.016)
female	0.001	-0.042**	0.007	-0.021
	(0.013)	(0.016)	(0.013)	(0.015)
ELL	0.028	0.047	0.011	0.046
	(0.029)	(0.038)	(0.029)	(0.039)
Asian/other	0.111**	0.115**	0.129**	0.052
	(0.027)	(0.032)	(0.026)	(0.029)
black	0.000	-0.029	-0.016	-0.052
	(0.029)	(0.034)	(0.027)	(0.032)
white	0.050	0.030	0.038	-0.026
	(0.031)	(0.036)	(0.029)	(0.034)
ever poor	-0.024	0.024	0.006	-0.029
	(0.027)	(0.032)	(0.027)	(0.032)
observations:	48,862	46,015	50,956	48,576

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

Table 2.6 Placebo

	Math		Reading	
	parameter	s.e.	parameter	s.e.
Peer Effect	0.004	0.032	0.016	0.032
Own Effect:				
lag test score	0.734**	0.003	0.661**	0.003
male	-0.012**	0.003	0.053**	0.004
ELL	-0.094**	0.007	-0.193**	0.009
Asian/other	0.183**	0.006	0.158**	0.007
black	-0.018**	0.006	-0.048**	0.007
white	0.074**	0.006	0.087**	0.008
ever poor	-0.056**	0.006	-0.056**	0.007
Contextual Effect:				
lag test score	0.075	0.038	0.029	0.041
male	-0.020	0.042	0.008	0.050
ELL	-0.011	0.083	0.045	0.105
Asian/other	0.031	0.071	-0.055	0.082
black	0.010	0.076	-0.132	0.087
white	-0.068	0.076	-0.034	0.089
ever poor	0.061	0.068	0.026	0.079
observations:	100,156		94,838	

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

Table 2.7 Removing classrooms with little variance in observed line order

	Cutoff is 0.05		Cutoff is 0.1		Cutoff is 0.15	
	Math	Reading	Math	Reading	Math	Reading
Peer Effect	0.090**	0.053**	0.096**	0.056**	0.100**	0.061**
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
Own Effect:						
lag test score	0.731**	0.660**	0.730**	0.660**	0.730**	0.661**
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
female	-0.012**	0.055**	-0.011**	0.053**	-0.012**	0.052**
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
ELL	-0.093**	-0.193**	-0.093**	-0.192**	-0.093**	-0.190**
	(0.006)	(0.007)	(0.006)	(0.008)	(0.006)	(0.008)
Asian/other	0.176**	0.157**	0.176**	0.157**	0.176**	0.156**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
black	-0.018**	-0.043**	-0.018**	-0.042**	-0.017**	-0.041**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
white	0.075**	0.087**	0.075**	0.087**	0.075**	0.087**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
ever poor	-0.058**	-0.056**	-0.057**	-0.054**	-0.057**	-0.053**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
Contextual Effect:						
lag test score	-0.008	0.048**	-0.009	0.058**	-0.005	0.068**
	(0.011)	(0.012)	(0.011)	(0.012)	(0.012)	(0.013)
female	0.009	-0.030**	0.009	-0.027*	0.009	-0.030*
	(0.009)	(0.011)	(0.010)	(0.012)	(0.010)	(0.012)
ELL	0.021	0.045	0.025	0.042	0.023	0.049
	(0.021)	(0.028)	(0.022)	(0.029)	(0.024)	(0.031)
Asian/other	0.129**	0.090**	0.131**	0.099**	0.140**	0.102**
	(0.019)	(0.022)	(0.020)	(0.023)	(0.021)	(0.024)
black	-0.009	-0.047	0.009	-0.053*	0.002	-0.056*
	(0.020)	(0.024)	(0.021)	(0.025)	(0.023)	(0.026)
white	0.047*	0.006	0.053*	0.012	0.061**	0.012
	(0.021)	(0.025)	(0.022)	(0.026)	(0.023)	(0.027)
ever poor	-0.007	-0.001	-0.008	0.003	-0.016	0.007
	(0.019)	(0.023)	(0.020)	(0.023)	(0.021)	(0.024)
observations:	99,275	94,022	97,029	91,920	92,962	88,121

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

Table 2.8 Multiple Distance levels

	Distance=2		Distance=3		Distance=4		Distance=5		Distance=6	
	Math	Reading	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Peer Effect:										
D=1	0.077**	0.038**	0.076**	0.039**	0.076**	0.038**	0.076**	0.038**	0.075**	0.039**
	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)
D=2	0.013	0.019	0.012	0.019	0.013	0.018	0.013	0.019	0.014	0.019
	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)
D=3			0.012	0.000	0.015	0.001	0.015	0.000	0.014	-0.001
			(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
D=4					-0.017	-0.008	-0.016	-0.01	-0.015	-0.01
					(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
D=5							-0.002	0.006	-0.002	0.005
							(0.011)	(0.011)	(0.011)	(0.011)
D=6									-0.004	0.01
									(0.011)	(0.011)
Own Effect:										
lag test score	0.731**	0.661**	0.731**	0.660**	0.731**	0.661**	0.731**	0.661**	0.731**	0.660**
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
female	-0.009*	0.056**	-0.009*	0.055**	-0.009*	0.054**	-0.008*	0.055**	-0.008*	0.055**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
ELL	-0.093**	-0.194**	-0.092**	-0.197**	-0.090**	-0.200**	-0.091**	-0.201**	-0.091**	-0.203**
	(0.006)	(0.007)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Asian/other	0.177**	0.158**	0.176**	0.156**	0.175**	0.156**	0.174**	0.157**	0.173**	0.158**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
black	-0.020**	-0.044**	-0.018**	-0.043**	-0.017**	-0.040**	-0.015*	-0.040**	-0.014*	-0.040**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
white	0.076**	0.089**	0.076**	0.088**	0.075**	0.088**	0.075**	0.088**	0.075**	0.088**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
ever poor	-0.057**	-0.055**	-0.057**	-0.057**	-0.056**	-0.058**	-0.056**	-0.059**	-0.057**	-0.058**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
observations	100,156	94,838	100,156	94,838	100,156	94,838	100,156	94,838	100,156	94,838

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

Table 2.9 Distances greater than one

	Distance=2		Distance=3		Distance=4		Distance=5		Distance=6	
	Math	Reading	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Peer Effect	0.092**	0.060**	0.102**	0.065**	0.101**	0.064**	0.094**	0.067**	0.091**	0.073**
	(0.011)	(0.011)	(0.014)	(0.014)	(0.016)	(0.016)	(0.018)	(0.018)	(0.019)	(0.020)
Own Effect:										
lag test score	0.732**	0.661**	0.732**	0.662**	0.733**	0.662**	0.733**	0.662**	0.733**	0.662**
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
male	-0.010*	0.055**	-0.010**	0.056**	-0.010**	0.056**	-0.011**	0.054**	-0.011**	0.054**
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
ELL	-0.093**	-0.193**	-0.093**	-0.195**	-0.092**	-0.197**	-0.092**	-0.198**	-0.092**	-0.200**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Asian/other	0.180**	0.159**	0.182**	0.160**	0.182**	0.161**	0.183**	0.161**	0.183**	0.162**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
black	-0.019**	-0.044**	-0.019**	-0.044**	-0.018**	-0.043**	-0.017**	-0.042**	-0.017**	-0.042**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.007)
white	0.077**	0.088**	0.078**	0.089**	0.078**	0.089**	0.079**	0.089**	0.078**	0.090**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
ever poor	-0.058**	-0.056**	-0.057**	-0.057**	-0.057**	-0.058**	-0.057**	-0.058**	-0.057**	-0.058**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
Contextual Effect:										
lag test score	-0.001	0.050**	-0.004	0.038*	0.006	0.040*	0.013	0.041	0.009	0.031
	(0.013)	(0.014)	(0.016)	(0.018)	(0.018)	(0.020)	(0.021)	(0.022)	(0.023)	(0.024)
male	0.000	-0.038**	0.003	-0.047**	0.003	-0.056**	0.006	-0.056**	0.005	-0.054**
	(0.011)	(0.013)	(0.012)	(0.015)	(0.014)	(0.016)	(0.015)	(0.018)	(0.017)	(0.020)
ELL	0.029	0.021	0.039	-0.026	0.071	-0.072	0.064	-0.093	0.055	-0.123
	(0.027)	(0.036)	(0.034)	(0.044)	(0.039)	(0.051)	(0.045)	(0.058)	(0.050)	(0.065)
Asian/other	0.102**	0.064*	0.101**	0.059	0.087**	0.058	0.074*	0.061	0.062	0.077
	(0.023)	(0.026)	(0.027)	(0.032)	(0.031)	(0.036)	(0.035)	(0.040)	(0.038)	(0.044)
black	-0.021	-0.062*	-0.007	-0.059	0.028	-0.024	0.054	-0.005	0.056	0.013
	(0.026)	(0.030)	(0.032)	(0.037)	(0.037)	(0.043)	(0.041)	(0.048)	(0.046)	(0.053)
white	0.041	0.019	0.059*	0.015	0.037	0.016	0.044	0.032	0.034	0.053
	(0.025)	(0.029)	(0.030)	(0.035)	(0.034)	(0.039)	(0.037)	(0.043)	(0.040)	(0.047)
ever poor	0.014	-0.001	0.034	-0.026	0.040	-0.033	0.045	-0.041	0.046	-0.042
	(0.022)	(0.026)	(0.026)	(0.031)	(0.030)	(0.035)	(0.033)	(0.038)	(0.035)	(0.041)
observations:	100,156	94,838	100,156	94,838	100,156	94,838	100,156	94,838	100,156	94,838

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

Table 2.10 Evolution of friendship importance over the school year

	Periods=2		Periods=3		Periods=4	
	Math	Reading	Math	Reading	Math	Reading
Peer Effect:						
Period 1	0.045**	0.030**	0.034**	0.032**	0.026**	0.023**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Period 2	0.038**	0.021*	0.028**	0.001	0.029**	0.006
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Period 3			0.027**	0.025**	0.024**	0.019*
			(0.009)	(0.009)	(0.009)	(0.009)
Period 4					0.01	0.008
					(0.008)	(0.008)
Own Effect:						
lag test score	0.730**	0.660**	0.730**	0.659**	0.730**	0.659**
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
female	-0.010*	0.055**	-0.010*	0.056**	-0.010*	0.056**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ELL	-0.093**	-0.193**	-0.093**	-0.193**	-0.093**	-0.192**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
asian/other	0.177**	0.157**	0.177**	0.156**	0.176**	0.156**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
black	-0.019**	-0.043**	-0.019**	-0.044**	-0.019**	-0.043**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
white	0.075**	0.087**	0.075**	0.087**	0.074**	0.086**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
ever poor	-0.058**	-0.055**	-0.057**	-0.056**	-0.056**	-0.056**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
observations	100,156	94,838	100,156	94,838	100,156	94,838

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

Table 2.11 Students must participate together more than one day

	At least 2 days		At least 3 days		At least 4 days	
	Math	Reading	Math	Reading	Math	Reading
Peer Effect	0.079**	0.048**	0.081**	0.049**	0.081**	0.049**
Own Effect:	(0.008)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)
lag test score	0.730**	0.660	0.730	0.660	0.730	0.660
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
female	-0.010*	0.055**	-0.010*	0.055**	-0.010*	0.055**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ELL	-0.093**	-0.193**	-0.093**	-0.193**	-0.093**	-0.193**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
black	-0.018**	-0.043**	-0.019**	-0.043**	-0.019**	-0.043**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
white	0.076**	0.087**	0.076**	0.087**	0.076**	0.087**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
Asian/other	0.177**	0.157**	0.177**	0.157**	0.177**	0.157**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
ever poor	-0.058**	-0.056**	-0.058**	-0.057**	-0.058**	-0.057**
Contextual Effect:	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
lag test score	0.000	0.041	-0.001	0.042	0.000	0.042
	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)
female	0.000	-0.027*	0.003	-0.027*	0.003	-0.029**
	(0.009)	(0.011)	(0.009)	(0.011)	(0.009)	(0.011)
ELL	0.025	0.033	0.021	0.033	0.019	0.032
	(0.020)	(0.027)	(0.020)	(0.027)	(0.020)	(0.027)
black	-0.007	-0.039	-0.009	-0.049*	-0.010	-0.051*
	(0.019)	(0.022)	(0.019)	(0.022)	(0.019)	(0.022)
white	0.031	-0.007	0.033	-0.026	0.013	-0.014
	(0.019)	(0.023)	(0.020)	(0.023)	(0.019)	(0.023)
Asian/other	0.096**	0.070**	0.093**	0.058**	0.088**	0.060**
	(0.018)	(0.020)	(0.018)	(0.020)	(0.018)	(0.020)
ever poor	0.009	-0.008	0.016	-0.008	0.012	-0.005
	(0.017)	(0.020)	(0.017)	(0.020)	(0.017)	(0.020)
observations:	100,156	94,838	100,156	94,838	100,156	94,838

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

Table 2.12 Students must participate together more than one day (continued)

	At least 5 days		At least 8 days		At least 10 days	
	Math	Reading	Math	Reading	Math	Reading
Peer Effect	0.080**	0.049**	0.083**	0.054**	0.079**	0.054**
Own Effect:	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
lag test score	0.730**	0.660	0.730	0.659	0.730	0.659
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
female	-0.011**	0.055**	-0.011**	0.056**	-0.011**	0.057**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ELL	-0.093**	-0.193**	-0.093**	-0.193**	-0.093**	-0.194**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
black	-0.019**	-0.043**	-0.019**	-0.043**	-0.019**	-0.043**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
white	0.076**	0.087**	0.076**	0.087**	0.076**	0.087**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
Asian/other	0.178**	0.157**	0.177**	0.157**	0.178**	0.157**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
ever poor	-0.058**	-0.056**	-0.058**	-0.057**	-0.058**	-0.057**
Contextual Effect:	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
lag test score	-0.004	0.041	-0.006	0.030	-0.006	0.028
	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)
female	0.004	-0.029**	0.004	-0.034**	0.004	-0.035**
	(0.009)	(0.011)	(0.009)	(0.011)	(0.009)	(0.011)
ELL	0.017	0.033	0.017	0.021	0.016	0.022
	(0.020)	(0.026)	(0.020)	(0.026)	(0.020)	(0.026)
black	-0.010	-0.047*	-0.006	-0.038	-0.009	-0.038
	(0.019)	(0.022)	(0.019)	(0.022)	(0.018)	(0.022)
white	0.015	-0.012	0.018	-0.001	0.018	0.004
	(0.019)	(0.023)	(0.019)	(0.022)	(0.019)	(0.022)
Asian/other	0.090**	0.066**	0.095**	0.076**	0.097**	0.081**
	(0.017)	(0.020)	(0.017)	(0.020)	(0.017)	(0.020)
ever poor	0.024	-0.003	0.022	-0.021	0.021	-0.014
	(0.017)	(0.020)	(0.017)	(0.020)	(0.017)	(0.020)
observations:	100,156	94,838	100,156	94,838	100,156	94,838

Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

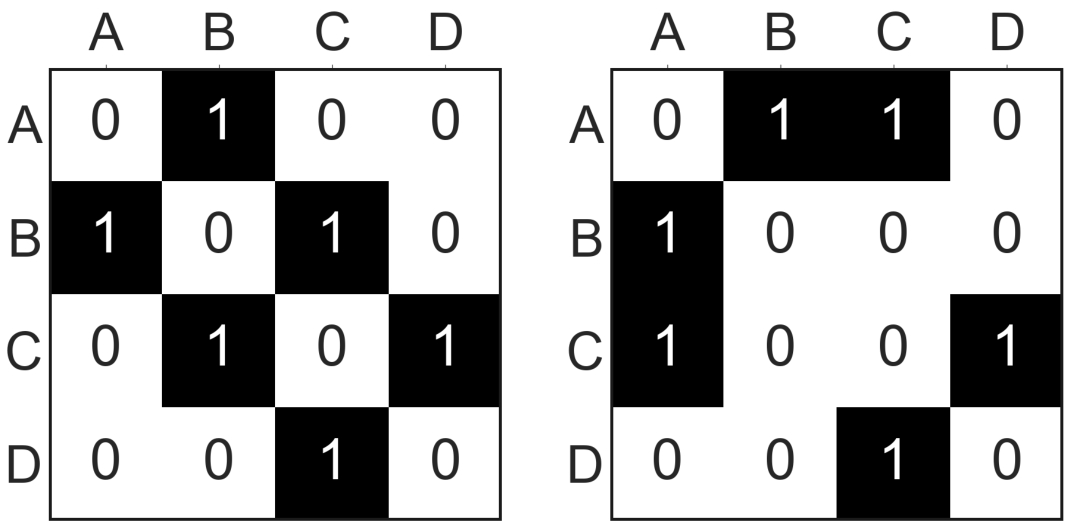
Table 2.13 Temporal proximity

	Math		Reading	
Peer Effect	0.025	(0.015)	0.046**	(0.015)
Own Effect:				
lag test score	0.731**	(0.002)	0.659**	(0.003)
female	-0.011**	(0.003)	0.051**	(0.004)
ELL	-0.094**	(0.006)	-0.196**	(0.007)
Asian/other	0.181**	(0.005)	0.161**	(0.006)
black	-0.019**	(0.005)	-0.042**	(0.006)
white	0.077**	(0.006)	0.088**	(0.007)
ever poor	-0.058**	(0.005)	-0.057**	(0.006)
Contextual Effect:				
lag test score	-0.007	(0.013)	-0.033*	(0.013)
female	-0.011	(0.011)	-0.034**	(0.013)
ELL	-0.023	(0.018)	-0.014	(0.025)
Asian/other	-0.029	(0.017)	0.019	(0.020)
black	-0.010	(0.018)	0.002	(0.021)
white	-0.024	(0.019)	0.003	(0.023)
ever poor	0.008	(0.017)	0.013	(0.020)
observations:	100,156		94,838	

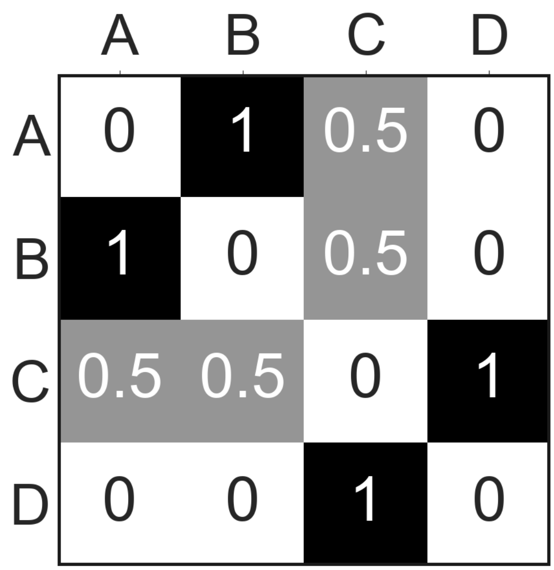
This table shows results using temporal proximity rather than physical (line order) proximity. Models include classroom fixed effects, own zip code fixed effects, and zip code exogenous fixed effects. Parameters with * are significant at the 5% level and ** at the 1% level.

