An Empirical and Economic Analysis of High School Peer Effects

by

Andrew J. Hill

B.Sc., The University of Cape Town, 2003
B.Com.(Hons), The University of Cape Town, 2004
M.Com., The University of Cape Town, 2006
M.A., The University of British Columbia, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2013

Abstract

Parents are concerned about the influence of friends during adolescence. Using the gender composition of schoolmates in an individual's close neighbourhood as an instrument for the gender composition of an individual's self-reported friendship network, Chapter 2 of this dissertation finds that the share of opposite gender friends has a sizeable negative effect on high school GPA. The effect is found across all subjects for students over the age of sixteen, but is limited to mathematics and science for younger students. Self-reported difficulties getting along with the teacher and paying attention in class are important mechanisms through which the effect operates. The subject-specific effects for younger students and larger estimates for females in general are consistent with a gender socialization hypothesis in which young females conform to traditional gender roles in the presence of males.

Chapter 3 investigates the extent to which course repeaters in high school mathematics courses exert negative externalities on their course-mates. Using individual and school-specific course fixed effects to control for ability and course selection, it shows that doubling the number of repeaters in a given course (holding the number of course-takers constant) results in a 0.15 reduction in GPA scores for first-time course-takers. Further results suggest that the negative effect is only evident when the share of repeaters reaches a threshold of five to ten percent of the total number of course-takers.

Chapter 4 provides evidence that part-time work during high school affects the college attendance and labour market entry decisions of young adults: 8-10th grade students working more than five hours per week are less likely to attend college and more likely to enter the labour market upon high school graduation than other students. The part-time working behaviour of same-grade schoolmates is used as an instrument for individual part-time working behaviour.

Preface

All chapters of this dissertation are original, unpublished and sole-authored.

This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524.

Ethics approval under the project title "Adolescent Experiences and Early Adult Outcomes" was obtained through the Behavioural Research Ethics Board of the University of British Columbia (H10-02708).

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

I wish to thank Nicole Fortin for countless hours of advice and supervision. Your honest and thoughtful feedback greatly improved my work. Dissertation committee members Thomas Lemieux and Craig Riddell provided insightful suggestions and much encouragement. Thank you.

I am grateful to University of British Columbia (UBC) faculty members David Green, Joshua Gottlieb, Yoram Halevy, Vadim Marmer, Kevin Milligan, Dotan Persitz, Marit Rehavi, and Francesco Trebbi for valuable input along the way. Seminar participants at UBC, the Canadian Economics Association 2012 Conference and the European Association of Labour Economists 2012 Conference provided useful feedback on selected chapters.

UBC classmates, officemates and labmates were a constant source of constructive criticism and support. I benefited greatly from sharing the graduate school experience with residents of St John's College, many of whom are now great friends.

I also wish to acknowledge classmates and faculty members at the University of Cape Town, particularly Johannes Fedderke, who encouraged my pursuit of a graduate education in economics.

Funding from the Canadian Labour Market and Skills Researcher Network (CLSRN), the Skye Foundation, and the Len Baumann Trust is greatly appreciated.

Dedication

To my parents, for challenging me to think and showing me how to love;
To my brothers, for leading the way;
And to my friends, for seeking adventure.

Chapter 1

Introduction

Individuals are affected by their peers. In particular, friends and schoolmates influence the decisions high school students make, the activities they perform, and the attitudes they possess. This dissertation investigates the effects of three specific dimensions of high school peer group composition on a variety of academic and labour market outcomes. First, it asks how the gender composition of an individual's high school friendship network affects school performance; second, it explores whether course repeaters in high school mathematics courses exert negative externalities on their coursemates; and, third, it investigates how the part-time working behavior of schoolmates in the same grade affects an individual's own part-time working behavior, and, consequently, how this peer-induced variation in part-time working behaviour affects the individual's subsequent decisions and outcomes.

Adolescent experiences have the potential to affect both an individual's contemporaneous and subsequent outcomes in economically meaningful ways. The first two chapters of this dissertation focus on academic attainment and achievement in high school as the outcomes of interest. Essentially, these chapters characterize how peers enter the education production function. Economists are interested in understanding the determinants of education because of the overwhelming evidence of positive returns to education in both the labour market (Card, 1999) and various quality of life measures[1]. The outcome of interest in the third chapter is the school-leaving decision to enter the labour market or go to college. This is important because of the path dependence typically observed in labour markets[2]; the effects of either entering the labour market or going to college after graduating high school are likely to persist throughout an individual's life.

High school students belong to a variety of peer groups. Peers may be friends, neighbours, classmates or schoolmates, and there are several finer dimensions within each of these groupings that are likely to exert an independent influence on incentives and outcomes. Investigating composition effects for each of these peer groups provides unique challenges, but an overarching concern is dealing with non-random selection into the relevant peer group. The first two chapters of this dissertation introduce novel empirical strategies to overcome this problem; the first chapter uses the gender composition of schoolmates in an individual's close neighbourhood as an instrument for the gender composition of an individual's friendship network, and the second chapter extends an existing fixed effects strategy to longitu-

[1] See, for example, Oreopoulos and Salvanes (2009).
[2] See, for example, Keane and Wolpin (1997).

dinal transcript data in a way that allows for a variety of additional controls not feasible in previous analyses. The third chapter uses the familiar across-cohort within-school variation introduced by Hoxby (2000), applies it to a new outcome, and then goes one step further by using the estimated peer composition effect as the first stage in an instrumental variables' estimation.

An important feature of each of the analyses is exploring the mechanisms through which the estimated peer composition effects operate. This is important because although there are several policy instruments available to affect high school peer group composition directly, an understanding of the mechanisms is likely to both improve policy predictions and provide a more fundamental description of how peers affect incentives and behaviour.

This dissertation builds on the existing peer effects and economics of education literature by investigating a new set of peer composition effects. Deepening our understanding of how high school peer composition enters the education production function is important in the continuing quest to improve and ultimately optimize education policy.

Chapter 2

The Girl Next Door: The Effect of Peer Gender Composition on High School Achievement

2.1 Introduction

Peer effects are an important concern in education (Sacerdote, 2011). The questions of whether classrooms should be single-gendered or mixed and whether students should be tracked into classes based on ability are largely based on the premise that peer effects matter. The clear importance of these questions for parents and policy-makers has stimulated an extensive literature investigating peer effects in education. Exogenous peer effects[3] are difficult to identify because peer groups are typically selected; parents choose schools for their children, and teenagers choose their friends. This paper introduces an innovative identification strategy to overcome this selection problem and estimate exogenous peer effects within friendship networks, a peer group that is intensively selected on both observable and unobservable dimensions.

Individuals in school spend a large amount of time with their friends (Fuligni and Stevenson, 1995; Gager et al, 1999), and school friends are generally considered to exert considerable influence over each other's incentives and actions. Peer gender composition is also believed to have a considerable effect on teenage behaviour. Beyond the general questions around single-sex education, there are debates around how gender-specific social interactions affect the development of cognitive skills, the development of social and interpersonal skills, and the propensity to engage in risky behaviour. This paper takes a new step by considering the effect of the share of opposite gender school friends on academic achievement.[4,5] An instrumental variables approach overcomes the endogeneity of peer composition arising from selection into friendship groups: the gender composition of schoolmates in an individual's close neighbourhood induces plausibly exogenous variation in the gender composition of an individual's friendship network.

The study makes four contributions to the economics of education liter-

[3] Manski (1993) discusses the different types of peer effects. Endogenous peer effects operating through the actions and decisions of friends are not modeled in this paper. Estimating these peer effects is the focus of several papers in the economics of education literature (Bramoullé et al, 2009; Cooley, 2010; De Giorgi et al, 2010; Lin, 2010).

[4] The effect of the share rather than the number of opposite gender friends is modelled in this paper. This is motivated by considering that many adolescent activities are conducted as friendship groups rather than separately as friendship pairs. The paper does not preclude the potential for the number of opposite gender friends to have a separate effect, but the analysis of this is left for future work; the empirical strategy in this paper cannot identify a number of friends effect without several additional assumptions.

[5] Waddell (2012) investigates the role of opposite gender peer drinking on adolescent sexual behaviour.

ature. First, it shows that an increase in the share of opposite gender school friends reduces academic achievement. To the best of my knowledge, there is no prior evidence of a causal effect associated with the gender composition of an individual's friendship network on academic outcomes.[6] A one standard deviation increase in the share of opposite gender friends results in a 0.4 (one half of a standard deviation) reduction in GPA scores. This is approximately twice the mean female-male achievement gap of 0.2 found in the data, suggesting a moderately-sized effect.

Second, this paper studies potential mechanisms through which friendship network gender composition effects operate. It finds that an increase in the share of opposite gender friends increases the reported frequencies of difficulties getting along with the teacher and difficulties paying attention in class, two effects occurring within the classroom and strongly associated with negative academic outcomes. No convincing evidence is found to support channels operating outside the classroom. Lavy and Schlosser (2011) investigate mechanisms through which peer gender composition effects operate for same-grade schoolmates, finding that a higher share of female peers lowers the level of classroom disruption, improves relationships in the classroom, increases students' overall satisfaction in school, and lessens teachers' fatigue. The mechanisms identified in this paper complement those found in their study, and provide further channels through which peer gender composition affects academic achievement.

Third, this paper provides suggestive empirical support for gender socialization effects (Galambos et al, 1990). Specifically, it presents suggestive evidence that young teenage girls may be incentivized to fulfil gender stereotypes in the presence of boys. The negative effect for younger students caused by an increase in the share of opposite gender friends is most evident for females in subjects traditionally considered the domain of males, mathematics and science, and are not found in English and history. These results are aligned with a small set of papers that have investigated gender socialization effects in the economics of education literature.[7] The findings in this

[6]Poulin et al (2011) attempt to identify an effect using longitudinal variation in the composition of friendship networks, but they cannot account for time-varying changes in unobservable characteristics. Cipollone and Rosolia (2007) use a policy change in southern Italy to show that increasing the schooling attainment of boys increases the schooling attainment of girls, but they do not observe friendship networks.

[7]Schneeweis and Zweimüller (2011) use natural variation in the gender composition of adjacent cohorts within schools to show that females with a higher share of male classmates are less likely to choose male-dominated vocational school types, and, in experimental settings, Gneezy et al (2003, 2009) and Booth and Nolen (2012a, 2012b) find that the behaviour of females responds to the gender composition of the group in which they are

paper complement the existing gender socialization literature by introducing an analysis at the friendship level, a peer group in which socialization pressures are likely to be considerable given the desire of adolescents to be accepted by their friends.

And, fourth, this paper distinguishes a socially-based classroom gender composition effect from other effects operating in the classroom. Previous grade gender composition estimates have not been able to separate the effect arising from peer interactions in the classroom with correlated effects that may be responding to classroom gender composition (such as teaching style and disciplining behaviour).

Existing studies examining exogenous peer effects on academic achievement have used two broad approaches to overcome peer selection. The first approach exploits the institutional random assignment of peers. Sacerdote (2001), Zimmerman (2003), Stinebrickner et al (2006) and Carrell et al (2011) use the random assignment of students to different residences at the same post-secondary institution to investigate the effects of peer characteristics on various student outcomes. This type of random assignment typically only occurs at the post-secondary level, so although it provides compelling identification, its application is limited to a subset of questions. Furthermore, assignment is typically not across genders, limiting the potential for studying gender composition effects.[8]

The second approach exploits some form of conditional exogenous variation in the composition of peer groups. Typically, this approach relies on the peer group being defined so that its composition along the dimension of interest is exogenous conditional on a set of observable characteristics. The primary application uses variation in the composition of students across grades within the same school to identify exogenous peer effects for same-grade schoolmates, and is based on selection into schools being a function of school characteristics rather than cohort-specific deviations from these characteristics. It has been used to investigate exogenous peer effects along multiple dimensions: race (Angrist and Lang, 2004; Hanushek et al, 2009), domestic violence (Carrell and Hoekstra, 2010), home language (Friesen and Krauth, 2011), and, related to this paper, gender (Hoxby, 2000; Lavy and Schlosser, 2011; Schneeweis and Zweimüller, 2011).[9] These studies provide compelling evidence that grade composition matters, but cannot inform our

interacting.

[8]Whitmore (2005) uses the class size randomization of Project STAR to investigate the effects of gender composition in elementary school classrooms.

[9]Bifulco et al (2011) draws attention to some concerns when interpreting these findings, and Krauth (2011) provides a treatment effect interpretation of these effects.

understanding of composition effects for finer peer groups in which observables cannot control for selection. Nonetheless, my finding that friendship network effects operate within the classroom provides further support and justification for the above papers that focus exclusively on peer composition effects at the grade level.

The paper is organized in the following way. Section 2 introduces the empirical methodology, particularly the strategy to overcome the endogeneity in the gender composition of high school friendship networks. The subsequent data section is divided into two subsections. First, Section 3.1 provides empirical support for the claim that distance is a significant determinant of friendship, the hypothesis on which the identification strategy relies, and Section 3.2 describes the data in detail. Section 4 begins by reporting the primary findings, and then considers a variety of potential mechanisms through which the effect may operate. It also includes a set of results related to long-term effects of friendship network gender composition. The conclusion provides an overall interpretation of the results.

2.2 Empirical Methodology and Specification

Individuals in a social network are often defined by their type (such as gender, race or age). The number of types, distribution of types and relative distribution of types in an individual's friendship network may all affect high school performance. This paper focuses on the effects associated with other-type friendships.[10] Models of network formation typically impose some additional cost for behaving like or interacting with other types (Bisin et al, (2011). This study defines type by gender, and proposes the idea that other-type or opposite gender friendships are associated with an additional cost

[10]Homophily in friendship networks (the tendency to form same-type friendships) has been extensively modeled in the network formation literature (Currarini et al, 2009). This paper exploits an aspect of the friendship formation process to obtain exogenous variation in network composition, but otherwise abstracts away from network formation to consider the effects of homophily on academic achievement and other outcomes (rather than understand why homophily arises).

or negative input in the education production function.[11,12]

The academic achievement Y of individual i in grade g and school s is modelled as a linear function of a female indicator F, a vector of remaining individual and background characteristics X, the share of opposite gender friends O, and grade and school fixed effects D:

$$Y_{igs} = \alpha F_i + \beta X_i + \gamma O_{is} + \delta D_g + \theta D_s + \epsilon_{igs} \tag{2.1}$$

The model is estimated on the combined sample of males and females, imposing gender symmetry in the effect. An interaction term $F_i \times O_{is}$ is included when gender symmetry is relaxed. This specification shows that the effect of opposite gender friends goes in the same direction for males and females, but is plausibly of different magnitudes (although not statistically different). The advantage of the symmetry restriction is an increase in the statistical power of the estimation.

The gender composition of an individual's high school friendship network is likely to be correlated with a variety of unobservable characteristics that affect academic achievement. Candidates include parental inputs, personality traits and non-cognitive skills. For example, supportive parents may encourage participation in a wide range of extra-mural activities (resulting in more gender-balanced friendship groups) as well as greater academic achievement. This introduces correlation between the gender composition variable and the error term in the absence of perfect controls for parental inputs.

Peer gender composition is also measured with error. Beyond the attenuation bias associated with potential classical measurement error, friendship networks are constructed from self-reported friendship nominations. This process may yield systematically biased measures of friendship gender composition. For example, certain students may nominate opposite gender class-

[11] The debate on whether classrooms should be single-gendered or mixed can also be interpreted in the context of same-type and other-type interactions. The effect of single-gendered classrooms is equivalent to the effect of same-type (own gender) classmates. To the extent that the grade gender composition literature informs the debate on single-sex education, this specification would actually have an easier interpretation than those using female grade share given the impossibility of increasing or decreasing the female grade share for all individuals.

[12] The primary outcome of interest in this paper is academic achievement. Opposite gender friends may also affect non-cognitive development, and given the recognized importance of soft skills in subsequent life outcomes, there may be a trade-off between the negative effects on school performance and potential positive effects on non-cognitive skills. A preliminary analysis found no convincing evidence of an effect of opposite gender friends on measures of soft skills in the Add Health data.

mates as friends in an effort to appear more popular, generating correlation between the share of opposite gender friends and high school performance if the selection of these students is correlated with their academic achievement.

The omitted variables problem and potential measurement error results in least squares estimation of Equation 2.1 providing biased estimates of γ, the effect of the gender composition of school friendship networks on academic achievement. A source of exogenous variation in gender composition is required to obtain consistent estimates of the effect. This paper exploits variation in the gender composition of schoolmates in the close neighbourhood to obtain exogenous variation in the gender composition of school friends.

The idea behind the identification strategy is introduced by an example. Two females, Alice and Barbara, attend the same school and are in the same grade. They share identical individual and background characteristics, and both live next door to someone who attends the same school. Alice lives next door to a male, Charles, and Barbara lives next door to a female, Debbie. Alice catches the bus with Charles, they are friends, and, in addition, Alice has become friends with some of Charles' (mostly male) friends.[13] Barbara and Debbie also catch the bus together, they are friends, and Barbara is also friends with some of Debbie's (mostly female) friends. As a result, Alice has a larger share of male friends than Barbara. This arose by chance, as both Alice's and Barbara's parents did not know the gender of their neighbours' children when they chose where to live even though both their choices may have been based on a variety of other factors, such as income and the proximity to a good school.

The relationship between distance and friendship is central to the identification strategy. The probability of Alice being friends with her neighbour Charles needs to be greater than the probability of Alice being friends with someone identical to Charles who lives on the other side of town. A further requirement is that there needs to be variation in the gender composition of schoolmates across neighbourhoods; the strategy does not work if everyone in the school has one male neighbour and one female neighbour as this would not generate variation in the gender composition of friendship networks. These conditions are discussed further in the data and descriptive statistics section. The strength of the relationship between distance and

[13] It is well-established that the probability of friendship increases with the existence of mutual friends (see, for example, Goodreau et al, 2009). This channel is not actually required for the identification strategy to work. It does, however, strengthen it, as the gender composition of Alice's friends is not only affected by her friendship with her male neighbour, Charles, but also by her friendships with Charles' mostly male friends.

friendship is also supported by existing empirical evidence. Using the same data as this paper, Mouw and Entwisle (2005) find that friends are more than five times as likely as non-friends to live within 0.25km of one another after conditioning on several observable characteristics.

The exact instrument used in this paper is a weighted average of the gender composition of someone's nearest twenty same-school neighbours (the set denoted by $J20$ in the specification below). Each neighbour j is identified as being of opposite gender to i by the indicator O_{ijs}, and their contribution to the mean function is weighted by an inverse function of the distance between the relevant individual and the neighbour D_{ijs} (the nearer the neighbor, the greater the weight). The weighting function $w(D_{ijs})$ takes the form of the standard Epanechnikov kernel with bandwidth equal to the distance to the twentieth nearest neighbor D_{J20}.

Variations of the instrument based on the gender composition of a different number of the nearest neighbours or neighbours within a specified radius, as well as weighted and unweighted versions, were considered. The chosen measure was found to have the strongest results, although the result does not depend on the functional form of the instrument or the weighting function.[14] The below equations specify the first stage for investigating the effect of the share of opposite gender friends.

$$O_{is} = \alpha_0 F_i + \beta_0 X_{is} + \gamma_0 \frac{\sum_{j \in J20} w(D_{ijs}) O_{ijs}}{\sum_{j \in J20} w(D_{ijs})} + \delta_0 D_g + \theta_0 D_s + \nu_{igs} \tag{2.2}$$

$$w(D_{ijs}) = \frac{3}{4}\left(1 - \left(\frac{D_{ijs}}{D_{J20is}}\right)^2\right) \tag{2.3}$$

The causal parameter of interest γ in Equation 2.1 is identified if the gender composition of an individual's close neighbourhood is restricted to affect academic performance only through the gender composition of the individual's friendship network.[15] This claim is supported by two arguments.

First, the gender of an individual's neighbour is very likely random. Essentially, parents do not choose the locations of their homes based on the

[14] The estimates associated with other instruments and weighting functions are less precise, but very similar in magnitude. Weights simply reflect the empirical observation that the probability of friendship is inversely related to spatial proximity. Several of these results are reported in the appendix.

[15] In addition to this exclusion restriction, the monotonicity assumption required for instrument validity is very likely satisfied. An individual exposed to an increase in the share of opposite gender close neighbours is unlikely to decrease their share of opposite gender friends.

gender of school-going neighbours. Figure 2.1 reports the distribution of parent motivations for housing locations in the data. The gender of children in the neighbourhood was not an available option, but the age of children in the neighbourhood was, and was infrequently cited (around five percent). This suggests that locational choice is rarely influenced by the composition of neighbourhood children. Potential correlation with observables is investigated in the empirical section by performing balance tests in which the instrument is regressed on a set of individual and background characteristics; these are shown to have no systematic effect.[16]

Second, the friendship network is defined as a set of weak ties rather than strong friendships.[17] Neighbours may exert influence without being strong friends, but are likely to be included in a weak friendship network. Two individuals are defined as friends if either nominated the other individual as a friend rather than the mutual nomination that would be indicative of a strong friendship. (The procedure for nominating and matching friends is discussed in more detail in the data section.) An alternative way of thinking about this would be defining the weak friendship network as the set of schoolmates with whom friend-like social interactions occur, and interpreting the friendship network in the data as a proxy for this network.[18]

There are two primary threats to identification. The first is that the gender of schoolmates in the close neighbourhood may affect another dimension of an individual's friendship network, and this other dimension of the friendship network may affect achievement. Two primary candidates are friendship network age composition and number of friends. For example, a male with only female schoolmates in the close neighbourhood may be more

[16]Several papers (such as Angrist and Evans, 1998) find and exploit the fact that girls are more likely than boys to come from larger families, particularly for lower income families. This phenomenon does not appear to be sufficiently large to have an effect on the instrument in this paper. Furthermore, the conditioning on extensive controls for family structure in this paper is likely to alleviate potential biases arising through this channel.

[17]Granovetter (1973) is the seminal paper on the importance of weak ties. Several papers have recognized the independent importance of strong ties (Card and Guiliano, 2011). Patacchini et al (2012) find that both strong and weak friendships have a contemporaneous effect on high school grades, but only strong friendship effects persist in the long run. This supports using both strong and weak ties when analyzing short run education production. Lavy and Sand (2012) find that different types of friends in the classroom have different effects on learning outcomes for Israeli students transitioning from elementary to middle school, a younger population than that studied in this paper.

[18]Same-school neighbours are not required to be (at least) weak friends for the empirical strategy to be valid; it just requires that neighbour's gender be orthogonal to achievement if they are not friends.

likely to befriend older (or younger) teenagers in the neighbourhood, as well as have fewer friends, both of which may affect school performance. These hypotheses cannot be ruled out, but are challenged by the finding that the gender composition of close neighbours does not affect the age composition of school friends or the number of friends.

The second concern is that the friendship nomination process may be affected by the gender composition of the close neighbourhood. For example, low-performing individuals may disproportionately nominate opposite gender neighbours as friends (to appear more popular) without them actually being friends. The consequence of this would be measurement error in friendship network gender composition (arising from self-reporting bias) being correlated with the instrument in a way that biases results. The appendix provides a more formal discussion of this problem, and reports results that contest this hypothesis by showing that a constructed proxy for measurement error in the gender composition of the self-reported friendship network is uncorrelated with the gender composition of the close neighbourhood.

It is worth noting for exposition that the instrument would be invalid for investigating the race composition of high school friendship networks. This is because race and neighbourhood characteristics are not independent. An individual with mostly black same-school neighbours is likely to be different along a number of dimensions to another individual in the same school with mostly white neighbours, even if they share the same observable characteristics.

The negative effect of opposite gender friends on education production may arise from a variety of (non-exclusive) sources. Equation 2.1 can be interpreted as the reduced form of a simple linear model in which the share of opposite gender friends affects a vector of intermediate mechanisms W, which, in turn, affects academic achievement.[19]

$$W_{m,is} = \alpha_m F_i + \beta_m X_{is} + \gamma_m O_{is} + \delta_m D_g + \theta_m D_s + \eta_{m,igs} \quad \text{for } m = 1, \ldots, M \tag{2.4}$$

[19]The parameter of interest in the primary specification $\gamma = \sum_{m=1,\ldots,M} \gamma_{S,m}\gamma_m$, the effect of the share of opposite gender gender friends on achievement, is obtained by summing over each mechanism the products of the effect of that mechanism on the outcome and the effect of the share of opposite gender friends on that mechanism. This model does not allow feedback from the academic outcome to the mechanisms.

$$Y_{igs} = \alpha_S F_i + \beta_S X_i + \sum_{m=1}^{M} \gamma_{S,m} W_{m,is} + \delta_S D_g + \theta_S D_s + \phi_{igs} \tag{2.5}$$

Using the instrument for the share of opposite gender friends in Equation 2.4 identifies the parameter γ_m, the effect of peer gender composition on the candidate mechanism W_m. Evidence that $\gamma_m \neq 0$ (gender composition affects the mechanism) and $\gamma_{S,m} \neq 0$ (the mechanism affects achievement) indicates an operating mechanism W_m, while $\gamma_m = 0$ rejects a candidate mechanism for the peer gender composition effect (although the mechanism could still affect achievement). The parameters $\gamma_{S,m}$ cannot be identified without additional exclusion restrictions (we cannot identify the effects of the mechanisms on academic outcomes), so $\gamma_{S,m} \neq 0$ is inferred from non-causal correlations or taken from existing empirical literature. A series of equations taking the form of Equation 2.4 are estimated to investigate the set of mechanisms that may be in operation.

An understanding of the mechanisms through which peer gender composition effects operate is useful for deepening our understanding of adolescent behaviour, as well as potentially informing policy. Friendship networks cannot be directly regulated, but policy instruments may be available to act on the channels through which these effects operate. Furthermore, they may also inform out-of-sample predictions. The effects of manipulating peer gender composition beyond what was originally observed are predictable only if the mechanisms continue operating in the same way. This is particularly relevant given the evidence in Carrell et al (2011) that reduced-form peer effects estimates do not inform out-of-sample predictions.

This paper broadly groups candidate mechanisms into those operating within and outside the classroom. First, opposite gender friends may reduce the quality of classroom inputs in the education production function. Abstracting away from the friendship formation process, consider that maintaining (the utility associated with) friendships requires regular interactions. Outside of the classroom, high school students typically engage in a range of gender-specific activities.[20] These activities provide ample opportunities for the own gender interactions that characterize and maintain own gender friendships. The mixed gender classroom provides a relatively scarce opportunity for interactions with opposite gender friends. As a result, individuals

[20] Fuligni and Stevenson (1995) and Gager et al (1999) document typical time use of American teenagers in the 1990s. Fuligni and Stevenson find that studying, part-time work, extracurricular activities (such as sports), watching television and socializing with friends each consume between 10 and 20 hours per week.

in class may distract or be distracted by opposite gender friends more than own gender friends, reducing the quality of classroom inputs for individuals with a greater share of opposite gender friends.