2.10 Figures

Figure 2.1 Constructing Classroom Weighting Matrices: A Basic Example



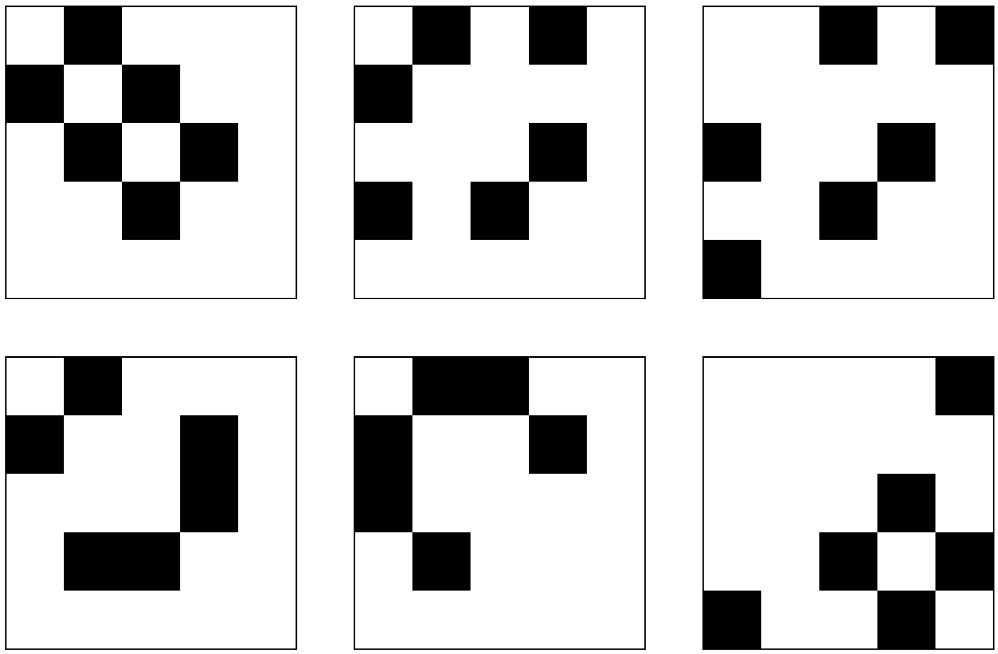
(a) Two example daily lunch line observations



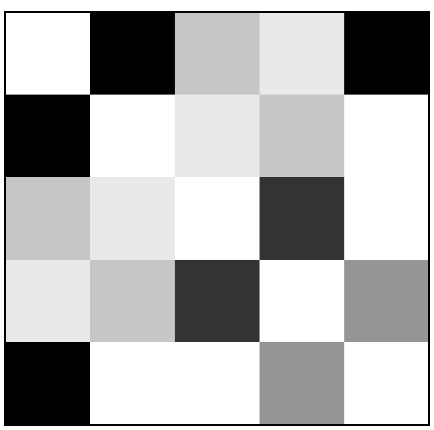
(b) Resulting weighting matrix

Figure 2.2 Example daily observations in (a) are constructed from example line orders [A,B,C,D] and [D,C,A,B]. (b) shows the resulting weighting matrix, constructed according to equation 2.1. The darker the square, the stronger the connection. A and B are constructed to have a strong connection, as are C and D. Notice that the "incidental" connection between the other students results in a weaker connection. As the number of days increases, the strength of these incidental connections becomes small relative to intentional connections as students choose to stand in line next to their friends.

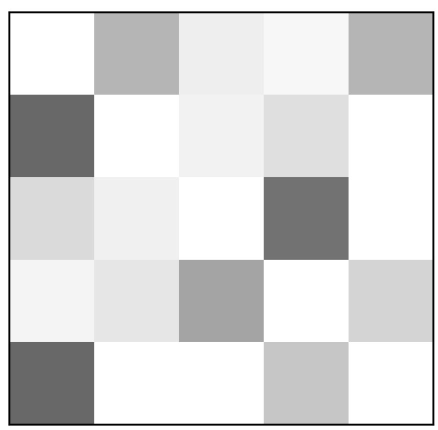
Figure 2.3 Constructing Proximity and Weighting Matrices



(a) Six example daily lunch line observations



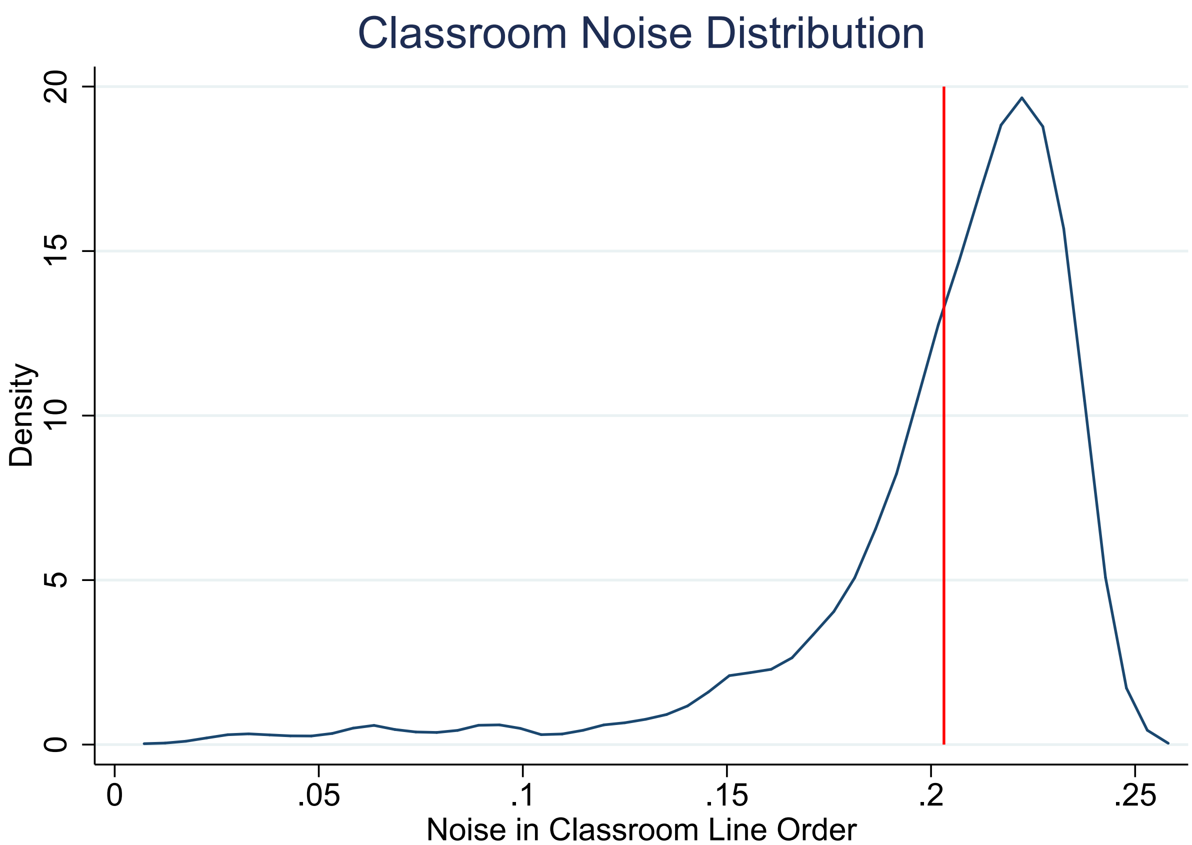
(b) Resulting proximity matrix



(c) Resulting weighting matrix

Example daily observations in (a) are constructed from example line orders [A,B,C,D]. [C,D,A,B], [D,C,A,E], [C,D,B,A], [D,B,A,C], and [A,E,D,C]. Notice that each day includes one absence, so that we can observe how this affects our measure of connection. (b) Shows the resulting proximity matrix, which is constructed according to equation 2.1 and (c) shows the result when we row-normalize this into a weighting matrix. Notice that students C and D are constructed to have a strong connection (near one another 5 of 6 days, and students A and B are constructed to have a weaker but still strong connection. Student E is next to A on the rare occasion E is present. We see these connections bear out in both the raw and normalized weighting matrices.

Figure 2.4 Classroom Noise Distribution



Vertical line indicates mean, which is equal to 0.2031616. Includes one entry for each of 4,077 classrooms. Density constructed using an Epanechnikov kernel with bandwidth = 0.004.

2.11 Appendix

2.11.1 Estimation procedures

We estimate a spatial autoregressive (SAR) model of the form:

$$Y = \lambda WY + \theta WX + X\beta + U \tag{2.5}$$

Where X contains a constant and the fixed effects for simplicity of notation. We estimate the model using maximum likelihood estimation (MLE), and so assume U is $\text{iid}(0, \sigma^2)$. Note that if we do not assume normality of the error term, this becomes quasi-maximum likelihood esetimation (QMLE).

In order to maintain the interdependencies of the error terms and incorporate classroom fixed effects, we follow the transformation approach discussed in L.-f. Lee and Yu (2010) and used in W. Horrace et al. (2019). This method involves the deviation from the classroom mean operator $Q = \iota\iota'/n$ an $n \times n$ matrix where n is classroom size. We define the orthonormal within transformation matrix Q as $[P, \iota_n/\sqrt{n}]$. Following L.-f. Lee and Yu (2010), we premultiply our model by P':

$$P'Y = \lambda P'WY + P'\theta WX + P'X\beta + P'U \tag{2.6}$$

This means our log likelihood function takes the form:

$$ln\mathscr{L}(\lambda, \beta, \sigma^2) = -\Big(\frac{n-1}{2}\Big)\Big[ln(2\pi) + ln(\sigma^2)\Big] + ln|I - \lambda P'WP| - \frac{\bar{e}'Q\bar{e}}{2\sigma^2} \tag{2.7}$$

Where $\bar{e} = P'Y - \lambda P'WQY - P'WQX\theta - P'X\beta$ After some algebra, we can rewrite this with only Q (and not P):

$$ln\mathscr{L}(\lambda, \beta, \sigma^2) = -\Big(\frac{n-1}{2}\Big)\Big[ln(2\pi) + ln(\sigma^2)\Big] - ln(1-\lambda) + ln|I - \lambda W| - \frac{e(\lambda, \xi)'Qe(\lambda, \xi)}{2\sigma^2} \tag{2.8}$$

where $\xi = (\theta, \beta)'$, $\mu = (WX, X)$ and $e(\xi) = Y - \lambda WY - WX\theta - X\beta = Y - \lambda WY - \mu\xi$. Notice that the parameter space for λ must be restricted such that its magnitude is less than one in order to guarantee that both $|I - \lambda W|$ will be strictly positive and $\ln(1-\lambda)$ is well defined.

We simplify estimation by concentrating out the ξ and σ^2 using first order conditions. Thus:

$$\xi^{\star}(\lambda) = (\mu' Q \mu)^{-1} \mu' Q (Y - \lambda W Y) \tag{2.9}$$

and

$$\sigma^{2\star}(\lambda, \xi) = \frac{e'(\lambda, \xi) Q e(\lambda, \xi)}{n-1} \tag{2.10}$$

This simplifies estimation substantially, as we need only maximize in one dimension. We get some cancellation from the $\sigma^{2\star}$ and our likelihood function for an individual class becomes:

$$ln\mathscr{L}(\lambda) = -\Big(\frac{n-1}{2}\Big)\Big[ln(2\pi) + 1 + ln[\sigma^{2\star}(\lambda)]\Big] - ln(1-\lambda) + ln|I - \lambda W| \tag{2.11}$$

We sum the likelihoods over all classrooms to obtain the complete likelihood, analogous to the way L.-f. Lee and Yu (2010) sum over the time periods.

In order to calculate the standard errors, we again follow L.-f. Lee and Yu (2010) and estimate the asymptotic variance matrix V_{ML} as in the block matrix below:

$$V_{ML} = \begin{pmatrix} a & d & e \\ d & b & 0 \\ e & 0 & c \end{pmatrix}^{-1} \tag{2.12}$$

Such that:

$$\begin{aligned}
a &= \frac{\partial \ln \mathscr{L}(\lambda, \xi, \sigma^2)}{\partial \lambda_k \partial \lambda_l} = (W_k G)' Q W_l G / \sigma^2 + \mathrm{tr}[W_k G Q W_l G] \\
a &= \frac{\partial \ln \mathscr{L}(\lambda, \xi, \sigma^2)}{\partial \xi^2} = \mu' Q \mu / \sigma^2 \\
c &= \frac{\partial \ln \mathscr{L}(\lambda, \xi, \sigma^2)}{\partial \sigma^4} = (n-1)/(2\sigma^4) \\
d &= \frac{\partial \ln \mathscr{L}(\lambda, \xi, \sigma^2)}{\partial \lambda_k \partial \xi} = (W_k G)' Q \mu \sigma^2 \\
e &= \frac{\partial \ln \mathscr{L}(\lambda, \xi, \sigma^2)}{\partial \lambda_k \partial \sigma^2} = \mathrm{tr}[Q W_k G] / \sigma^2
\end{aligned} \tag{2.13}$$

where $G = (I - \sum_k \lambda_k W_k)^{-1}$. The standard errors are then the square roots of the diagonal of V_{ML}.

2.11.2 Measuring Order Noise

When analyzing a social network based off the observed lunch line order, we may be concerned that students do not have agency over their place in line. If the students are ordered by some external factor, such as a teacher, then the interpretation of our results changes. As such, we attempt to determine whether there is a large set of classrooms in which students are ordered. The measure we intend to create will be able to determine whether a consistent order is used throughout the period of observation. If a teacher orders students alphabetically (for example) for lunch every day, we will detect this line order as having little noise.

In determining a good measure of noise, the measure must have two specific characteristics. First, the measure needs to be invariant to classroom size so that we can compare noise across classrooms without concern that the driving factor is number of students. The second issue is that students do not participate every day, so the measure must be able to contend with varying student combinations and line sizes. Thus any bias in our measure cannot be a function of class size or lunch participation rate. For the purpose of developing intuition, we discuss first an intuitive measure that does not meet these criteria, and then its relationship to a measure that does.

Consider every pair of students (i, j) within a classroom. For these students, we define two quantities A_{ij} and B_{ij}. Let $r(i)$ be the rank of student i and $\mathbb{1}_d(i, j)$ be an indicator function for both students i and j being present at lunch on day d. Then $A_{ij} = \sum_D \mathbb{1}_d\big(r(i) < r(j)\big)$ and $B_{ij} = \mathbb{1}_d(i, j)$. These quantities allow us to determine the number of switches $C_{ij} = \min\big(A_{ij}, B_{ij} - A_{ij}\big)$ if we assume that the most common order is the "true" order of the students. The quantity $S = \sum_{i<j} C_{ij}$ gives the total number of inversions, and we normalize this by the number of observed pairs $B = \sum_{i<j} B_{ij}$, so that our noise measure is $M = S/B$. This has a nice interpretation as the chance that a given pair is swapped. However the measure does not quite have the properties we would like - the measure varies by size and participation rate.

To understand the issue, we look at the bias of our measure. For all pairs (i, j), there exists some probability q of swapped order, conditioning on the appearance of both students (i,j). This results in the expectation of the total number of times i and j swap order equal to $q \cdot B_{ij}$. We consider the expectation of our estimator: $\mathbb{E}[C_{ij}] = \sum_{k=0}^{B_{ij}} \binom{B_{ij}}{k} q^k (1-q)^{(B_{ij}-k)} \min(k, B_{ij}-k)$. This expectation varies with B_{ij}. When $B_{ij} \in \{0,1\}$, $\mathbb{E}[C_{ij}] = 0$; when $B_{ij} \in \{2,3\}$, $\mathbb{E}[C_{ij}] = B_{ij} q(1-q)$; and more complex objects as B_{ij} increases. We look for a $\tilde{C}_{ij}$ where $\mathbb{E}[\tilde{C}_{ij}] = B_{ij} q(1-q)$ regardless of B_{ij} (being invariant in B_{ij} should meet the requirements of invariance to class size and participation rate). To do this, we replace $\min(k, B_{ij}-k)$ with $\binom{B_{ij}-1}{k}^{-1} \binom{B_{ij}-2}{k-1} = \frac{k(B_{ij}-k)}{B_{ij}-1} = \varphi$. What is nice about φ is that it keeps much of the meaning of $\min(k, B_{ij}-k)$. Without loss of generality, we can say that $\min(k, B_{ij}-k) = k$. Notice that in both measures, $C_{ij=0}$ when $k=0$, so we consider only $k > 0$. Then $1 \le k \le \frac{B_{ij}}{2}$. Thus $\frac{B_{ij}}{2} \le B_{ij} - k \le B_{ij} - 1$. This implies $\frac{k}{2} < \frac{kB_{ij}}{2(B_{ij}-1)} \le \frac{k(B_{ij}-k)}{B_{ij}-1} = \varphi \le k$, and we see that φ is bounded by k and $\frac{k}{2}$, although it loses the nice interpretation of our estimator being the chance i and j are swapped. We do however gain invariance by size and absences, which we will show when we finish constructing the measure. As before, we normalize this by dividing by the number of observed pairs. The result is:

$$M = \frac{\sum_{i<j;B_{ij}>1} \tilde{C}_{ij}}{\sum_{i<j;B_{ij}>1} B_{ij}} \quad \text{where} \quad \tilde{C}_{ij} = \frac{A_{ij}(B_{ij}-A_{ij})}{B_{ij}-1} \tag{2.14}$$

The omission we are left with is the case for $B_{ij} = 1$. Given probability p that a pair of students participate in lunch, the expectation that the students participate in lunch together only one time is $\mathbb{E}[B_{ij} = 1] = Dp(1-p)^{(D-1)}$. Average lunch participation is 65.8% over a school year of 180 days.[19] If two students participated 25% of the time (such that $p = .0625$), the chance of $B_{ij} = 1$ is less than e^{-9} if the participation of students i and j is independent. Given such a small chance, we ignore the scenario $B_{ij} = 1$ and forcibly remove such occurrences from the measure.