Second, opposite gender friendships may reduce the quantity and quality of non-classroom inputs in the education production function, such as homework. For example, high school social activities may be more fun if opposite gender friends are present. This increases the time spent socializing, possibly at the expense of homework, and may reduce both the quantity and quality of homework produced the subsequent day. The idea behind this class of mechanisms is that, holding all else equal, opposite gender friends increase the marginal utility of leisure, resulting in an equilibrium characterized by leisure increasing (and homework decreasing) with an increase in the share of opposite gender friends.

The evidence in the empirical section supports the first set of mechanisms over the second set of mechanisms, suggesting that friendship network gender composition effects operate within rather than outside the classroom.

2.3 Data

This paper uses data from the National Longitudinal Study of Adolescent Health (Add Health). The Add Health is a school-based longitudinal study of a nationally representative sample of US adolescents who were in grades 7 to 12 during the 1994-1995 school year. The selected schools were representative of the US with respect to region of country, urbanicity, size, type, and ethnicity. Students in each school were stratified by grade and sex, and an average of 200 students were selected from each school to form the core sample. This sample was interviewed between April and December 1995 in the first wave of the study. The second wave of the study was conducted the subsequent year, and there have been two further in-home interviews, the most recent being in 2008. This paper primarily uses data from the first wave of the study. The fourth wave of the study is used to investigate the effect of the gender composition of high school friendship networks on long-term outcomes.

2.3.1 Distance and Friendship

There are two aspects of the data that are both unique to the Add Health study and of particular importance: spatial locations[21] and friendship networks. The Euclidean distance between individuals' homes can be calculated from spatial locations recorded in the data. These locations are reported in terms of X and Y-coordinates for each individual in a school relative to an arbitrary origin. Figure 2.2 provides an example of the spatial distribution of individuals in a small school in the Add Health data. Individuals are clustered in the centre of the map, presumably near the location of the school. The grey lines connecting nodes reveal the friendship networks within the school. Females are denoted by red circles and males by blue triangles.

Another example of the spatial distribution of individuals within a school is provided by Figure 2.3. This figure highlights the identification strategy. The friendship network for an arbitrarily chosen female individual in the school is shown in grey. Her nearest twenty schoolmates are circled. The gender composition of the individual's friendship network is instrumented by the distance-weighted gender composition of the circled individuals.[22] Six of the selected individual's eleven friends are included in the set of the twenty nearest neighbours; the proportion of matched friends within the twenty nearest neighbours is larger than the proportion of matched friends outside the nearest twenty neighbours, suggesting the role of distance in friendship formation for this individual.

Friendship networks are constructed using data from the first wave of the study. Surveyed individuals were asked to nominate up to five male friends and five female friends.[23] Individuals could leave nominations blank, but could not exceed the limit of five nominations per gender. These nominations were matched to other individuals in the same school using school rosters. Sixty-eight percent of friendship nominations by individuals in the sample are matched. Unmatched nominations typically arise from two sources:

[21] Spatial locations have been exploited in a small number of papers in the peer effects and education literature. Helmers and Patnam (2011), for example, incorporate spatial peer interaction into a production function of child cognitive development.

[22] The share of opposite gender friends and the distance-weighted share of opposite gender neighbours are both slightly above 0.3 for this individual.

[23] Some individuals were asked to nominate only one male and one female friend. The gender compositions of friendship networks computed for these individuals are interpreted as noisier proxies for the gender composition of the underlying friendship network. Results in which the sample is split by the number of friendship nominations or restricted to those with at least two friends show that restricting the number of nominations does not affect the primary conclusions.

nominations to individuals in another school or nominations using names that could not be matched on the school roster (for example, nominations using nicknames). The effect investigated in this paper is for the gender composition of matched friends.

Figure 2.4 plots the distribution of the share of matched friendship nominations (ratio of matched nominations to total nominations) by gender, showing that nominations are fully matched for about forty percent of individuals in the sample. The share of matched friendship nominations is orthogonal to the gender composition of schoolmates in the close neighbourhood; the correlation coefficient between the instrument and the share of matched nominations is -0.02. This ensures that the empirical strategy deals with potential bias introduced by the matching process (in addition to other biases), such as nominations from weak students being less likely to be matched.

Friendship networks can be defined in a variety of ways using these data. This paper primarily defines any nomination or receipt of nomination as a friendship, generating a network of weak friendships. This is the preferred definition of friendship as it includes the largest set of potential influences. Results with alternative definitions of friendships are reported in the appendix to show that the findings are generally robust to the definition of friendship that is chosen. Effects associated with the gender composition of nominated friends and the gender composition of nominating friends (individuals from whom friend nominations are received) are also considered.

The friendship nomination and sampling processes are graphically illustrated in Figure 2.5. The first and second panels show a hypothetical school with nine students of which five are randomly sampled to complete the detailed survey. The third and fourth panels show the friendship nomination process, including D nominating an individual who could not be matched to an individual in the school. The fifth panel drops the individuals who were not sampled to show the observed school friendship network. Note that this network distinguishes the direction of nominations from which alternative types of friendships can be defined. The sixth panel shows the weak friendship network used in the analysis. I is dropped as peer gender composition is not well-defined for an individual with no matched friends. The direction of nominations is no longer distinguished as any nomination defines a weak friendship.

Each friendship nomination generates a dyadic pair. Table 2.1 describes the 13,142 friendship pairs generated by matched nominations in the analyzed sample. Reciprocated nominations appear as two observations in these data, but the reported interactions may differ as they depend on the response of the surveyed individual. The first row of the table provides simple evidence that distance is a significant determinant of friendship. The mean distance between friends in a school is significantly smaller than the mean distance between two randomly-drawn individuals in a school.[24]

Males nominate a higher share of opposite gender friends than females. Slightly fewer than forty percent of friendship pairs go to each other's home, about half meet after school, and over forty percent spend time together during the weekend. Forty-four percent of friendship pairs in the data "talk about a problem"; interestingly, but perhaps unsurprisingly, this activity is much more likely in friendship pairs nominated by females. The most common activity among friendship pairs is talking on the phone, which occurs in about sixty percent of friendships in the data.

Table 2.2 provides evidence that the distance between individuals affects the intensity of their social interactions. This table reports results from regressing binary indicators of each of the interactions with nominated friends discussed above on the distance between the two individuals, the gender (and relative gender) of the nominated friend, and a vector of individual characteristics.[25]

Conditional on being friends, individuals are more likely to go to a friend's house, meet after school and spend time together on the weekend if they live closer together. Distance also affects the likelihood of talking on the phone despite this activity being largely independent of spatial proximity. Under the hypothesis that talking on the phone is correlated with the strength of the friendship, this provides suggestive evidence that friends who are geographically proximate have stronger relationships. Females are less likely to meet after school or during weekends and more likely to talk on the phone or about a problem with their nominated friends. All interactions are less likely with opposite gender friends. Given that interactions are more probable with close neighbours, and that these interactions vary by the gender of the nominated friend, the variation in friendship network gender composition induced by the gender composition of close neighbours

[24]This comparison does not control for characteristics that may be correlated with the distance between individuals, such as the probability of being the same race. For example, in a school neighborhood in which everyone west of the school is white and everyone east of the school is black, the mean distance between friends may be smaller just because friends are more likely to be of the same race. The evidence that distance affects friendships provided by Mouw and Entwisle (2005) conditions on several observable characteristics including race.

[25]This analysis is only possible within friendship pairs as individuals were not asked about their potential interactions with all other individuals in the school; this would be prohibitively costly in terms of data collection.

will affect an individual's weekly interactions in a meaningful manner.

2.3.2 Descriptive Statistics

The Add Health Wave I dataset samples 20,769 individuals from 80 schools[26]. Individuals without core demographic information, GPA scores and spatial locations are dropped.[27] The gender composition of an individual's friendship is only well-defined when the individual has at least one friend. Individuals with no matched friends are therefore dropped from the data. Finally, schools in which fewer than twenty students remain in the sample after this process are also dropped. This leaves a final sample of 8,435 individuals from 76 schools.[28]

Descriptive statistics of the variables used in the paper are reported in Table 2.3. The primary outcome variable considered in the paper is an overall mean of self-reported grades across four subjects: English, Mathematics, Science and History. Letter-grades are converted to numerical grades by assigning fours to As and ones to Ds or lower. The overall mean grade is computed by equally weighting all non-missing subject grades for each individual.

Figure 2.6 shows the full distribution of overall high school grades by gender. There are two striking differences in the grade distributions for males and females. First, the female distribution is centred at a higher grade (the mode is 3 for females and 2.5 for males), and, second, there is a spike in the distribution for females at scores of 4 (As in all subjects). The mass of females scoring at the top of the distribution is also noted by Fortin et al (2011) and Bertrand and Pan (2011). Measurement error in self-reported GPA is computed using transcript GPA scores for a subset of the sample for whom these are available.

The next set of variables describes the gender composition of high school friendship networks. Recall that individuals must be linked to at least one other individual to be included in the data. The type of network for which

[26] About half of the 80 schools are school pairs. School pairs are created to represent one school when the sampled high school does not have lower grades (such as ninth grade). This is done by probabilistically matching high schools without lower grades to one feeder school in the area based on the likelihood with which students come from the set of candidate feeder schools.

[27] Individuals with and without GPA scores and spatial locations appear similar along observable dimensions, reducing the concern of sampling bias.

[28] The mean number of students per school is 422, the median is 100, and the smallest and largest schools have 20 and 1515 students, respectively. The sample includes two Catholic schools and five private schools.

the variable mean is calculated is denoted in parentheses after the variable. The weak friendship network in which any nomination is considered to establish a friendship is the primary focus of this paper, but a description of the strong friendship network in which reciprocated nominations define a friendship is included for comparison purposes.

The mean share of opposite gender friends is slightly below 0.4 in the weak friendship network. This confirms the tendency towards nominating friends of the same gender. The distribution of the mean share of opposite gender friends is plotted in Figure 2.7. This is done for both the full sample and for a restricted sample in which only individuals matched to at least two friends are included. There are mass points at one and zero in the full sample. This is because about sixty percent of the sample were only matched to same-gender friends or were only matched to one friend. The distribution for those with at least two friends shows the modal share of opposite gender friends to be 0.5, but retains the strong feature of a tendency towards same gender friendships.[29]

The share of opposite gender friends in strong friendship networks (defined by reciprocated nominations) is considerably lower than that found in weak friendship networks. This shows that opposite gender friends are less likely to reciprocate nominations than friends of the same gender.

Exogenous variation in the neighbourhood gender composition of schoolmates is used to obtain identifying variation in the gender composition of friendship networks. The exact instrument is based on the distance-weighted gender composition of each individual's nearest twenty neighbours (in the same school). The next row of Table 2.3 shows that the mean share of opposite gender close neighbours is very close to the expected 0.5 in the full sample and for males and females.

The distribution of this variable is important for two reasons. First, there is a concern that all individuals may have a similar share of male and female schoolmates in their close neighbourhoods. Under this scenario, even if distance were a significant determinant of friendship, it would not generate variation in the gender composition of friendship networks. The distribution of the weighted gender composition of the nearest twenty neighbours for the full sample, as well as by gender, is plotted in Figure 2.8, confirming variation in neighbourhood gender composition.

Second, we can test whether the distribution of the share of opposite

[29]The appendix reports estimation results for the restricted sample that are similar to those for the full sample. This provides evidence that results are not driven by individuals only matched to one friend.

gender neighbours is consistent with a data-generating process in which location decisions are independent of the gender composition of the close neighbourhood. Parents favouring gender-balanced neighbourhoods, for example, would be evident if the standard deviation of the share of opposite gender neighbours were smaller than a comparable series based on random location decisions (although the mean share of opposite gender neighbours may be unaffected and remain 0.5). This would be a concern if these parents also systematically affect their children's school performance. We perform a Kolmogorov-Smirnov test for equality of distributions on the observed measure of neighbourhood gender composition and a constructed pseudo-measure of neighbourhood gender composition in which gender is randomly reassigned to households.[30] The null hypothesis of equality of distributions cannot be rejected, supporting the claim that location decisions and the neighbourhood gender composition are orthogonal.

Self-reported measures describing the extent to which individuals have behavioural troubles at school are used to support the hypothesis that parts of the socialization effects identified in this paper operate within the classroom. Ordinal scales for these variables are converted to numerical scales by assigning zeroes to responses of "Never" and fours to responses of "Every day", the most frequent of five categories. Males are more likely than females to report having both troubles getting along with the teacher and paying attention in school. Both types of troubles occur within the classroom. They are infrequently reported. Considering behaviours outside the classroom, males are more likely to report trouble completing homework and interacting with other students.

Socialization effects may also operate outside the school environment if opposite gender friends affect the marginal utility of leisure. The number of friends' variable in Table 2.3 corresponds to the number of matched friends in the weak friendship network. Conditional on being matched to at least one other individual in the data, individuals have an average of 2.6 friends. This is slightly greater for males than females. The reported number of friends is likely to be less than the true number of friends in an individual's school friendship network. This is due to the imposition of a maximum number of nominations and some nominations being unmatched[31]. The composition

[30] This assignment is based on birth months being odd or even, which is assumed truly random.

[31] It is probable that unmatched nominations occur more frequently within schools in which individuals were less likely to be sampled. Results (not reported) in which the sample is limited to the set of schools in which all individuals were sampled reveal the same pattern of effects.

measure used in this paper is therefore interpreted as a proxy for the true gender composition of the friendship network.

Over one half of the sample report being in a relationship in the last 18 months. Females are more likely to report a previous romantic relationship.[32]

The gender composition of an individual's friendship network may also affect smoking and drinking behaviour (see, for example, Clark and Loheac, 2007), and smoking and drinking may affect academic achievement. About one quarter of the sample report smoking at least one day in the past month, and this does not differ by gender. Males are more likely than females to report being drunk at least one day in the past year; just under one third of males and just over one quarter of females report this behaviour. Various other measures of smoking and drinking behaviour were also considered; they convey essentially the same information as these measures.

Finally, we are interested in the persistence of gender composition effects. The long-term outcomes of subsequent-year GPA, graduated high school, attended college and ever married are taken from the fourth wave of the Add Health study in which individuals are asked about their educational and relationship histories. This wave was conducted in 2008 when individuals were 24 to 32 years old. Ninety-five percent of the sample graduates high school and sixty-eight percent of the sample completes at least one year of post-secondary education, the definition of attending college used in this paper. The probability of males attending college is ten percentage points lower than that for females. Almost one half of the sample report being married (or previously being married).

Core demographic characteristics reported in Appendix Table A.1 provide information on the composition of the sample. Just over half the sample is white and one fifth of the sample is black.[33] Ninety percent of the sample is born in the US and the mean age is 16, corresponding approximately to the tenth grade. The means of all other control variables are also reported. These include variables describing parent characteristics, home language,

[32]The behaviours associated with "being in a romantic relationship" are likely to vary considerably across individuals in high school. Finer measures of relationship-type behaviour would be required to obtain a fuller picture of the potential effects of peer gender composition.

[33]The Add Health study over-sampled black students. Sample weights are not used in this analysis because their application to friendship pairs is ambiguous; it is not clear how friendships with over-sampled individuals should affect the gender composition of that individual's friendship network. At the same time, it is noted that results are insensitive to the inclusion of sample weights at the estimation stage, although they do affect the precision of some of the estimates.

household income, family structure and grade repetition[34].

2.4 Results

The objective of this paper is to empirically investigate the effect of gender homophily in high school friendship networks on academic achievement. Opposite gender friends are shown to have a negative effect on high school performance. Subsequent results explore whether the effect differs by gender, across school subjects, and by age. Finally, results investigating the mechanisms through which peer gender composition affects achievement are reported. Errors are clustered at the school level throughout the analysis.[35]

The first two columns of Table 2.4 report OLS results from regressing GPA scores on friendship network gender composition measures. The first column reports results from the model that imposes gender symmetry, and the second column reports results from the model that includes a gender interaction on the explanatory variable of interest.[36] Results in these columns show that males with a higher share of opposite gender friends are associated with better school performance, while the correlation for females (the sum of the coefficients) is close to zero. As discussed in the methodology section, this correlation could arise from unobserved parental inputs, bias in self-reported friendship nominations or other unobserved characteristics. The causal effects subsequently reported are consistently of the opposite sign. This is consistent with individuals with large shares of opposite gender friends being positively selected, and confirms the importance of an empirical strategy to overcome the endogeneity bias in the gender composition of friendship networks.

The third and fourth columns of Table 2.4 reports the direct effect of the instrument on academic achievement. The gender composition of same-school neighbours is considered exogenous, so these estimates have a causal interpretation. An increase in the share of opposite gender schoolmates in

[34]Grade repetition controls are included throughout the paper to control for potential differences in friendship network formation and effects for repeating students. Results are similar if these students are excluded from the analysis.

[35]This is conservative given the inclusion of school fixed effects and that the level of variation exploited in this paper is cross-sectional and at the individual level. The precision of estimates obtained without clustering are generally very similar.

[36]Note that the effect for males in the gender interaction model is the coefficient on the share of opposite gender friends, and the effect for females is the sum of this coefficient and the coefficient on the interaction. Only the first coefficient (the male effect) and the sum (the female effect) are reported in subsequent tables for ease of interpretation.

the close neighbourhood reduces school performance.[37] The fourth column shows that the sign of the effect does not differ by gender, justifying the statistically more powerful gender-symmetric model. This paper interprets these effects to be operating through weak friendship networks.

Goux and Maurin (2007) exploit the institutional environment in France to conclude that an adolescent's outcomes in junior high school are strongly influenced by (and not just correlated with) the performance of neighbours. Foley (2012) finds that neighbourhoods affect university participation. The reduced form result in this paper supports the hypothesis that close neighbours matter. It provides a potential mechanism for these findings and suggests that part of the neighbourhood effect may be driven by the set of close neighbours that are in an individual's weak friendship network.

The primary causal estimates from the instrumental variable (IV) specification are reported in Table 2.5. The top panel reports results for the model in which gender symmetry in the effect is imposed. The bottom panel reports results for the model that relaxes gender symmetry, confirming that the effect is the same sign and not statistically different for males and females.

Results for the preferred model include the coefficient of interest and F-statistic from the first stage, as well as a weak IV-robust confidence interval (Moreira and Pan, 2001). The first stage coefficients are precise with reasonably-sized F-statistics across specifications. The weak IV-robust confidence intervals allow potential nonnormality in GMM statistics arising from weak identification (as discussed in Stock et al, 2002). Andrews and Stock (2005) advocate inference based on this confidence interval given its robustness properties; opposite gender friends affect academic achievement if the confidence interval is bounded away from zero.

The negative effect of opposite gender friends is evident both without (first column) and with (second column) controls. The point estimate in the second column is negative and the corresponding weak IV-robust confidence interval does not include zero. The standard deviation of the share of opposite gender friends is 0.4, so the estimate of -1.0 means that a one standard deviation increase in the opposite gender friend share causes a 0.4 decline in GPA. In comparison to other variables in the model, this is twice the coefficient on the female indicator. This is interpreted as a moderately-sized effect.

[37] This reduced form analysis is also performed on the original sample before individuals with no matched friends are dropped. The estimated coefficient of -0.063 (0.035) is not statistically different from the estimated coefficient of -0.132 (0.054) reported in the table.

Results in the bottom panel provide suggestive evidence that the effect is larger for females than males. The magnitude of the effect for females is consistently around three times larger than the effect for males (with the caveat that a lack of power prevents statistically distinguishing these estimates).

The subject-specific results reported in the third and fourth columns of Table 2.5 suggest that the overall effect is larger in mathematics and science than English and history.[38] This is shown to be driven by the absence of an effect in English and history for individuals under the age of sixteen in Table 2.6. The negative effect of opposite gender friends for girls in their early teenage years in traditionally male-dominated school subjects is consistent with gender socialization effects in the existing developmental psychology and economics of education literature: adolescent females conform to traditional gender roles (such as not doing well in mathematics and science) in the presence of males. The reduced socialization effects on males are also consistent with previous studies and the commonly-held view that females may have more to gain from reduced gender socialization pressures (such as single-sex classrooms).

Results in Table 2.6 are constructed by splitting the sample at the age of sixteen. As discussed above, the effect for younger individuals is limited to mathematics and science, while the effect for older individuals is larger and equally prevalent across all school subjects. This table also provides insight into the operation of the instrument. Older high school students in the US are likely to be more mobile; the driving age in most states is sixteen. This suggests that the instrument is likely to be less effective for older students as geographic distance becomes a less important determinant of friendship. The smaller and less precise first stage coefficients for older students are consistent with this hypothesis.

The mechanisms through which friendship network gender composition affects academic achievement are important for understanding how high school peers affect incentives and actions. An investigation of these mechanisms provides a fuller description of the education production function and indirectly informs policy related to gender composition in the school environment. Results in Tables 2.7 and 2.8 explore possible channels through which the gender composition of friendship groups affects school performance. Table 2.7 considers a set of school behavioural troubles and Table 2.8 investigates social behaviours.[39]

[38] Subjects are not considered individually as there are some individuals with missing subject GPA scores and grouping increases the sample sizes.

[39] A more direct mechanism may operate through the academic ability of peers. As a

Individuals were asked the frequencies with which they have troubles getting along with the teacher, paying attention in class, getting homework done and relating to other students on a five-point scale (from zero to four with four being the most frequent). The first row of Table 2.7 (OLS coefficient in GPA regression) reports strong negative correlations between the frequencies of these troubles and GPA scores. Results in the respective columns show that an increase in the share of opposite gender friends increases the frequency of trouble getting along with the teacher and paying attention in class, while the effects on trouble getting homework done and trouble with other students are similarly positive, but imprecisely measured.[40]

The significant effects on the first two school troubles are sizeable. A one standard deviation increase of 0.4 in the opposite gender friend share is associated with a 0.5 increase in the reported frequencies getting along with the teacher and paying attention in class. These are large effects given both these series have approximate means and standard deviations of one.[41]

The set of mechanisms in Table 2.8 relate to effects most likely occurring outside the classroom. The only significant gender composition effect among this set of mechanisms operates through the probability of being in a romantic relationship. An exogenous increase in the share of opposite gender friends increases the likelihood of reporting being in a romantic relationship in the past 18 months. Being in a romantic relationship is negatively associated with GPA scores when included independently of the other mechanisms (not reported), but is essentially uncorrelated with GPA scores when all four social mechanisms are included (first row). The existing literature finds a

result of the gender gap in school performance, females with a larger share of opposite gender friends will, on average, have a larger share of less academically able friends. This, in turn, may reduce the school performance of these females. No empirical support for this hypothesis was found; the gender composition of the close neighbourhood had no effect on the ability composition of friends for males and females separately, as well as the combined sample.

[40] The gender composition of friends taking the same classes is highly correlated (0.7) with the gender composition of all friends. This correlation is computed for a small subsample of individuals for whom indices indicating the extent of shared courses with schoolmates was available. It supports using the measure of gender composition based on all friends when investigating friendship network gender composition effects inside the classroom.

[41] Note that the correlations between the reported troubles and GPA scores are not causal. There are many factors correlated with these troubles and not included in the set of controls that may affect school performance. An alternative interpretation of this result is considering the first two reported school troubles as proxies for general classroom behaviour and that the gender composition effect operates more generally through classroom behaviour.

correlation but does not make a strong case for a causal relationship between romantic relationships or sexual activity and high school achievement (Halpern, 2000; Sabia, 2007). The fifth column shows that the behavioural troubles found in the classroom are not evident in the home; individuals with greater shares of opposite gender friends are not significantly more likely to have had serious arguments with their mothers in the four weeks preceding the survey.

Results in Table 2.9 provide further analysis of the three potential mechanisms for which precise estimates were obtained. The increased troubles in the classroom and probability of being in a romantic relationship are strongly evident for older high school students, but not for younger students. This is consistent with the negative effect in mathematics and science for younger students being a consequence of broader gender socialization effects rather than any of the direct effects considered here.

The gender composition of an individual's friendship network is likely to fluctuate during high school as individuals move in and out of friendship groups. The Add Health study included friendship nominations during both the initial in-school interview (in which a brief survey was admitted to all individuals in each sampled school) and the subsequent Wave 1 interview (conducted on a subset of individuals at each school). The correlations between the friendship network gender compositions are positive and significant, varying between 0.3 and 0.5. This correlation confirms the presence of an enduring component in peer gender composition, suggesting the potential for long run effects. Note that Wave 1 friendship nominations generate the opposite gender friend shares used in this paper as outcomes and spatial locations are obtained from this wave.

Results in Table 2.10 consider the effect on four long-term outcomes of interest: subsequent-year GPA scores, graduated high school, attended college and ever married. These outcomes are measured in Wave 4 of the Add Health study conducted in 2008; the sample is substantially smaller due to attrition.[42] Estimates suggest an imprecise, negative effect of the opposite gender share of friends on the three academic outcomes, and a significant positive effect on the probability of ever being married. The latter finding is not surprising given that an increase in the share of opposite gender friends increases the probability of being in a romantic relationship in high school, and the likely correlation between this and ever being married. These results suggest persistence in the peer gender composition effects associated with high school friendship groups.

[42] Results using imputed values are similar to those reported here.

The identification strategy relies on the gender composition of same-school close neighbours being orthogonal to all factors affecting academic achievement other than the gender composition of weak friendship networks. Correlation with unobservables is inherently untestable, but correlation with observables is investigated graphically and statistically by regressing the instrument on observable characteristics.[43] Thinking about the gender composition of an individual's close neighbourhood as a random treatment, this loosely investigates whether assignment was truly random.

Figure 2.9 plots the mean share of opposite gender close neighbours for the four categories of mother's and father' education, as well as four bins of annual household income. Plots suggest the absence of a systematic pattern in the gender composition of close neighbours; means are 0.5 across categories for each variable for both males and females. Plots for the share of white same-school neighbours are included for comparison, and, as expected, means are no longer constant across categories for each variable. This indicates that the share of white same-school neighbours is correlated with the selected socio-economic indicators, and could not be interpreted as a random treatment.

Appendix Table A.2 confirms that the instrument is balanced across observable individual characteristics. The correlation with the female indicator variable reflects mild gender imbalance in the sample, while remaining correlations are not systematic. There are some significant correlations in the gender-specific results in the second and third columns, but these are also not systematic.

The fourth column reports correlations between the share of white same-school neighbours and individual characteristics. This column is included for comparison purposes. The significant negative correlations with mother's education and public assistance receipt confirm that the race composition of close neighbours is not balanced; white neighbourhoods differ systematically from non-white neighbourhoods along observable dimensions. This suggests that these neighbourhoods may also differ along unobservable dimensions under the assumption of correlated selection on observables and unobservables, indicating the invalidity of this approach to investigating the race composition of friendship networks.

Remaining robustness checks are reported and discussed in the appendix. These include showing that the pattern of results is not affected by the functional form of the instrument, the school-specific nomination process or

[43] Altonji and Taber (2005) combine estimated selection on observables with various assumptions about selection on unobservables to bound estimates of a treatment effect.

the chosen definition of friendship.

2.5 Conclusion

Parents are typically concerned about the composition of their high school children's peer groups. High school years are considered to be particularly formative, and the general view is that friends exert considerable influence over their peers during this period. This paper supports this hypothesis by finding that an increase in the share of opposite gender friends causes a reduction in high school academic achievement. The magnitude of the effect is moderate: a one-standard deviation increase in the share of opposite gender friends is associated with a 0.4 reduction in GPA.

An abundance of existing papers in the economics of education find that classroom gender composition matters (Hoxby, 2000; Lavy and Schlosser, 2011; Schneeweis and Zweimüller, 2011). This paper provides evidence that the gender composition of school friends plays an important role. Part of the effect is shown to operate within the classroom environment through increased troubles getting along with the teacher and paying attention in class. These mechanisms are similar in type to those proposed by Lavy and Schlosser (2011) for the positive effect of female classmates. They find that female classmates lower the level of classroom disruption and improve relationships in the classroom through changes in classroom composition and not individual behaviour. Taken together with the results in this paper in which individual behaviour *is* affected by friendship network gender composition, evidence is increasingly supportive of a general hypothesis in which social interactions between genders affect classroom education production.

The negative effect of opposite gender friends for younger students is found in mathematics and science and not in English and history. Interpreted alongside existing experimental and non-experimental studies, the exclusivity of the effect in the traditionally male-dominated science subjects, and the consistently larger estimates for females, are supportive of a gender socialization hypothesis in which young adolescent females conform to traditional gender roles in the presence of males. Given the sub-optimality of socially-constructed constraints on achievement, this suggests there may be efficiency gains from policy interventions limiting these effects.[44]

[44] One example of such a strategy may be single-sex mathematics and science classrooms. According to the National Association for Single Sex Public Education (www.singlesexschools.org), the number of coeducational schools offering single-sex classrooms has increased from around a dozen in 2002 to 390 in the 2011-2012 school year.

This study can also be interpreted in the context of the continuing debate around single-sex and mixed gender education (Halpern et al, 2011). The difficulties getting along with the teacher and paying attention in class caused by an increase in the share of opposite gender friends are unlikely to occur in single-sex classrooms that exclude opposite gender friends. These effects on classroom behaviour are also indicative of the type of channels through which achievement may be affected by reorganizing classroom gender composition, and suggest that effects may operate through more than better matches between teaching styles and the gender composition of the class.[45]

[45]There are several other factors not captured by this analysis that would need to be considered for an evaluation of single-sex education, but the evidence in this paper contributes to the debate in the absence of random assignment to single-sex and mixed gender classrooms in North America. Jackson (2011) exploits quasi-random assignment to single-sex and mixed gender high schools in Trinidad and Tobago to investigate the effects of single-sex education.

2.6 Figures

Figure 2.1: Parent motivation for housing location

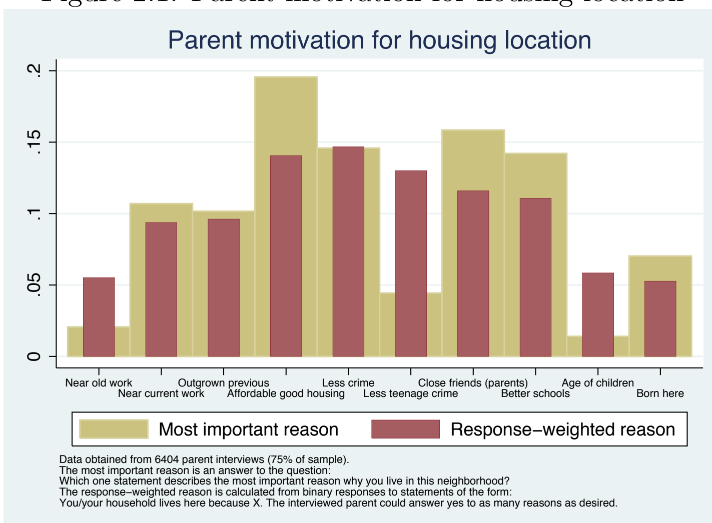


Figure 2.2: Spatial distribution within selected school 1

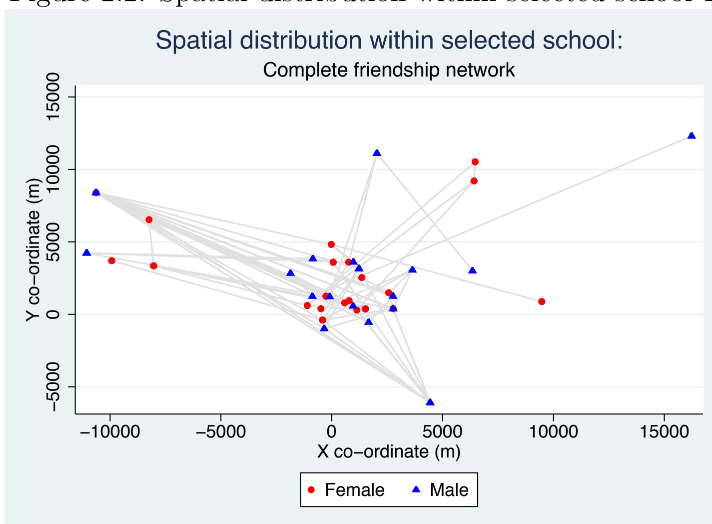


Figure 2.3: Spatial distribution within selected school 2

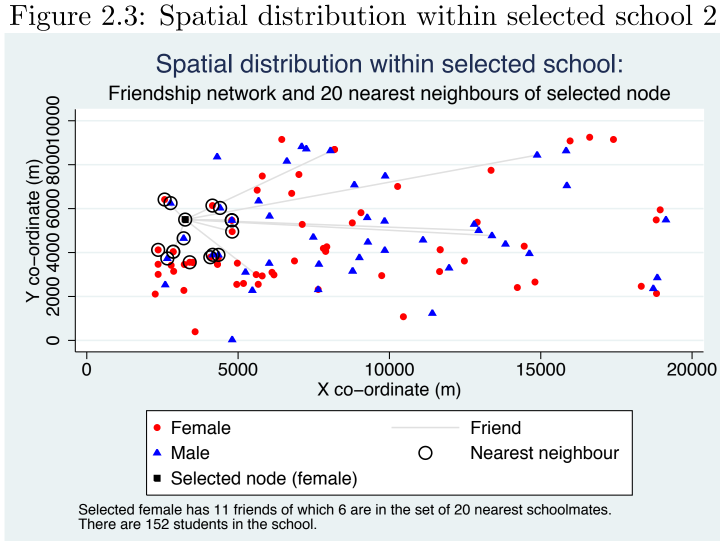


Figure 2.4: Share of matched friendship nominations

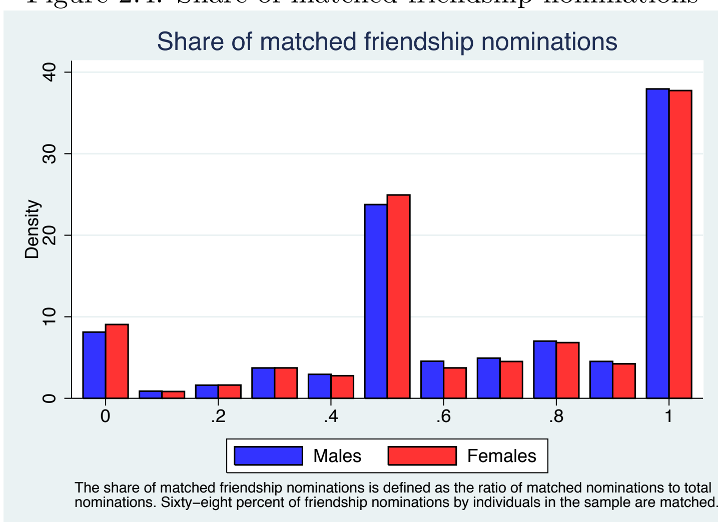
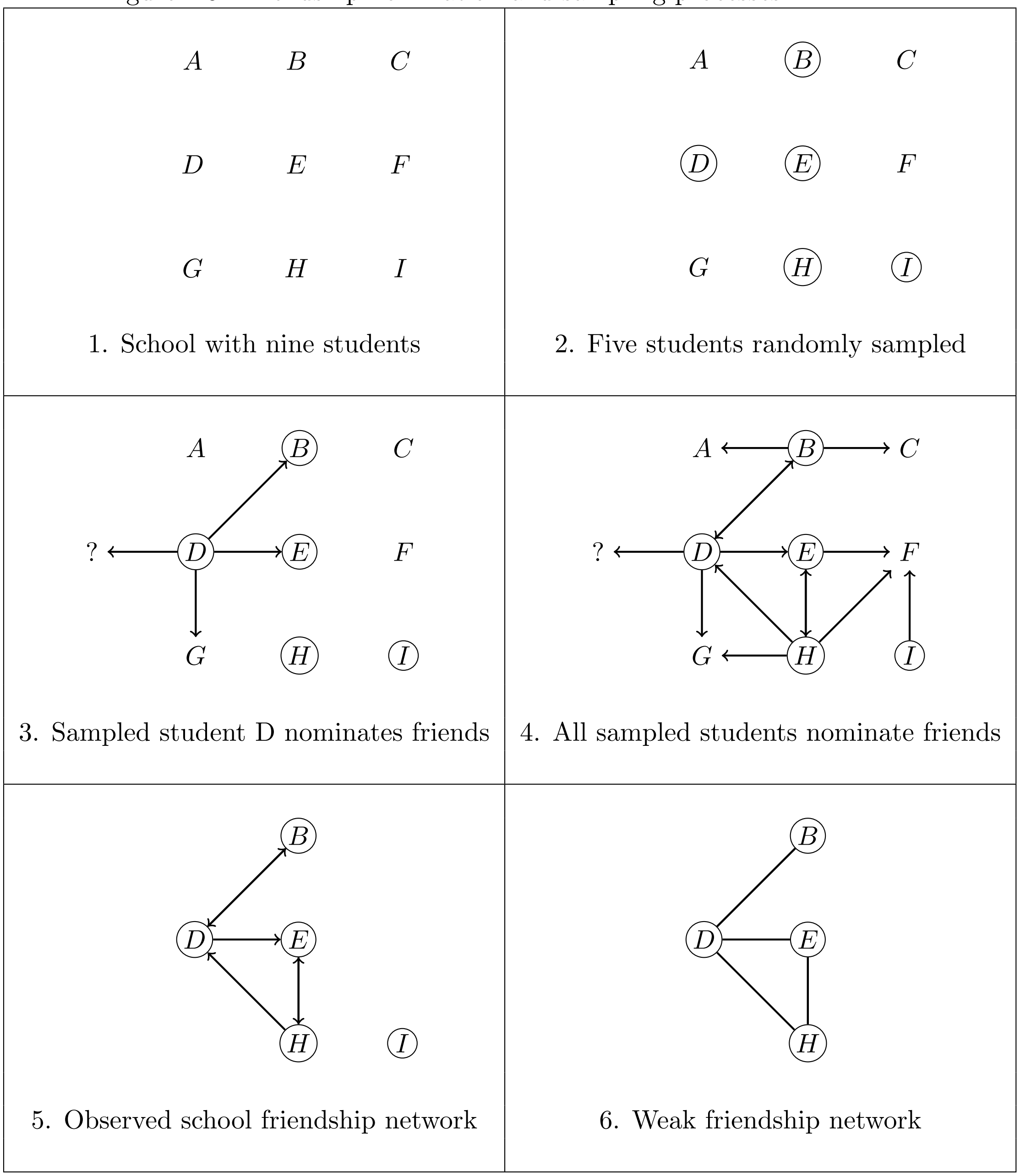


Figure 2.5: Friendship nomination and sampling processes



1. School with nine students
2. Five students randomly sampled
3. Sampled student D nominates friends
4. All sampled students nominate friends
5. Observed school friendship network
6. Weak friendship network

Figure 2.6: Distribution of grades

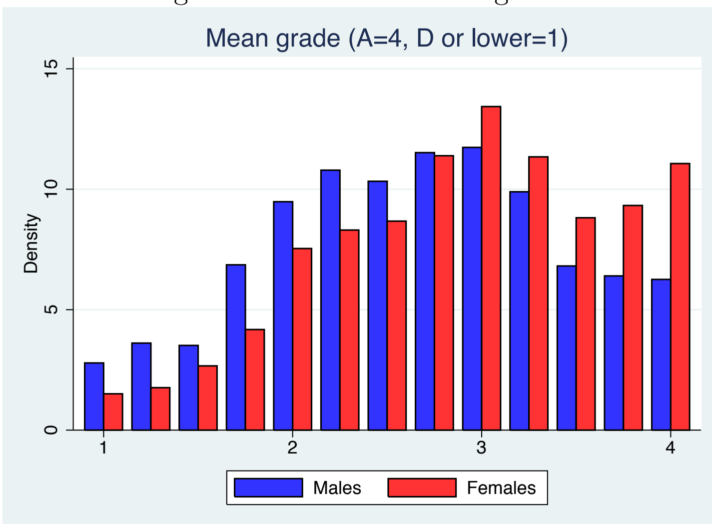


Figure 2.7: Distribution of friendship network gender composition

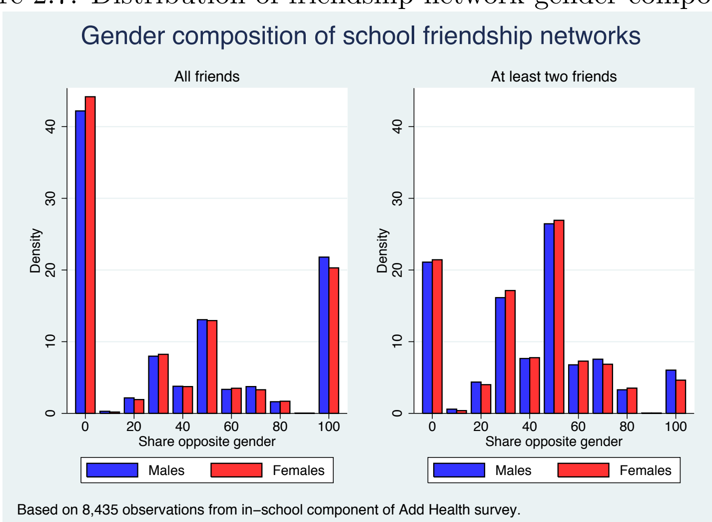


Figure 2.8: Distribution of close neighbourhood gender composition

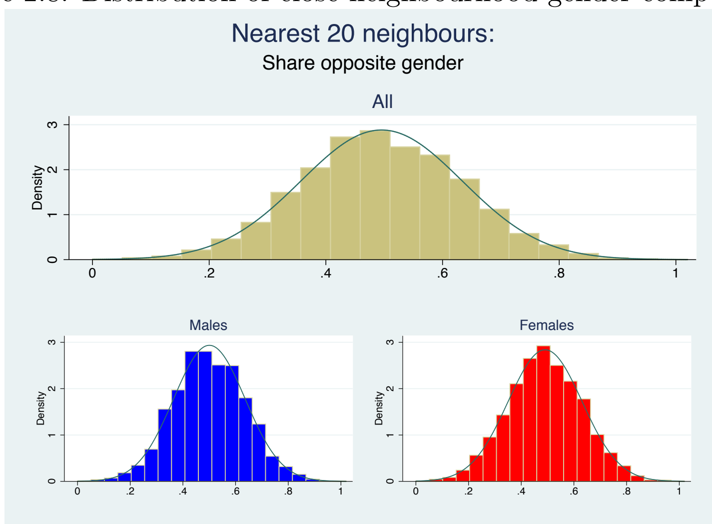
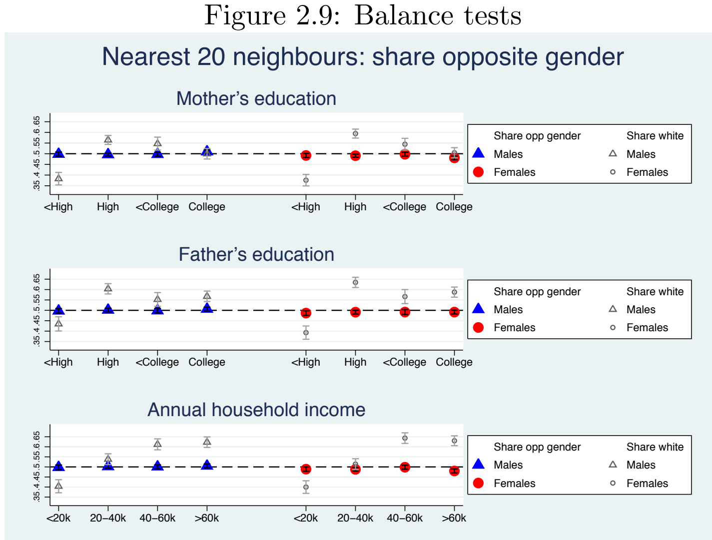


Figure 2.9: Balance tests



2.7 Tables

Table 2.1: Descriptive statistics: dyadic pairs

	Mean (standard deviation)		
	All	**Males**	**Females**
Distance			
Distance between friends (m)	5,273	5,043	5,506
	(8,404)	(8,199)	(8,602)
	[98]	[136]	[139]
Distance between randomly-drawn schoolmates (m)	7,140	7,038	7,237
	(6,317)	(6,007)	(6,599)
	[68]	[94]	[101]
Gender of nominated friend			
Opposite gender friend	0.37	0.39	0.35
	(0.48)	(0.49)	(0.48)
Female friend	0.52	0.39	0.65
	(0.50)	(0.49)	(0.48)
Interactions with nominated friend			
Go to friend's house	0.38	0.41	0.35
	(0.48)	(0.49)	(0.48)
Meet after school to hang out	0.51	0.53	0.50
	(0.50)	(0.50)	(0.50)
Spend time together during weekend	0.42	0.44	0.41
	(0.49)	(0.50)	(0.49)
Talk about a problem	0.44	0.33	0.55
	(0.50)	(0.47)	(0.50)
Talk on the phone	0.60	0.57	0.64
	(0.49)	(0.50)	(0.48)
Observations	13,142	6,612	6,530
Share	1.00	0.50	0.50

Reciprocated nominations appear twice in these data. Standard deviations in parentheses. Standard errors in square brackets.

Table 2.2: OLS estimates of friend interactions on distance

	(1)	(2)	(3)	(4)	(5)
	Dyadic data: all nominations				
	Go to friend's house	Meet after school	Spend time during w/end	Talk about a problem	Talk on phone
Distance quantiles and interactions					
Omitted category: Large distance between friends					
Small distance between friends	0.26***	0.11***	0.14***	0.01	0.05**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Medium distance between friends	0.08***	0.01	0.03	-0.00	0.05**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Female x small distance	-0.05**	-0.00	-0.01	-0.02	-0.05**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Female x medium distance	-0.03	0.03	-0.01	-0.02	-0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Opposite gender friend x small distance	-0.10***	-0.03	-0.06**	-0.00	-0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Opposite gender friend x medium distance	-0.02	0.02	-0.01	0.04	-0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Female x opposite gender friend x small	0.01	-0.03	0.01	0.01	0.06
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Female x opposite gender friend x medium	-0.01	-0.05	-0.00	-0.07*	-0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Gender and friend gender					
Female	-0.03	-0.04*	-0.04**	0.31***	0.14***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Opposite gender friend	-0.20***	-0.21***	-0.18***	-0.02	-0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Female x opposite gender friend	0.01	0.03	0.03	-0.18***	-0.16***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Observations	13,141	13,141	13,140	13,140	13,140
R-squared	0.14	0.10	0.09	0.13	0.09

Table 2.3: Descriptive statistics: key variables

	Mean (standard deviation)		
	All	**Males**	**Females**
GPA (A=4, D or lower=1; self-reported)			
Overall mean	2.8	2.7	2.9
	(0.8)	(0.8)	(0.8)
Mathematics and Science	2.7	2.6	2.8
	(0.9)	(0.9)	(0.9)
English and History	2.9	2.7	3.0
	(0.9)	(0.9)	(0.8)
School friends			
Share opposite gender	0.38	0.38	0.37
(any nomination)	(0.39)	(0.39)	(0.39)
Share opposite gender	0.17	0.19	0.16
(reciprocated nomination)	(0.33)	(0.34)	(0.32)
Nearest 20 schoolmates (weighted)			
Share opposite gender	0.49	0.50	0.49
	(0.14)	(0.14)	(0.14)
School behavioural troubles (Never=0, every day=4)			
Trouble getting along with teacher	0.8	0.9	0.7
	(0.9)	(1.0)	(0.9)
Trouble paying attention in class	1.2	1.3	1.1
	(1.0)	(1.0)	(1.0)
Trouble getting homework done	1.2	1.3	1.1
	(1.1)	(1.1)	(1.0)
Trouble with other students	0.8	0.9	0.8
	(1.0)	(1.0)	(1.0)
Friends, relationships, smoking and drinking behaviour			
Number of friends	2.6	2.7	2.6
	(2.6)	(2.6)	(2.6)
Relationship in past 18 months	0.56	0.54	0.57
	(0.50)	(0.50)	(0.49)
Smoked at least one day in past 30 days	0.26	0.26	0.25
	(0.44)	(0.44)	(0.43)
Drunk at least one day in past year	0.28	0.31	0.26
	(0.45)	(0.46)	(0.44)
Long-term outcomes (reduced samples)			
Graduated high school	0.94	0.93	0.95
	(0.23)	(0.25)	(0.22)
Attended college	0.68	0.62	0.72
	(0.47)	(0.48)	(0.45)
Ever married	0.45	0.43	0.46
	(0.50)	(0.49)	(0.50)
Observations	8,435	4,124	4,311
Share	1.00	0.49	0.51

Table 2.4: OLS estimates of GPA on gender composition of schoolmates and close neighbours

	(1)	(2)	(3)	(4)
	Overall GPA (A=4,D or lower=1)			
School friends				
Share opposite gender	0.07***	0.11***		
	(0.02)	(0.03)		
Female x share opposite gender		-0.08**		
		(0.04)		
Nearest 20 schoolmates				
Share opposite gender			-0.13**	-0.08
			(0.06)	(0.09)
Female x share opposite gender				-0.11
				(0.12)
Controls				
Female	0.20***	0.23***	0.19***	0.25***
	(0.02)	(0.02)	(0.02)	(0.06)
Other controls[a]	x	x	x	x
School and grade fixed effects	x	x	x	x
Gender-specific correlations				
Share opposite gender: males		0.11***		-0.08
		(0.03)		(0.09)
Share opposite gender: females[b]		0.03		-0.18**
		(0.02)		(0.08)
Observations	8,435	8,435	8,435	8,435
R-squared	0.24	0.24	0.23	0.23

[a]Other controls include individual demographics, parent demographics and education, household income and family structure. [b]Estimate obtained by summing share opposite gender and female x share opposite gender coefficients. Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.5: IV estimates of GPA on gender composition of high school friendship networks

	Overall GPA (A=4, D or lower=1)		Math & Science GPA	English & History GPA
	Gender-symmetric effects			
	(1)	(2)	(3)	(4)
School friends				
Share opposite gender	-0.84*	-1.05**	-1.58**	-0.67
	(0.51)	(0.53)	(0.72)	(0.54)
weak IV-robust 95% CI	[-2.3, 0.1]	[-2.6, -0.2]	[-4.0, -0.5]	[-2.1, 0.3]
Controls				
Female	0.21***	0.18***	0.13***	0.24***
	(0.02)	(0.02)	(0.03)	(0.02)
Other controls[a]		x	x	x
School and grade fixed effects	x	x	x	x
First-stage coefficients[b]				
Share opposite gender in close neighbourhood	0.13***	0.13***	0.12***	0.13***
	(0.03)	(0.03)	(0.03)	(0.03)
Diagnostics				
F-stat on excluded instrument	16.12	15.32	12.17	15.03
	Gender-specific effects[c]			
	(5)	(6)	(7)	(8)
School friends				
Share opposite gender: males	-0.47	-0.54	-0.89	-0.33
	(0.88)	(0.80)	(1.14)	(0.79)
Share opposite gender: females	-1.28	-1.63	-2.37	-1.05
	(1.05)	(1.15)	(1.79)	(1.04)
p-value of gender difference	0.61	0.50	0.56	0.63
Controls				
Female	0.52	0.60	0.68	0.51
	(0.61)	(0.61)	(0.95)	(0.55)
Other controls	x	x	x	x
School and grade fixed effects	x	x	x	x
Observations	8,435	8,435	8,169	8,410

[a]Other controls include individual demographics, parent demographics and education, household income and family structure. [b]Each coefficient is from the corresponding first stage regression for that column. [c]These models include interaction female x share opposite gender, so female estimate obtained by summing share opposite gender and female x share opposite gender coefficients. Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6: IV estimates of academic achievement by age

	Overall GPA	Math and Science GPA	English and History GPA
		Age ≤ 16	
	(1)	(2)	(3)
School friends			
Share opposite gender	-0.70	-1.39*	0.09
	(0.60)	(0.82)	(0.61)
Controls			
Female	0.19***	0.14***	0.26***
	(0.03)	(0.04)	(0.03)
Other controls	x	x	x
School and grade fixed effects	x	x	x
First-stage coefficients			
Share opposite gender	0.14**	0.14***	0.14***
	(0.05)	(0.05)	(0.05)
F-stat on excl instrument	9.15	8.76	9.07
Observations	4,142	4,133	4,134
		Age > 16	
	(4)	(5)	(6)
School friends			
Share opposite gender	-1.60**	-1.83*	-1.88*
	(0.81)	(1.01)	(1.00)
Controls			
Female	0.17***	0.12***	0.23***
	(0.03)	(0.03)	(0.03)
Other controls	x	x	x
School and grade fixed effects	x	x	x
First-stage coefficients			
Share opposite gender	0.11**	0.09**	0.11***
	(0.04)	(0.04)	(0.04)
F-stat on excl instrument	8.85	5.94	8.82
Observations	4,293	4,036	4,276

Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.7: IV estimates of potential mechanism - school and classroom behaviours

	Trouble getting along with teacher	Trouble paying attention in class	Trouble getting homework done	Trouble with other students
OLS coefficient in GPA regression[a]	-0.10***	-0.06***	-0.15***	0.02
	(0.01)	(0.01)	(0.01)	(0.01)
Gender-symmetric effects[b]				
	(1)	(2)	(3)	(4)
School friends				
Share opposite gender	1.22*	1.27*	1.11	0.69
	(0.73)	(0.75)	(0.76)	(0.58)
Controls				
Female	-0.18***	-0.13***	-0.20***	-0.03
	(0.03)	(0.03)	(0.03)	(0.03)
All other controls	x	x	x	x
Observations	8,493	8,492	8,491	8,491

[a]Estimates in this row from OLS regression of GPA on potential mechanisms. [b]Results from gender-specific regressions not reported as no significant gender differences. Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.8: IV estimates of potential mechanism - social and home behaviours

	Number of friends	Relationship in past 18 months	Smoked in past 30 days	Drunk in past year	Serious argument with mom
OLS coefficient in GPA regression[a]	0.02***	0.00	-0.29***	-0.13***	-0.08***
	(0.01)	(0.01)	(0.03)	(0.02)	(0.02)
Gender-symmetric effects[b]					
	(1)	(2)	(3)	(4)	(5)
School friends					
Share opposite gender	-1.73	0.89**	-0.08	0.31	0.46
	(1.13)	(0.38)	(0.28)	(0.26)	(0.30)
Controls					
Female	-0.03	0.05***	-0.00	-0.04***	0.08***
	(0.05)	(0.02)	(0.02)	(0.01)	(0.01)
All other controls	x	x	x	x	x
Observations	8,497	8416	8,440	8,483	7,984

[a]Estimates in this row from OLS regression of GPA on potential mechanisms. [b]Results from gender-specific regressions not reported as no significant gender differences. Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.9: IV estimates of selected mechanisms by age

	Trouble getting along with teacher	Trouble paying attention in class	Relationship in past 18 months
		Age ≤ 16	
	(1)	(2)	(3)
School friends			
Share opposite gender	0.76	0.36	0.09
	(0.95)	(0.79)	(0.47)
Controls			
Female	-0.19***	-0.12***	0.03
	(0.04)	(0.03)	(0.02)
All other controls	x	x	x
Observations	4,141	4,141	4,133
Dependent variable			
Mean	0.94	1.19	0.46
[standard deviation]	[0.99]	[1.00]	[0.50]
		Age > 16	
	(4)	(5)	(6)
School friends			
Share opposite gender	1.60**	2.41**	1.89***
	(0.77)	(1.11)	(0.66)
Controls			
Female	-0.18***	-0.14***	0.07**
	(0.03)	(0.04)	(0.03)
All other controls	x	x	x
Observations	4,293	4,292	4,283
Dependent variable			
Mean	0.74	1.27	0.65
[standard deviation]	[0.89]	[1.03]	[0.48]

Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.10: IV estimates of long-term effects of peer gender composition

	Subsequent year GPA	Graduated high school	Attended college	Ever married
Gender-symmetric effects[a]				
	(1)	(2)	(3)	(4)
School friends				
Share opposite gender	-0.68	-0.13	-0.13	0.52**
	(0.52)	(0.11)	(0.25)	(0.26)
Controls				
Female	0.19***	0.01	0.08***	0.04***
	(0.02)	(0.01)	(0.01)	(0.01)
All other controls	x	x	x	x
Observations	5,822	6,646	6,647	5,894

[a]Results from gender-specific regressions not reported as no significant gender differences. Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 3

If At First You Don't Succeed: Negative Externalities in High School Course Repetition

3.1 Introduction

The questions of whether low-achieving students should be retained in a grade or required to repeat a failed course are answered by the extent to which grade or course repetition affects the retained or repeating individual *and* the extent to which grade or course repeaters affect their classmates. An extensive literature has investigated the effect of grade or course repetition on the individual, but there is a surprising lack of evidence on the potential effects of grade or course repeaters on their classmates.[46] This paper addresses the gap in the literature by investigating whether course repeaters in high school mathematics courses exert significant negative externalities on their course-mates. Using individual and school-specific course fixed effects to control for ability and course selection, it shows that doubling the number of repeaters in a given course (holding the number of course-takers constant) results in a 0.6 reduction in GPA scores for first-time course-takers.

Many US states have both increased the number of mathematics credits required for high school graduation and specified particular mathematics courses that need to be passed (Reys et al, 2007). Media reports indicate that this has increased the likelihood of repetition for students who fail high school mathematics courses (Helfand, 2006). Seven percent of students in the sample are repeating a failed mathematics course. For Algebra I, this increases to fifteen percent. The effects of repetition in high school mathematics course are therefore important to understand. The state-specific policies as of 2006 are summarized in Tables 3.1 and 3.2, and confirm that a majority of US states have specific mathematics requirements for high school graduation. The negative externalities exerted by repeaters on their classmates found in this paper suggest a cost to course repetition ignored by previous analyses, and, to the extent that the above policies encourage course repetition, a cost to these policies that has been overlooked by policy-makers.[47]

[46] Lavy, Paserman and Schlosser (2011) come closest by investigating how the share of students who are old for their grade (having been retained) affects their same-grade schoolmates.

[47] There may, of course, be a benefit or cost experienced by the repeating individual. This is not the focus of this paper, but is clearly important for a complete policy analysis. Rose and Betts (2004) find that advanced high school mathematics courses have greater effects on students' earnings a decade after graduation than less advanced courses. This may be interpreted as suggesting a possible benefit to repeating and passing a difficult mathematics course. A recent article in the New York Times (Hacker, 2012) criticizing policies that require algebra for high school graduation evoked substantial debate and strong opinions on both sides, although little in the way of convincing empirical evidence.

Understanding the externalities imposed by repeaters in high school mathematics courses may also inform the grade retention debate. This is because both grade retention and course repetition result in students being exposed to a set of low-achieving classmates who are likely to share similar characteristics.[48] To the extent that repeating and retained students exert similar externalities on their classmates, this paper suggests grade retention analyses should include effects exerted on classmates of the retained individual.

Course repeaters may exert externalities on their course-mates in a variety of ways. These course composition effects can be grouped into two categories: general effects arising from repeaters being low-achievers and specific repeater effects not exerted by other low-achievers. Low-achieving students are likely to disproportionately extract teacher inputs or redirect teacher inputs away from first-time course-takers. They may need more time to understand concepts, slowing the pace of the class, and may also be more likely to misbehave in the classroom given that disruptive behavior is generally correlated with classroom ability, requiring teacher intervention. Low-achieving classmates may also be more likely to directly distract their classmates, lowering education production even without affecting teacher inputs.

In addition to these low-achiever effects, course repeaters may exert additional externalities specifically related to failing and retaking a course. They may be bored and inattentive when encountering course material for the second time, increasing the likelihood of disruptive behavior. Repeaters may also have a poor attitude or be uncooperative because they failed the course the previous year, and this may negatively affect both their classmates and the teacher.[49]

Course repeaters may also exert externalities through course size and class assignment (for courses with more than one class). Course size effects

[48]The effects of grade retention and course repetition on the individual, however, are likely to differ along several dimensions. This is primarily because retained and repeating students are likely to be of different ages and maturities (retention typically occurs in junior and middle schools while course repetition typically occurs in high school). In addition, retained and repeating students are exposed to a different peer group shock (retained students repeat all courses associated with a particular grade so are exposed to a completely new set of peers while repeating students are only exposed to new peers in the course they repeat).

[49]Another potential repeater mechanism operates in the other direction; repeaters may provide examples of the consequences of failure, incentivizing more effort from first-time course-takers at risk of failing. This paper finds an overall negative effect, so this channel is at most a mitigating factor.

are fully controlled in the estimation procedure, but are unlikely to be a factor given the large changes in class sizes required to observe effects[50]. Class assignment may matter if repeaters are assigned to classes non-randomly. For example, repeaters may be assigned to the best teacher for a particular course if failing for a second time is particularly costly (either from the perspective of the school or the student). This may increase the likelihood of first-time course-takers being assigned to another class with a worse teacher, leading to poorer performance for first-time course-takers.

The primary focus of this study is an analysis of the combined low achiever and repeater effects that course repeaters exert on their course-mates. This is the appropriate level of analysis for an overall evaluation of course repetition effects. Secondary results attempt to separate the general low-achiever and specific repeater externalities.

The paper uses a fixed effects strategy on longitudinal transcript data for multiple cohorts of US high school students to estimate the causal effect of course repeaters on their classmates. Essentially, the study compares the achievement of first-time course-takers in the same mathematics course (such as Algebra I) in the same high school in different years using year-to-year variation in the number of repeaters in the course to identify the effect. It is assumed that unobserved year-specific shocks to classroom education production in the previous year provide variation in the number of course repeaters. An example of this is an absent teacher causing a higher course failure rate.

Holding the number of students in a course constant (either parametrically or using course size fixed effects), the academic achievement of first-time course-takers is shown to be negatively correlated with the number of repeaters in the course that year. This relationship is robust to a variety of different specifications. The effect is concentrated in the lower and middle parts of the achievement distribution, and males and females are similarly affected. Suggestive evidence that the negative externalities exerted by course repeaters are due to their being low-achieving and repeating is provided.

These results are best compared with those obtained by Lavy, Paserman and Schlosser (2011). Defining low-ability students as students who are old for their grade (most likely having repeated kindergarten or first grade), they find that the proportion of low-ability peers is negatively correlated with the academic achievement of regular students. Variation in the composition of seven adjacent cohorts of 10th grade students in Israeli high schools (from 1994 to 2000) is used to identify the effect. It is argued that the majority of

[50]See, for example, Hanushek (1999).

students had little experience with their peers prior to entering high school, so results are not driven by common cohort-specific shocks.

This paper has three key distinctions from Lavy et al (2011). First, we observe course enrolment and achievement for all students in a set of high schools for multiple years allowing the inclusion of both individual and school-specific course fixed effects. This approach deals with potentially confounding individual effects (such as cohort-specific shocks and ability differences) and course effects (such as repeaters being more likely to repeat difficult courses) that cannot be dealt with using repeated cross-sectional data. Second, it isolates the effects of course-mates rather than grade-mates. Students in the same grade may have little interaction and may not take many of the same courses, which would attenuate effects for analyses performed at the grade level. And, third, it focuses on high school mathematics courses in the US, which is particularly relevant given policies stipulating minimum mathematics requirements for graduation in US high schools increasing the likelihood of mathematics course repetition.

Repeaters are low-achieving peers for first-time course-takers. Results can therefore be compared with the literature investigating ability peer effects in high school. These papers exploit a variety of identification strategies and typically find moderately-sized, negative achievement effects for individuals exposed to low-ability peers.[51]

The externalities exerted by course repeaters may also be placed in the context of the related literature investigating the effects of grade retention. Babcock and Bedard (2011) investigate the long run effects of primary school retention rates on both the retained and the promoted. They cannot separate the effects of retention for the retained and promoted, but find that a one standard deviation increase in early grade retention is associated with a 0.7 increase in mean male hourly wages that is evident throughout the wage distribution. They do, however, find that retention rates and educational attainment are statistically insignificant and economically small, and acknowledge the possibility that retention rates may affect the retained and promoted in opposite directions. In the same way, it is plausible that repetition may be beneficial to the repeating student and costly to the classmates of the repeating student.

The literature investigating the causal effect of retention on the retained has exploited a variety of policies to overcome selection into retention. It provides evidence of both positive and negative effects. Positive achievement effects of retention for third grade students are found by Jacob and Lefgren

[51] See, for example, Lavy et al (2012) and Burke and Sass (2013).

(2004) and Greene and Winters (2007). These papers use Chicago and Florida accountability policies respectively to obtain exogenous variation in grade retention. Ding (2010) finds that holding children back in kindergarten has positive but diminishing effects on their academic performance up to third grade. Eide and Showalter (2001) use kindergarten entry dates as an instrument for retention and find that retention reduces the probability of dropping out of high school for white students. [52] These findings suggest generally positive effects of retention on students retained up to the third grade.

The effects for older students (more like those studied in this paper) appear to be more nuanced. Jacob and Lefgren (2004, 2009) find that retention in the sixth grade does not significantly affect achievement or high school graduation, while retention in the eighth grace reduces the probability of high school graduation. Using data from junior high schools in Uruguay and a policy of automatic grade failure for certain low-achieving students, Manacorda (2012) shows that grade failure increases dropout rates and lowers educational attainment.

Fruehwirth, Navarro and Takahashi (2011) recognize that retention effects are likely to differ by the grade at which the student is retained and the unobservable behavioural and cognitive abilities of the student. They allow for heterogeneous effects in their econometric model and obtain generally negative effects from retention, suggesting grade retention is not an effective policy for raising the performance of low-ability students.

This remainder of this paper is organized in the usual way: methodology, data, results and then interpretation.

3.2 Empirical Methodology

The academic achievement of first-time course-taker i in course j, high school s, cohort c and year t is modeled as a linear function of the natural logarithm of one plus the number of repeaters in the course R_{jst}[53], the natural logarithm of the number of students in the course C_{jst} (course size), and a composite error term:

$$GPA_{ijsct} = \beta \ln(1 + R_{jst}) + \gamma \ln C_{jst} + \epsilon_i + \epsilon_{js} + \epsilon_{sc} + \epsilon_t + \epsilon_{jst} + \mu_{ijsct} \quad (3.1)$$

[52] Estimates for black students were uninformative.

[53] The addition of one to the number of repeaters ensures that the natural logarithm of zero is avoided. Results are qualitatively similar when courses with zero repeaters are dropped from the sample.

The coefficient β represents the level change in student GPA score for a percentage change in the number of course repeaters. The $\ln C_{jst}$ term controls for the potential negative effect of course size on achievement and is important because of the mechanical relationship between the number of course repeaters and course size. Without controlling for course size, estimates of the negative externalities exerted by repeaters would exaggerate the effect.

This parametrization of the education production function is chosen so estimated coefficients are easy to interpret. Results from a variety of other specifications in which course repeaters enter linearly, quadratically and as shares are reported in the appendix. The interpretation of results are consistent across specifications.

The error term is modeled to consist of individual ability ϵ_i, school-specific course difficulty ϵ_{js}, a school-specific cohort effect ϵ_{sc}, a general time trend ϵ_t, a school-specific course time trend ϵ_{jst}, and a remaining idiosyncratic shock to achievement μ_{ijsct}.

Several components of this error term may be correlated with the number of course repeaters, which would bias estimates of the effect. A variety of fixed effects and time trends are included in the estimation to remove these potential biases.[54] Several of these rely on observing multiple years of student achievement in high school for multiple cohorts, representing an advantage over repeated cross-sectional analyses or longitudinal analyses of one cohort.

$$GPA_{ijsct} = \beta \ln(1 + R_{jst}) + \gamma \ln C_{jst} + \theta_{js} + \theta_t + \theta_{jst} + \theta_i + \mu_{ijsct} \quad (3.2)$$

School-specific course fixed effects θ_{js} control for course difficulty as well as any other course-specific factor affecting both the achievement of first-time course-takers and the number of course repeaters. A positive correlation between course difficulty and the number of course repeaters is expected if students are more likely to fail and repeat difficult courses. Alternatively, low-ability students who consider themselves more likely to repeat a course may select out of difficult courses (if the course is not required for graduation). This would generate a negative correlation between course difficulty and the number of repeaters. The net direction of the correlation between

[54] These are implemented using a two-stage procedure in which fixed effects are applied to demean variables in the first stage before the analysis is performed on the demeaned variables in the second stage. This is because the final estimation is only performed on first-time course-takers, but the fixed effects need to capture the influence of repeaters.

course difficulty and the number of course repeaters could be either positive or negative, which respectively would bias the effect upwards or downwards in the absence of these fixed effects.

Year fixed effects θ_t and linear school-specific course trends θ_{jst} control for any correlated trends in student achievement and course repetition. Consider grade inflation. Every subsequent year, fewer students fail a given course, resulting in fewer course repeaters every subsequent year. At the same time, first-time course-takers perform better every year. This generates a pattern of increased achievement associated with fewer course repeaters that has nothing to do with course repetition. In the absence of this set of controls, estimates of the effect of course repeaters would be upwardly-biased.

School-specific cohort fixed effects θ_{sc} may be included to control for cohort effects. An alternative approach to dealing with cohort effects is the inclusion of individual fixed effects θ_i. [55] Individual fixed effects are preferred as they improve precision by controlling for individual ability. They also control for other forms of individual selection not considered in the above discussion.

Finally, it is noted that grading to a curve would bias the estimated effects. Repeaters are low-achieving students, so maintaining a constant course average in the presence of an increase in the number of repeaters would necessitate higher GPA scores for first-time course-takers. This would attenuate estimates of the externalities exerted by course repeaters on first-time course-takers, so results would be a lower bound of the true effect. However, variation in the unconditional means of school-specific course GPA scores in different years suggest that year-to-year grading to a curve may not be that pervasive.

Descriptive statistics include results from ordinary least squares regressions of current achievement on an individual's past mathematics course achievement (such as failing and repeating the course). Estimates from this equation do not have a causal interpretation as we cannot control for non-random selection into course repetition, but are included to describe what happens to individual students when they repeat a course.[56]

Placebo tests in which achievement depends on the number of repeaters in the same course and same school but in different years are conducted

[55] Individual fixed effects nest cohort fixed effects as individuals belong to one cohort.

[56] Existing studies have used a variety of natural experiments and policies to obtain causal estimates of this relationship. These are discussed in the introduction.

using the following equation where $p \in \{t-1, t, t+1, t+2\}$:

$$GPA_{ijsct} = \beta_p \ln(1 + R_{jsp}) + \gamma_p \ln C_{jsp} + \theta_{js} + \theta_t + \theta_{jst} + \theta_i + \mu_{ijsct} \quad (3.3)$$

The number of repeaters at time $t-1$ and time $t+2$ should be uncorrelated with the achievement of first-time course-takers at time t, so it is expected that $\beta_{t-1} = \beta_{t+2} = 0$. In addition, there may be a negative relationship between the number of repeaters at time $t+1$ and the achievement of first-time course-takers at time t. This is because there may be fewer repeaters when first-time course-takers perform well the previous year and more repeaters when first-time course-takers perform poorly.

Separating the general low achiever and specific repeater effects is investigated by including separate variables for the number of course repeaters who failed the course the previous year (the variable used in the primary specification above), the number of course-mates who failed their mathematics course the previous year F (but are not necessarily repeating the course that they failed), and the number of students who are repeating the course even though they may not have failed it the previous year Q.[57]

$$GPA_{ijsct} = \beta_R \ln(1+R_{jst}) + \beta_F \ln(1+F_{jst}) + \beta_Q \ln(1+Q_{jst}) + \gamma \ln C_{jst} + \epsilon_{ijsct} \quad (3.4)$$

The effect of the number of students who failed their mathematics course the previous year β_F captures externalities exerted by low-achieving course-mates, while specific repeater externalities are reflected in the coefficient on the number of students who are repeating without necessarily having failed β_Q. These effects are contrasted with the externalities associated with course-mates repeating the course after failing the course the previous year β_R.

Results from this specification need to be interpreted with some caution. First, the variables R, F and Q are highly collinear, increasing specification sensitivity and reducing out-of-sample performance. And, second, the variation that identifies the coefficients is generated by endogenous student choices. We do not have a policy or natural experiment that determines whether a student who fails a mathematics course choose to repeat it or chooses to do another mathematics course, and this may affect the externalities they exert.

[57] There are a surprisingly large number of students repeating a course after passing it the previous year. This phenomenon is discussed in more detail in the data section.

Mechanisms through which course repeaters exert negative externalities are investigated by considering how the number of repeaters in a course affects the self-reported educational experience of first-time course-takers. Students were surveyed during high school and asked about the frequency with which they experienced a set of difficulties in the classroom D_{isct}. These data are only available for two years (there were only two survey waves conducted while students were in high school) and for a small number of students, so these estimations do not include individual fixed effects:

$$D_{isct} = \beta \ln(1 + R_{jst}) + \gamma \ln C_{jst} + \theta_{js} + \theta_t + \theta_{jst} + \mu_{ijsct} \tag{3.5}$$

The subsequent section provides a full description of the data used in the analysis.

3.3 Data and Descriptive Statistics

This paper uses data from the National Longitudinal Study of Adolescent Health (Add Health). The Add Health is a school-based longitudinal study of a nationally representative sample of US adolescents who were in grades 7 to 12 during the 1994-1995 school year. A core sample was selected to participate in a series of detailed surveys, the most recent being in 2008 when individuals were aged between 24 and 32.

Complete high school transcript data (grades 9 to 12) are available for individuals selected for the core sample. For all of these individuals, the transcript data include a categorization of the mathematics course taken in every year of high school (or an indication that no mathematics courses were taken that year), the GPA score obtained in each of these courses, and a failure index variable describing whether the student passed or failed these courses.[58] This information is required for all students in a school in order to accurately compute the course composition measures used in the analysis. If transcript information is only available for a subset of students in the school, information is only available for a subset of a student's course-mates. The study is therefore restricted to fifteen schools in which all students in the school were selected for the core sample. This is known as the saturated

[58] A subset of students may have taken more than one mathematics course in a given year. For these students, the provided course categorization is for the highest level mathematics course taken that year, the reported GPA score is the mean GPA score over all mathematics courses taken, and the failure index describes the share of mathematics courses failed.

sample. Figure 3.1 plots the number of students who enrolled in at least one mathematics course in each of these schools, showing that there are two large schools and thirteen smaller schools.

The analysis is restricted to the years between 1992 and 1996 to ensure that courses are mostly populated by students included in the above sample. The pooled sample includes 6341 student-years. Appendix Table B.1 reports the demographics of the sample. There are 2270 unique students in the sample, so course achievement data is observed an average of 2.8 times per student. There are 3191 student-year observations describing the achievement of male students and 3150 student-year observations describing the achievement of female students. Descriptive statistics are provided in Table 3.3. Female students consistently perform better than male students across all measures of academic achievement.

Past achievement for each student in each year is described by three variables: an indicator for repeating a mathematics course that was failed the previous year, an indicator for failing any mathematics course the previous year (without necessarily repeating it the subsequent year), and an indicator for repeating the same mathematics course (without necessarily having failed it the previous year). There are some student-year observations with missing past achievement information. These students are considered first-time course-takers, although removing them from the sample does not change the results. Figure 3.2 reports the distribution of mathematics GPA scores by previous performance. As expected, first-time course-takers perform considerably better than repeat course-takers, although there are some repeat course-takers who obtain the maximum GPA score of 4.

The first column of Table 3.3 indicates that 7 percent of students in the sample are repeating a failed mathematics course.[59] The externalities exerted by these students on first-time course-takers are the primary focus of this study. The other two achievement indicators provide secondary evidence to distinguish the externalities associated with general low-achievement and specific repetition: 17 percent of students failed their previous mathematics course and 14 percent of students are repeating a mathematics course.

Course composition measures are obtained by averaging the individual achievement indicators of students in the same course in the same school in the same year. Course-mates may not be classmates if courses are divided into multiple classes within a school. The mean number of students per mathematics course is 112[60], indicating that an average course consists of

[59] Technically, these are student-years, so 7 percent of student-year observations describe students repeating a failed math course.

[60] Note that these means are computed by equally weighting student-year observations

more than one class. On average, first-time course-takers are exposed to five students who are repeating the course after failing it the previous year. Course composition is also described in terms of shares rather than counts. The distribution of course sizes for all of the course-years included in the analysis is plotted in Figure 3.3. The identification relies on variation in the number of repeaters in the larger courses; effects are imprecisely estimated if these courses are excluded. Figure 3.4 plots the variation in the number of students repeating a failed course per school-course-year. Thirty percent of student-course observations correspond to school-course-years in which no students are repeating a failed course, and the median and mean number of students repeating a failed course per school-course-year are 3 and 5.5, respectively.

Course-specific descriptive statistics are provided in Table 3.4. Mathematics courses are categorized into nine different groupings by survey administrators.[61] These are loosely ordered by difficulty from Basic/Remedial Mathematics to Calculus. The three most popular high school mathematics courses (by enrolment) are Algebra I, Geometry and Algebra II. Results are largely driven by variation in the number of repeaters across these three courses.

Fifteen percent of students in Algebra I are repeating the course after failing it the previous year. The shares of students repeating the more advanced courses of Geometry and Algebra II are smaller. This indicates that low-ability students who are most likely to fail and repeat select out of mathematics courses after taking Algebra I. This may be by choice or because they are not allowed to progress given their achievement in Algebra I.

The transition of students between mathematics courses is described in the two panels of Table 3.5. The first panel is based on 3741 student-year observations and describes the course transition of students who passed their previous mathematics course. The second row of this panel indicates that of the 450 students who passed General/Applied Mathematics, 66 percent take Algebra I the following year. Seventy-one percent of students follow Algebra I with Geometry while seven percent follow with Algebra II. Somewhat surprisingly, 16 percent of students repeat Algebra I after passing it. The

and not by equally weighting course-year observations, so large course-years with many students receive a greater weight. This also explains why the mean shares are not simply the ratios of the mean counts.

[61]The actual categorization process is not important for this paper given that students in the same course in the same school in the same year are necessarily categorized as taking the same course.

primary source of this irregularity appears to be one large school. This school is excluded from the analysis in a sensitivity check to confirm that this anomaly is not affecting results.[62] Ninety percent of students who pass Geometry follow it with Algebra II and 74 percent of students who pass Algebra II follow it with Calculus. A typical progression for passing students is a subset of the path General/Applied Mathematics to Algebra I to Geometry to Algebra II to Calculus, although several other course paths are also observed.

The second panel of Table 3.5 describes the course transitions of students who failed their previous mathematics course. It indicates that, for most courses, repetition is the modal behavior of students who failed. Interestingly, a nontrivial number of students still progress. Twenty-three percent of students who fail General/Applied Mathematics take Algebra I the next year, 29 percent of students who fail Algebra I take Geometry, and 37 percent of students who fail Geometry take Algebra II.

The final set of descriptive statistics is provided in Table 3.6. This table describes how current student achievement is associated with past achievement in a series of OLS regressions. The negative coefficients in the first three columns reflect that students repeating a failed math course are lower achievers (and likely of lower ability) than first-time course-takers. The coefficient drops from -0.8 to -0.4 in the third column when school-specific course fixed effects control for course difficulty. This suggests that part of the reduced achievement of repeaters is because they are repeating more difficult courses than those taken by other students.