We can see that our measure is invariant to class size and participation rate in Table 2.14, which reports results of Monte-Carlo simulations under changes in class size and participation rate. Line orders are generated randomly for each of 180 days to simulate observation throughout the

19 The school year is required to be at least 180 days, and is sometimes longer than this. In 2013, the school year was exactly 180 days.

school year. Standard errors decrease in participation rate.

The measure is meant to detect noise, so we also simulate increases in randomness to show that the measure works as promised. Table 2.15 reports results of Monte-Carlo simulations on the measure of classroom noise in response to increasing levels of randomness. Line orders are generated randomly for X% of the 180 days, where X is in the percent random column. In all of these simulations, students have a 70% chance of participation in the line on any simulated day. The measure increases as randomness increases. We also show what may be apparent from the previous discussion, which is that when students are perfectly ordered in the classroom, the measure is zero. The average measure observed in the data is 0.203, which is consistent with between 50% and 60% randomness in the lunch line. This makes sense, as we expect variation in the line order, but as students reveal their preferences for line location and who to be in line with, we expect the sorting to be less than random. It is likely that classroom geography also plays a part in which groups of students are most likely in the front of the line on a consistent basis, further reducing the number of inversions detected by the measure. We expect that the lower level of randomness is not restricted to specific days of the year, as in the simulations, but rather each day has non-random variation.

It is important to note that this measure will be limited if the teacher attempts to be more equitable and alternates (for example) lining their students up alphabetically one day and reverse alphabetically another day, we would not detect this as an ordering (because there will be many line switches even without large changes in relative position). While limited in this way, we argue that the majority of orderings we might be concerned with (ex: alphabetical, height, location with the classroom, etc) will be detected by this measure.

2.11.3 Classroom Assignment

There are two forms of selection that are important to address. The first is the assignment of students to their set of potential peers. This occurs through two channels: assignment of students to school and then to classrooms. The second is the selection of friends within the classroom.[20] In this section I provide evidence that, conditional on school attended, student assignment into classrooms is consistent with randomness over most observed characteristics.

A key assumption for our estimates to be causal is that that assignment of students to their choice of peers (classroom assignment) is random. We have no insight into the assignment process, but we do show that over most observed characteristics, assignment is consistent with randomness. The objective of this test is to show that classroom assignment is not a function of the observed characteristics along which sorting into friendships might occur. To do this, we consider a series of multinomial logits as follows:

$$Class_i = \alpha + X_i \beta_{gst} + \varepsilon_i \tag{2.15}$$

For each iteration of equation (2.15) we include a single grade g within a single school s, during a single year t. We exclude all school-grades for which there is only a single classroom, as these schools by definition assign their students to classrooms randomly (less than 5% of our sample are in cohorts with only one classroom). $Class_i$ indicates the classroom assignment for student i, and the number of options varies by school-grade.[21] X_i is a binary indicator variable for a characteristic of student i. Each iteration of equation (2.15) gives us an estimate β_{gst} and a t-statistic. The t-statistic tells us the significance of the characteristic for assignment at that school-grade, and we collect the t-statistic for all school-grades. We then conduct our own random assignment

20 In the current version, this within-classroom friendship selection is dealt with primarily through group fixed effects. Homophily plays an important role in who students select as their friends, and the models used in this paper include a large set of demographic characteristics to control for these sorting avenues - including gender, ethnicity, residential zip code, and others. This is in line with other literature which uses fixed effects for networks of importance to control for these sorting effects. However, I include additional information in my measure of connection strength. To the extent that this additional information is the result of sorting which is not controlled for by these avenues, further work needs to be done. Future versions of this paper will include a more thorough examination of this within-classroom sorting.

21 There are between 2 and 11 classrooms in a school-grade-year. The mean is 4.2 classrooms.

of students to classrooms and run the same set of models, again collecting these t-statistics. We then compare the distributions of t-statistics from the observed and simulated models.

Figures 2.5 and 2.6 show the results of these tests. Each dot in these figures compares equal ranked t-statistics from the simulated and observed populations. If these distributions are the same, we should expect a 45 degree line. Most of the observed characteristics remain reasonably close to the 45 degree line, with the largest deviation at the tails of the distribution. A notable exception is the English language learner characteristic, which appears to deviate significantly from the 45 degree line. This indicates that classroom assignment may group English language learners into classrooms together. It is important to note that this test behaves best when the group sizes are similar in size, such as when we compare female and male students. English language learners make up only 12.5% of the population. That said, this is similar (slightly smaller than) the size of both white and Asian/other students in our sample, and both of these groups appear to behave better in this visual test. Thus, with the exception of English language learners, we conclude that we do not need to be concerned about selection into classrooms based on observed characteristics.

2.11.4 Appendix Tables and Figures

Table 2.14 Monte Carlo simulations of classroom size and participation rates

Participation rate:	Class size 20		Class size 25		Class size 30	
25%	0.2501	(0.0027)	0.25	(0.0023)	0.2501	(0.0019)
50%	0.25	(0.0008)	0.25	(0.0007)	0.25	(0.0006)
75%	0.25	(0.0005)	0.25	(0.0004)	0.25	(0.0004)
100%	0.25	(0.0003)	0.25	(0.0003)	0.25	(0.0003)

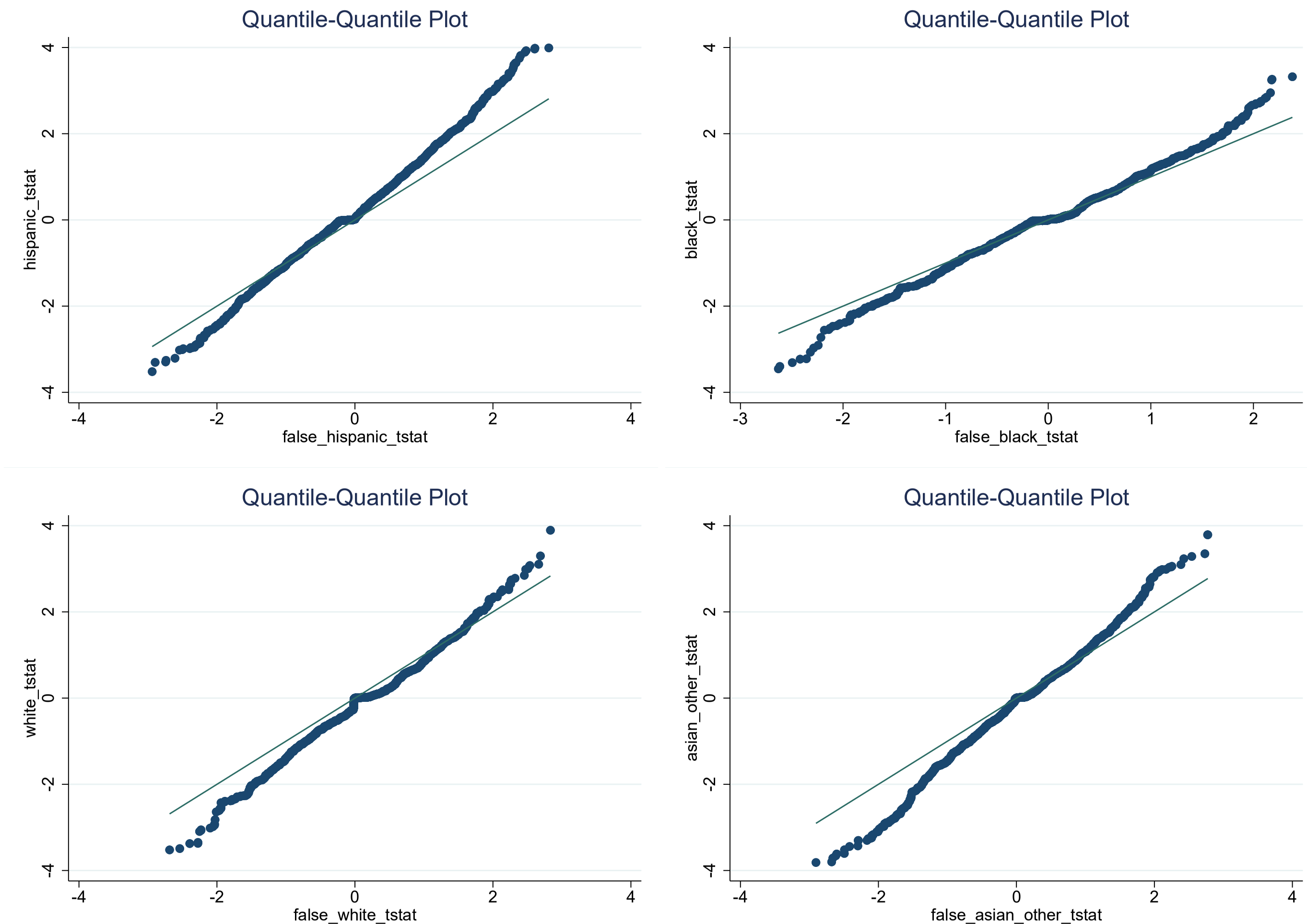
Table reports the results of Monte-Carlo simulations on the measure of classroom noise. Simulation is for 1,000 classrooms at each combination of class size and participation rate. Standard errors are in parenthesis. Line orders are generated randomly for each of 180 days to simulate observation throughout the school year. The measure is invariant to class size and participation rate, although standard errors increase as participation rate decreases.

Table 2.15 Monte Carlo simulations for different levels of randomness

Percent random:	mean	s.e.
0%	0.0000	(0.0000)
10%	0.0489	(0.0033)
20%	0.0902	(0.0040)
30%	0.1287	(0.0042)
40%	0.1603	(0.0039)
50%	0.1886	(0.0034)
60%	0.2104	(0.0029)
70%	0.2282	(0.0022)
80%	0.2402	(0.0015)
90%	0.2478	(0.0008)
100%	0.2500	(0.0004)
N	10,000	

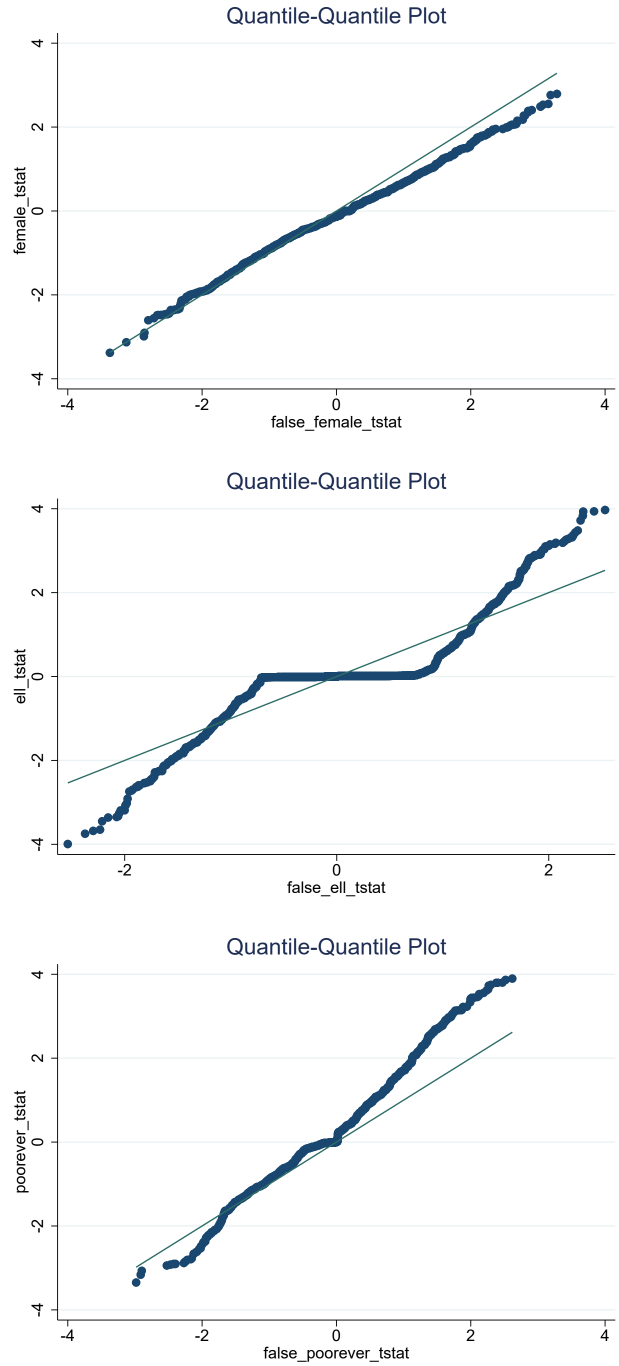
Table reports results of Monte-Carlo simulations on the measure of classroom noise. Simulation is for 10,000 classrooms at each combination of class size and participation rate. Line orders are generated randomly for X% of the 180 days, where X is in the percent random column. Students have a 70% chance of participation in the line on any simulated day. The measure increases as randomness increases.

Figure 2.5 Quantile-Quantile Plots



A series of multinomial logits are run to estimate the importance of each ethincity indicator in class assignment. In each pair of graphs, the left plots the t-statistics from these against the t-statisitics from a similar exercise in which we randomly assign students to classrooms. Thus we are plotting these two distributions against one another, and if the distributions are the same, we should expect a straight line of slope one. We argue that these provide evidence that class assignment is consistent with a random process.

Figure 2.6 Quantile-Quantile Plots



A series of multinomial logits are run to estimate the importance of gender, whether students are English language learners, or whether the student rides is poor. In each pair of graphs, the left plots the t-statistics from these against the t-statisitics from a similar exercise in which we randomly assign students to classrooms. Thus we are plotting these two distributions against one another, and if the distributions are the same, we should expect a straight line of slope one. We argue that these provide evidence that class assignment is consistent with a random process for most observed characteristics. This test performs best when groups are large (ex: females are about half the student population) and is more noisy when the network is small (ex: English language learners are only about 12.5% of the student population).

Chapter Three

Obesity Peer Effects in NYC Elementary Schools

Jonathan Presler

Syracuse University

3.1 Abstract

Obesity rates continue to climb in the US, as they have been doing for the past several decades. This phenomenon is evident in both adults and children, and the obesity epidemic has profound individual and public costs. Obesity during a child's early years are particularly problematic because obesity is difficult to reverse. This is a multi-faceted problem, and an area of growing literature focuses on the social contagion around obesity and overweight. Most of this work focuses on adolescents and college students - missing crucial early years of development. This paper focuses on obesity peer effects in New York City elementary school classrooms. Using reduced form models common in the peer effects literature, I find evidence of significant causal social effects both in BMI and exposure to overweight and obese students. Further exploration into the heterogeneity of these effects finds that elementary boys are more affected by peer weight than girls, and I find little difference by race.

Acknowledgements

Support for this project comes from the National Institutes of Health (R01HD070739 and R01DK108682-01). I would like to thank the NYC Department of Education's Policy Research and Support Group and the New York City Department of Health and Mental Hygiene for providing data. For data support, advice, and suggestions, I also thank Meryle Weinstein, Amy Ellen Schwartz, William C. Horrace, Alfonso Flores-Lagunes, and Joanna Bailey. I thank seminar participants at the Maxwell School at Syracuse University for useful comments on previous drafts. The opinions expressed are those of the author and do not represent the views of the NYC Department of Education or NIH.

3.2 Introduction

Obesity rates in the United States have been climbing steadily over the past several decades in both adults and children. This obesity epidemic has profound individual and public costs. One troubling component of this epidemic is the persistent increase in childhood obesity (Anderson, Butcher, and Schanzenbach, 2019, Flegal et al., 2010, and Hedley et al., 2004). Early years during childhood are particularly important because obesity is difficult to reverse. Overweight children as young as five years old have a twenty percent chance of becoming obese within the year, which will likely persist (Cunningham, Kramer, and Narayan, 2014). There are many contributing factors to the obesity epidemic, including changes in diet and physical exercise over time. One mechanism of interest is the social effects of weight gain among children, as social multipliers may enhance or mitigate policy interventions. Causal estimates of peer effects are notoriously difficult to obtain and studies looking at crucial early years are sparse. This paper addresses these issues by employing reduced form models common in the peer effects literature to test for the existence of weight and obesity social spillovers in New York City elementary school classrooms. The nature of these models allow me to explore heterogeneity and determine sub-populations with higher peer impact.

3.2.1 Problems of Obesity

Overweight and obesity cause a number of individual health complications and public externalities that make a reduction in these rates desirable. Childhood obesity is a strong predictor of adult obesity (Serdula et al., 1993, Must and Strauss, 1999, Kindblom et al., 2009, and Whitlock et al., 2005). Obesity in adults is linked with a number of health concerns including asthma (Sutherland, 2014), depression and other psychiatric disorders (Mustillo et al., 2003), type 2 diabetes, hyperlipidemia, hypertension, and heart disease (Thorpe et al., 2004 and Park et al., 2012). In children, being overweight has been linked to a number of immediate problems during childhood, including heightened risk of orthopedic, neurological, pulmonary, gasteroenterological, and endocrine conditions, joint problems, allergies (Must and Strauss, 1999, E. D. Taylor et al., 2006, and Halfon, Larson, and Slusser, 2013), type 2 diabetes (Ludwig and Ebbeling, 2001), depression (Morrison et al., 2015 and Halfon, Larson, and Slusser, 2013), and an overall decrease in quality of life (V. H. Taylor et al., 2013 and Schwimmer, Burwinkle, and Varni, 2003). Obesity in children leads to problems in adulthood including increased risk of type 2 diabetes, hypertension, heart disease, and mortality (Must and Strauss, 1999 and Park et al., 2012). Obese school children also are more likely to develop ADHD, learning disabilities, developmental delays, grade repetition, and missed school days (Halfon, Larson, and Slusser, 2013).