The remaining three columns in Table 3.6 include individual fixed effects to control for individual ability. The correlations between repeating a failed course and current achievement are imprecise but positive in the fourth and fifth columns, suggesting an increase in achievement for students repeating a failed course relative to when they took it for the first time. The sixth column includes the other past achievement indicators to separate the associations with failing and repeating. Students perform better after failing their previous mathematics course and when repeating the same mathematics course, but there is no additional improvement for specifically repeating a failed course. It is emphasized that these associations are descriptive and non-causal.

[62] One possible hypothesis is that two different courses at this school were categorized as Algebra I.

3.4 Results

Table 3.7 reports the primary set of results. Controls are added sequentially across columns. The coefficient on the log number of students failed and repeating of -0.25 in the first column is estimated without controlling for course size and excluding school-course and individual fixed effects. The coefficient falls in magnitude to -0.20 when course size is controlled in the second column, confirming that the previous estimate exaggerated the effect. Course size and GPA are shown to be negatively correlated. The specification in the third column includes school-course fixed effects to control for course difficulty, and the magnitude of the effect falls further to -0.17. This indicates that the net effect of course difficulty biased the estimates downwards; first-time course-takers systematically perform worse in more difficult courses with more repeaters.

The fourth column reports results from the preferred specification, controlling for course size and including the full set of fixed effects. The coefficient of -0.15 is the level change in GPA scores for first-time course-takers caused by a doubling of the number of repeaters in a course (a 100 percentage point increase in the number of course repeaters or a 1 unit increase in the natural logarithm of the number of course repeaters). For the average mathematics course, this is an increase from five to ten repeaters in a course of around 100 students. Course repeaters do exert negative externalities.

The relationship between course size and GPA is no longer significant. Without information on class assignment (within courses), class size effects cannot be directly investigated with these data. This result does, however, suggest that repetition effects may be a more important concern than class size effects (which have received considerable attention).

Results from placebo tests in Table 3.8 support the empirical strategy. The first column indicates that doubling the number of repeaters in the course the year before it was taken is associated with a -0.02 (no) change in GPA scores for first-time course-takers. The second column is the original specification, while the third column reveals that the achievement of first-time course-takers is negatively correlated with the number of repeaters the next year, although the estimate is not significant. This is expected as course repeaters are course-takers from the previous year that performed poorly.

Distributional effects are investigated in Figure 3.5. This graph plots estimates from a series of linear probability models in which binary indicators for attaining at least the specified GPA score are the dependent variables. Results in this figure partially inform our understanding of the negative externalities exerted by repeaters. Negative effects at the top of the distribution may indicate teachers transferring inputs from high achievers to low achievers (such as slowing the pace of the class), negative effects throughout the ability distribution may indicate repeaters being generally disruptive, while negative effects concentrated at the bottom of the distribution may indicate repeaters specifically distracting other low achievers. The negative externalities exerted by repeaters are evident in the middle and lower parts of the distribution. This is evidence against the hypothesis that teachers transfer inputs away from high achievers when there are more repeaters in a course, and suggests repeaters may specifically distract other students in similar parts of the achievement distribution.

Course repeaters may exert negative externalities on first-time course-takers only when they reach a threshold share of the course. This form of nonlinearity cannot be captured by the above specifications. Figure 3.6 investigates threshold effects by plotting coefficients from a series of regressions taking the form of Equation 3.2, but with the explanatory variable being a binary indicator of whether the share of repeaters exceeds the specified level. The estimated effect is the difference in GPA scores between first-time course-takers exposed to a share of repeaters above the specified level and first-time course-takers exposed to a share of repeaters below the specified level. The plot suggests the negative effect is already evident when the share of repeaters reaches five percent of course-takers, although it is only statistically significant when the share reaches nine percent of course-takers. The negative effect on first-time course-takers remains relatively flat until the share of repeaters reaches fifteen percent after which it becomes very imprecise.

Gender and race heterogeneity in the effect is investigated by interacting the number of repeaters with gender and race indicators. These results are reported in Table 3.9. The third column includes both gender and race interactions. Doubling the number of repeaters in a course reduces the GPA scores of white males (the omitted category) by 0.27. Females are slightly less affected than males, but the gender difference is not statistically different. The negative externalities exerted by repeaters on black first-time course-takers are significantly smaller than those exerted on white students, while other differences by race are imprecisely estimated. Descriptive statistics in Appendix Table B.1 indicate that black students are more likely to fail and repeat mathematics courses. The smaller effect for black students suggests smaller effects in schools with more black students, and, given that black students are more likely to repeat, may indicate a declining effect for each additional percentage point increase in the number of repeaters.[63]

Results in Table 3.10 attempt to distinguish the externalities exerted by course repeaters because they are low achievers and the externalities exerted specifically because they are repeating. The number of students who failed their previous mathematics course is considered a proxy for the number of low achievers in the course. All repeaters should exert the specific externalities associated with course repetition to some extent, so including the number of repeaters who previously passed or failed therefore captures the specific repeater externality. (Recall from Table 3.5 that a surprising number of students repeat a passed course.) As discussed in the empirical methodology section, these measures are highly correlated and results are somewhat sensitive. They are interpreted as suggestive rather than conclusive.

The second and third columns reveal that both the number of low achievers and the number of repeaters are negatively correlated with the GPA of first-time course-takers when included in separate regressions. The third column includes both of these measures and the original variable. Only the coefficient of -0.14 on the log number of students failed and repeating is negative and statistically significant, although the number of low achievers (as measured by the number of students who failed their previous mathematics course) also enters the estimated GPA production function negatively. This suggests that both specific repeater and general low achiever effects may be in operation. One implication of this is that encouraging low-achieving students to progress to higher-level mathematics course rather than repeat may not fully address the issue as the negative externalities exerted by these students would persist in the higher-level courses. A more appropriate policy for negating these externalities may be to direct failing students away from mathematics courses or towards less cognitively-demanding numeracy courses.

3.5 Conclusion

Mathematics is difficult for many students, and course repetition in high school mathematics courses is a common occurrence. This repetition is promoted by policies in several US states that stipulate a minimum level of mathematics to graduate high school. Mathematics is also generally

[63] The logarithmic functional form captures some nonlinearity in the effect, but actual nonlinearities may be more pronounced or take a different form. The small sample and the related absence of statistical power do not allow a fuller investigation of this; a non-parametric analysis in which a series of bins for the number of repeaters were included as explanatory variables was uninformative.

considered important for future job market success, acting as further encouragement for students to repeat failed mathematics courses. Previous discussions around the benefits and costs of course repetition have focused on the potentially-repeating individual student.

This paper takes a new step by considering the externalities exerted by course repeaters on other students taking the course for the first time. A doubling of the number of repeaters in a mathematics course leads to a 0.15 reduction (approximately equal to the mean female-male achievement gap) in GPA scores for first-time course-takers. The effect appears to dominate course size effects, and, given the relationship between course size and class size and the extensive literature on class size, warrants more attention.

Using Israeli data, Lavy et al (2011) finds that higher proportions of low-ability students in a grade are associated with reductions in the general quality of the classroom environment. This provides a candidate mechanism through which the negative externalities reported in this paper may operate. The estimated distributional effects indicate that course repeaters negatively affect students at the middle and lower parts of the achievement distribution. This suggests that course repeaters may be more likely to distract classmates who are located in similarly-low parts of the achievement distribution rather than high achievers, which is particularly concerning given these students are already at risk. The effect does not appear to operate through teachers redirecting resources to low-ability students from high-ability students, so policies that promote maintaining a constant level of teacher inputs irrespective of the classroom distribution of repeaters may not be effective in alleviating the negative externalities.

Results also suggest that the negative externalities exerted by course repeaters arise because these students are both low-achieving *and* repeating. This is important because policies that reduce course repetition may not deal with the low-achiever effects. If the negative externalities exerted by course repeaters outweigh the potential benefits of repetition for the repeating student, a more fitting solution may be promoting numeracy courses rather than Algebra and Geometry for high school students who do not display an aptitude for mathematics.

Finally, suggestive evidence indicates that the negative effect is mitigated if the share of repeaters remains below five percent. This presents a possible policy response of stipulating a maximum share of repeaters permitted in a course. The overall finding of negative externalities emphasizes the need to include the effect of repeaters on their classmates when considering optimal grade retention, course repetition and high school graduation policies.

3.6 Figures

Figure 3.1: Number of students enrolled in math courses per school

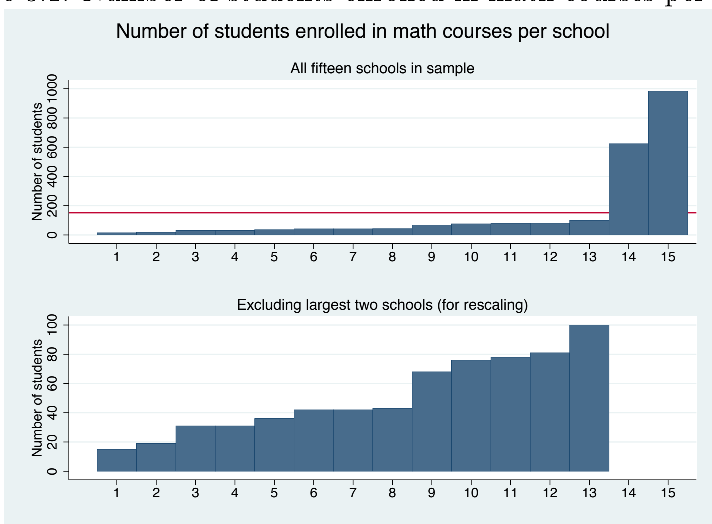


Figure 3.2: Distribution of math GPA scores by past achievement

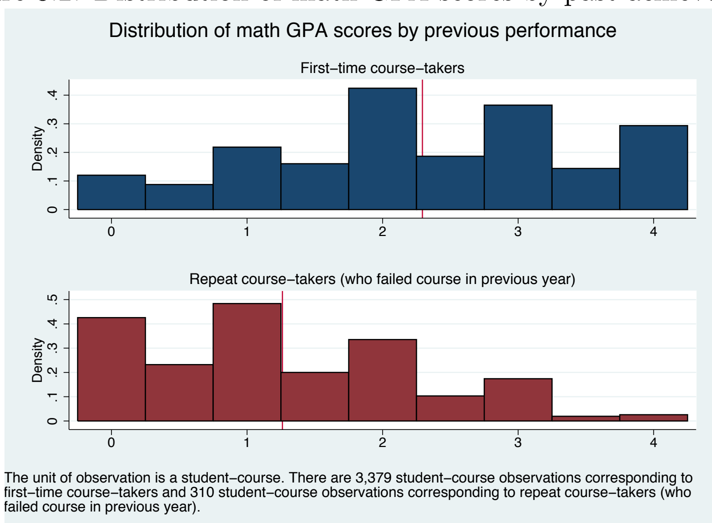


Figure 3.3: Number of students per school-course-year (class)

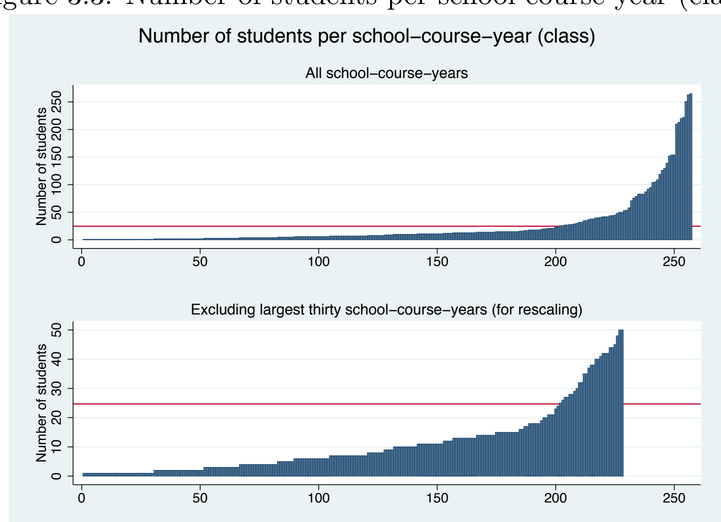
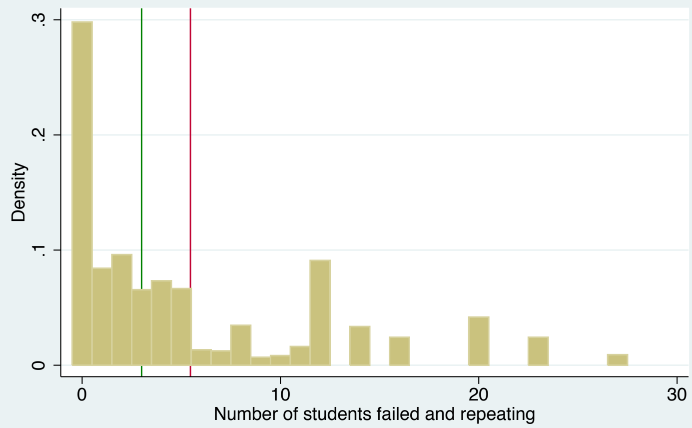
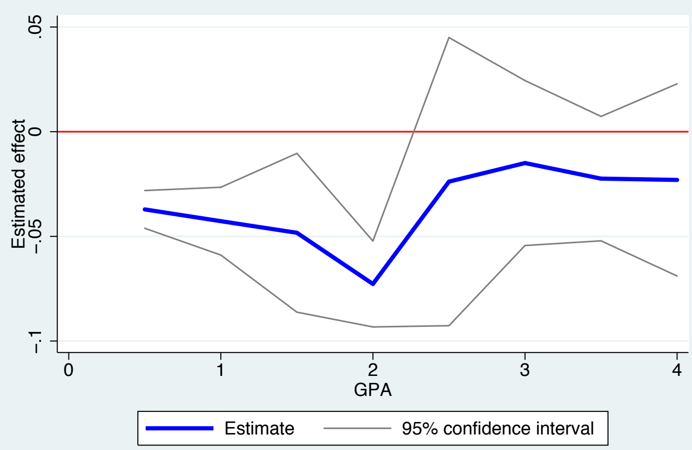


Figure 3.4: Distribution of number of students repeating a failed course per school-course-year



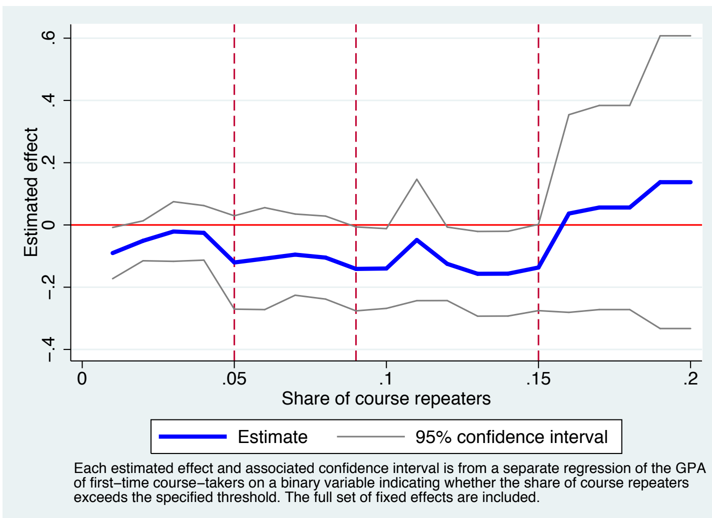
The unit of observation is a student–course. Thirty percent of student–course observations correspond to school–course–years (classes) in which no students are repeating a failed course. The median and mean number of students repeating a failed course per school–course–year are 3 and 5.5, respectively indicated by the green and red vertical lines.

Figure 3.5: Distributional effects



Each estimated effect and associated confidence interval is from a separate linear probability model. The dependent variable is an indicator for attaining the specified GPA and the independent variable is the natural logarithm of the number of course repeaters. The full set of fixed effects are included.

Figure 3.6: Threshold effects of share repeaters on GPA of first-time course-takers



3.7 Tables

Table 3.1: Number of years of high school mathematics courses/credits required for graduation

Years	States	Total
Specified at local level	CO, IA, ME, MA, NE	5
1 year		0
2 years	AK, AZ, CA, ID, MT, ND, WI	7
3 years	CT, DC, DoDEA, HI, IL, KS, KY, LA, MD, MN, MO, NH, NM, NJ, NV, NY, OH, OK, OR, PA, TN, UT, VT, WY	24
4 years	AL, AR, DE, FL, MI, MS, RI, SC, TX, WA, WV	11
Varies by diploma	IN (2-4 yrs), GA (3-4 yrs), NC (3-4 yrs), SD (3-4 yrs), VA (3-4 yrs)	5

Source: Reys et al, 2007

Table 3.2: Courses required for high school graduation/diploma

Course	States	Total
Algebra I	AL, AR, CA, DoDEA, DC*, FL*, GA*, IL, KY, MD, MI, MS, ND, NH, NM**, OK**, SD, TX, UT*	19
Algebra I Integrated Mathematics I	IN, LA*, NC, TN*	4
Geometry	AL, AR, DoDEA, IL, KY, MD, MI, TX, UT*	9
Geometry or Integrated Mathematics II		0
Algebra II	AR, MI	2
Algebra II Integrated Mathematics III	DE*	1
Algebra I, Geometry, Algebra II OR Integrated Mathematics I-III	LA, TN*, VA	3

* Or an equivalent course, ** Minimum requirement.
Source: Reys et al, 2007

Table 3.3: Descriptive statistics - Pooled (Units: student-years)

	Mean (standard deviation)		
	All	Males	Females
Academic outcomes:			
Math GPA score (transcript)[a]	2.17	2.05	2.28
	(1.17)	(1.16)	(1.17)
Individual past achievement - binary indicators:[b]			
Failed[c] and repeating math course	0.07	0.08	0.06
Failed math course in previous year	0.17	0.19	0.15
Repeating math course from previous year	0.14	0.16	0.13
Course-mates:[d]			
Course size (number of students)	111.63	113.11	110.12
	(88.10)	(87.47)	(88.72)
Number of students failed and repeating	5.46	5.55	5.37
	(6.55)	(6.57)	(6.52)
Number of students failed	13.22	13.48	12.96
	(14.34)	(14.29)	(14.38)
Number of student repeating	12.44	12.72	12.15
	(18.36)	(18.47)	(18.25)
Share of students failed and repeating	0.08	0.08	0.08
	(0.10)	(0.10)	(0.11)
Share of students failed	0.18	0.19	0.18
	(0.19)	(0.19)	(0.19)
Share of students repeating	0.19	0.19	0.19
	(0.24)	(0.24)	(0.24)
Observations	6341	3191	3150
Share	1	0.50	0.50

[a]The math GPA score is the mean GPA over all math courses taken in a given year if more than one course is taken in the year. [b]Means for these binary indicators based on smaller samples due to missing past achievement for some individuals. [c]*Failed* is a binary indicator that equal to one if any failure in previous year's math courses. [d]Course-mates are students in the same school, taking the same course, in the same year.

Table 3.4: Descriptive statistics by math course - Pooled (student-years)

	Basic/ Remedial	General/ Applied	Pre-algebra	Algebra I	Geometry	Algebra II	Advanced	Pre-calculus	Calculus
Academic outcomes:									
Math GPA score (transcript)[a]	1.66	2.07	1.96	1.92	2.26	2.32	3.00	2.65	3.04
Individual past achievement - binary indicators:[b]									
Failed[c] and repeating math course	0.11	0.08	0.12	0.15	0.05	0.04	0.00	0.01	0.00
Failed math course in previous year	0.54	0.36	0.51	0.21	0.13	0.10	0.00	0.05	0.00
Repeating math course from previous year	0.22	0.28	0.20	0.36	0.07	0.05	0.04	0.04	0.00
Individual current achievement - binary indicators:									
Fail and repeat math course	0.10	0.03	0.08	0.08	0.04	0.05	0.00	0.02	0.00
Fail math course	0.31	0.19	0.25	0.23	0.17	0.18	0.04	0.09	0.05
Repeat math course	0.20	0.12	0.15	0.20	0.07	0.07	0.25	0.12	0.20
Course-mates:[d]									
Course size (number of students)	51.12	89.09	41.61	159.69	130.24	100.96	8.07	58.28	24.85
Number of students failed and repeating	3.70	1.92	2.18	10.94	4.48	4.34	0.00	0.79	0.00
Number of students failed	21.08	6.11	11.32	16.11	16.72	12.57	0.21	3.43	0.00
Number of student repeating	8.51	7.24	4.17	29.04	7.01	6.11	0.25	2.41	0.00
Share of students failed and repeating	0.12	0.12	0.13	0.14	0.04	0.04	0.00	0.01	0.00
Share of students failed	0.53	0.34	0.40	0.20	0.12	0.10	0.01	0.05	0.00
Share of students repeating	0.22	0.47	0.19	0.35	0.06	0.05	0.04	0.03	0.00
Observations	340	571	313	1790	1469	1160	72	483	143
Share	0.05	0.09	0.05	0.28	0.23	0.18	0.01	0.08	0.02

[a]The math GPA score is the mean GPA over all math courses taken in a given year if more than one course is taken in the year. [b]Means for these binary indicators based on smaller samples due to missing past achievement for some individuals. [c]*Failed* is a binary indicator that equal to one if any failure in previous year's math courses. [d]Course-mates are students in the same school, taking the same course, in the same year.

Table 3.5: Transition matrices - shares: Mathematics (student-years)

Panel A: No math course failure in previous year

	Current course									
Previous course	1	2	3	4	5	6	7	8	9	**Total**
1 - Basic/Remedial	0.19	0.10	0.21	0.47	0.02	0	0	0	0	**163**
2 - General/Applied	0.12	0.11	0.06	0.66	0.03	0.01	0	0.00	0	**450**
3 - Pre-algebra	0.06	0.05	0.08	0.78	0.02	0.00	0	0	0	**202**
4 - Algebra I	0.01	0.02	0.00	0.16	0.71	0.07	0	0.01	0	**1,264**
5 - Geometry	0.00	0.01	0.00	0.02	0.03	0.90	0.01	0.02	0	**947**
6 - Algebra II	0.00	0.04	0.00	0.01	0.12	0.03	0.05	0.74	0.01	**539**
7 - Advanced	0	0	0	0	0.11	0	0.33	0.44	0.11	**9**
8 - Pre-calculus	0	0	0	0	0	0.01	0.09	0.07	0.83	**167**
9 - Calculus	0	0	0	0	0	0	0	0	0	**0**
Total	**115**	**139**	**89**	**758**	**1,012**	**972**	**64**	**449**	**143**	**3,741**

Panel B: Any math course failure in previous year

	Current course									
Previous course	1	2	3	4	5	6	7	8	9	**Total**
1 - Basic/Remedial	0.41	0.07	0.28	0.24	0	0	0	0	0	**68**
2 - General/Applied	0.46	0.18	0.12	0.23	0.01	0	0	0	0	**95**
3 - Pre-algebra	0.30	0.11	0.39	0.18	0	0.02	0	0	0	**56**
4 - Algebra I	0.10	0.08	0.07	0.44	0.29	0.02	0	0.00	0	**323**
5 - Geometry	0.06	0.09	0.09	0.05	0.34	0.37	0	0.01	0	**163**
6 - Algebra II	0.04	0.11	0.04	0.03	0.08	0.53	0	0.18	0	**79**
7 - Advanced	0	0	0	0	0	0	0	0	0	**0**
8 - Pre-calculus	0	0.13	0	0	0.13	0.13	0	0.63		**8**
9 - Calculus	0	0	0	0	0	0	0	0	0	**0**
Total	**134**	**78**	**93**	**199**	**156**	**110**	**0**	**22**	**0**	**792**

Table 3.6: Correlation between previous and current mathematics achievement

Dependent variable: GPA score	(1)	(2)	(3)	(4)	(5)	(6)
Previous year academic achievement:						
Failed and repeating course	-0.89***	-0.81***	-0.39***	0.36	0.30	-0.11
	(0.04)	(0.02)	(0.08)	(0.25)	(0.21)	(0.12)
Failed course in previous year						0.34***
(not necessarily repeating)						(0.06)
Repeating course from						
previous year						0.18***
(not necessarily having failed)						(0.04)
Fixed effects:						
Year (5)		x	x	x	x	x
School-cohort (56)		x	x	x	x	x
School-course (84)			x		x	x
Individual (2047)				x	x	x
Observations (student-years)	4533	4533	4533	4533	4533	4533
Number of students	2047	2047	2047	2047	2047	2047

Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.7: Effect of course repeaters on academic performance of first-time course-takers

	Sample: First-time course-takers			
Dependent variable: Math GPA score	(1)	(2)	(3)	(4)
Course-mates:				
Log number of students failed	-0.25***	-0.20***	-0.17***	-0.15**
and repeating	(0.04)	(0.03)	(0.02)	(0.04)
Log number of students in course		-0.18***	-0.11***	-0.03
		(0.03)	(0.04)	(0.15)
Fixed effects:				
Year (5) and school-cohort (53)	x	x	x	x
School-course (78)			x	x
School-course trends (78)			x	x
Individual (1810)				x
Observations (student-years)	3379	3379	3379	3379
Number of students	1810	1810	1810	1810

Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.8: Placebo tests: Pseudo course-mate achievement at time $t-1$ to $t+2$

	Sample: First-time course-takers at time t			
Dependent variable: Math GPA score	(1)	(2)	(3)	(4)
Pseudo course-mate achievement at time:	$t-1$	t	$t+1$	$t+2$
Pseudo course-mates:				
Log number of students failed	-0.02	-0.15**	-0.09	0.10
and repeating	(0.04)	(0.04)	(0.08)	(0.09)
Fixed effects[a]	x	x	x	x
Observations (student-years)	3160	3379	3324	3337
Number of students	1739	1810	1790	1783

[a]Year, school-cohort, school-course, school-course trends and individual fixed effects, as well as log number of students in course included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.9: Gender and race heterogeneity in effect of course repeaters

	Sample: First-time course-takers		
Dependent variable: Math GPA score	(1)	(2)	(3)
Course-mates:			
Log number of students failed	-0.21*	-0.21***	-0.27***
and repeating	(0.11)	(0.05)	(0.08)
x Female	0.11		0.12
	(0.14)		(0.13)
x Black		0.21*	0.22*
		(0.11)	(0.10)
x Hispanic		-0.01	-0.01
		(0.11)	(0.11)
x Asian		0.19	0.19
		(0.06)	(0.07)
x Other		0.25	0.25
		(0.41)	(0.40)
Fixed effects[a]	x	x	x
Observations (student-years)	3377	3377	3377
Number of students	1808	1808	1808

[a]Year, school-cohort, school-course, school-course trends and individual fixed effects, as well as log number of students in course included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.10: Separating effects of course-mates' course failure and course repetition

	Sample: First-time course-takers			
Dependent variable: Math GPA score	(1)	(2)	(3)	(4)
Course-mates:				
Log number of students failed	-0.15**			-0.14***
and repeating	(0.04)			(0.04)
Log number of students failed		-0.18***		-0.12
		(0.09)		(0.09)
Log number of students repeating			-0.09***	0.05
			(0.15)	(0.04)
Fixed effects[a]	x	x	x	x
Observations (student-years)	3379	3379	3379	3379
Number of students	1810	1810	1810	1810

[a]Year, school-cohort, school-course, school-course trends and individual fixed effects, as well as log number of students in course included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 4

To Go To College or Get A Job? The Effects of Part-Time Work During High School

4.1 Introduction

Over half the students in US high schools engage in some form of market work during the school year.[64] Part-time work during high school may affect subsequent labour market outcomes in a variety of ways. The focus of this paper is an exploration of the extent to which high school work experience incentivizes labour market entry and college attendance after high school.

Part-time work during high school may increase the opportunity cost of attending college. Working during high school increases the likelihood of employment after high school, and may also increase initial wages if there are returns to high school work experience. This would increase the probability of labour market entry after high school. At the same time, working in an unskilled occupation during high school may provide motivation to pursue postsecondary education as a means to greater job satisfaction and higher wages in the future, encouraging college attendance.

The effects of part-time work during high school are also likely to be heterogeneous. The effect of an additional hour of market work is probably of different magnitude, and possibly of different sign, for a high school student who works relatively few hours per week and a high school student who works relatively many hours per week. It is also likely to differ by the grade and age during which the part-time work occurs. An important contribution of this paper is providing a joint analysis of grade or age and work intensity heterogeneity in the effects of part-time work.

This paper contributes to the literature by finding that part-time work during high school reduces college attendance and lowers the age of entry into the full-time labour market for 8-10th grade students with high work intensities. There is no effect on the probability of dropping out. These outcomes have not been fully investigated in existing studies. It also considers the effect on subsequent self-reported job satisfaction, finding no effect of part-time work irrespective of the grade in which the work occurred and the intensity of the work. Noting the concerns associated with using self-reported job satisfaction as a dependent variable (Bertrand and Mullainathan, 2001), this suggests that part-time work during high school may affect the career paths of individuals but not their subsequent wellbeing.

There are likely to be several unobservable factors that influence both high school part-time working behaviour and the decision to work or study

[64] Pabilonia (2001) provides a full description of part-time working behaviour during high school in the 1990s using data from the National Longitudinal Survey of Youth 1997. The paper also includes a discussion of the Federal Fair Labor Standards Act, the law that governs the ages at and intensities with which children are allowed to work.

after high school. This endogeneity problem is overcome by exploiting peer-induced variation in part-time hours worked. The basic idea is that conditionally random variation in the working behaviour of an individual's peers induces exogenous variation in the individual's own working behaviour. Essentially, the presence of peer effects allows peer behaviour to be used as an instrument for an individual's own behaviour. The exclusion restriction is that peer working behaviour does not affect the outcome of interest through any channel other than individual working behaviour. The first part of the empirical strategy applies an existing methodology to provide evidence of peer effects in part-time working behaviour, and the second part of the paper uses this estimated peer effect as the first-stage of an instrumental variables' estimation.

There is an established literature investigating the effects of part-time work during high school on both academic achievement and labor market outcomes. The primary challenge in investigating the effect of working during school is controlling for the endogeneity in the decision to work during high school. The endogeneity problem arises because unobserved factors (such as ability, motivation or parental inputs) affect both the decision to work and the respective outcome variable.

The methodology employed in this paper partly exploits the fact that contemporaneous academic achievement effects associated with part-time work are zero or close to zero. This conclusion is supported by several papers using a variety of datasets and empirical strategies.[65] One notable exception to finding no or negligible effects is Tyler (2003). This study uses variation in the labour supply of 12th grade students generated by interstate variation in child labour laws to find that decreasing work intensity improves mathematics scores. Oettinger (1999) and Montmarquette (2007) find some evidence of negative academic achievement effects for individuals with very high work intensities. Given the local nature of these estimated effects, it seems reasonable to conjecture the general absence of an effect with some underlying heterogeneity.

Effects of part-time work on subsequent labour outcomes are mixed. The existing literature has focused on wage and employment effects. Ruhm (1997) finds that hours worked during an individual's senior year in high school and future earnings are correlated. The paper argues that an extensive set of controls are sufficient to overcome the endogeneity problem.

[65]Dustmann and van Soest (2007) uses the UK National Child Development Study, Rothstein (2007) uses the National Longitudinal Survey of Youth 1997, Sabia (2009) uses the National Longitudinal Study of Adolescent Health, and Buscha et al (2012) uses the National Education Longitudinal Study of 1988.

Other papers find no effects. Light (1999) uses an instrumental variables strategy and concludes that the direct effect of high school employment on subsequent wages is small and relatively short-lived, while Hotz (2002) uses dynamic selection methods to reach a similar conclusion.

Related papers investigate the effects of part-time work during college on academic and employment outcomes. These are useful for comparison purposes, although working during high school and college are likely to have quite different effects. Stinebrickner and Stinebrickner (2003) compare ordinary least squares, fixed effects and instrumental variable approaches, stressing the importance of dealing with the endogeneity of hours worked. They find that working during college has a small, negative effect on academic achievement. Hakkinen (2006) shows that working during college has various short-term effects on earnings and time-to-degree, but concludes that there are ultimately no significant returns to student employment.

The structure of this paper is as follows. The second section of the paper introduces the empirical strategy. After explaining the identification strategy, a detailed exposition of the first stage and second stage regressions needed for identifying the effect are presented. The first-stage subsection outlines the conditional effect of the instrument on the explanatory variable of interest. This is non-trivial as there are various complexities that need to be considered when deriving a causal peer effect. The second- stage subsection considers the causal relationship of interest: the effect of high school working behavior on school performance. The third section of the paper explains the data used in the analysis and the fourth section reports the estimation results.

4.2 Empirical Methodology

Several unobservable factors affect both part-time working behaviour during high school and future labour market outcomes. For example, parents who promote academic achievement during high school may discourage market work, and these may be the same parents who encourage their children to pursue the postsecondary education that leads to positive job market outcomes. An observed negative correlation between the intensity of part-time work during high school and future labour market outcomes may then be driven by differences in parental inputs across students. As a result, an ordinary least squares regression of educational or labour market outcome on hours worked during high school is likely to yield biased estimates of the effect of part-time work.

This paper uses an instrumentation strategy to overcome the endogeneity problem. The part-time working behavior of an individual's peers is employed as an instrument for an individual's own part-time working behavior. The key idea is that students are induced to work varying numbers of hours during high school by variation in the working intensity of their peers. Grade fixed effects, school fixed effects and using past rather than contemporaneous peer behavior support the claim of instrument exogeneity.

There are a variety of channels through which the hours worked by students in the same grade may be correlated. First, individuals in the same grade are likely to share information. This information may be about the general costs and benefits of working during high school, or about actual job opportunities. For example, an individual working at a 10-hour per week job may inform schoolmates of job opportunities for similar work at the same employer. This would introduce positive correlation in hours worked during school. This channel is supported by the extensive literature on the role of job information networks in job search.[66]

Second, individuals may work similar hours to other students in the same grade because of similarities in their recreational activities and expenditure patterns. Individuals may have an incentive to work more hours (and obtain more disposable income) if they want to engage in costly social activities with schoolmates who work more hours (and therefore have more disposable income). Examples of costly social activities include anything from watching movies to drinking alcohol or smoking cigarettes.

And, third, there are various local neighborhood effects that may result in positive correlation among the hours worked by students in the same grade and school. An example of this may be the proximity of the high school to employers of high school workers (such as fast food chains).

4.2.1 The First Stage: Peer Effects in High School Working Behaviour

The first requirement for using the empirical strategy outlined above is establishing a causal link between the hours worked by an individual's schoolmates in the same grade and an individual's own hours worked. A standard specification in the empirical social interactions and peer effects literature is the linear-in-means model in which an individual's weekly hours worked

[66] Ioannides and Loury (2004) provide a survey.

is a function of the mean weekly hours worked by the peer group:

$$H_i^t = \alpha_0 + \alpha_1 \frac{1}{n-1} \sum_{j \neq i} H_j^t + \epsilon_i^t \tag{4.1}$$

Individual i's peers are indexed by j and n denotes the size of the peer group. This model is simple to estimate, but the limitations associated with interpreting the relationship between individual behavior and mean group behavior as causal are well-understood.

Manski (1993, 2000) considers three reasons why we may observe correlations between individual and group behavior. These are termed endogenous, exogenous and correlated effects in the peer effects literature. Endogenous effects arise when individuals respond to the actions of other group members. This is the nature of the causal relationship considered in this paper, and is typically difficult to identify.

Exogenous effects (also known as contextual effects) arise when an individual's behavior is a function of an individual's exogenous or background characteristics, and these exogenous characteristics are shared by group members. For example, white males are more likely to work in high school than other demographic groups. Under the empirical regularity that the peer groups of white males are more likely to be constituted of other white males (Currarini et al, 2009), observed correlation in peer group working behavior may be due to this exogenous effect.

Finally, correlated effects arise when group members respond to environmental or institutional factors or shocks that are common to members of the group. This is particularly relevant in the school setting considered in this paper. Observed correlation in the working behavior of schoolmates may be a consequence of school location if, for example, some schools are close to potential employers of high school workers and other schools are not. This would be considered a correlated effect.

One of the primary difficulties in identifying endogenous effects arises because an individual's own behavior and the mean behavior of that individual's peer group are simultaneously determined. In other words, the behavior of an individual both affects and is affected by the behavior of group members. This is known as the reflection problem.[67]

In order to separate the different effects and deal with the reflection

[67] This is less of a concern when the relationship between individual behavior group and group behavior is nonlinear (see, for example, Bramoulle, Djebbari and Fortin, 2009, and Brock and Durlauf, 2001). A developing strand of the social interactions literature exploits nonlinearity to identify peer effects, but this is not considered in this paper.

problem, the standard linear-in-means model is amended in three ways: the reference behavior of the group is lagged by one period, the mean characteristics of the reference group are included as explanatory variables, and school fixed effects are included. The amended specification is as follows:

$$H_{igs}^{t} = \alpha_0 + \alpha_1 X_{igs} + \alpha_2 \frac{1}{n_i - 1} \sum_{j \neq i} H_{jgs}^{t-1} + \alpha_3 \frac{1}{n_i - 1} \sum_{j \neq i} X_{jgs} + D_g + D_s + \epsilon_{igs}^{t} \tag{4.2}$$

The peer group in this paper is defined to be individuals in the same grade in the same school, same-grade schoolmates. This is a natural definition of a high school peer group if we consider that individuals in high school spend most of their time with other individuals of similar ages who are likely to be in the same grade. Classmates may be a finer measure of the relevant peer group, but this introduces bias due to nonrandom selection into classes (see, for example, Hoxby, 2000). H_{igs}^{t} denotes the hours worked by individual i in grade g and school s. Other students in the same grade are indexed by j, and n_i denotes the size of individual i's peer group (the number of students in individual i's school and grade). Exogenous individual characteristics are denoted by X_{igs} and grad and school fixed effects are denoted by D_g and D_s.

The inclusion of mean group characteristics controls for potential contextual peer effects under the assumption that the effect of group characteristics on the outcome variable is linear. Bifulco et al (2011) considers the effect of classmate characteristics on various economic and social outcomes, and finds that only mother's education plays a statistically significant role in determining these early adult outcomes. This paper includes a variety of group characteristics in addition to mother's education.

Grade fixed effects control for grade-specific variation in part-time work during high school. This controls for the correlation induced by older students both working more hours per week and being in higher grades. School fixed effects control for correlated effects arising from school-specific factors or shocks that may affect the working behavior of members of the same school. These include the proximity of the school to employers of high school workers, as well as local labor market conditions.

Note that the identifying variation is cross-sectional. Time-varying changes in local labour market conditions would simultaneously affect peer working hours and the decision to attend college, which would invalidate the instrument. Relying on across-cohort within-school variation in peer working hours at a particular point in time (with fixed local labour market conditions) eliminates this concern.

The reflection problem is solved by using weekly hours worked the previous year by students currently in the same grade; an individual's current working behavior is affected by but cannot affect the previous working behavior of students in the same grade. Several papers have used this approach to deal with the reflection problem (see, for example, Clark and Loheac, 2007). The lag length of one year is chosen based on data availability.[68]

One caveat in using this approach to solve the reflection problem is that hours worked during school cannot be static over time. This would be the case, for example, if hours worked were fixed throughout high school. This is because we cannot disentangle the simultaneity in the determination of an individual's own hours worked and the hours worked by students in the same grade if hours worked do not change over time.

4.2.2 The Second Stage: The Effect of High School Working Behaviour

The causal relationship between an individual's same-grade schoolmates' weekly market hours worked and an individual's own weekly market hours worked is interesting in its own right. This provides evidence of peer effects in high school, and informs understanding of adolescent decision-making in general. Peer effects relating to alcohol and substance use in high school have been documented in the social interactions literature, but to my knowledge there is no prior evidence of peer effects in market working behavior during high school. This paper takes an additional step and uses the exogenous variation in hours worked induced by classmates' hours worked to identify the effect of part-time work during high school on future educational and labour market outcomes.

The proposed empirical strategy identifies the causal effect of part-time work during high school under the assumption that the part-time working behavior of an individual's same-grade schoolmates only affects individual outcomes through the part-time working behaviour of the considered individual and not through any other channel. There are scenarios in which this exclusion restriction would be violated. Consider a world in which part-time work during high school negatively affects school performance and the academic achievement of same-grade schoolmates affects an individual's own academic achievement. Now consider a student in this world exposed to same-grade schoolmates who work an above average number of hours. The

[68]Richer data would allow an analysis of the role played by lag length.

school performance of this student will be reduced both because she is induced to work more hours by her peers and because she has lower-achieving peers (because they work more hours). The suggested instrumentation strategy would combine these effects, exaggerating the estimated effect of part-time working behavior on academic achievement.

Two arguments support the above identifying assumption. First, the contemporaneous effect of part-time work on school performance is zero or close to zero. This claim is supported by both the existing literature and by empirical results in this paper. Rothstein (2007) and Sabia (2009) find no effects on academic achievement, Tyler (2003) and Dustmann (2007) find small effects, and, using the instrumentation strategy proposed above (that would exaggerate the estimates), my study finds no effect.

And, second, classroom peer effects on achievement are known to be small (see, for example, Zimmerman, 2003). This means that even in the presence of small effects of part-time work on school performance, the compounded effects operating through peer achievement are likely to be negligibly small. Essentially, this study shows that part-time work during high school only affects an individual's future educational and labour market decisions, and these decisions are not affected by an individual's previous same-grade schoolmates.

The effect of hours worked during high school on future educational and labour market outcomes is modeled by the following equation:

$$Y_{igs}^{t+1} = \beta_0 + \beta_1 X_{igs} + \beta_2 \frac{1}{n_i - 1} \sum_{j \neq i} X_{jgs} + \beta_3 H_{igs}^{t} + D_g + D_s + \upsilon_{igs}^{t+1} \quad (4.3)$$

The outcome of interest at time $t+1$ (some future time) for individual i in grade g and school s at time t is denoted by Y_{igs}^{t+1}. Hours worked at time t is instrumented by classmates' hours worked at time $t-1$. The need to instrument is made explicit in the above equation because we expect that own hours worked H_{igs}^{t} is correlated with the error term ϵ_{igs}^{t+1} through some unobserved characteristic that affects both hours worked during school and the future educational or labour market outcome. Controls for both the characteristics of the student and the student's same-grade schoolmates are included, and grade and school fixed effects capture grade and school-specific differences in the outcome of interest and high school working behavior.

Results from the initial specification in which the coefficient on the hours worked term β_3 is not allowed to vary by grade are interpreted as an average effect of high school market work on the outcome of interest. Subsequent results explore grade (or age) heterogeneity by separately estimating the equation for 9th and 10th grade students and 11th and 12th grade students.[69]

Heterogeneity is also explored along the dimension of part-time work intensity. These results are obtained using a two-step procedure. The first-stage relationship is estimated on the full sample (as above), but the second-stage analysis is performed separately on individuals working fewer than 5 hours per week and individuals working at least 5 hours per week (during the school term). These results have a specific interpretation. For example, the β_3 coefficient from the equation estimated on the sample of individuals working fewer than 5 hours per week is the effect of an additional hour of work on the outcome of interest conditional on having chosen to work fewer than 5 hours per week. These results cannot account for the initial decision to work few or many hours.

4.3 Data and Descriptive Statistics

This paper uses data from the National Longitudinal Study of Adolescent Health (Add Health). The Add Health is a longitudinal study of a nationally representative sample of US adolescents who were in grades 7 to 12 during the 1994-1995 school year. The second wave of the study was conducted the subsequent year, and there have been two further in-home interviews, the most recent being in 2008. This paper uses data from the first, second and fourth waves of the study.

Descriptive statistics of the explanatory and control variables are provided in Table 4.1. The core sample consists of 8,429 students (after dropping observations with missing information). The part-time work information was obtained asking survey participants how many hours they spend working for pay during a typical non-summer week. In the second wave of the study, students in the sample worked an average of 8.5 hours per week during high school. The means for 8-10th grade students and 11-12th grade students are 5 and 12.5 hours, confirming that older students work more than younger students in high school. Over 40 percent of the students in the sample report no market work during high school. The share increases to 55 percent when adding students working less than four hours per week,

[69]Note that the corresponding first-stage relationship is also estimated separately, allowing the structure of peer effects in market work during high school to vary across the two groups.

leaving 45 percent of the sample working over five hours per week.[70] The distribution of part-time hours worked for all surveyed students working 40 hours or less is plotted in Figure 4.1. The cut-off of five hours is chosen when investigating heterogeneity as it lies approximately midway between the median and mean, ensuring enough variation in hours worked above and below the cut-off to estimate effects. Results remain qualitatively similar with small variations in this cut-off. Ruhm (1997) reports lower means from the NLSY, suggesting an increase in the intensity of part-time work during high school in the US from the early 1980s (NLSY) to the mid 1990s (Add Health).

The subsequent sets of variables describe demographic differences in part-time working behavior. Males engage in more market work than females during high school. There is a disproportionately high share of white students working over five hours per week while the share of black students working over five hours per week is disproportionately low. Evidence suggests that individuals with less educated parents work more intensively.

Table 4.2 describes the outcome variables. Most of these were measured during the fourth wave of the study when individuals were aged between 24 and 32. The first set of variables relate to education. Mean overall GPA scores (measured at the same time as the explanatory variable during Wave 2) vary little by grade or work intensity. 11-12th grade students are more likely to have graduated high school by Wave 4 of the study than 8-10th grade students. This is because some of the students in the 8-10th grade sample will choose to drop out by the time they would have been in the 11-12th grade sample. This form of selection is more evident when looking at college attendance; 68 percent of 8-10th grade students in the sample attend college while 74 percent of 11-12th grade students do so.

The second set of variables describes labour market outcomes in early adulthood. Differences in labour market outcomes across grades (or ages) and work intensities are mostly small and generally statistically insignificant. Older students earn more (which is somewhat mechanical as they are older and more experienced at the time of Wave 4 study), they are less likely to do physical work, and are more likely to be satisfied with their jobs. In terms of work intensity, students working over five hours per week earn more, work longer hours, are more likely to do physical work, and are more likely to be satisfied with their job than other students.

These correlations include the effects of several observed and unobserved

[70] Note that individuals working zero hours per week are included in the group working zero to four hours per week.

factors that vary with part-time working behavior and the outcome variable. The subsequent section reports causal effects.

4.4 Results

Results from the first stage of the instrumental variables estimation are reported in Table 4.3. Specifications include the full set of controls with errors clustered at the school-grade level. The first column reports that an increase of one hour in the mean hours worked by same-grade schoolmates the previous year is associated with a 0.51 increase in individual hours worked for the full sample. The coefficient is precisely estimated and the F-statistic is over 40, confirming a suitably strong relationship for implementation of the instrumental variables strategy. The second and third columns report the first-stage results for the samples of 8-10th and 11-12th grade students, respectively. The effect of mean peer hours worked on own hours worked drops to 0.31 and 0.36 for the two groups. It remains precise, although the F-statistic falls. The estimated coefficients on the controls show that females and black students work fewer hours per week (than males and white students), while older students (within a grade) work more than younger students.

Table 4.4 reports results from balance tests in which the instrument is regressed on the full set of controls. The purpose of this table is to show that observable controls are uncorrelated with the instrument after conditioning on the grade and school fixed effects necessary for identification. This suggests that the same is true for unobservable characteristics to the extent that observable and unobservable characteristics are correlated (Altonji et al, 2005), providing some support for the claim of instrument exogeneity. Grade and school fixed effects are included sequentially. The p-value of 0.26 on the F-statistic associated with the full specification in the third column indicates that we cannot reject the hypothesis that the coefficients on the individual controls are jointly equal to zero; the instrument is conditionally independent of observables.[71]

The effects of part-time work on four educational outcomes are reported in Table 4.5. All regressions include controls for individual characteristics, mean grade characteristics, grade fixed effects and school fixed effects. The

[71]The set of observable characteristics are considered non-identifying controls. They are included to increase the precision of the estimates rather than control for some form of selection. The school and grade fixed effects are considered necessary for identification of the effect. They deal with grade-specific and school-specific factors that may otherwise bias the estimates.

first column shows that the number of hours worked during high school has no effect on contemporaneous GPA scores. This supports the findings of Rothstein (2007) and Sabia (2009) in which the achievement effects of part-time work are zero. The absence of an effect on high school graduation reported in the second column is broadly consistent with this result.

There are, however, negative effects on the number of years of education and the probability of attending college. These are shown in the third and fourth columns. An additional hour of work reduces education by 0.06 years and reduces the probability of attending college by one percentage point. These results suggest that part-time work during high school affects the choices individuals make after graduating from high school without having affected achievement during high school. Essentially, students who engage in market work during high school appear more likely to enter the job market and less likely to pursue further education after high school. A variety of reasons for this are proposed.

First, part-time work during high school may lower job search costs upon high school graduation. Students may be able to continue working in their high school jobs after school or be given other opportunities with the same employer. The reduced uncertainty of finding work increases the expected returns from pursuing market work. Second, the opportunity cost of attending college may be greater for students who worked part-time during high school. This is because giving up a paying job to study is more costly than giving up staying at home and watching television (for example). And, third, students who work part-time during high school may have acquired more independence and therefore be more attached to the job market than other students. They may have developed spending habits and other behaviours that encourage working rather than studying.

These mechanisms cannot be directly investigated with the available data, but their plausibility is explored by analyzing the effects of part-time work on college expectations and a variety of subsequent labour market outcomes. The remaining results are all presented in the form of three-by-three tables for each outcome to reflect heterogeneous effects. Each cell reports the estimated coefficient from a regression estimated on the specified subsample. The columns consider grade (or age) heterogeneity and correspond to the full sample, 8-10th grade students, and 11-12th grade students, while the rows consider work intensity heterogeneity and correspond to the full sample, students working fewer than five hours per week, and students working at least five hours per week.

Results in Table 4.6 show that the only nonzero effect of part-time work on high school outcomes is for 11-12th grade students working fewer than five hours per week. For these students, there is a 0.1 reduction in GPA for an additional hour of work.

The negative effects of part-time work on years of education and college attendance are driven by 8-10th grade students working more than five hours per week. An additional hour of work for these students reduces education by 0.44 years and reduces the probability of attending college by eleven percentage points. These students are likely to have a strong attachment to the job market by the time they finish high school and their opportunity cost of studying may be greater than students working fewer or zero hours. Results in the table also indicate that these students have a reduced desire to attend college and expectation of attending college. These effects on expectations are somewhat consistent with those found by Neumark and Joyce (2011) in which school-to-work programs increased the perceived likelihood of future labor market activity. Interestingly, 11-12th grade students working fewer than five hours per week are more likely to both expect to attend and attend college. Working during high school may provide information that re-enforces the desire to pursue postsecondary education for these students.

The final table of results investigates the effect of working during high school on labour market outcomes. Recall that these were measured during the fourth wave of the study when individuals were aged between 24 and 32. Table 4.7 indicates that students who work more during high school have their first full-time job at younger ages than other students. This is driven by 8-10th grade students working more than five hours per week, the same group who were less likely to attend college. They enter the full-time labour market 0.23 years younger for every additional weekly hour of part-time work during high school. This supports the hypothesis that 8-10th grade students working more than five hours per week during high school choose are incentivized to enter the labour market rather than study after high school.

Working during high school is associated with increases in income for 11-12th grade students irrespective of their work intensity, although the results for hours worked also indicate that these students work more hours. For students working less than five hours, results in Table 4.5 indicate that this could be due to an increased probability of attending college. Generally, students working more than five hours per week during high school remain more hard-working than their same-grade schoolmates in early adulthood.

The final two variables describe the type of work and job satisfaction. These are proxies for job quality. Individuals who study rather enter the labour market after school may be employed in higher quality jobs when aged between 24 and 32, so given that part-time work during high school encourages early entry in the labour market, students who work more during high school may have lower quality jobs. Alternatively, students who enter the labour market early may have accumulated sufficient work experience to be promoted into more satisfying and higher quality jobs by their late twenties.

Results are not conclusive. There is no effect of part-time work during high school on self-reported job satisfaction, and the only significant effects on the probability of doing physical labour are for 11-12th grade students working more than five hours per week. For these students, there is some evidence that part-time work during high school results in subsequent selection out of jobs requiring physical labour.