In addition to these health concerns, overweight individuals (particularly white females) may see lower wages than healthy-weight counterparts (Cawley, 2004 and Moro, Tello-Trillo, and Tempesti, 2018), as well as lower chance of being employed at all (Biener, Cawley, and Meyerhoefer, 2018).

Problems with overweight and obesity are not limited to private health or wage concerns. Health care costs in the US continue to rise, and a sizable piece of these expenditure increases can be attributed to continuing growth in overweight and obesity (Thorpe et al., 2004, Trogdon et al., 2012, and Biener, Cawley, and Meyerhoefer, 2018). These medical costs are borne both privately and publicly. Trogdon et al. (2012) find that between about a quarter and half of these costs are financed publicly through medicare and medicaid (variation by state). In addition, Cawley (2008) estimates that voters are interested in reducing childhood obesity and that the state of New York has a willingness to pay of approximately $690 million to reduce childhood obesity by half in the state (which is higher than current spending levels). Medical expenditures in the US are estimated to be between 6 and 11 percent lower in the absence of obesity (Trogdon et al., 2012 and Biener, Cawley, and Meyerhoefer, 2018), with an estimated total cost of obesity in the US of nearly $150 billion in 2008 (Finkelstein et al., 2009). In addition, Lakdawalla, Bhattacharya, and Goldman (2004) links obesity to higher levels of disability insurance, medicare, and medical costs.

3.2.2 Causes of Obesity

The obesity epidemic is a multi-faceted problem with many contributing factors. Lack of exercise (Troiano et al., 2000) and high levels of television watching (Gordon-Larsen, Adair, and Popkin, 2002) are contributing factors to the increase in childhood obesity. Additionally, childrens' diets have been changing, and the consumption of higher fat foods (Donahoo et al., 2008) and fast food consumption (Pereira et al., 2005) are associated with childhood overweight and obesity. Fast food advertising and availability appear to be contributing factors (Robinson et al., 2007).

Social factors may also be at play both in the causes of overweight and obesity, as well as in the ways that children deal with the condition. Lack of family and peer support may cause overweight adolescents to use unhealthy weight control measures (Vander Wal, 2012). To my knowledge, N. A. Christakis and J. H. Fowler (2007) are the first to explore the importance of social networks in spreading obesity, work that continued in Cohen-Cole and Jason M. Fletcher (2008b) and J. Fowler and N. Christakis (2008). These results provide evidence of correlations between peer weight. Additional work used random assignment of students to peer groups in college to explore peer effects on fitness and weight gain (Carrell, Hoekstra, and West, 2011 and Yakusheva, Kapinos, and Weiss, 2011). Several papers use the Add Health survey, finding strongest effects among female students (Renna, Grafova, and Thakur, 2008) and the strongest correlations for those most overweight (Halliday and Kwak, 2009). Asirvatham, Nayga Jr., and Thomsen (2014) specifically looks at elementary school students and finds evidence of role model effects where the weight of older students in the school influences weight gain among younger students.

Social networks may influence obesity through a number of vectors. Some of these may be contextual - proximity to individuals of lower (higher) BMI may shift student perspectives on what defines a normal weight, causing a healthy-weight student to feel overweight (underweight). Students with a minority of heavy or light classmates may be attracted to or repulsed by these students, and alter behavior accordingly (shining light or bad apple stories).[1] Other vectors may

1 See Hoxby and Weingarth (2005) for a discussion of common peer effects models

be more directly social, as student eating habits are shared, students learn from their peers which foods are socially acceptable, and students participate (or do not participate) in physical activities and sports with one another. This paper remains agnostic to the specific mechanisms at play, but will employ a linear model of peer effects as a baseline. The linear model will detect whether there is a net effect from a composite of these peer mechanisms.

3.3 Data and Sample Construction

3.3.1 Data

Data come from the New York City Department of Edcuation (NYCDOE) administrative database and the NYC Department of Health and Mental Hygiene Office of School Health. I use data from the 2018 academic year, with lag outcomes from 2017. I observe student characteristics and demographics including sex, ethnicity, grade, zip code, poverty indicator, and whether a student is an English language learner. In addition, I observe student's classroom assignment.

3.3.2 Body Mass Index

This paper uses Body Mass Index (BMI) as a measure of student overweight. BMI is defined by dividing an individual's weight (kilograms) by the square of their height (meters). BMI is relied upon in the medical literature for use in determining whether children are overweight (Whitlock et al., 2005). For adults there are specific benchmarks defining underweight, healthy weight, overweight, and three classes of obesity. In children, interpreting BMI is less straightforward and depends upon both age and sex. Figure 3.1 uses data from the CDC to illustrate how BMI changes with respect to age in elementary age children.

I use eight measures of body composition as outcomes to explore the social effects of obesity in elementary school classrooms. These eight measures make up two categories of variables. The first set of four variables are continuous measures of body fat composition. These include the body mass index (BMI), a BMI z-score normalized at the NYC level, a BMI z-score normalized at the national level, and a BMI percentile measure, normalized at the national level. The last three are calculated from BMI, but together they give a better picture of the effect. The second set of four variables are categorical variables denoting whether a student is underweight, obese, overweight, or severely obese. These categories are coded to be inclusive, such that an obese student is also overweight. The reference group is normal-weight students. I use the CDC's definitions of underweight, overweight, obese, and severely obese. The CDC defines a child as underweight if the child's BMI is below the fifth percentile. Healthy weight is defined between the fifth and eighty-fifth percentiles. Overweight is between the eighty-fifth and ninety-fifth percentile. Obese is above the ninety-fifth percentile. The CDC classifies a child as severely obese if their BMI is at or above 120% of the ninety-fifth percentile BMI. The measure of student BMI used in this paper comes from the NYC Fitnessgram.

3.3.3 Fitnessgram

The Fitnessgram is an annual health and fitness assessment designed for students grades K-12 in order to determine whether students display a health risk. This data is collected by the Office of School Health, and while the Fitnessgram measures five areas of student health, our focus in this paper is on body weight composition. Data reported are student height and weight, which we use to calculate student BMI. This height and weight assessment is collected for all students grade K-12 in NYC public schools. The date of test varies but is typically concentrated at the beginning of the year. I use the test date when calculating student age in order to correctly place students into weight categories according to CDC guidelines.

I also use several other outcomes calculated using student BMI. The first is a normalized z-score of BMI for students in New York City (NYC). I also include a z-score normalized to the CDC's national percentiles. Finally, I calculate each student's percentile, again based on the national numbers from the CDC.

3.3.4 Sample

My sample comes from the universe of general education students in the NYC public schools during the 2018 academic year. I further restrict my sample to students in classrooms with at least 20 students. The purpose is to eliminate the risk of including students in an integrated co-teaching (ICT) classroom, as ICT classrooms are structured differently[2] and include students with disabilities who may have different weight profiles than general education students. In order to avoid potential correlated effects due to this type of classroom, I remove these students from my sample.[3] I then restrict to students who have both a current year measure of BMI and a lag measure of BMI.

3.4 Methodology

There are a number of challenges that make identifying peer effects difficult, which have been well documented in the literature. Primary among these are the reflection problem and the selection of peers into their peer group. The reflection problem (Manski, 1993) is an identification issue that arises due to the feedback loop in which a student affects their peers, but is in turn affected by these same peers. In this application, reflection means that we cannot disentangle whether an overweight student causes their peers to be overweight or whether their peers caused the student

2 ICT classrooms combine general education students and students with disabilities where they are taught general education curriculum by a team of two teachers: one general education teacher and one special education teacher. ICT classrooms typically have a ratio of 40% students with disabilities and 60% general education students. Because my sample only includes general education students, a minimum class size of 20 students should remove this type of classroom.

3 Manski (1993) points out that correlated effects have the potential to contaminate estimates of peer effects. Correlated effects are when peers are exposed to the same treatment, so that their outcomes are correlated, but this is not due to any sort of social interaction or peer effect. For example, students living in a food desert may have higher BMI on average, but this is due to an unhealthy diet rather than social interaction.

to be overweight themselves. In order to mitigate the reflection problem, I use lagged peer outcomes rather than current year outcomes to measure classmate weight, as in B. Sacerdote (2001) and Carrell, Fullerton, and West (2009). Because the measures are not contemporaneous, this breaks the feedback loop by estimating the effect of classmate lag weight measures on current year weight measures. Thus students are explicitly affected by their classmates, but students cannot in turn affect their classmates.

If students select their peer group, then it becomes unclear whether correlations between student outcomes are due to peer effects or simply because students select similar peers. To avoid this issue altogether, much of the literature focuses on school-grade cohorts as the peer group (Hoxby, 2000, Carrell, Hoekstra, and Kuka, 2018). The argument is that while students select their school, they do not select their cohort. Thus the variation within a school and between years leaves variation that is arguably random. However, this is clearly not the optimal peer group. Elementary students interact primarily with classmates, not cohort-mates. Burke and Sass (2013) supports this intuition and show that the classroom peer effect is much more relevant. In this paper, I focus on classmates and use the variation between classrooms within cohorts. The key assumption for a causal interpretation of these results is that students are not assigned to classrooms based on measures of overweight and obesity. I provide evidence that this assumption is reasonable in Section 3.6.1.

In this paper, I ask two types of questions. In the first, I ask whether student BMI increases with classmate BMI.[4] The second question asks whether exposure to certain types of classmates (such as overweight classmates) affects student BMI. In both of these questions the classroom is the environment in which these social spillovers occur. The classroom is the space for which school administrators control the assignment process, making it the primary policy-relevant space for

4 I use several measures of body composition related to BMI, but for simplicity I refer to these collectively as BMI in this section. Specifically, these are BMI, a z-score BMI standardized to the NYC level, a z-score BMI standardized to the national level using the CDC's numbers, and student's national BMI percentile. These outcomes will be used to answer both types of questions. Additionally, I use categorical variables of student body weight (overweight, obese, severely obese, and underweight) as weight classification variables.

social spillovers within the school.

To measure the influence of classmates, I construct a leave-out-mean (LOM) of classmate weight measures. The LOM is an unweighted average of classmate BMI, leaving out the student's own contribution to the classroom average. This captures the variation of interest, but also includes signal from confounders. For example, schools with high rates of poverty might be composed of students with higher average BMI than schools with lower rates of poverty.

In order to isolate the variation between classrooms within cohorts, I follow previous work in the literature (Hoxby, 2000 and Carrell, Hoekstra, and Kuka, 2018). School and grade fixed effects isolate cohorts, eliminating correlated effects associated with student environment, cafeteria, and neighborhood while leaving variation from three sources. The first source of variation is the idiosyncratic variation between classrooms, which is the key variation of interest. This captures variation in exposure to student types and classmate BMI, but could also include correlated effects related to (for example) teacher effects or the assigned lunch period. Ideally, I would be able to add controls for classroom characteristics, or classroom fixed effects to control for unobserved characteristics. However, data limitations mean I cannot include classroom controls, and classroom fixed effects would eliminate the very source of variation in which we are interested. To address this, I assume that these classroom correlated effects are orthogonal to student body weight measures. This is very similar to the assumption that students are not assigned to classrooms based on measures of overweight and obesity, which I provide evidence for in Section 3.6.1.
The second and third sources of variation occur of within the classroom. The structure of a LOM implies variation due to the student for which it is calculated. The student with the highest BMI in the classroom will tautologically be exposed to a lower average BMI than the student's peers. The extent to which the classroom LOM changes for each student is a function of classroom size. Because I use lagged outcomes in the LOM, the two sources of within-classroom variation act as desired. Heavy students are exposed to lighter students, allowing me to explore this peer effect. Considering typical mechanisms behind the peer effect, the impact of each peer is also typically

a function of the group size. For example, more overweight classmates may normalize being overweight. This mechanism assumes some linearity in group size. Future work will explore non-linearities in group size.

I use two similar reduced form models to test for the existence of weight-related peer effects. The first is a leave-out-mean (LOM) model, which results in a simple regression as follows:

$$y_{icsg} = \beta \frac{\sum_{k \neq i} y_{kc-1}}{n_c t - 1} + \alpha_0 X_{isg} + \alpha_{1s} + \alpha_{2r} + \alpha_{3g} + \varepsilon_{iscg} \tag{3.1}$$

where i, k are students, s is school, c is classroom, n_c is number of students in the classroom, g is grade, r is residential zip code, β measures the effect of the classroom LOM, X_{icsg} are student characteristics including sex, ethnicity, lag outcome, age in months at the time of the Fitnessgram, square of age, and indicators for English language learner status, student with disabilities, and our poverty measure (whether a student ever qualified for free or reduced price lunch). α_{1s} is a school fixed effect, α_{2r} is a residential zip code fixed effect, α_{3g} is a grade fixed effect, and ε_{iscg} is an error term.

In addition to the LOM model, I am also interested in the effect of exposure to certain types of students in the classroom (Hoxby (2000) and Carrell, Hoekstra, and Kuka (2018)). Here I employ a very similar model, but instead of regressing on the LOM of the lagged variable, I regress on the proportion of other students in the classroom who fall into certain weight categories (overweight, obese, severely obese, and underweight) in order to see the effect on measures of student BMI. The idea here is that exposure to an obese student might have an effect on student weight, but it might not have enough effect to move normal weight students to obesity. The group of students that might be affected in Model 3.1 may be smaller than the group of students affected in the following model:

$$y_{icsg} = \beta \frac{\sum_{k \neq i} \delta_{kc-1}}{n_c t - 1} + \alpha_0 X_{isg} + \alpha_{1s} + \alpha_{2r} + \alpha_{3g} + \varepsilon_{iscg} \tag{3.2}$$

where δ_{kc-1} indicates whether student k in classroom c had the characteristic of interest in the previous year.

It is important to note that the purpose of these models is to test for the existence of peer effects in overweight and obesity. The reduced form models used here will not be able to distinguish between the endogenous peer effect and the contextual effect Manski, 1993. That said, both of these effects are social effects and of interest when attempting to detect the existence of obesity peer effects. Future work may attempt to disentangle these effects, which imply different policy prescriptions in order to harness or mitigate their effects.

3.5 Results

The baseline model uses the LOM construction from equation 3.1. Table 3.2 shows these results for our eight outcomes of interest. Column (1) shows the impact of average classmate lagged BMI on a student's own current BMI. The LOM coefficient is a significant 0.028, meaning that an increase in classmate BMI of one unit leads to own BMI increasing by 2.8% of a BMI unit. The standard devation of BMI in the sample is 4.03, meaning that a one standard deviation increase in peer BMI leads to over an 11% increase in own BMI. A student's lag weight measure is the best predictor of their current weight measure; I estimate the coefficient on the lag BMI measure is 1.010, which has the largest t-value of any estimate.[5] This means that current year BMI is slightly larger than the lag BMI, on average. When we compare student ethnicities, the baseline group is Hispanic students. Asian/other (-0.146) and white (-0.117) students have much lower BMI than Hispanic students, and this is significant at the one percent level. Black students have a lower BMI than Hispanic students by 0.121 BMI units, significant at the ten percent level. Female students have BMI that is 0.038 BMI units lower than their male counterparts. English Language Learners (ELL) have BMI that is 0.046 BMI units higher than those who are not ELL. BMI appears to change with grade in a non-linear fashion. The comparison group is first graders, and we see second graders have a higher BMI than first graders, third graders have an indistinguishable BMI

5 Models using only lagged weight measures as regressors show R-squared values nearly as high as the models presented in Table 3.2.

from first graders, and subsequent years have lower BMI with grade. However, this masks the true pattern because I also control for age (in months) of students in the model. Student age is important to allow comparison between students in the same grade, as the Fitnessgram measurements are taken at slightly different dates in the school year (although most are concentrated at the start of the school year). Age is quadratic, with a significant increase in the linear portion and a decrease in the squared term. Additionally, students with disabilities (SWD) are typically heavier than those without disabilities. Recall that the analysis is restricted to general education students, so the cause of this disability is typically physical and may reflect lower mobility and ability to exercise for SWD in the sample. The poverty measure used in this is whether a student ever meets the free or reduced price lunch (FRPL) requirements. The estimate indicates that poor students are 0.062 BMI units heavier than their less poor peers.

Column (2) shows the results of the same model using z-bmi normalized nationally, and we estimate an effect of 0.201. This means that an increase of one standard deviation (defined nationally) in weight increases own weight by 20% of a standard deviation. Notice that this estimate is larger than the estimate in column (3), which uses a z-score standardized within NYC. This is because the mean weight in NYC public schools is higher than the national average. Model (4) uses a students' national BMI percentile as the dependent variable. When exposed to a classroom with average BMI that is 10 percentile points above the mean, own BMI increases by 1.18 percentage points.

Columns (5) through (8) show the results of linear probability models for each of the four weight categories. Model (5) shows the result for the obese category. A student who moves from a class with no obese students to one in which every student is obese faces a 6.6% chance of becoming obese. Column (6) shows the result for overweight, and column (7) shows the result for severe obesity. Notice that the coefficient remains significant in each model, but the magnitude declines in severity. Column (6), the model exploring overweight, shows the largest effect at nearly 8%, and column (7) shows the smallest effect for severe obesity at under 2%. This is intuitive, as it takes a larger shock to move someone from a normal weight into more severe categories. Column (8) shows the results for the underweight outcome. Exposure to underweight students appears to have an effect that is much larger in magnitude. This is particularly interesting because underweight is the category with the fewest students. Going from a classroom with no underweight students to all underweight students leads to an increase of 18.3% chance that a student becomes underweight themselves.

All eight columns show positive and significant effects for the coefficient of interest, and most coefficients have the same sign in all models. It is not surprising that the coefficient of interest is positive in all models. This shows that student weight measures moves in the same direction as their peers. Poverty is always significant and leads to increases in weight over all eight models. Females are, on average, lighter than males.

Notice that the linear probability models for the categorical variables of interest (whether a student is obese, overweight, severely obese, or underweight) for the most part display smaller effects than the continuous measures of body mass. This is likely because even if there is a social effect, it is a more extreme case to observe students who are moved enough by their peers so that they end up in a different weight category. The exception is the underweight category, which has a significant effect on classmates also being underweight.