4.5 Conclusion

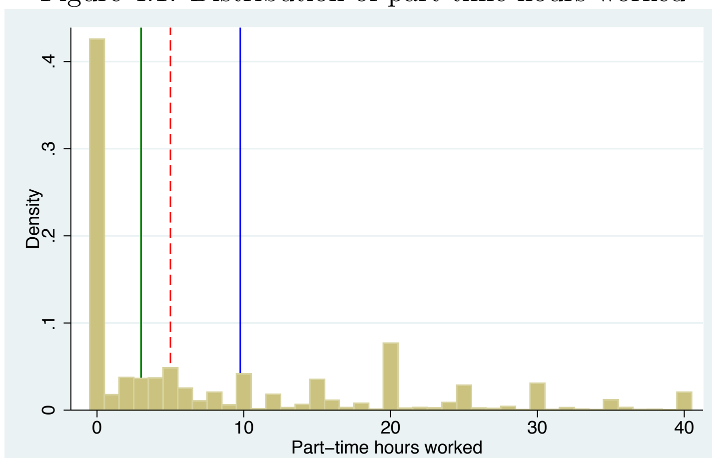
This paper contributes to our understanding of the effects of part-time work during high school by exploring grade and work intensity heterogeneity and focusing on early adult outcomes. In doing so, it provides an alternative narrative on the benefits and costs of market work during high school. Consistent with the existing literature (Rothstein, 2007; Sabia, 2009), there appears to be no effect on contemporaneous academic achievement. This paper rather focuses on effects on post-high school decision-making with respect to college attendance and entry into the full-time labour market.

These effects of part-time work during high school on subsequent labour outcomes differ by the grade in which the work occurred and the time intensity of the work. This paper finds that 8-10th grade students working more than five hours per week are both less likely to attend college than other students and begin full-time work at a younger age than other students. There is no effect on the college attendance decision and the age of first full-time job for 8-10th grade students working less than five hours per work, as well as 11-12th grade students working any number of hours. The effects for 8-10th grade students with high work intensity may be because they are strongly attached to the labour market by the time they graduate high school and have higher opportunity costs of postsecondary education than other students.

The effects on subsequent income appear to operate through other channels. An additional hour of part-time work during high school increases subsequent income for 11-12th grade students working any number of hours, while 8-10th grade students working less than five hours per week experience a negative income shock from part-time work. The positive effects on income for older students may be due to information or motivation gained from working during high school, or, more directly, the acquisition of skills and work experience that yield subsequent returns in the labour market. There is some evidence that these students are also more likely to attend college, which would also increase subsequent income.

4.6 Figures

Figure 4.1: Distribution of part-time hours worked



The median and mean number of hours worked are 3 and 9.8, respectively indicated by the solid gree and blue vertical lines. The dashed vertical line indicates the sample split at 5 hours used in the regression analysis.

4.7 Tables

Table 4.1: Descriptive statistics I - weekly hours worked and controls

	Mean (standard deviation)					
		8-10th	*11-12th*	*Hours worked (Wave 2)*		
	All	*grades*	*grades*	*0*	*0-4*	*Over 5*
Part-time work during high school						
Weekly hours worked	8.55	5.10	12.63			
in Wave 2	(11.47)	(8.84)	(12.82)			
Weekly hours worked	5.78	3.31	8.71			
in Wave 1	(9.62)	(6.99)	(11.34)			
Individual characteristics						
Female	0.50	0.50	0.51	0.53	0.51	0.48
White	0.67	0.68	0.64	0.57	0.61	0.73
Black	0.15	0.15	0.16	0.20	0.18	0.12
Hispanic	0.12	0.11	0.12	0.15	0.13	0.10
Asian	0.04	0.03	0.05	0.05	0.04	0.03
Other	0.03	0.03	0.03	0.03	0.03	0.03
Age (years and months)	16.27	15.34	17.83	15.94	15.80	16.83
Not born in US	0.05	0.04	0.07	0.06	0.06	0.05
Mother's education						
Less than high school	0.16	0.17	0.15	0.18	0.17	0.16
High school	0.35	0.36	0.33	0.32	0.32	0.37
Some college	0.18	0.18	0.19	0.18	0.18	0.19
College	0.26	0.25	0.27	0.26	0.27	0.24
Father's education						
Less than high school	0.13	0.13	0.14	0.14	0.13	0.14
High school	0.24	0.25	0.22	0.22	0.22	0.27
Some college	0.13	0.12	0.14	0.12	0.12	0.14
College	0.24	0.23	0.25	0.23	0.25	0.22
Household income (Wave 2)						
Less than $20k	0.15	0.17	0.12	0.17	0.16	0.14
$20k - $40k	0.24	0.24	0.23	0.23	0.23	0.25
$40k - $60k	0.20	0.21	0.19	0.19	0.20	0.20
More than $60k	0.22	0.21	0.25	0.21	0.23	0.22
Observations	8429	4570	3859	3558	4632	3797
Share	1	0.54	0.46	0.42	0.55	0.45

Table 4.2: Descriptive statistics II - outcomes (Wave 4 unless otherwise stated)

	Mean (standard deviation)					
		8-10th	*11-12th*	*Hours worked (Wave 2)*		
	All	*grades*	*grades*	*0*	*0-4*	*Over 5*
Educational outcomes						
Mean GPA score	2.83	2.83	2.84	2.82	2.85	2.81
(Wave 2)	(0.75)	(0.77)	(0.73)	(0.76)	(0.76)	(0.74)
Graduated high school	0.95	0.94	0.97	0.95	0.95	0.95
	(0.21)	(0.25)	(0.16)	(0.21)	(0.22)	(0.21)
Years of education	14.46	14.26	14.71	14.48	14.51	14.40
	(2.17)	(2.21)	(2.09)	(2.21)	(2.21)	(2.12)
Desire to attend college[a]	4.44	4.46	4.42	4.48	4.49	4.37
(Wave 2)	(1.03)	(1.02)	(1.04)	(0.99)	(0.98)	(1.08)
Expectation of attending	4.22	4.20	4.24	4.24	4.25	4.18
college (Wave 2)	(1.13)	(1.11)	(1.15)	(1.10)	(1.09)	(1.17)
Attended college	0.71	0.68	0.74	0.70	0.71	0.70
	(0.46)	(0.47)	(0.44)	(0.46)	(0.45)	(0.46)
Labour market outcomes						
Age at first full-time job	20.42	20.04	20.95	20.52	20.49	20.32
	(2.56)	(2.53)	(2.50)	(2.62)	(2.59)	(2.51)
Number of jobs	3.42	3.66	3.15	3.47	3.55	3.27
	(2.66)	(2.62)	(2.67)	(2.79)	(2.79)	(2.48)
Income	35088	31661	39095	33609	33617	36858
	(42198)	(37419)	(46864)	(45811)	(45462)	(37832)
Hours worked per week	41.11	40.93	41.31	40.62	40.70	41.60
	(11.20)	(11.41)	(10.96)	(11.16)	(11.32)	(11.04)
Do physical work	0.57	0.61	0.54	0.55	0.56	0.58
	(0.49)	(0.49)	(0.50)	(0.50)	(0.50)	(0.49)
Satisfied with job	0.74	0.72	0.75	0.71	0.72	0.76
	(0.44)	(0.45)	(0.43)	(0.45)	(0.45)	(0.43)

[a]The desire to attend college and expectation of attending college are self-reported rankings from 1 to 5. These were obtained during Wave 2 of the study.

Table 4.3: First-stage results - peer effects in part-time work during high school

Dependent variable: Part-time hours worked during high school	*All grades* (1)	*8-10th grades* (2)	*11-12th grades* (3)
Hours worked by same-grade schoolmates	0.51***	0.31**	0.36***
(in previous year)	(0.08)	(0.12)	(0.12)
Individual characteristics			
Female	-0.98***	-0.56*	-1.81***
Black	-1.84***	-1.34	-2.55**
Hispanic	0.04	0.42	-0.54
Asian	-0.31	-0.27	0.11
Other	0.68	0.40	1.44
Age (years and months)	1.87***	2.22***	1.54**
Not born in US	-0.42	-0.77	-0.17
Mother's education *(Omitted: high school)*			
Less than high school	-0.44	-0.09	-1.00
Some college	-0.22	-0.76*	0.55
College	-0.81*	-0.46	-1.38
Father's education *(Omitted: high school)*			
Less than high school	0.45	-0.22	1.74*
Some college	-0.22	-0.33	0.05
College	-1.08**	-0.98*	-1.06
Household income: *(Omitted: >$60k)*			
<$20k	0.82	0.63	0.73
$20k - $40k	1.01**	0.74	1.28
$40k - $60k	0.55	0.36	0.65
Other controls[a]	x	x	x
Diagnostics			
F-statistic on excluded instrument	43.33	6.34	9.38
Number of school-grade clusters	506	287	219
Observations	8429	4570	3859

[a]Other controls include indicators describing household structure, grade repetition history, school year in progress and school in saturated sample, as well as school-grade characteristics and grade and school fixed effects. Robust standard errors clustered by school-grade in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.4: Balance tests - OLS results from regressing instrument on controls

Dependent variable: hours worked by same-grade schoolmates (previous period)	(1)	(2)	(3)
Individual characteristics			
(Omitted: male, white, born in US)			
Female	0.09	-0.05	-0.05
Black	-1.25***	-1.32***	-0.15*
Hispanic	-0.84***	-0.99***	-0.06
Asian	-1.73***	-1.96***	-0.17
Other	-0.49*	-0.52*	0.09
Age (years and months)	2.07***	0.29***	0.04
Not born in US	-0.23	-0.20	-0.28**
Mother's education *(Omitted: high school)*			
Less than high high school	0.03	0.13	0.12
Some college	-0.05	-0.17	-0.08
College	-0.28**	-0.31***	-0.01
Father's education *(Omitted: high school)*			
Less than high high school	0.14	0.04	0.09
Some college	0.23	0.17	0.21**
College	-0.09	-0.19	0.06
Household income *Omitted: >$60k*			
<$20k	-0.23	-0.06	-0.07
$20k - $40k	-0.14	-0.11	-0.05
$40k - $60k	-0.07	-0.02	-0.10
Other controls[a]	x	x	x
Identifying controls			
Grade fixed effects		x	x
School fixed effects			x
Diagnostics			
F-statistic on non-identifying controls	22.04	3.55	1.17
p-value	0.00	0.00	0.26
Observations	8429	8429	8429

[a]Other controls include indicators describing household structure, grade repetition history, school year in progress and school in saturated sample, as well as school-grade characteristics. Robust standard errors clustered by school-grade in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.5: IV results - effect of part-time work during high school on educational outcomes

	Mean GPA score (1)	Graduated high school (2)	Years of education (3)	Attended college (4)
Hours worked	-0.004	-0.001	-0.06**	-0.01**
(instrumented)	(0.010)	(0.003)	(0.03)	(0.01)
Individual characteristics				
Female	0.19***	0.01	0.36***	0.06***
Black	-0.14***	0.01	-0.12	0.02
Hispanic	-0.21***	-0.04*	-0.24*	-0.03
Asian	0.14**	0.01	0.18	0.01
Other	-0.05	-0.03	-0.20	-0.01
Age (years and months)	-0.03	-0.04***	-0.20***	-0.04**
Not born in US	0.08	0.02	0.34**	0.08**
Mother's education				
Less than high school	-0.04	-0.05***	-0.25***	-0.07***
Some college	0.10***	0.01	0.41***	0.08***
College	0.15***	0.01	0.68***	0.10***
Father's education				
Less than high school	-0.03	-0.03**	-0.18*	-0.05*
Some college	0.14***	0.02*	0.42***	0.08***
College	0.21***	0.01*	0.76***	0.11***
Household income				
<$20k	-0.11***	0.00	-0.61***	-0.09
$20k - $40k	-0.11***	0.00	-0.41***	-0.05
$40k - $60k	-0.02	0.01	-0.21**	-0.02
Other controls[a]	x	x	x	x
Identifying controls				
Grade fixed effects	x	x	x	x
School fixed effects	x	x	x	x
Observations	8339	8429	8429	8429

[a]Other controls include indicators describing household structure, grade repetition history, school year in progress and school in saturated sample, as well as school-grade characteristics. Robust standard errors clustered by school-grade in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.6: IV results - educational outcomes

	All grades	*8-10th grades*	*11-12th grades*
		Mean GPA score (Wave 2)	
All hours	-0.004	-0.01	0.004
Hours worked<5	-0.01	0.02	-0.09***
Hours worked≥5	-0.003	-0.04	0.03
		Graduated high school	
All hours	-0.001	0.002	0.002
Hours worked<5	-0.003	0.010	-0.004
Hours worked≥5	0.002	-0.011	0.006
		Years of education	
All hours	-0.06**	-0.23**	0.03
Hours worked<5	0.01	0.04	0.13
Hours worked≥5	-0.11***	-0.44***	-0.04
		Desire to attend college (Wave 2)	
All hours	-0.01	-0.11**	0.08
Hours worked<5	0.01	-0.04	0.04
Hours worked≥5	-0.04*	-0.24***	0.02
		Expect to attend college (Wave 2)	
All hours	0.001	-0.12*	0.05
Hours worked<5	0.02	-0.02	0.14**
Hours worked≥5	-0.02	-0.20***	-0.04
		Attended college	
All hours	-0.01**	-0.07**	0.01
Hours worked<5	-0.002	-0.02	0.03**
Hours worked≥5	-0.02**	-0.11***	-0.003

Robust standard errors clustered by school-grade in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

Table 4.7: IV results - labour market outcomes

	All grades	*8-10th grades*	*11-12th grades*
		Age at first full-time job	
All hours	-0.08*	-0.38**	-0.02
Hours worked<5	-0.05	-0.13	-0.07
Hours worked≥5	-0.11**	-0.23*	-0.003
		Log(number of jobs)	
All hours	-0.002	0.01	0.04***
Hours worked<5	0.01	0.04	0.05*
Hours worked≥5	-0.01	-0.06*	0.04*
		Log(income)	
All hours	0.01	-0.01	0.09*
Hours worked<5	0.004	-0.08*	0.10*
Hours worked≥5	0.01	0.09	0.10***
		Log(hours worked per week)	
All hours	0.009**	0.01	0.03**
Hours worked<5	-0.003	-0.01	-0.01
Hours worked≥5	0.02***	0.04**	0.05***
		Do light or hard physical work	
All hours	-0.01**	-0.03	-0.02
Hours worked<5	-0.01	-0.02	0.02
Hours worked≥5	-0.01	0.01	-0.04***
		Satisfied with job	
All hours	0.002	-0.003	0.004
Hours worked<5	-0.01	-0.01	0.00
Hours worked≥5	0.01	0.03	0.01

Robust standard errors clustered by school-grade in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

Bibliography

[1] Joseph G Altonji, Todd E Elder, and Christopher R Taber. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1):151–184, 2005.

[2] Donald W.K. Andrews and James H. Stock. Inference with weak instruments. Working Paper 313, National Bureau of Economic Research, August 2005.

[3] Joshua D Angrist and William N Evans. Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *American Economic Review*, 88(3):450–477, 2011.

[4] Joshua D Angrist and Kevin Lang. Does School Integration Generate Peer Effects? Evidence from Bostons Metco Program. *American Economic Review*, 2004.

[5] Philip Babcock and Kelly Bedard. The Wages of Failure: New Evidence on School Retention and Long-Run Outcomes. *Education Finance and Policy*, 6(3):293–322, July 2011.

[6] M Bertrand and Sendhil Mullainathan. Do people mean what they say? Implications for subjective survey data. *The American Economic Review*, 91(2), 2001.

[7] Marianne Bertrand and Jessica Pan. The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, 5(1):32–64, 2013.

[8] R Bifulco and JM Fletcher. The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy*, 3(February):25–53, 2011.

[9] Alberto Bisin, Eleonora Patacchini, Thierry Verdier, and Yves Zenou. Formation and persistence of oppositional identities. *European Economic Review*, 55(8):1046–1071, 2011.

[10] Alison Booth and Patrick Nolen. Choosing to compete: How different are girls and boys? *Journal of Economic Behavior & Organization*, 81(2):542–555, February 2012.

[11] Alison L Booth and Patrick Nolen. Gender Differences in Risk Behaviour: Does Nurture Matter? *Economic Journal*, 122:56–78, 2012.

[12] Y. Bramoullé, H. Djebbari, and B. Fortin. Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55, 2009.

[13] W. a. Brock and S. N. Durlauf. Discrete Choice with Social Interactions. *The Review of Economic Studies*, 68(2):235–260, April 2001.

[14] MA Burke and TR Sass. Classroom peer effects and student achievement. *Journal of Labor Economics*, 31(1):51–82, 2013.

[15] Franz Buscha, Arnaud Maurel, Lionel Page, and Stefan Speckesser. The Effect of Employment while in High School on Educational Attainment: A Conditional Difference-in-Differences Approach. *Oxford Bulletin of Economics and Statistics*, 74(3):380–396, June 2012.

[16] David Card. The causal effect of education on earnings. *Handbook of Labor Economics*, 3:1801–1863, 1999.

[17] David Card and Laura Giuliano. Peer effects and multiple equilibria in the risky behavior of friends. *Review of Economics and Statistics*, (0), 2011.

[18] Scott E Carrell and Mark L Hoekstra. Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyones Kids. *American Economic Journal: Applied Economics*, 2(1):211–228, 2010.

[19] Scott E Carrell, Bruce I Sacerdote, and James E West. From natural variation to optimal policy? the lucas critique meets peer effects. Technical report, National Bureau of Economic Research, 2011.

[20] Piero Cipollone and Alfonso Rosolia. Social interactions in high school: Lessons from an earthquake. *American Economic Review*, 97(3):948–965, 2007.

[21] Andrew E Clark and Youenn Lohéac. "It wasn't me, it was them!" social influence in risky behavior by adolescents. *Journal of Health Economics*, 26(4):763–84, July 2007.

[22] Jane Cooley. Desegregation and the Achievement Gap: Do Diverse Peers Help? 2010.

[23] Sergio Currarini, Matthew Jackson, and Paolo Pin. An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, 77(4):1003–1045, 2009.

[24] Giacomo De Giorgi, Michele Pellizzari, and Silvia Redaelli. Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2):241–275, 2010.

[25] Yingying Dong. Kept back to get ahead? Kindergarten retention and academic performance. *European Economic Review*, 54(2):219–236, February 2010.

[26] Christian Dustmann and Arthur Soest. Part-time work, school success and school leaving. *Empirical Economics*, 32(2-3):277–299, August 2006.

[27] Eric R. Eide and Mark H. Showalter. The effect of grade retention on educational and labor market outcomes. *Economics of Education Review*, 20(6):563–576, December 2001.

[28] Kelly Foley. Can neighbourhoods change the decisions of youth on the margins of university participation? *Canadian Journal of Economics*, 45(1):167–188, February 2012.

[29] Nicole M Fortin, Philip Oreopoulos, and Shelley Phipps. Leaving boys behind: Gender disparities in high academic achievement. Technical report, National Bureau of Economic Research, 2013.

[30] Jane Friesen and Brian Krauth. Ethnic enclaves in the classroom. *Labour Economics*, 18(5):656–663, October 2011.

[31] JC Fruehwirth, Salvador Navarro, and Yuya Takahashi. How the timing of grade retention affects outcomes: Identification and estimation of time-varying treatment effects. 2011.

[32] Andrew J. Fuligni and Harold W. Stevenson. Time use and mathematics achievement among american, chinese, and japanese high school students. *Child Development*, 66(3):pp. 830–842, 1995.

[33] Constance Gager, Teresa Cooney, and Kathleen Thiede Call. The effects of family characteristics and time use on teenage girls and boys household labor. *Journal of Marriage and Family*, 61(4):982–994, 1999.

[34] Nancy L. Galambos, David M. Almeida, and Anne C. Petersen. Masculinity, femininity, and sex role attitudes in early adolescence: Exploring gender intensification. *Child Development*, 61(6):pp. 1905–1914, 1990.

[35] Uri Gneezy, Kenneth L. Leonard, and John A. List. Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society. *Econometrica*, 77(5):1637–1664, 2009.

[36] Uri Gneezy, Muriel Niederle, and Aldo Rustichini. Performance in competitive environments: gender differences. *Quarterly Journal of Economics*, (August):1049–1074, 2003.

[37] Steven M Goodreau, James A Kitts, and Martina Morris. Birds of a Feather, Or Friend of a Friend?: Using Exponential Random Graph Models to Investigate Adolescent Social Networks. *Demography*, 46(1):103–125, 2009.

[38] Dominique Goux and Eric Maurin. Close Neighbours Matter: Neighbourhood Effects on Early Performance at School. *Economic Journal*, 117:1193–1215, 2007.

[39] Mark S Granovetter. The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.

[40] Jay P. Greene and Marcus a. Winters. The effects of exemptions to Florida's test-based promotion policy: Who is retained? *Economics of Education Review*, 28(1):135–142, February 2009.

[41] Andrew Hacker. Is Algebra Necessary? *The New York Times*, July 28, 2012, 2012.

[42] Iida Häkkinen. Working while enrolled in a university: does it pay? *Labour Economics*, 13(2):167–189, April 2006.

[43] C T Halpern, K Joyner, J R Udry, and C Suchindran. Smart teens don't have sex (or kiss much either). *The Journal of Adolescent Health*, 26(3):213–25, March 2000.

[44] Diane F Halpern, Lise Eliot, Rebecca S Bigler, Richard A Fabes, Laura D Hanish, Janet Hyde, Lynn S Liben, and Carol Lynn Martin. The Pseudoscience of Single-Sex Schooling. *Science*, 333:1706–1707, 2011.

[45] E.A. Hanushek. The failure of input-based schooling policies. *Economic Journal*, 113(485):64–98, 2003.

[46] Eric A Hanushek, John F Kain, and Steven G Rivkin. New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement. *Journal of Labor Economics*, 27(3), 2009.

[47] Duke Helfand. A Formula for Failure in L . A . Schools. *The Los Angeles Times*, January 30, 2006, 2006.

[48] Christian Helmers and Manasa Patnam. Does the Rotten Child Spoil His Companion? Spatial Peer Effects Among Children in Rural India. 2010.

[49] V. Joseph Hotz, Lixin Colin Xu, Marta Tienda, and Avner Ahituv. Are There Returns to the Wages of Young Men from Working While in School? *Review of Economics and Statistics*, 84(2):221–236, May 2002.

[50] Caroline Hoxby. Peer effects in the classroom: Learning from gender and race variation. Technical report, National Bureau of Economic Research, 2000.

[51] Yannis M Ioannides and Linda Datcher Loury. Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, 42(4):1056–1093, 2004.

[52] C Kirabo Jackson. Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in trinidad and tobago. *Journal of Public Economics*, 96(1):173–187, 2012.

[53] BA Jacob and Lars Lefgren. Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics*, 86(February):226–244, 2004.

[54] Brian a Jacob and Lars Lefgren. The Effect of Grade Retention on High School Completion. *American Economic Journal: Applied Economics*, 1(3):33–58, June 2009.

[55] Charlene Marie Kalenkoski and Sabrina Wulff Pabilonia. Time to work or time to play: The effect of student employment on homework, sleep, and screen time. *Labour Economics*, 19(2):211–221, April 2012.

[56] Michael P Keane and Kenneth I Wolpin. The career decisions of young men. *Journal of Political Economy*, 105(3):473–522, 1997.

[57] Brian Krauth. Peers as treatments. 2011.

[58] Victor Lavy, MD Paserman, and Analia Schlosser. Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom*. *The Economic Journal*, 122:208–237, 2012.

[59] Victor Lavy and Edith Sand. The friends factor: How students social networks affect their academic achievement and well-being? Working Paper 18430, National Bureau of Economic Research, October 2012.

[60] Victor Lavy and Analía Schlosser. Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics*, 3(October 2006):1–33, 2011.

[61] Victor Lavy, Olmo Silva, and Felix Weinhardt. The good, the bad, and the average: evidence on ability peer effects in schools. *Journal of Labor Economics*, 30(2):367–414, 2012.

[62] Audrey Light. High school employment, high school curriculum, and post-school wages. *Economics of Education Review*, 18:291–309, 1999.

[63] X. Lin. Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28(4):825–860, 2010.

[64] Marco Manacorda. The Cost of Grade Retention. *Review of Economics and Statistics*, 94(2):596–606, May 2012.

[65] Charles F Manski. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60(3):531–542, 1993.

[66] Charles F Manski. Economic Analysis of Social Interactions. *Journal of Economic Perspectives*, 14(3):115–136, August 2000.

[67] Claude Montmarquette, Nathalie Viennot-Briot, and Marcel Dagenias. Dropout, school performance, and working while in school. *The Review of Economics ...*, 89(November):752–760, 2007.

[68] Marcelo J Moreira and Brian Poi. Implementing Conditional Tests with Correct Size in the Simultaneous Equations Model. *Stata Journal*, (1):1–15, 2001.

[69] Ted Mouw and Barbara Entwisle. Residential Segregation and Interracial Friendship in Schools. *American Journal of Sociology*, 112(2):394–441, 2006.

[70] David Neumark and Mary Joyce. Evaluating school-to-work programs using the new NLSY. *Journal of Human Resources*, 36(4):666–702, 2001.

[71] GS Oettinger. Does high school employment affect high school academic performance? *Industrial and Labor Relations Review*, 53(1):136–151, 1999.

[72] Philip Oreopoulos and Kjell G Salvanes. How large are the returns to schooling? hint: Money isn't everything. Working Paper 15339, National Bureau of Economic Research, September 2009.

[73] SW Pabilonia. Evidence on youth employment, earnings, and parental transfers in the National Longitudinal Survey of Youth 1997. *Journal of Human Resources*, 36(4), 2012.

[74] E. Patacchini, E. Rainone, and Y. Zenou. Student Networks and Long-Run Educational Outcomes: The Strength of Strong Ties. 2012.

[75] François Poulin, Anne-Sophie Denault, and Sara Pedersen. Longitudinal Associations Between Other-Sex Friendships and Substance Use in Adolescence. *Journal of Research on Adolescence*, 21(4):776–788, December 2011.

[76] Barbara J Reys, Shannon Dingman, Nevels Nevels, and Dawn Teuscher. High School Mathematics: State-Level Curriculum Standards and Graduation Requirements. Technical report, Center for the Study of Mathematics Curriculum, 2007.

[77] Heather Rose and Julian R. Betts. The Effect of High School Courses on Earnings. *Review of Economics and Statistics*, 86(2):497–513, May 2004.

[78] DS Rothstein. High school employment and youths' academic achievement. *Journal of Human Resources*, (October 2003), 2007.

[79] CJ Ruhm. Is high school employment consumption or investment? *Journal of Labor Economics*, 15(4):735–776, 1997.

[80] Joseph J. Sabia. Reading, writing, and sex: The effect of losing virginity on academic performance. *Economic Inquiry*, 45(4):647–670, 2007.

[81] Joseph J. Sabia. School-year employment and academic performance of young adolescents. *Economics of Education Review*, 28(2):268–276, April 2009.

[82] B Sacerdote. Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics*, (May), 2000.

[83] Bruce Sacerdote. *Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?*, volume 3 of *Handbook of the Economics of Education*, chapter 4, pages 249–277. Elsevier, June 2011.