Because it is a significant jump between these categories for many students, I explore the effect of exposure to these four categories (obese, overweight, severely obese, and underweight) on the four continuous measures of weight. Even if exposure to obese students does not move a normal weight student to obesity, we may detect movement along a continuous measure of body weight composition. Tables 3.3, 3.4, 3.5, and 3.6 show the impacts of exposure to students in the four weight categories on continuous measures of body mass (respectively BMI, BMI percentile, Z-BMI normalized at the National level, and Z-BMI normalized at the NYC level). The story is reassuringly similar across all four outcomes. Notice that exposure to underweight students elicits a significantly stronger effect in all specifications. In Table 3.3 we see that as we move from exposure to overweight then obese and then to severely obese, the effect on own BMI only increases. This is consistent with a story of normalizing weight rather than distaste for higher weight.

Table 3.3, column (1) shows the effect of exposure to obese classmates on own BMI. The coefficient of interest is the proportion of overweight students in the classroom, which has a statistically significant estimate of 0.779. This means that going from a classroom with zero overweight students to a classroom composed entirely by overweight students leads to an increase in BMI of 0.78 units. Consider an average student has 26 classmates and a representative 10 of these students are overweight (over 38% of students in the sample are overweight). Halving the number of overweight classmates reduces own BMI by 0.15 units. This is nearly a 1% reduction in the student's BMI due to exposure to just 5 fewer overweight classmates in the average student's exposure to overweight students. Cawley (2008) estimates taxpayers have a willingness to pay of $690.6 million to reduce childhood obesity by half in the state of New York (a willingness to pay that exceeds current spending on childhood obesity reduction). Considering the costs involved in effectively reducing childhood obesity, these spillovers suggest a way to reduce these costs – as well as benefits that extend beyond the target children.

Column (2) of Table 3.3 is similar to model (1), but the coefficient of interest estimates the effect of the proportion of classmates who are obese. The point estimate is 0.867, statistically significant, and larger than the point estimate for overweight in column (1). This suggests that exposure to an obese student has a larger effect than exposure to an overweight student. That said, the net effect of obese students is less than overweight, as about 21% of students in the sample who are overweight. Column (3) repeats this process for severe obesity, and we again see an increase in the point estimate, suggesting that the relationship to heavy students is increasing in weight over most of the weight distribution. Column (4) presents the results for exposure to underweight students. The estimate shows a stronger effect than the other weight measures. Exposure to classmates who are entirely underweight reduces own BMI by 1.90 units, relative to exposure to classmates who are normal weight.

Tables 3.4, 3.5, and 3.6 are very similar to Table 3.3, but they report the effect of exposure to classmates in the four weight categories on BMI percentile, z-BMI normalized at the national level, and z-BMI normalized at the NYC level (respectively). The story in each of these sets of results remains the same as in Table 3.3.

3.5.1 Heterogeneous Effects

Table 3.7 interacts the baseline model from Table 3.2 and equation 3.1 with the sex of students. Female students are typically equal or lower in weight than their male peers. There seems to be little difference between boys and girls in the continuous measures, but girls are less likely than boys to be affected by their obese, overweight, and severely obese peers. Further analysis should be done to determine possible mechanisms. For example, it could be because obese, overweight, and severely obese classmates are typically of same or opposite sex. I detect no difference by gender in the effect of exposure to underweight peers.

Table 3.8 displays the results when I interact the baseline model with student race and ethnicity. Hispanic students are the reference group, as they are the modal student in New York City. Asian/other and White students are on average lower weight than Black and Hispanic students. Black students are of similar or slightly lower weight than Hispanic students. Under most specifications, there is little difference in body mass composition by ethnicity. The primary exception is that Asian/other and White students seem to be more affected by underweight peers.

3.5.2 Comparisons to the Literature

While this paper fills a gap in the literature by estimating causal obesity peer effects for classmates, it is still useful to compare these effects to other estimates in the literature. I compare with work exploring obesity spillovers among adolescents (using Add Health) and college students, as well as one paper which is, to my knowledge, the only other paper looking at obesity spillovers in elementary schools.

Asirvatham, Nayga Jr., and Thomsen (2014) explores at obesity spillovers in elementary schools. They find evidence of role model effects such that for every 20 highest grade students in the school who become obese, one younger grade student becomes obese. The mechanism is different than the classroom peer effects I estimate, but it is worth noting that the effects I estimate are comparable. Changing an average sized classroom (26 peers, see Table 3.1) from no obese students to twenty obese students leads to just over 5% increase in probability of obesity.

Carrell, Hoekstra, and West (2011) finds that one standard deviation increase in high school fitness scores (lag fitness score) increases college fitness by 16.5% of a standard deviation. This estimate is between the two z-scores used in my estimates (one normalized nationally and the other at the NYC level). The context is quite different (college and elementary school), but it suggests a mechanism through which obesity spillovers might occur - how students play during recess or how much effort students exert during physical education. Further work is necessary to understand the mechanisms behind the obesity spillovers estimated in this paper and the extent to which physical activity affects these spillovers.

Halliday and Kwak (2009) uses the Add Health data to estimate social effects of obesity among adolescent friends. They find that an increase of average friend's BMI by one unit is associated with an increase of own BMI of 0.19 units. Renna, Grafova, and Thakur (2008) also uses the Add Health data and find that an increase of 6 points in average friend BMI leads to an increase in own BMI of 1 point, and a slightly lower estimate for the BMI percentile. These are substantially higher than my own estimates, which I argue to be causal. This discrepancy is due to a combination of reasons. First, estimates using friendship networks are likely higher than the classroom estimates presented in this paper, as friends may better target which peers are influential (Presler, 2020). Second, it is difficult to obtain causal estimates using the Add Health network because of the endogenous network formation which is unobserved, and the inability to control for the correlated effects of students in the same friendship group. Finally, the context is again different, as I estimate effects for elementary students and Halliday and Kwak (2009) and Renna, Grafova, and Thakur (2008) estimate these effects for middle and high school adolescents. Additional evidence of the difference in context is evident when I estimate higher effects for males than females, but Halliday and Kwak (2009) estimates the opposite. Yakusheva, Kapinos, and Weiss (2011) supports the finding that female students are more affected than male students by their peers in a college environment. This suggests that there may be a point at which the sex that is more affected by their peers switches between elementary and late middle school and is a potential avenue for further exploration.

3.6 Robustness Checks

3.6.1 Random Assignment

In order to ascribe a causal interpretation to these results, one of the key assumptions presented in Section 3.4 is that classroom assignment is orthogonal to weight measures. The actual assignment process is a black box, but I show that classroom assignment is consistent with a random process relative to the four weight categories used in this paper (underweight, overweight, obese, and severely obese). To do this, I consider a series of multinomial logits as follows:

$$Class_i = \alpha + X_i\beta_{gst} + \varepsilon_i \tag{3.3}$$

For each iteration of equation (3.3) we include a single grade g within a single school s. I exclude all school-grades for which there is only a single classroom, as these schools by definition assign their students to classrooms randomly (less than 5% of our sample are in cohorts with only one classroom. $Class_i$ indicates the classroom assignment for student i, and the number of options varies by school-grade.[6] X_i is a binary indicator of weight category for student i. Each iteration of equation (3.3) gives an estimate β_{gst} and a t-statistic. The t-statistic represents the significance of the weight category in the classroom assignment process at that school-grade, and I collect the

6 There are between 2 and 11 classrooms in a school-grade-year. The mean is 4.2 classrooms.

t-statistic for all classroom assignments in all school-grades. I then simulate random assignment of students to classrooms within their same school-grade and run the same set of models, again collecting these t-statistics. I compare the distributions of t-statistics between the observed and simulated models.

Figure 3.2 shows the results of these comparisons. Each dot in these figures compares equal ranked t-statistics from simulated and observed populations. When these distributions are the same, the result is a 45 degree line. Each weight category remains very close to this 45 degree line. Thus, I conclude that classrooms assignment based on the four weight categories of interest is consistent with a random process, and the assumption that classroom assignment is orthogonal to these measures of weight is reasonable.

3.6.2 A Falsification Test

When using reduced form models of peer effects, there is a concern that the models are not actually detecting peer effects, but some spurious correlations - perhaps due to correlated effects not correctly controlled for by the fixed effects. Borrowing from Cohen-Cole and Jason M Fletcher (2008a), I run a model of height peer effects as a falsification test. I use the same empirical strategy as used to produce Table 3.2. The results of this test are found in Table 3.9. The height peer effect is 0.004, which is statistically significant. The interpretation is that a shock increasing average classmate height by a foot (12 inches) leads to an increase in own height that is less than a sixteenth of an inch. Given the sample size, I interpret this as a precise zero estimate. Height should not be contagious, and this supports my claim that the models presented here are detecting actual peer effects.

3.7 Conclusion

In this paper, I present evidence of obesity peer effects among elementary school students. Despite the crucial role that a child's early years play in childhood and adolescent obesity, this remains an understudied age group. Exposure to a classroom of peers who are overweight increases an elementary student's own BMI by 0.78 BMI points relative to a classroom with no overweight students - an effect that increases for exposure to obese and severely obese students. Exposure to underweight students produces the strongest effect, with exposure to two underweight student in the average classroom reducing BMI by nearly 0.15 BMI points relative to a student who has no underweight classmates.

These results suggest that policies aimed at reducing childhood obesity have spillover effects beyond their target population. Further work is necessary in order to develop specific policy recommendations. Effects appear to be non-linear along BMI percentiles, as students who are underweight have an outsized influence relative to students who are severely obese. This suggests that even within the normal-weight category (a large group of students between the fifth and eighty-fifth percentiles) there is potential benefit for moving students towards the median. LOM estimates using continuous measures of body weight composition such as BMI support this idea. Additional study is necessary to explore how different pieces of the BMI distribution are affected by their peers.

3.8 Tables

Table 3.1 Sample Summary Statistics

variable	mean	sd	variable	mean	sd
BMI	18.475	4.027	Grade:		
zBMI (NYC)	-0.005	0.989	First	0.186	0.389
zBMI (National)	0.544	1.231	Second	0.197	0.397
BMI Percentile	64.748	31.169	Third	0.204	0.403
Underweight	0.047	0.212	Fourth	0.204	0.403
Overweight	0.382	0.486	Fifth	0.209	0.407
Obese	0.210	0.407	Ethnicity:		
Severely Obese	0.061	0.239	Hispanic	0.410	0.492
Class Size	26.955	3.439	Black	0.184	0.388
Female	0.503	0.500	White	0.176	0.381
Ever FRPL	0.818	0.386	Asian/other	0.230	0.421
ELL	0.162	0.368			

The sample contains from 264,733 observations from the 2018 academic year. BMI indicates the body mass index. zBMI (NYC) indicates z-scores for BMI normalized to the population of New York City public school students. zBMI (National) indicates z-scores for BMI normalized nationally using the CDC's percentiles. The BMI Percentile is the measure of percentile based off the CDC's numbers. Underweight, overweight, obese, and severely obese are indicator variables for whether a student fits into the category. Categories are inclusive, such that an obese student is also considered overweight. Female, ever FRPL, and ELL are also indicator variables. Ever FRPL indicates whether a student is ever observed as qualifying for free or reduced price lunch. ELL indicates whether a student is an English language learner. Grade shows the breakdown of the sample by grade, and Ethnicity does this for each ethnic group. Notice that the modal ethnicity is Hispanic, so this will be the left out group in subsequent regressions.

Table 3.2 Baseline Reduced Form Regressions

Dep. Var.	(1) bmi	(2) zbmi	(3) zbmi2	(4) bmipct	(5) overweight	(6) obese	(7) sevobese	(8) underweight
LOM	0.028***	0.201***	0.140***	0.118***	0.078***	0.066***	0.017**	0.183***
	(0.002)	(0.005)	(0.004)	(0.004)	(0.006)	(0.006)	(0.007)	(0.008)
Lag	1.010***	0.802***	0.886***	0.814***	0.745***	0.772***	0.770***	0.413***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Asian/other	-0.146***	-0.071***	-0.039***	-1.627***	-0.035***	-0.029***	-0.015***	0.023***
	(0.010)	(0.004)	(0.003)	(0.106)	(0.002)	(0.002)	(0.001)	(0.001)
Black	-0.021*	-0.020***	-0.006**	-0.411***	-0.015***	-0.013***	-0.004***	0.003**
	(0.011)	(0.004)	(0.003)	(0.113)	(0.002)	(0.002)	(0.001)	(0.001)
White	-0.117***	-0.050***	-0.031***	-1.222***	-0.029***	-0.021***	-0.011***	0.009***
	(0.011)	(0.005)	(0.003)	(0.119)	(0.002)	(0.002)	(0.001)	(0.001)
Female	-0.038***	-0.054***	-0.002	-1.436***	-0.027***	-0.020***	-0.010***	0.002***
	(0.006)	(0.003)	(0.002)	(0.066)	(0.001)	(0.001)	(0.001)	(0.001)
ELL	0.046***	0.013***	0.007***	0.535***	0.005***	0.004**	-0.001	0.000
	(0.009)	(0.004)	(0.002)	(0.096)	(0.002)	(0.002)	(0.001)	(0.001)
Grade 2	0.070***	0.030***	0.023***	0.796***	0.009***	0.010***	0.002	-0.002
	(0.016)	(0.006)	(0.004)	(0.165)	(0.003)	(0.003)	(0.002)	(0.002)
Grade 3	-0.028	-0.003	0.024***	0.091	0.011**	0.008**	-0.000	-0.002
	(0.023)	(0.009)	(0.006)	(0.240)	(0.005)	(0.004)	(0.002)	(0.003)
Grade 4	-0.110***	-0.020*	0.025***	-0.146	0.008	0.004	-0.002	0.000
	(0.029)	(0.012)	(0.008)	(0.297)	(0.006)	(0.005)	(0.003)	(0.003)
Grade 5	-0.207***	-0.021	0.023**	-0.267	0.004	0.001	-0.003	-0.004
	(0.034)	(0.014)	(0.009)	(0.346)	(0.007)	(0.005)	(0.003)	(0.004)
SWD	0.036***	0.004	0.009***	-0.003	0.005***	0.007***	0.006***	0.001
	(0.009)	(0.004)	(0.002)	(0.093)	(0.002)	(0.001)	(0.001)	(0.001)
Ever FRPL	0.062***	0.030***	0.015***	0.865***	0.016***	0.010***	0.003***	-0.006***
	(0.010)	(0.004)	(0.003)	(0.104)	(0.002)	(0.002)	(0.001)	(0.001)
Age (months)	0.047***	0.005***	-0.003***	0.090**	0.003***	0.002***	0.000	-0.002***
	(0.003)	(0.001)	(0.001)	(0.036)	(0.001)	(0.001)	(0.000)	(0.000)
Age^2 (months)	-0.000***	-0.000***	0.000***	-0.001***	-0.000***	-0.000***	-0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-2.651***	-0.091	0.174***	3.744**	-0.038	-0.023	0.009	0.110***
	(0.180)	(0.073)	(0.048)	(1.859)	(0.035)	(0.029)	(0.018)	(0.021)
Observations	263,966	263,966	263,966	263,966	263,966	263,966	263,966	263,966
R-squared	0.842	0.713	0.807	0.711	0.567	0.582	0.530	0.226
Res. Zip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	18.48	0.544	-0.00504	64.75	0.382	0.210	0.0608	0.0470

This table shows the results of the reduced form model in equation 3.1. Under each column number is the dependent variable of interest, regressed on the classroom mean of the lag dependent variable for the student's classmates. The lag is always the lag of the dependent variable, so in column (1) this is lag BMI. Hispanic students are the comparison group for ethnicity because they are the modal student.

Standard errors are in parentheses.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

Table 3.3 Effect of Exposure to Body Types on BMI

Exposure Variable:	(1) overweight	(2) obese	(3) sevobese	(4) underweight
Proportion	0.779***	0.867***	1.064***	-1.901***
	(0.031)	(0.039)	(0.069)	(0.072)
Lag BMI	1.010***	1.010***	1.010***	1.010***
	(0.001)	(0.001)	(0.001)	(0.001)
Asian/other	-0.140***	-0.140***	-0.143***	-0.143***
	(0.010)	(0.010)	(0.010)	(0.010)
Black	-0.020*	-0.020*	-0.021*	-0.021*
	(0.011)	(0.011)	(0.011)	(0.011)
White	-0.113***	-0.113***	-0.115***	-0.114***
	(0.011)	(0.011)	(0.011)	(0.011)
Female	-0.038***	-0.038***	-0.038***	-0.036***
	(0.006)	(0.006)	(0.006)	(0.006)
ELL	0.036***	0.033***	0.031***	0.025***
	(0.009)	(0.009)	(0.009)	(0.009)
Grade 2	0.069***	0.075***	0.089***	0.085***
	(0.016)	(0.016)	(0.016)	(0.016)
Grade 3	-0.033	-0.016	0.013	-0.002
	(0.023)	(0.023)	(0.023)	(0.023)
Grade 4	-0.107***	-0.081***	-0.047*	-0.060**
	(0.028)	(0.028)	(0.028)	(0.028)
Grade 5	-0.187***	-0.153***	-0.120***	-0.142***
	(0.033)	(0.033)	(0.033)	(0.033)
SWD	0.034***	0.034***	0.037***	0.044***
	(0.009)	(0.009)	(0.009)	(0.009)
Ever FRPL	0.056***	0.056***	0.058***	0.057***
	(0.010)	(0.010)	(0.010)	(0.010)
Age (months)	0.047***	0.047***	0.047***	0.045***
	(0.003)	(0.003)	(0.003)	(0.003)
Age^2 (months)	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-2.414***	-2.358***	-2.260***	-2.013***
	(0.177)	(0.177)	(0.177)	(0.177)
Observations	263,966	263,966	263,966	263,966
R-squared	0.843	0.843	0.843	0.843
Res. Zip FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	18.48	18.48	18.48	18.48

This table shows the results of the reduced form model in equation 3.2 with BMI as the dependent variable. Under each column number is the exposure variable. The proportion is the proportion of classmates who were in the exposure variable category the previous year. Thus in column (1) the exposure variable is overweight, and the coefficient on proportion indicates the effect on BMI of exposure to a classroom in which all classmates are overweight, relative to a classroom in which all classmates are normal weight. Hispanic students are the comparison group for ethnicity because they are the modal student.

Standard errors are in parentheses.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

Table 3.4 Effect of Exposure to Body Types on BMI Percentile

Exposure Variable:	(1) overweight	(2) obese	(3) sevobese	(4) underweight
Proportion	7.794***	7.883***	7.994***	-23.742***
	(0.330)	(0.405)	(0.723)	(0.753)
Lag BMI Percentile	0.815***	0.815***	0.815***	0.814***
	(0.001)	(0.001)	(0.001)	(0.001)
Asian/other	-1.615***	-1.619***	-1.652***	-1.635***
	(0.106)	(0.106)	(0.106)	(0.106)
Black	-0.415***	-0.415***	-0.420***	-0.419***
	(0.113)	(0.113)	(0.113)	(0.113)
White	-1.213***	-1.216***	-1.238***	-1.220***
	(0.119)	(0.119)	(0.119)	(0.119)
Female	-1.433***	-1.433***	-1.429***	-1.418***
	(0.066)	(0.066)	(0.066)	(0.066)
ELL	0.376***	0.339***	0.316***	0.252***
	(0.096)	(0.096)	(0.096)	(0.096)
Grade 2	0.954***	1.037***	1.190***	1.073***
	(0.165)	(0.165)	(0.164)	(0.164)
Grade 3	0.270	0.478**	0.772***	0.496**
	(0.240)	(0.239)	(0.239)	(0.238)
Grade 4	0.013	0.326	0.667**	0.398
	(0.298)	(0.297)	(0.296)	(0.296)
Grade 5	-0.036	0.346	0.676*	0.311
	(0.346)	(0.345)	(0.345)	(0.344)
SWD	0.020	0.027	0.058	0.116
	(0.093)	(0.093)	(0.093)	(0.093)
Ever FRPL	0.830***	0.835***	0.859***	0.837***
	(0.104)	(0.104)	(0.105)	(0.104)
Age (months)	0.091**	0.097***	0.094***	0.072**
	(0.036)	(0.036)	(0.036)	(0.036)
Age^2 (months)	-0.001***	-0.001***	-0.001***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	7.847***	8.559***	9.571***	12.306***
	(1.850)	(1.850)	(1.850)	(1.847)
Observations	263,966	263,966	263,966	263,966
R-squared	0.711	0.711	0.710	0.711
Res. Zip FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	64.75	64.75	64.75	64.75

This table shows the results of the reduced form model in equation 3.2 with BMI percentile as the dependent variable. Under each column number is the exposure variable. The proportion is the proportion of classmates who were in the exposure variable category the previous year. Thus in column (1) the exposure variable is overweight, and the coefficient on proportion indicates the effect on BMI percentile of exposure to a classroom in which all classmates are overweight, relative to a classroom in which all classmates are normal weight. Hispanic students are the comparison group for ethnicity because they are the modal student.

Standard errors are in parentheses.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

Table 3.5 Effect of Exposure to Body Types on National Z-BMI

Exposure Variable:	(1) overweight	(2) obese	(3) sevobese	(4) underweight
Proportion	0.213***	0.233***	0.275***	-0.538***
	(0.009)	(0.010)	(0.019)	(0.020)
Lag z-BMI	0.886***	0.886***	0.887***	0.886***
	(0.001)	(0.001)	(0.001)	(0.001)
Asian/other	-0.040***	-0.040***	-0.041***	-0.041***
	(0.003)	(0.003)	(0.003)	(0.003)
Black	-0.006**	-0.006**	-0.006**	-0.006**
	(0.003)	(0.003)	(0.003)	(0.003)
White	-0.032***	-0.032***	-0.033***	-0.033***
	(0.003)	(0.003)	(0.003)	(0.003)
Female	-0.002	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
ELL	0.010***	0.009***	0.008***	0.006**
	(0.002)	(0.002)	(0.002)	(0.002)
Grade 2	0.016***	0.017***	0.022***	0.020***
	(0.004)	(0.004)	(0.004)	(0.004)
Grade 3	0.010	0.015**	0.023***	0.018***
	(0.006)	(0.006)	(0.006)	(0.006)
Grade 4	0.007	0.015*	0.024***	0.020**
	(0.008)	(0.008)	(0.008)	(0.008)
Grade 5	0.003	0.013	0.022**	0.015*
	(0.009)	(0.009)	(0.009)	(0.009)
SWD	0.010***	0.010***	0.010***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)
Ever FRPL	0.016***	0.016***	0.017***	0.016***
	(0.003)	(0.003)	(0.003)	(0.003)
Age (months)	-0.004***	-0.004***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Age^2 (months)	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.135***	0.151***	0.179***	0.247***
	(0.048)	(0.048)	(0.048)	(0.048)
Observations	263,966	263,966	263,966	263,966
R-squared	0.807	0.807	0.807	0.807
Res. Zip FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	-0.00504	-0.00504	-0.00504	-0.00504

This table shows the results of the reduced form model in equation 3.2 with BMI z-scores (normalized at the national level) as the dependent variable. Under each column number is the exposure variable. The proportion is the proportion of classmates who were in the exposure variable category the previous year. Thus in column (1) the exposure variable is overweight, and the coefficient on proportion indicates the effect on BMI z-scores of exposure to a classroom in which all classmates are overweight, relative to a classroom in which all classmates are normal weight. Hispanic students are the comparison group for ethnicity.

Standard errors are in parentheses.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

Table 3.6 Effect of Exposure to Body Types on NYC Z-BMI

Exposure Variable:	(1) overweight	(2) obese	(3) sevobese	(4) underweight
Proportion	0.318***	0.328***	0.332***	-1.034***
	(0.013)	(0.016)	(0.029)	(0.030)
Lag z-BMI	0.802***	0.802***	0.803***	0.802***
	(0.001)	(0.001)	(0.001)	(0.001)
Asian/other	-0.073***	-0.073***	-0.075***	-0.074***
	(0.004)	(0.004)	(0.004)	(0.004)
Black	-0.020***	-0.020***	-0.020***	-0.020***
	(0.004)	(0.004)	(0.004)	(0.004)
White	-0.053***	-0.053***	-0.054***	-0.053***
	(0.005)	(0.005)	(0.005)	(0.005)
Female	-0.054***	-0.054***	-0.054***	-0.054***
	(0.003)	(0.003)	(0.003)	(0.003)
ELL	0.013***	0.012***	0.011***	0.008**
	(0.004)	(0.004)	(0.004)	(0.004)
Grade 2	0.039***	0.042***	0.048***	0.043***
	(0.006)	(0.006)	(0.006)	(0.006)
Grade 3	0.011	0.019**	0.031***	0.019**
	(0.009)	(0.009)	(0.009)	(0.009)
Grade 4	-0.005	0.007	0.021*	0.009
	(0.012)	(0.012)	(0.012)	(0.012)
Grade 5	-0.004	0.011	0.025*	0.009
	(0.014)	(0.014)	(0.014)	(0.014)
SWD	0.005	0.005	0.007*	0.009**
	(0.004)	(0.004)	(0.004)	(0.004)
Ever FRPL	0.033***	0.033***	0.034***	0.033***
	(0.004)	(0.004)	(0.004)	(0.004)
Age (months)	0.005***	0.005***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Age^2 (months)	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.130*	-0.102	-0.059	0.058
	(0.073)	(0.073)	(0.073)	(0.073)
Observations	263,966	263,966	263,966	263,966
R-squared	0.711	0.711	0.711	0.712
Res. Zip FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.544	0.544	0.544	0.544

This table shows the results of the reduced form model in equation 3.2 with BMI z-scores (normalized at the NYC level) as the dependent variable. Under each column number is the exposure variable. The proportion is the proportion of classmates who were in the exposure variable category the previous year. Thus in column (1) the exposure variable is overweight, and the coefficient on proportion indicates the effect on BMI z-scores of exposure to a classroom in which all classmates are overweight, relative to a classroom in which all classmates are normal weight. Hispanic students are the comparison group for ethnicity.

Standard errors are in parentheses.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

Table 3.7 Reduced Form Interaction with Sex

Dep. Var.	(1) bmi	(2) zbmi	(3) zbmi2	(4) bmipct	(5) overweight	(6) obese	(7) sevobese	(8) underweight
LOM	0.030***	0.201***	0.137***	0.118***	0.098***	0.082***	0.046***	0.175***
	(0.002)	(0.006)	(0.005)	(0.005)	(0.008)	(0.008)	(0.009)	(0.011)
Female	0.016	-0.054***	-0.002	-1.485***	-0.013***	-0.015***	-0.007***	0.001
	(0.044)	(0.004)	(0.002)	(0.344)	(0.003)	(0.002)	(0.001)	(0.001)
Female*LOM	-0.003	-0.000	0.007	0.001	-0.041***	-0.030***	-0.057***	0.015
	(0.003)	(0.007)	(0.006)	(0.006)	(0.009)	(0.010)	(0.012)	(0.014)
Lag	1.010***	0.802***	0.886***	0.814***	0.744***	0.772***	0.770***	0.413***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Asian/other	-0.146***	-0.071***	-0.039***	-1.627***	-0.035***	-0.029***	-0.015***	0.023***
	(0.010)	(0.004)	(0.003)	(0.106)	(0.002)	(0.002)	(0.001)	(0.001)
Black	-0.021*	-0.020***	-0.006**	-0.411***	-0.015***	-0.013***	-0.004***	0.003**
	(0.011)	(0.004)	(0.003)	(0.113)	(0.002)	(0.002)	(0.001)	(0.001)
White	-0.117***	-0.050***	-0.031***	-1.222***	-0.029***	-0.021***	-0.011***	0.009***
	(0.011)	(0.005)	(0.003)	(0.119)	(0.002)	(0.002)	(0.001)	(0.001)
ELL	0.046***	0.013***	0.007***	0.535***	0.005***	0.004**	-0.001	0.000
	(0.009)	(0.004)	(0.002)	(0.096)	(0.002)	(0.002)	(0.001)	(0.001)
Grade 2	0.069***	0.030***	0.023***	0.796***	0.009***	0.010***	0.002	-0.002
	(0.016)	(0.006)	(0.004)	(0.165)	(0.003)	(0.003)	(0.002)	(0.002)
Grade 3	-0.028	-0.003	0.024***	0.091	0.011**	0.008**	-0.000	-0.002
	(0.023)	(0.009)	(0.006)	(0.240)	(0.005)	(0.004)	(0.002)	(0.003)
Grade 4	-0.110***	-0.020*	0.025***	-0.146	0.008	0.004	-0.002	0.000
	(0.029)	(0.012)	(0.008)	(0.297)	(0.006)	(0.005)	(0.003)	(0.003)
Grade 5	-0.207***	-0.021	0.023**	-0.267	0.004	0.001	-0.003	-0.004
	(0.034)	(0.014)	(0.009)	(0.346)	(0.007)	(0.005)	(0.003)	(0.004)
SWD	0.036***	0.004	0.009***	-0.003	0.005**	0.007***	0.006***	0.001
	(0.009)	(0.004)	(0.002)	(0.093)	(0.002)	(0.001)	(0.001)	(0.001)
Ever FRPL	0.062***	0.030***	0.015***	0.865***	0.016***	0.010***	0.003***	-0.006***
	(0.010)	(0.004)	(0.003)	(0.104)	(0.002)	(0.002)	(0.001)	(0.001)
Age (months)	0.048***	0.005***	-0.003***	0.090**	0.003***	0.002***	0.000	-0.002***
	(0.003)	(0.001)	(0.001)	(0.036)	(0.001)	(0.001)	(0.000)	(0.000)
Age^2 (months)	-0.000***	-0.000***	0.000***	-0.001***	-0.000***	-0.000***	-0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-2.681***	-0.091	0.174***	3.770**	-0.046	-0.026	0.007	0.110***
	(0.181)	(0.073)	(0.048)	(1.868)	(0.035)	(0.029)	(0.018)	(0.021)
Observations	263,966	263,966	263,966	263,966	263,966	263,966	263,966	263,966
R-squared	0.842	0.713	0.807	0.711	0.567	0.582	0.530	0.226
Res. Zip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	18.48	0.544	-0.00504	64.75	0.382	0.210	0.0608	0.0470

This table shows the results of the reduced form model in equation 3.1 in which the variable of interest is interacted with the sex of the student. Under each column number is the dependent variable of interest, regressed on the classroom mean of the lag dependent variable for the student's classmates. The lag is always the lag of the dependent variable, so in column (1) this is lag BMI. Hispanic students are the comparison group for ethnicity because they are the modal student.

Standard errors are in parentheses.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

Table 3.8 Reduced Form Interaction with Ethnicity

Dep. Var.	(1) bmi	(2) zbmi	(3) zbmi2	(4) bmipct	(5) overweight	(6) obese	(7) sevobese	(8) underweight
LOM	0.027***	0.186***	0.150***	0.105***	0.078***	0.084***	0.030***	0.106***
	(0.002)	(0.007)	(0.006)	(0.005)	(0.008)	(0.009)	(0.010)	(0.013)
Asian/other	-0.127**	-0.097***	-0.038***	-3.950***	-0.035***	-0.021***	-0.014***	0.012***
	(0.064)	(0.006)	(0.003)	(0.493)	(0.005)	(0.003)	(0.001)	(0.001)
Black	-0.069	-0.005	-0.004	0.611	-0.009*	-0.009**	-0.003**	0.004**
	(0.061)	(0.008)	(0.003)	(0.541)	(0.006)	(0.004)	(0.002)	(0.002)
White	-0.275***	-0.063***	-0.031***	-3.750***	-0.033***	-0.014***	-0.009***	0.006***
	(0.071)	(0.006)	(0.003)	(0.532)	(0.005)	(0.003)	(0.001)	(0.002)
Asian/other*LOM	-0.001	0.060***	-0.008	0.040***	0.000	-0.043***	-0.017	0.215***
	(0.004)	(0.010)	(0.009)	(0.008)	(0.013)	(0.014)	(0.018)	(0.018)
Black*LOM	0.003	-0.026**	-0.029***	-0.017*	-0.016	-0.020	-0.021	-0.016
	(0.004)	(0.011)	(0.009)	(0.009)	(0.014)	(0.015)	(0.016)	(0.023)
White*LOM	0.010**	0.023**	-0.013	0.044***	0.013	-0.038**	-0.046**	0.061***
	(0.004)	(0.011)	(0.010)	(0.009)	(0.014)	(0.016)	(0.020)	(0.021)
Lag	1.010***	0.802***	0.886***	0.814***	0.745***	0.772***	0.770***	0.413***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Female	-0.038***	-0.054***	-0.002	-1.436***	-0.027***	-0.020***	-0.010***	0.002***
	(0.006)	(0.003)	(0.002)	(0.066)	(0.001)	(0.001)	(0.001)	(0.001)
ELL	0.045***	0.013***	0.007***	0.548***	0.005***	0.003**	-0.001	0.000
	(0.009)	(0.004)	(0.002)	(0.096)	(0.002)	(0.002)	(0.001)	(0.001)
Grade 2	0.069***	0.030***	0.023***	0.804***	0.009***	0.010***	0.002	-0.002
	(0.016)	(0.006)	(0.004)	(0.165)	(0.003)	(0.003)	(0.002)	(0.002)
Grade 3	-0.029	-0.003	0.024***	0.098	0.011**	0.008**	-0.000	-0.002
	(0.023)	(0.009)	(0.006)	(0.240)	(0.005)	(0.004)	(0.002)	(0.003)
Grade 4	-0.111***	-0.020*	0.025***	-0.146	0.008	0.004	-0.002	0.000
	(0.029)	(0.012)	(0.008)	(0.297)	(0.006)	(0.005)	(0.003)	(0.003)
Grade 5	-0.208***	-0.021	0.023**	-0.278	0.004	0.001	-0.003	-0.004
	(0.034)	(0.014)	(0.009)	(0.346)	(0.007)	(0.005)	(0.003)	(0.004)
SWD	0.036***	0.004	0.009***	-0.005	0.005***	0.007***	0.006***	0.001
	(0.009)	(0.004)	(0.002)	(0.093)	(0.002)	(0.001)	(0.001)	(0.001)
Ever FRPL	0.063***	0.029***	0.015***	0.847***	0.016***	0.010***	0.004***	-0.006***
	(0.010)	(0.004)	(0.003)	(0.104)	(0.002)	(0.002)	(0.001)	(0.001)
Age (months)	0.047***	0.004***	-0.003***	0.083**	0.003***	0.002***	0.000	-0.002***
	(0.003)	(0.001)	(0.001)	(0.036)	(0.001)	(0.001)	(0.000)	(0.000)
Age^2 (months)	-0.000***	-0.000***	0.000***	-0.000***	-0.000***	-0.000***	-0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-2.622***	-0.069	0.173***	4.874***	-0.037	-0.028	0.008	0.110***
	(0.182)	(0.073)	(0.048)	(1.875)	(0.035)	(0.029)	(0.018)	(0.021)
Observations	263,966	263,966	263,966	263,966	263,966	263,966	263,966	263,966
R-squared	0.842	0.713	0.807	0.711	0.567	0.582	0.530	0.226
Res. Zip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	18.48	0.544	-0.00504	64.75	0.382	0.210	0.0608	0.0470

This table shows the results of the reduced form model in equation 3.1 in which the variable of interest is interacted with the ethnicity of the student. Under each column number is the dependent variable of interest, regressed on the classroom mean of the lag dependent variable for the student's classmates. The lag is always the lag of the dependent variable, so in column (1) this is lag BMI. Hispanic students are the comparison group for ethnicity because they are the modal student.

Standard errors are in parentheses.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

Table 3.9 Falsification Test: Height Peer Effects

	par.	s.e.
LOM of Height	0.004***	(0.001)
Lag	0.944***	(0.001)
Asian/other	0.028***	(0.007)
Black	0.112***	(0.008)
White	0.011	(0.008)
Female	0.123***	(0.004)
ELL	-0.021***	(0.007)
Grade 2	-0.149***	(0.011)
Grade 3	-0.189***	(0.016)
Grade 4	-0.177***	(0.020)
Grade 5	-0.191***	(0.024)
SWD	-0.006	(0.006)
Ever FRPL	0.020***	(0.007)
Age (months)	0.019***	(0.002)
Age^2 (months)	-0.000**	(0.000)
Constant	3.334***	(0.130)
Observations	263,966	
R-squared	0.932	
Res. Zip FE	Yes	
School FE	Yes	
Dep. Var. Mean	52.43	

This table shows the results of the reduced form model in equation 3.2 with student height (inches) as the dependent variable, regressed on the LOM of classmate lagged height. Hispanic students are the comparison group for ethnicity because they are the modal student. This serves as a falsification test, as student height should not be contagious.

Standard errors are in parentheses.
*** Significant at the 1% level
** Significant at the 5% level
* Significant at the 10% level

3.9 Figures

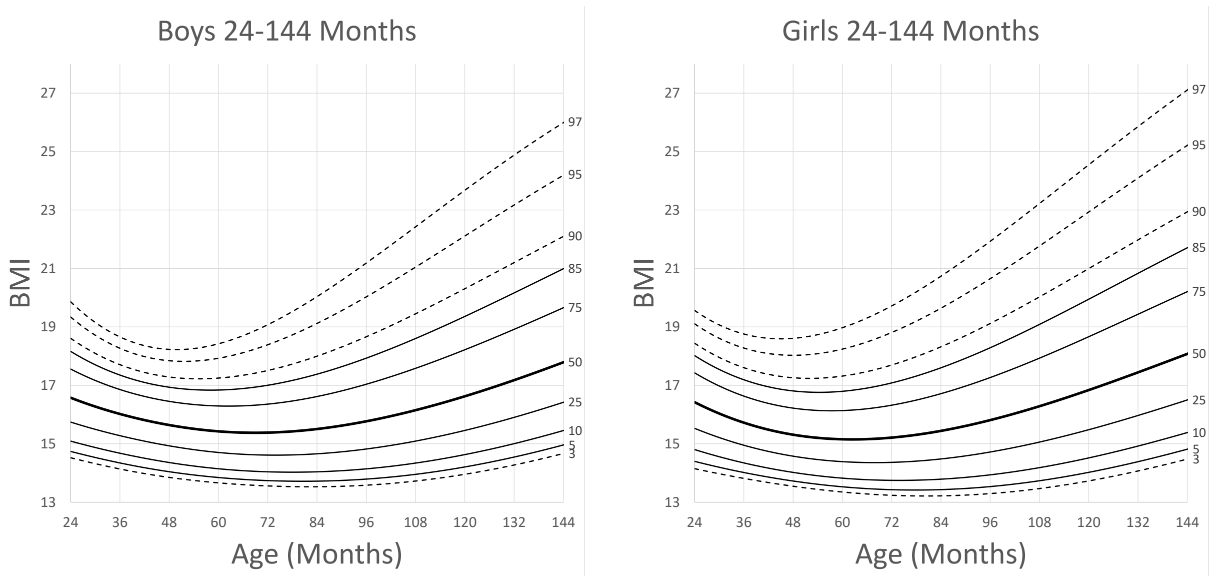
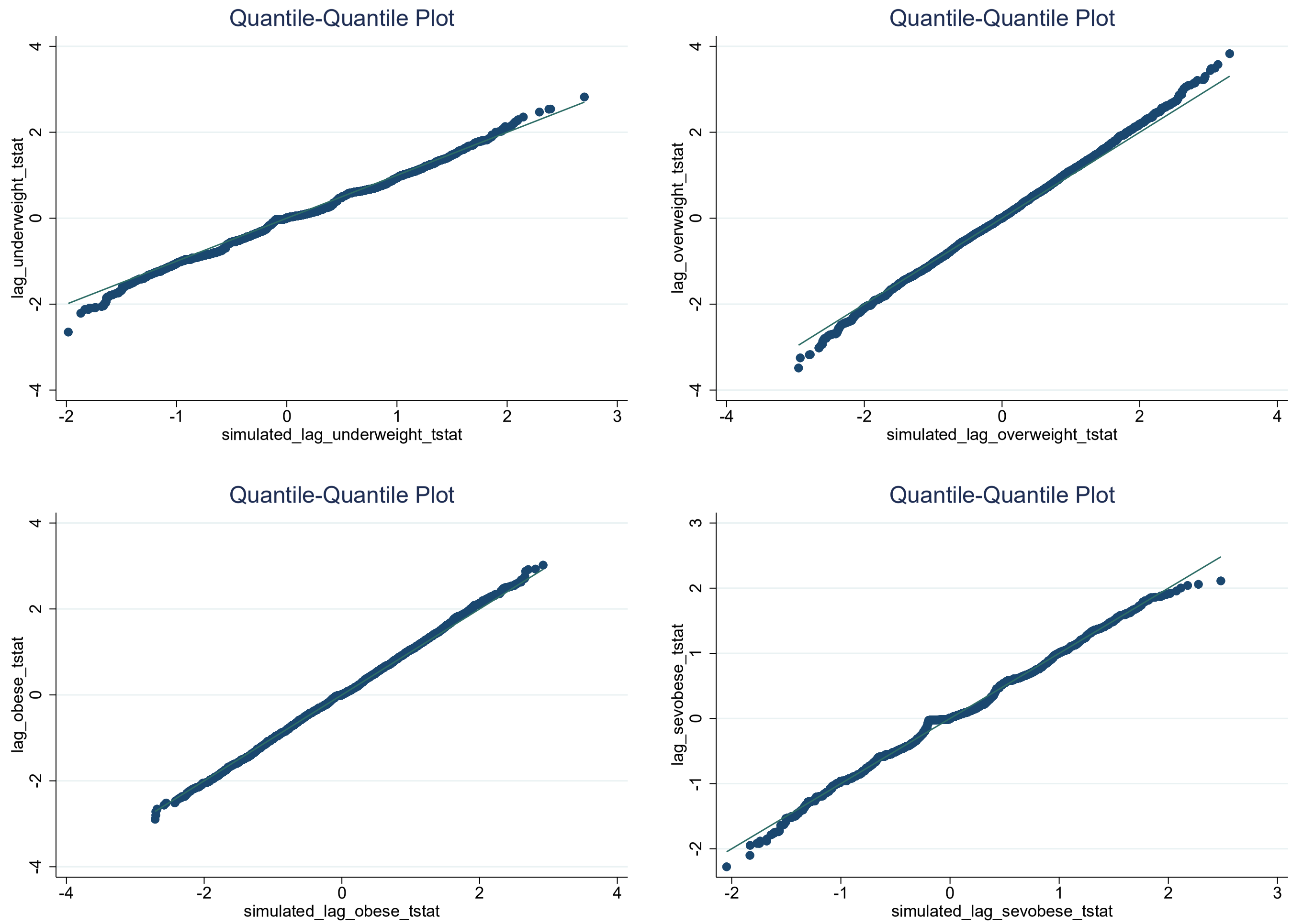


Figure 3.1 BMI-Age Charts

Figure uses BMI-for-age data tables from the CDC (https://www.cdc.gov/growthcharts/clinical_charts.htm), trimmed to show ages 2 through the end of elementary school. Ages below elementary school are shown to illustrate the quadratic nature of normal weight. The bold line shows the national median BMI, with curves above and below denoting ranges for children in other percentile groups of BMI for each age. Dashed lines indicate percentiles that are outside the healthy BMI range.

Figure 3.2 Quantile-Quantile Plots



A series of multinomial logits are run to estimate the importance of each weight category in class assignment. Lagged values are used, as this is what would be observed at time of classroom assignment. In each quantile-quantile plot, I plot the t-statistics from observed classroom assignment against the t-statisitics from a similar exercise in which we randomly assign students to classrooms. We are plotting these two distributions against one another, and if the distributions are the same, we should expect a straight line of slope one. I argue that these provide evidence that class assignment is consistent with a random process for each weight category.

Bibliography

Ajilore, Olugbenga et al. (2014). "Uncovering peer effects mechanisms with weight outcomes using spatial econometrics". In: *The Social Science Journal* 51.4, pp. 645–651. ISSN: 0362-3319. DOI: https://doi.org/10.1016/j.soscij.2014.07.008. URL: http://www.sciencedirect.com/science/article/pii/S0362331914000834.

Anderson, Patricia M., Kristin F. Butcher, and Diane Whitmore Schanzenbach (2019). "Understanding recent trends in childhood obesity in the United States". In: *Economics & Human Biology* 34. New Findings in Economics and Human Biology:A special Issue in Honor of Founding Editor John Komlos, pp. 16–25. ISSN: 1570-677X. DOI: https://doi.org/10.1016/j.ehb.2019.02.002. URL: http://www.sciencedirect.com/science/article/pii/S1570677X18302429.

Asirvatham, Jebaraj, Rodolfo M. Nayga Jr., and Michael R. Thomsen (2014). "Peer-Effects in Obesity among Public Elementary School Children: A Grade-Level Analysis". In: *Applied Economic Perspectives and Policy* 36.3, pp. 438–459. DOI: 10.1093/aepp/ppu011. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1093/aepp/ppu011. URL: https://onlinelibrary.wiley.com/doi/abs/10.1093/aepp/ppu011.

Bennett, Magdalena and Peter Bergman (2018). *Better Together? Social Networks in Truancy and the Targeting of Treatment*. IZA Discussion Papers 11267. Institute of Labor Economics (IZA). URL: https://EconPapers.repec.org/RePEc:iza:izadps:dp11267.

Biener, Adam, John Cawley, and Chad Meyerhoefer (Jan. 2018). "The Impact of Obesity on Medical Care Costs and Labor Market Outcomes in the US". In: *Clinical Chemistry* 64.1, pp. 108–117. ISSN: 0009-9147. DOI: 10.1373/clinchem.2017.272450. eprint: https://academic.oup.com/clinchem/article-pdf/64/1/108/32641764/clinchem0108.pdf. URL: https://doi.org/10.1373/clinchem.2017.272450.

Bifulco, Robert, Jason M. Fletcher, and Stephen L. Ross (Feb. 2011). "The Effect of Classmate Characteristics on Post-secondary Outcomes: Evidence from the Add Health". In: *American Economic Journal: Economic Policy* 3.1, pp. 25–53. DOI: 10.1257/pol.3.1.25. URL: http://www.aeaweb.org/articles?id=10.1257/pol.3.1.25.

Blume, Lawrence E. et al. (2011). "Identification of social interaction". In: *Handbook of Social Economics*. Ed. by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson. Vol. 1B. San Diego, CA: Elsevier, pp. 853–964.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2009). "Identification of peer effects through social networks". In: *Journal of Econometrics* 150.1, pp. 41–55. ISSN: 0304-4076. DOI: https://doi.org/10.1016/j.jeconom.2008.12.021. URL: http://www.sciencedirect.com/science/article/pii/S0304407609000335.

Burke, Mary A. and Tim R. Sass (2013). "Classroom Peer Effects and Student Achievement". In: *Journal of Labor Economics* 31.1, pp. 51–82. DOI: 10.1086/666653. eprint: https://doi.org/10.1086/666653. URL: https://doi.org/10.1086/666653.

Carrell, Scott E., Richard L. Fullerton, and James E. West (2009). "Does Your Cohort Matter? Measuring Peer Effects in College Achievement". In: *Journal of Labor Economics* 27.3, pp. 439–464. DOI: 10.1086/600143. eprint: https://doi.org/10.1086/600143. URL: https://doi.org/10.1086/600143.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka (Nov. 2018). "The Long-Run Effects of Disruptive Peers". In: *American Economic Review* 108.11, pp. 3377–3415. DOI: 10.1257/aer.20160763. URL: http://www.aeaweb.org/articles?id=10.1257/aer.20160763.

Carrell, Scott E., Mark Hoekstra, and James E. West (2011). "Is poor fitness contagious?: Evidence from randomly assigned friends". In: *Journal of Public Economics* 95.7, pp. 657–663. ISSN:

0047-2727. DOI: https://doi.org/10.1016/j.jpubeco.2010.12.005. URL: http://www.sciencedirect.com/science/article/pii/S0047272710002070.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West (2013). "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation". In: *Econometrica* 81.3, pp. 855–882. DOI: 10.3982/ECTA10168. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10168. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA10168.

Cawley, John (2004). "The Impact of Obesity on Wages". In: *The Journal of Human Resources* 39.2, pp. 451–474. ISSN: 0022166X. URL: http://www.jstor.org/stable/3559022.

— (2008). "Contingent valuation analysis of willingness to pay to reduce childhood obesity". In: *Economics & Human Biology* 6.2, pp. 281–292. ISSN: 1570-677X. DOI: https://doi.org/10.1016/j.ehb.2008.05.003. URL: http://www.sciencedirect.com/science/article/pii/S1570677X08000257.

Christakis, Nicholas A. and James H. Fowler (2007). "The Spread of Obesity in a Large Social Network over 32 Years". In: *New England Journal of Medicine* 357.4. PMID: 17652652, pp. 370–379. DOI: 10.1056/NEJMsa066082. eprint: https://doi.org/10.1056/NEJMsa066082. URL: https://doi.org/10.1056/NEJMsa066082.

Cohen-Cole, Ethan and Jason M Fletcher (2008a). "Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis". In: *BMJ* 337. ISSN: 0959-8138. DOI: 10.1136/bmj.a2533. eprint: https://www.bmj.com/content/337/bmj.a2533.full.pdf. URL: https://www.bmj.com/content/337/bmj.a2533.

— (2008b). "Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic". In: *Journal of Health Economics* 27.5, pp. 1382–1387. ISSN: 0167-6296. DOI: https://doi.org/10.1016/j.jhealeco.2008.04.005. URL: http://www.sciencedirect.com/science/article/pii/S0167629608000362.

Coleman, J.S. et al. (1966). *Equality of Educational Opportunity*. Tech. rep. OE-38001. U.S. Government Printing Office, Washington, DC: National Center for Educational Statistics.

Cunningham, Solveig A., Michael R. Kramer, and K.M. Venkat Narayan (2014). “Incidence of Childhood Obesity in the United States”. In: *New England Journal of Medicine* 370.5. PMID: 24476431, pp. 403–411. DOI: 10.1056/NEJMoa1309753. eprint: https://doi.org/10.1056/NEJMoa1309753. URL: https://doi.org/10.1056/NEJMoa1309753.

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli (Apr. 2010). “Identification of Social Interactions through Partially Overlapping Peer Groups”. In: *American Economic Journal: Applied Economics* 2.2, pp. 241–75. DOI: 10.1257/app.2.2.241. URL: http://www.aeaweb.org/articles?id=10.1257/app.2.2.241.

Donahoo, William et al. (2008). “Dietary Fat Increases Energy Intake Across the Range of Typical Consumption in the United States”. In: *Obesity* 16.1, pp. 64–69. DOI: 10.1038/oby.2007.31. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1038/oby.2007.31. URL: https://onlinelibrary.wiley.com/doi/abs/10.1038/oby.2007.31.

Elliott, Marc N. et al. (Apr. 2009). “Using the Census Bureau’s surname list to improve estimates of race/ethnicity and associated disparities”. In: *Health Services and Outcomes Research Methodology* 9.2, p. 69. ISSN: 1572-9400. DOI: 10.1007/s10742-009-0047-1. URL: https://doi.org/10.1007/s10742-009-0047-1.

Finkelstein, Eric A. et al. (2009). “Annual Medical Spending Attributable To Obesity: Payer-And Service-Specific Estimates”. In: *Health Affairs* 28.Supplement 1, w822–w831. DOI: 10.1377/hlthaff.28.5.w822. eprint: https://doi.org/10.1377/hlthaff.28.5.w822. URL: https://doi.org/10.1377/hlthaff.28.5.w822.

Flegal, Katherine M. et al. (Jan. 2010). “Prevalence and Trends in Obesity Among US Adults, 1999-2008”. In: *JAMA* 303.3, pp. 235–241. ISSN: 0098-7484. DOI: 10.1001/jama.2009.2014. eprint: https://jamanetwork.com/journals/jama/articlepdf/185235/joc90148_235_241.pdf. URL: https://doi.org/10.1001/jama.2009.2014.

Fowler, J.H. and N.A. Christakis (2008). “Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais”. In: *Journal of Health Economics* 27.5, pp. 1400–1405. ISSN: 0167-6296. DOI: https://doi.org/10.

1016/j.jhealeco.2008.07.001. URL: http://www.sciencedirect.com/science/article/pii/S0167629608000854.

Goldsmith-Pinkham, Paul and Guido W. Imbens (2013). "Social Networks and the Identification of Peer Effects". In: *Journal of Business & Economic Statistics* 31.3, pp. 253–264. DOI: 10.1080/07350015.2013.801251. eprint: https://doi.org/10.1080/07350015.2013.801251. URL: https://doi.org/10.1080/07350015.2013.801251.

Gordon-Larsen, Penny, Linda S. Adair, and Barry M. Popkin (2002). "Ethnic Differences in Physical Activity and Inactivity Patterns and Overweight Status". In: *Obesity Research* 10.3, pp. 141–149. DOI: 10.1038/oby.2002.23. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1038/oby.2002.23. URL: https://onlinelibrary.wiley.com/doi/abs/10.1038/oby.2002.23.

Halfon, Neal, Kandyce Larson, and Wendy Slusser (2013). "Associations Between Obesity and Comorbid Mental Health, Developmental, and Physical Health Conditions in a Nationally Representative Sample of US Children Aged 10 to 17". In: *Academic Pediatrics* 13.1, pp. 6–13. ISSN: 1876-2859. DOI: https://doi.org/10.1016/j.acap.2012.10.007. URL: http://www.sciencedirect.com/science/article/pii/S1876285912002847.

Halliday, Timothy J. and Sally Kwak (July 2009). "Weight gain in adolescents and their peers". In: *Economics & Human Biology* 7.2, pp. 181–190. URL: https://ideas.repec.org/a/eee/ehbiol/v7y2009i2p181-190.html.

Hedley, Allison A. et al. (June 2004). "Prevalence of Overweight and Obesity Among US Children, Adolescents, and Adults, 1999-2002". In: *JAMA* 291.23, pp. 2847–2850. ISSN: 0098-7484. DOI: 10.1001/jama.291.23.2847. eprint: https://jamanetwork.com/journals/jama/articlepdf/198912/jbr40059.pdf. URL: https://doi.org/10.1001/jama.291.23.2847.

Horrace, William et al. (2019). "What makes a classmate a peer? Examining which peers matter in NYC". Working Paper.

Horrace, William C., Xiaodong Liu, and Eleonora Patacchini (2016). "Endogenous network production functions with selectivity". In: *Journal of Econometrics* 190.2. Endogeneity Problems in Econometrics, pp. 222–232. ISSN: 0304-4076. DOI: https://doi.org/10.1016/j.jeconom.2015.06.005. URL: http://www.sciencedirect.com/science/article/pii/S0304407615001748.

Horrace, William C. and Christopher F. Parmeter (2017). "Accounting for Multiplicity in Inference on Economics Journal Rankings". In: *Southern Economic Journal* 84.1, pp. 337–347. DOI: 10.1002/soej.12219. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/soej.12219. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/soej.12219.

Hoxby, Caroline (Aug. 2000). *Peer Effects in the Classroom: Learning from Gender and Race Variation*. Working Paper 7867. National Bureau of Economic Research. DOI: 10.3386/w7867. URL: http://www.nber.org/papers/w7867.

Hoxby, Caroline and Gretchen Weingarth (2005). *Taking race out of the equation: School reassignment and the structure of peer effects*. Working Paper. URL: http://www.pseweb.eu/ydepot/semin/texte0607/HOX2007SCH.pdf.

Hsieh, Chih-Sheng and Lung Fei Lee (2016). "A Social Interactions Model with Endogenous Friendship Formation and Selectivity". In: *Journal of Applied Econometrics* 31.2, pp. 301–319. DOI: 10.1002/jae.2426. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.2426. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2426.

Kelejian, Harry H. and Gianfranco Piras (2014). "Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes". In: *Regional Science and Urban Economics* 46.C, pp. 140–149. URL: https://EconPapers.repec.org/RePEc:eee:regeco:v:46:y:2014:i:c:p:140-149.

Kelejian, Harry H. and Ingmar R. Prucha (2004). "Estimation of simultaneous systems of spatially interrelated cross sectional equations". In: *Journal of Econometrics* 118.1. Contributions to econometrics, time series analysis, and systems identification: a Festschrift in honor of Manfred Deistler, pp. 27–50. ISSN: 0304-4076. DOI: https://doi.org/10.1016/S0304-4076(03)00133-7. URL: http://www.sciencedirect.com/science/article/pii/S0304407603001337.

Kindblom, Jenny M. et al. (2009). "BMI Changes During Childhood and Adolescence as Predictors of Amount of Adult Subcutaneous and Visceral Adipose Tissue in Men". In: *Diabetes* 58.4, pp. 867–874. ISSN: 0012-1797. DOI: 10.2337/db08-0606. eprint: https://diabetes.diabetesjournals.org/content/58/4/867.full.pdf. URL: https://diabetes.diabetesjournals.org/content/58/4/867.

Lakdawalla, Darius N., Jayanta Bhattacharya, and Dana P. Goldman (2004). "Are The Young Becoming More Disabled?" In: *Health Affairs* 23.1. PMID: 15002639, pp. 168–176. DOI: 10.1377/hlthaff.23.1.168. eprint: https://doi.org/10.1377/hlthaff.23.1.168. URL: https://doi.org/10.1377/hlthaff.23.1.168.

Laschever, Ron A. (June 2013). *The Doughboys Network: Social Interactions and the Employment of World War I Veterans*. Tech. rep. URL: https://ssrn.com/abstract=1205543.

Lee, Lung-Fei (2007). "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects". In: *Journal of Econometrics* 140.2, pp. 333–374. URL: https://EconPapers.repec.org/RePEc:eee:econom:v:140:y:2007:i:2:p:333-374.

Lee, Lung-fei and Jihai Yu (2010). "Estimation of spatial autoregressive panel data models with fixed effects". In: *Journal of Econometrics* 154.2, pp. 165–185. ISSN: 0304-4076. DOI: https://doi.org/10.1016/j.jeconom.2009.08.001. URL: http://www.sciencedirect.com/science/article/pii/S030440760900178X.

Lin, Xu (2010). "Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables". In: *Journal of Labor Economics* 28.4, pp. 825–860. DOI: 10.1086/653506. eprint: https://doi.org/10.1086/653506. URL: https://doi.org/10.1086/653506.

— (Nov. 2015). "Utilizing spatial autoregressive models to identify peer effects among adolescents". In: *Empirical Economics* 49.3, pp. 929–960. ISSN: 1435-8921. DOI: 10.1007/s00181-014-0897-4. URL: https://doi.org/10.1007/s00181-014-0897-4.

Ludwig, David S. and Cara B. Ebbeling (Sept. 2001). "Type 2 Diabetes Mellitus in ChildrenPrimary Care and Public Health Considerations". In: *JAMA* 286.12, pp. 1427–1430. ISSN: 0098-7484. DOI: 10.1001/jama.286.12.1427. eprint: https://jamanetwork.com/journals/jama/articlepdf/194216/jct10011.pdf. URL: https://doi.org/10.1001/jama.286.12.1427.

Manski, Charles F. (July 1993). “Identification of Endogenous Social Effects: The Reflection Problem”. In: *The Review of Economic Studies* 60.3, pp. 531–542. ISSN: 0034-6527. DOI: 10.2307/2298123. eprint: http://oup.prod.sis.lan/restud/article-pdf/60/3/531/4468725/60-3-531.pdf. URL: https://doi.org/10.2307/2298123.

Mayer, Adalbert and Steven Puller (2008). “The old boy (and girl) network: Social network formation on university campuses”. In: *Journal of Public Economics* 92.1-2, pp. 329–347. URL: https://EconPapers.repec.org/RePEc:eee:pubeco:v:92:y:2008:i:1-2:p:329-347.

Moro, Andrea, D. Sebastian Tello-Trillo, and Tommaso Tempesti (2018). “The Impact of Obesity on Wages: the Role of Personal Interactions and Job Selection”. In: *working paper*. URL: https://andreamoro.net/perm/papers/obesity.pdf.

Morrison, Katherine M. et al. (2015). “Association of depression & health related quality of life with body composition in children and youth with obesity”. In: *Journal of Affective Disorders* 172, pp. 18–23. ISSN: 0165-0327. DOI: https://doi.org/10.1016/j.jad.2014.09.014. URL: http://www.sciencedirect.com/science/article/pii/S0165032714005680.

Must, A. and R. S. Strauss (1999). “Risks and consequences of childhood and adolescent obesity”. In: *International Journal of Obesity* 23.2, S2–S11. ISSN: 1476-5497. DOI: 10.1038/sj.ijo.0800852. URL: https://doi.org/10.1038/sj.ijo.0800852.

Mustillo, Sarah et al. (2003). “Obesity and Psychiatric Disorder: Developmental Trajectories”. In: *Pediatrics* 111.4, pp. 851–859. ISSN: 0031-4005. DOI: 10.1542/peds.111.4.851. eprint: https://pediatrics.aappublications.org/content/111/4/851.full.pdf. URL: https://pediatrics.aappublications.org/content/111/4/851.

Nye, Barbara, Spyros Konstantopoulos, and Larry V. Hedges (2004). “How Large Are Teacher Effects?” In: *Educational Evaluation and Policy Analysis* 26.3, pp. 237–257. DOI: 10.3102/01623737026003237. eprint: https://doi.org/10.3102/01623737026003237. URL: https://doi.org/10.3102/01623737026003237.

Park, M. H. et al. (2012). “The impact of childhood obesity on morbidity and mortality in adulthood: a systematic review”. In: *Obesity Reviews* 13.11, pp. 985–1000. DOI: 10.1111/j.1467-

789X.2012.01015.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-789X.2012.01015.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-789X.2012.01015.x.

Patacchini, Eleonora, Edoardo Rainone, and Yves Zenou (2017). "Heterogeneous peer effects in education". In: *Journal of Economic Behavior and Organization* 134, pp. 190–227. ISSN: 0167-2681. DOI: https://doi.org/10.1016/j.jebo.2016.10.020. URL: http://www.sciencedirect.com/science/article/pii/S0167268116302499.

Pereira, Mark A et al. (2005). "Fast-food habits, weight gain, and insulin resistance (the CARDIA study): 15-year prospective analysis". In: *The Lancet* 365.9453, pp. 36–42. ISSN: 0140-6736. DOI: https://doi.org/10.1016/S0140-6736(04)17663-0. URL: http://www.sciencedirect.com/science/article/pii/S0140673604176630.

Presler, Jonathan (2020). *You are who you eat with: Evidence on academic peer effects from school lunch lines*. Working Paper.

Qu, Xi and Lung-fei Lee (2015). "Estimating a spatial autoregressive model with an endogenous spatial weight matrix". In: *Journal of Econometrics* 184.2, pp. 209–232. DOI: 10.1016/j.jeconom.2014.08. URL: https://ideas.repec.org/a/eee/econom/v184y2015i2p209-232.html.

Renna, Francesco, Irina B. Grafova, and Nidhi Thakur (2008). "The effect of friends on adolescent body weight". In: *Economics & Human Biology* 6.3. Symposium on the Economics of Obesity, pp. 377–387. ISSN: 1570-677X. DOI: https://doi.org/10.1016/j.ehb.2008.06.005. URL: http://www.sciencedirect.com/science/article/pii/S1570677X08000312.

Robinson, Thomas N. et al. (Aug. 2007). "Effects of Fast Food Branding on Young Children's Taste Preferences". In: *Archives of Pediatrics & Adolescent Medicine* 161.8, pp. 792–797. ISSN: 1072-4710. DOI: 10.1001/archpedi.161.8.792. eprint: https://jamanetwork.com/journals/jamapediatrics/articlepdf/570933/poa70028_792_797.pdf. URL: https://doi.org/10.1001/archpedi.161.8.792.

Ryan, Ronan et al. (Jan. 2012). "Use of name recognition software, census data and multiple imputation to predict missing data on ethnicity: application to cancer registry records". In:

BMC medical informatics and decision making 12, p. 3. ISSN: 1472-6947. DOI: 10.1186/1472-6947-12-3. URL: http://europepmc.org/articles/PMC3353229.

Sacerdote, Bruce (May 2001). "Peer Effects with Random Assignment: Results for Dartmouth Roommates*". In: *The Quarterly Journal of Economics* 116.2, pp. 681–704. ISSN: 0033-5533. DOI: 10.1162/00335530151144131. eprint: http://oup.prod.sis.lan/qje/article-pdf/116/2/681/5375285/116-2-681.pdf. URL: https://doi.org/10.1162/00335530151144131.

— (2011). "Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" In: ed. by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann. Vol. 3. Handbook of the Economics of Education. Elsevier, pp. 249–277. DOI: https://doi.org/10.1016/B978-0-444-53429-3.00004-1. URL: http://www.sciencedirect.com/science/article/pii/B9780444534293000041.

Schwimmer, Jeffrey B., Tasha M. Burwinkle, and James W. Varni (Apr. 2003). "Health-Related Quality of Life of Severely Obese Children and Adolescents". In: *JAMA* 289.14, pp. 1813–1819. ISSN: 0098-7484. DOI: 10.1001/jama.289.14.1813. eprint: https://jamanetwork.com/journals/jama/articlepdf/196343/joc21985.pdf. URL: https://doi.org/10.1001/jama.289.14.1813.

Serdula, M.K. et al. (1993). "Do Obese Children Become Obese Adults? A Review of the Literature". In: *Preventive Medicine* 22.2, pp. 167–177. ISSN: 0091-7435. DOI: https://doi.org/10.1006/pmed.1993.1014. URL: http://www.sciencedirect.com/science/article/pii/S0091743583710145.

Sutherland, E. Rand (2014). "Linking obesity and asthma". In: *Annals of the New York Academy of Sciences* 1311.1, pp. 31–41. DOI: 10.1111/nyas.12357. eprint: https://nyaspubs.onlinelibrary.wiley.com/doi/pdf/10.1111/nyas.12357. URL: https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1111/nyas.12357.

Taylor, Erica D. et al. (2006). "Orthopedic Complications of Overweight in Children and Adolescents". In: *Pediatrics* 117.6, pp. 2167–2174. ISSN: 0031-4005. DOI: 10.1542/peds.2005-1832. eprint: https://pediatrics.aappublications.org/content/117/6/2167.full.pdf. URL: https://pediatrics.aappublications.org/content/117/6/2167.

Taylor, Valerie H. et al. (2013). "The impact of obesity on quality of life". In: *Best Practice & Research Clinical Endocrinology & Metabolism* 27.2. Complications of Obesity, pp. 139–146. ISSN: 1521-690X. DOI: https://doi.org/10.1016/j.beem.2013.04.004. URL: http://www.sciencedirect.com/science/article/pii/S1521690X13000365.

Thorpe, Kenneth E. et al. (2004). "The Impact Of Obesity On Rising Medical Spending". In: *Health Affairs* 23.Suppl1, W4-480-W4–486. DOI: 10.1377/hlthaff.W4.480. eprint: https://doi.org/10.1377/hlthaff.W4.480. URL: https://doi.org/10.1377/hlthaff.W4.480.

Trogdon, Justin G. et al. (2012). "State- and Payer-Specific Estimates of Annual Medical Expenditures Attributable to Obesity". In: *Obesity* 20.1, pp. 214–220. DOI: 10.1038/oby.2011.169. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1038/oby.2011.169. URL: https://onlinelibrary.wiley.com/doi/abs/10.1038/oby.2011.169.

Troiano, Richard P et al. (Nov. 2000). "Energy and fat intakes of children and adolescents in the United States: data from the National Health and Nutrition Examination Surveys". In: *The American Journal of Clinical Nutrition* 72.5, 1343s–1353s. ISSN: 0002-9165. DOI: 10.1093/ajcn/72.5.1343s. eprint: https://academic.oup.com/ajcn/article-pdf/72/5/1343s/23962147/1343s.pdf. URL: https://doi.org/10.1093/ajcn/72.5.1343s.

Vander Wal, Jillon S. (2012). "The relationship between body mass index and unhealthy weight control behaviors among adolescents: The role of family and peer social support". In: *Economics & Human Biology* 10.4. Obesity and the Family across the Life Course, pp. 395–404. ISSN: 1570-677X. DOI: https://doi.org/10.1016/j.ehb.2012.04.011. URL: http://www.sciencedirect.com/science/article/pii/S1570677X12000676.

Vigdor, Jacob and Thomas Nechyba (2007). "Peer effects in North Carolina public schools". In: *In Schools and the equal opportunities problem, ed. Ludger Woessmann and Paul*. MIT Press.

Whitlock, Evelyn P. et al. (2005). "Screening and Interventions for Childhood Overweight: A Summary of Evidence for the US Preventive Services Task Force". In: *Pediatrics* 116.1, e125–e144. ISSN: 0031-4005. DOI: 10.1542/peds.2005-0242. eprint: https://pediatrics.aappublications.

org/content/116/1/e125.full.pdf. URL: https://pediatrics.aappublications.org/content/116/1/e125.

Xu, Yilan and Linlin Fan (2018). "Diverse friendship networks and heterogeneous peer effects on adolescent misbehaviors". In: *Education Economics* 26.3, pp. 233–252. DOI: 10.1080/09645292.2017.1406458. eprint: https://doi.org/10.1080/09645292.2017.1406458. URL: https://doi.org/10.1080/09645292.2017.1406458.

Yakusheva, Olga, Kandice Kapinos, and Marianne Weiss (2011). "Peer effects and the Freshman 15: Evidence from a natural experiment". In: *Economics & Human Biology* 9.2, pp. 119–132. ISSN: 1570-677X. DOI: https://doi.org/10.1016/j.ehb.2010.12.002. URL: http://www.sciencedirect.com/science/article/pii/S1570677X10001103.

Jonathan Presler

110 Eggers Hall, Syracuse, NY 13244-1020 jlpresle@syr.edu
Website : jonathanpresler.com

Education

Syracuse University - Syracuse, NY
PhD in Economics, Expected June 2020

North Dakota State University - Fargo, ND
Non-degree Student in Economics and Mathematics, 2015

Washington University in Saint Louis - Saint Louis, MO
Bachelor of Arts, Architecture and English, 2009

Specialization and Interests

Fields: Public, Labor, and Urban Economics

Research Interests: Peer Effects, Education, Discrimination

Employment and Teaching Experience

Research Associate, *Center for Policy Research,* Syracuse University 2017-2020

Teaching Assistant, *Dept. of Economics,* Syracuse University 2016-2019
Economic Principles, Intermediate Microeconomics, Intermediate Math Microeconomics, Game Theory

Instructor for Economic Principles, *Dept. of Economics,* Syracuse University Summer 2017

Teaching Assistant, *Dept. of Mathematics,* North Dakota State University 2014-2015

Instructor at *More than Carpentry*, a job training program in Saint Louis, MO 2009-2012

Working Papers

"You Are Who You Eat With: Evidence on Academic Peer Effects in School Lunch Lines"
Job Market Paper.

"What Makes a Classmate a Peer? Examining Which Peers Matter in NYC Elementary Schools"
with William C. Horrace, Hyunseok Jung, and Amy Ellen Schwartz

"Obesity Peer Effects in NYC Elementary Schools"

"Is Discrimination Scarring? Effects of the September 11, 2001 Terror Attacks on New Entrants in the Labor Force"

Conference Presentations

"What Makes a Classmate a Peer? Examining Which Peers Matter in NYC Elementary Schools"

Association for Education Finance and Policy (AEFP) March 2019

Daniel Patrick Moynihan Summer Workshop in Education and Social Policy May 2019

"You Are Who You Eat With: Evidence on Academic Peer Effects in School Lunch Lines"

Association for Public Policy Analysis and Management (APPAM) November 2019

Fellowships and Awards

Syracuse University Graduate Fellowship	2015-2020
Syracuse University Graduate Assistantship	2016-2019
Maxwell School Summer Fellowship	2014-2019
Travel Grant, Syracuse University	2019

Skills

Stata
Python (SciPy, Numpy, Pandas)
LaTeX
MATLAB

Citizenship

United States

References

Amy Ellen Schwartz
(Primary Advisor)
Professor, Economics and Public Administration and International Affairs
Daniel Patrick Moynihan Chair in Public Affairs
Syracuse University
(315) 443-9362
amyschwartz@syr.edu

William C. Horrace
Distinguished Professor, Economics
Syracuse University
(315) 443-9061
whorrace@maxwell.syr.edu

Alfonso Flores-Lagunes
Professor, Economics
Syracuse University
(315) 443-9045
afloresl@maxwell.syr.edu

Donald H. Dutkowsky
(teaching reference)
Professor, Economics
Syracuse University
(315) 443-1918
dondutk@maxwell.syr.edu