[84] Nicole Schneeweis and Martina Zweimüller. Girls, girls, girls: Gender composition and female school choice. *Economics of Education Review*, 31(4):482–500, August 2012.

[85] Ralph Stinebrickner and Todd R. Stinebrickner. What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics*, 90(8-9):1435–1454, September 2006.

[86] Ralph Stinebrickner and TR Stinebrickner. Working during school and academic performance. *Journal of Labor Economics*, 21(2):473–491, 2003.

[87] James H Stock, Jonathan H Wright, and Motohiro Yogo. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics*, 20(4):518–529, 2002.

[88] Petra E. Todd and Kenneth I. Wolpin. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485):F3–F33, 2003.

[89] JohnH. Tyler. Using State Child Labor Laws to Identify the Effect of SchoolYear Work on High School Achievement. *Journal of Labor Economics*, 21(2):381–408, April 2003.

[90] Glen R. Waddell. Gender and the influence of peer alcohol consumption on adolescent sexual activity. *Economic Inquiry*, 50(1):248–263, 2012.

[91] Diane Whitmore. Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment. *American Economic Review*, 95(2):199–203, 2011.

[92] DJ Zimmerman. Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85(February):9–23, 2003.

Appendix A

The Effect of Peer Gender Composition on High School Achievement

The appendix describes results from a variety of robustness and sensitivity checks. It also provides an econometric framework from which the estimated parameters can be given a cumulative effect interpretation. Appendix Table A.1 describes the composition of the sample, and Appendix Table A.2 provides a balance test suggesting orthogonality of the instrument. These are discussed in the main body of the paper. The remaining appendix discussion is split into four subsections. The first subsection discusses the sensitivity and robustness of the instrument, the second subsection outlines the potential for non-classical measurement error and explains how the empirical strategy overcomes it, the third section shows that the first stage relationship is not driven by spatial outliers, and the fourth section derives the cumulative effect interpretation.

A.1 Sensitivity of Results to Instrument Specification

The instrument is constructed using the weighted gender composition of the nearest twenty schoolmates. Results in Appendix Table A.3 show that an unweighted version of this instrument, as well as weighted and unweighted instruments based on the gender composition of schoolmates within 2km, generate similar findings. The estimates using the distance-based measure are less precise but of similar magnitude to those reported in Table 2.5.

Appendix Table A.4 considers the sensitivity of results both to the density of friendship networks from which the gender composition measures are derived and the chosen definition of friendship. The first set of results splits the sample according to the number of friend nominations asked of surveyed individuals and restricts the sample to individuals with at least two friends. The second set of results repeats the primary analysis using different definitions of friendship. These results address concerns related to differences between true and observed friendship networks, and the friendship definition on which these networks are based.

Individuals were asked to nominate either one friend or five friends. The study was designed so that all individuals in the same school nominated the same number of friends. The gender composition measures derived from friendship networks based on five friend nominations are likely to be measured with less error than those based on single nominations. This table shows that potential biases introduced by this aspect of the design do not affect the initial findings. Results are consistent with those originally reported, although they are measured imprecisely due to the smaller sample sizes.

The friendship network gender composition of individuals matched to only one friend are extreme; observed opposite gender friend shares are either zero or one. To show that these individuals do not drive the result, the analysis is performed on a sample restricted to individuals with at least two friends. The point estimate of interest is very similar in this specification. The results are also similar when the analysis is performed on the restricted sample of individuals for whom at least seventy-five percent of friendship nominations were matched. The gender composition of friendship networks for these individuals is likely measured with less error, explaining the maintained precision of the estimates despite the smaller sample.

The nominating process discussed in the data section allows for different definitions of friendships. The preferred definition of friendship for this paper considers any friendship nomination to form a friendship. This is because the identification strategy relies on neighbours affecting outcomes only through the friendship network, and the weakest definition of the friendship network is most likely to satisfy this exclusion restriction. Two alternative friendship networks definitions based on the nomination process are directional: nominated and nominating friendship networks. These definitions only consider either sent or received nominations to form friendships, respectively. A fourth definition of the friendship network is the strong friendship network discussed in the body of the paper in which only reciprocated nominations form friendships.

Estimates in Appendix Table A.4 are similar to those reported in the paper for weak friendship networks, although they are less precisely measured due to the smaller sample sizes. (The sample sizes are smaller because the stronger definitions result in greater exclusion from the sample. Recall that individuals are excluded from the sample if they are not matched to any friends as the gender composition of friendship networks is not well-defined for these individuals.)

The urbanicity of the community in which the school is located may affect the results. This is both because the dependence of achievement on peer gender composition may vary between urban, suburban and rural schools, and because the first stage relationship between distance and friendship may have a different structure across these types of communities. Results in Table A.5 show that the first stage is weak in urban schools (the first column), so the instrument cannot inform our understanding of the effect in these communities. However, restricting the sample to suburban and rural schools (the fourth column) shows the negative effect of opposite gender friends on achievement.

The validity of the IV strategy is tested in two ways in Appendix Table A.6. First, it considers an experiment similar to randomly reassigning the gender of schoolmates and showing that the reassigned gender composition of close neighbours does not affect the original gender composition of friendship networks. And, second, it confirms the first stage relationship for a composition measure for which randomness is even less contestable than gender.

These tests are performed by introducing another composition measure: share even birth month. Consider an individual with a set of neighbours in the data. Now consider an experiment in which the gender of neighbours is reassigned so that neighbours with an even birth month are of the opposite gender. The share of reassigned opposite gender close neighbours (equivalent to the share of even birth month close neighbours) should not be correlated with the share of (true) opposite gender friends. The first column of Appendix Table A.6 reports results from regressing the share of opposite gender friends on the share of even birth month close neighbours and shows that they are uncorrelated.

The share of even birth months friends should have no effect on academic achievement, but, given that individuals are more likely to be friends with schoolmates living in the close neighbourhood, the share of even birth month close neighbours should affect an individual's share of even birth month friends. The second column of Appendix Table A.6 confirms the presence of this relationship. Using this as the first stage of an (unnecessary) instrumental variables strategy, the share of even birth month friends is shown to have no effect on school performance (as expected). This table supports the validity of the first stage by performing a placebo test on the first stage and due to the absence of plausible alternative explanations for the relationship between the share of even birth month neighbours and friends.

A.2 Non-classical Measurement Error Arising from Self-reporting Bias

The gender composition of an individual's friendship network is derived from self-nominated friends, and the measure of academic achievement is self-reported GPA. These data may suffer from self-reporting bias. This section outlines the potential for such biases, as well as showing that an instrumental variables strategy deals with these concerns under the assumption that the instrument is orthogonal to self-reporting bias in the outcome and explanatory variable. This result is somewhat obvious, but is useful for understanding the direction of the potential bias under various assumptions about the self-reporting bias and its correlation with other variables in the model.

Consider a simple model in which the true value of an outcome y^* is a linear function of the true value of an explanatory variable x^*.

$$y^* = x^*\beta + e \tag{A.1}$$

Allow for some form of endogeneity, so $\text{cov}(x^*, e) \neq 0$.

We can write the measured values of the outcome and explanatory variables y and x as the sum of the true values and a measurement error term. These error terms are not random, allowing for some form of systematic self-reporting bias.

$$y = y^* + u_y \tag{A.2}$$

$$x = x^* + u_x \tag{A.3}$$

We can substitute Equations A.2 and A.3 into the true model in Equation A.1.

$$y = x\beta + (e + u_y + u_x\beta) \tag{A.4}$$

The error term in Equation A.4 is clearly correlated with the explanatory variable. The correlation between x and e follows from the endogeneity, and there is a mechanical correlation between x and u_x from Equation A.3.

We can investigate the consistency of an OLS estimate $\hat{\beta}$ more formally.

$$\begin{aligned}
\text{plim}\,\hat{\beta} &= \tfrac{\text{cov}(y,x)}{\text{var}(x)} \\
&= \tfrac{\text{cov}(y^*+u_y,x^*+u_x)}{\text{var}(x^*+u_x)} \\
&= \tfrac{\text{cov}(x^*\beta+e+u_y,x^*+u_x)}{\text{var}(x^*+u_x)} \\
&= \tfrac{1}{\text{var}(x^*)+\text{var}(u_x)}\, \{\beta\,\text{var}(x^*) + \beta\,\text{cov}(x^*, u_x) + \text{cov}(x^*, e) \\
&\qquad + \text{cov}(u_x, e) + \text{cov}(x^*, u_y) + \text{cov}(u_x, u_y)\}
\end{aligned} \tag{A.5}$$

This equation breaks down the potential biases into six components (the terms of the sum inside the braces). The first and third components are the familiar attenuation and endogeneity biases. The remaining components are best understood in terms of the example of this paper. Consider the outcome to be GPA and the explanatory variable to be peer gender composition.

The second component $\text{cov}(x^*, u_x)$ is the bias introduced by correlation between true peer gender composition and self-reporting bias in peer gender composition. For example, individuals with few opposite gender friends may over-report opposite gender friendships, such that $\text{cov}(x^*, u_x) < 0$. This would either bias the estimate towards zero or change the sign of the estimate depending on the relative magnitude of the attenuation bias.

The fourth component $\text{cov}(u_x, e)$ relates to correlation between unobserved determinants of GPA and self-reporting bias in peer gender composition. This may involve some personality trait such as overconfidence. Overconfident individuals may both perform poorly academically and over-report opposite gender friendships, for example. This would also bias the point estimate downwards as $\text{cov}(u_x, e) < 0$.

The fifth component $\text{cov}(x^*, u_y)$ is bias introduced by the correlation between true peer gender composition and self-reporting bias in GPA. For example, males with a larger share of female friends may systematically over-report GPA if female friends are of higher ability (on average), and individuals have a propensity toward reporting the mean GPA of their friendship networks.

Finally, the sixth component $\text{cov}(u_x, u_y)$ relates to correlation between self-reporting bias in peer gender composition and self-reporting bias in GPA. This correlation would be generated by a world in which some people consistently tell the truth and others consistently distort the truth. For example, some individuals may systematically exaggerate all self-reported data in the direction that is perceived to be more socially-favourable.

The above analysis highlights the potential concerns of using self-reported data. The subsequent analysis shows that using an instrumental variables strategy deals with these concerns under some assumptions about the instrument. (Of course it is already well-known that instruments deal with the attenuation and endogeneity biases.)

Consider an instrument for the explanatory variable z^*, such that $\text{cov}(z^*, e) = 0$. It is assumed to be of the same scale for ease of exposition.

$$z^* = x^* + \epsilon \tag{A.6}$$

Now consider the covariance between the measured outcome variable y

and the instrument z^*.

$$\begin{aligned}
\operatorname{cov}(y, z^*) &= \operatorname{cov}(y^* + u_y, z^*) \\
&= \operatorname{cov}(x^*\beta + e + u_y, z^*) \\
&= \operatorname{cov}(x\beta + e + u_y - u_x\beta, z^*) \\
&= \beta \operatorname{cov}(x, z^*) + \operatorname{cov}(z^*, e) + \operatorname{cov}(z^*, u_y) - \beta \operatorname{cov}(z^*, u_x) \\
&= \beta \operatorname{cov}(x, z^*) + \operatorname{cov}(z^*, u_y) - \beta \operatorname{cov}(z^*, u_x)
\end{aligned} \tag{A.7}$$

In order for $\beta = \frac{cov(y,z^*)}{\operatorname{cov}(x,z^*)}$, the familiar exactly-identified univariate IV result, it is required that $\operatorname{cov}(z^*, u_y) = 0$ and $\operatorname{cov}(z^*, u_x) = 0$. In other words, the self-reporting biases in the outcome and explanatory variables need to be uncorrelated with the instrument.

In terms of this paper, self-reporting biases in GPA and peer gender composition need to be uncorrelated with neighbourhood gender composition. The only real concern is that neighbourhood gender composition may affect self-reporting bias in peer gender composition. For example, an individual with a large share of opposite gender close neighbours may over-report opposite gender friends.

Appendix Table A.7 reports results from regressing constructed proxies for measurement error in peer gender composition and GPA on the instrument, a female indicator, and self-reported GPA. Proxies for measurement error are constructed by differencing the observed measure from another measure that does not suffer from potential self-reporting biases. The gender composition measurement error proxy is the difference between the weak friendship network (in which all nominations generate friendships, so susceptible to self-reporting biases) and strong friendship network (reciprocated nominations generate friendships, so less susceptible to self-reporting bias) gender composition measures, and the achievement measurement error proxy is the difference between self-reported and transcript GPA scores (for the subsample for which transcript GPA scores are available). These results suggest that the instrument is uncorrelated with measurement error in self-reported GPA and peer gender composition, providing support for the empirical strategy.

A.3 Boundary Concerns

One remaining concern is that distance to the community origin may generate the first stage relationship between the gender composition of close neighbours and the gender composition of friends. Consider a world in which gender is spatially uniformly distributed throughout the school community and individuals are only friends with their neighbours. Individuals close to the community origin will have an equal share of own and opposite gender neighbours, and therefore an equal share of own and opposite gender friends. Individuals at the community boundary, though, may only have only neighbour, and therefore only one friend.

This type of community organization and friendship formation would generate the observed first stage, but variation would be driven completely by individuals at the community boundary. These individuals are likely to differ systematically from individuals at the community origin, and therefore reduce the generalizability of the estimated local effect.

Appendix Table A.8 shows that this is not a concern in the data. This table splits the sample into three groups according to the distance between individuals and the community origin (defined as the mean X- and Y-coordinates in a school community). The first stage is strongest for those in the middle third (in terms of distance to the community origin) and not the furthest third, showing that the effect is not driven by individuals at the boundary. Interestingly, the effect is weak for individuals closest to the community origin. This is consistent with the idea that the increased density of schoolmates in the neighbourhood close to the community origin may result in increased opportunities for friendship formation. This reduces the relationship between distance and friendship as individuals are able to choose among their close neighbours for matches that make better friends.

A.4 Cumulative Effect Interpretation

The discussion in the main body of the paper provides a contemporaneous interpretation of the friendship network gender composition effect. This section recognizes that peer gender composition affects education production in every period, allows education production to depend on past production[72], and, given persistence in friendship networks, shows that the parameter estimated in the paper can be interpreted as the cumulative effect of opposite gender friends. It subsequently provides a set of assumptions that allow separation of the cumulative and contemporaneous effects.

Consider the simplified education production of individual i at age t to be a a function of individual characteristics X and the share of opposite

[72] Hanushek (2003) and Todd and Wolpin (2003) discuss different formulations of education production functions, particularly the different assumptions underlying level and value-added specifications.

gender friends O:

$$Y_{it} = \beta_t X_{it} + \gamma_t O_{it} + \epsilon_{it} \tag{A.8}$$

In this formulation, the age-varying parameters β_t and γ_t can be interpreted as the cumulative effects of individual characteristics and opposite gender friends on achievement (up to age t). We can explicitly include lagged achievement to give these parameters a contemporaneous (or value-added) interpretation (dropping the i subscript for clarity).

$$Y_t = \beta_t X_t + \gamma_t O_t + \lambda_t Y_{t-1} + e_t \tag{A.9}$$

Education production at age t can then be expressed as a function of the history of individual characteristics $\{X_t, X_{t-1}, \ldots, X_1\}$, opposite gender friend shares $\{O_t, O_{t-1}, \ldots, O_1\}$, initial ability Y_0, and the history of production shocks $\{e_t, e_{t-1}, \ldots, e_1\}$:

$$\begin{aligned} Y_t =& \beta_t X_t + \sum_{j=1}^{t-1} \beta_{t-j} \prod_{k=1}^{j} \lambda_{t+1-k} X_{t-j} + \gamma_t O_t + \sum_{j=1}^{t-1} \gamma_{t-j} \prod_{k=1}^{j} \lambda_{t+1-k} O_{t-j} \\ &+ \prod_{k=1}^{t} \lambda_{t+1-k} Y_0 + e_t + \sum_{j=1}^{t-1} \prod_{k=1}^{j} \lambda_{t+1-k} e_{t-j} \end{aligned} \tag{A.10}$$

We now make two simplifying assumptions.

*A*1. $X_u = X_{u-1}$ for all u.

*A*2. The share of opposite gender friends evolves according to the following process:

$$\begin{aligned} O_1 &= \gamma_0 Z + u_1 \\ O_2 &= \kappa_2 O_1 + u_2 \\ &\ldots \\ O_t &= \kappa_t O_{t-1} + u_t \end{aligned} \tag{A.11}$$

The first assumption fixes individual characteristics as an individual ages. This assumption is not restrictive for characteristics such as gender, race and immigrant status. It may have some bite for characteristics such as parent education and household income that potentially vary for some school-going individuals over time.

The second assumption describes a simple evolution for the share of opposite gender friends. The initial share of opposite gender friends O_1 is a linear function of the share of opposite gender friends in the close neighbourhood Z with all other initial determinants of friendship composition in the error term u_1.[73] Opposite gender friend shares then follow an AR(1) process where friendship network gender composition depends on lagged friendship network gender composition and an additive shock (which is not necessarily orthogonal to other components of the model). This is reasonable given persistence in friendship networks as an individual ages (networks do not reset every period). The model does not allow families to relocate.

These assumptions allow us to express education production as a function of individual characteristics X_t, current share of opposite gender friends O_t, initial ability Y_0, the history of shocks to production $\{e_t, e_{t-1}, \ldots, e_1\}$, and the history of shocks to friendship network gender composition that affect current achievement through their affect on past achievement $\{u_t, u_{t-1}, \ldots, u_1\}$.

$$\begin{aligned} Y_t = &\left\{ \beta_t + \sum_{j=1}^{t-1} \beta_{t-j} \prod_{k=1}^{j} \lambda_{t+1-k} \right\} X_t + \left\{ \gamma_t + \sum_{j=1}^{t-1} \gamma_{t-j} \prod_{k=1}^{j} \frac{\lambda_{t+1-k}}{\kappa_{t+1-k}} \right\} O_t \\ &+ \prod_{k=1}^{t} \lambda_{t+1-k} Y_0 + e_t + \sum_{j=1}^{t-1} \prod_{k=1}^{j} \lambda_{t+1-k} e_{t-j} \\ &- \sum_{j=1}^{t-1} \gamma_{t-j} \prod_{k=1}^{j} \lambda_{t+1-k} \sum_{l=1}^{j} \left\{ \prod_{m=l}^{j} \kappa_{t+1-m} \right\}^{-1} u_{t+1-l} \end{aligned} \tag{A.12}$$

Along with the direct effects of characteristics and peer gender composition operating through β_t and γ_t, the additional term in the respective coefficients describe the indirect effects operating through past achievement. Under this framework, consistent estimates of the coefficients on X_t and O_t should therefore be interpreted as the cumulative effects of individual characteristics and the share of opposite gender friends on achievement.

The initial endogeneity problem in estimating the effect of opposite gender friends arose because of the correlation between friendship network gen-

[73] The effect of individual characteristics on friendship network gender composition is omitted at this stage for tractability; it would not change the econometric analysis if included. It is included when the model is taken to the data.

der composition and the contemporaneous unobservable determinants of education production. This formulation suggests additional concern due to potential correlation with the both the history of education production shocks, and the past shocks to friendship network composition that affect current achievement through past achievement.

The neighbourhood gender composition instrument, however, remains correlated with the current share of opposite gender friends and orthogonal to all components of the error term (initial ability and the shocks), allowing identification of the cumulative effect of opposite gender friends. The orthogonality with initial ability and unobservable determinants of achievement follows from the same arguments as provided in the initial exposition. Orthogonality with the shocks to friendship network gender composition are a consequence of the assumption that the only direct effect of close neighbourhood gender composition on friendship network gender composition is in the initial period.

The correlation between current share of opposite gender friends and share of opposite gender close neighbours is evident in the below expression for O_t, which is essentially the first stage for using Z as an instrument for O_t.

$$\begin{aligned} O_t &= \kappa_t O_{t-1} + u_t \\ &= \prod_{j=1}^{t-1} \kappa_{j+1}\gamma_0 Z + \sum_{j=1}^{t-1}\prod_{k=1}^{j} \kappa_{t+1-k} u_{t-j} + u_t \end{aligned} \tag{A.13}$$

The predicted friendship network gender composition $\hat{O}_t = \prod_{j=1}^{t-1} \widehat{\kappa_{j+1}\gamma_0} Z$ identifies the cumulative effect of opposite gender friends in Equation A.12.

The following two equations define the reduced form first and second stage parameters for an individual of age t. The dependence of the share of opposite gender friends on characteristics X is made explicit.

$$O_t = \pi_t X_t + \omega_t Z + v_t \tag{A.14}$$

$$Y_t = \psi_t X_t + \rho_t O_t + \mu_t \tag{A.15}$$

The reduced form parameters of interest can be expressed in terms of the underlying parameters.

$$\omega_t = \prod_{j=1}^{t-1} \kappa_{j+1}\gamma_0 \tag{A.16}$$

$$\rho_t = \left\{ \gamma_t + \sum_{j=1}^{t-1} \gamma_{t-j} \prod_{k=1}^{t} \frac{\lambda_{t+1-k}}{\kappa_{t+1-k}} \right\} \tag{A.17}$$

The time-varying underlying parameters are not identified even with data for individuals of different ages. Additional assumptions separate the contemporaneous and cumulative effects of opposite gender friends.

*A*3. $\beta_u = \beta_{u-1}$, $\gamma_u = \gamma_{u-1}$, $\lambda_u = \lambda_{u-1}$ and $\kappa_u = \kappa_{u-1}$ for all u.
*A*4. $\lambda \neq 1$.
*A*5. $\frac{\lambda}{\kappa} \neq 1$.

The third assumption imposes constancy in the parameters over all ages in both the education production function described by Equation A.12 and the friendship network gender composition process described by Equation A.13.[74] Assumption *A*3 identifies the underlying parameters if we observe individuals of different ages. The fourth and fifth assumptions simply allow us to use the formula for summation of a geometric sequence. These are easily relaxed, and an alternative derivation is provided below for when *A5* does not hold.

Given these assumptions, education production for an individual of high school age t is given by:

$$\begin{aligned} Y_t =& \beta \frac{1-\lambda^t}{1-\lambda} X_t + \gamma \frac{1-(\frac{\lambda}{\kappa})^t}{1-(\frac{\lambda}{\kappa})} O_t + \lambda^t Y_0 + e_t + \sum_{j=1}^{t-1} \lambda^j e_{t-j} \\ & - \gamma \sum_{j=1}^{t-1} \lambda^j \sum_{l=1}^{j} \kappa^{l-j-1} u_{t+1-l} \end{aligned} \tag{A.18}$$

The simplified first stage (omitting dependence on X as before) is given

[74] This may be less restrictive if we limit the model to describe individuals of middle and high school ages where initial ability is that accumulated by the beginning of middle school and middle school initiates a new friendship gender composition process.

by

$$O_t = \kappa^{t-1}\gamma_0 Z + \sum_{j=1}^{t-1} \kappa^j u_{t-j} \tag{A.19}$$

The parameters of interest are identified up to an indexing of a from the first and second stage estimates for two adjacent age groups.[75] Table 2.6 reports estimates for those above and below the age of sixteen, providing the necessary information. Finer sample splits provide overidentifying restrictions, but are costly in terms of statistical precision given the small samples.

The following four equations determine the contemporaneous effect of opposite gender friends γ, the effect of lagged achievement λ, the friendship network gender composition process κ, and the correlation between initial friendship network gender composition and close neighbourhood gender composition γ_0. The standard errors of these parameters need to be bootstrapped given that λ is the solution to a higher order polynomial for which an analytical solution may not exist.

$$\kappa = \frac{\omega_{a+1}}{\omega_a} \tag{A.20}$$

$$\gamma_0 = \frac{\omega_a^a}{\omega_{a+1}^{a-1}} \tag{A.21}$$

$$\rho_a(\frac{\lambda}{\kappa})^{a+1} - \rho_{a+1}(\frac{\lambda}{\kappa})^a + (\rho_{a+1} - \rho_a) = 0 \tag{A.22}$$

$$\gamma = \rho_a \frac{1 - (\frac{\lambda}{\kappa})}{1 - (\frac{\lambda}{\kappa})^a} \tag{A.23}$$

λ and κ describe AR(1) processes for education production and friendship network gender composition. It is not unreasonable to consider the case where these parameters are equal, violating $A5$ (the assumption that allowed us to express the coefficient as the summation of a geometric sequence). An alternative assumption results in the following formulation (where κ and γ_0 are defined as before).

[75] The age index a is a free parameter. For example, choosing $a = 1$ assumes that the modelled education production process begins at high school. Age increments are described by the integers, but need not correspond to the same period of time over the education process. For example, two years at high school could correspond to one year at primary school. This provides flexibility in the choice of a, so estimates for a wide range of possible values are reported.

A5′. $\frac{\lambda}{\kappa} = 1$

$$Y_t = \beta \frac{1-\lambda^t}{1-\lambda} X_t + \gamma t O_t + \lambda^t Y_0 + e_t + \sum_{j=1}^{t-1} \lambda^j e_{t-j} - \gamma \sum_{j=1}^{t-1} \lambda^j \sum_{l=1}^{j} \kappa^{l-j-1} u_{t+1-l} \tag{A.24}$$

$$\gamma = \frac{\rho_a}{a} \tag{A.25}$$

$$\lambda = \gamma \tag{A.26}$$

$$a = \frac{\rho_a}{\rho_{a+1} - \rho_a} \tag{A.27}$$

This formulation restricts the index a, which is useful given that the contemporaneous effect of friendship network gender composition γ would otherwise only be identified up to a multiplicative constant.

Appendix Table A.9 reports estimates of these parameters under the different sets of assumptions: the first six columns under $A5$ and assumptions on the age index, and the seventh column under $A5'$. The friendship network gender composition process κ is stable and indicates that two-thirds of the share of opposite gender friends persists as an individual ages. It is precisely estimated because it is the ratio of precise first stage estimates $\frac{\omega_{a+1}}{\omega_a}$. The dependence of initial friendship network gender composition on the gender composition of the close neighbourhood is only precisely estimated if it is assumed that the friendship process begins at high school ($a = 1$). It gets larger as the gender composition process is assumed to begin at earlier ages (which is mechanical given the AR(1) friendship process), but is less precisely estimated. This is because it is the ratio of exponential functions of imprecise first stage estimates. Similarly, the correlation between current and lagged GPA λ gets larger as GPA accumulation is assumed to begin at earlier ages. Finally, the contemporaneous effect of the share of opposite gender friends on achievement γ remains relatively precise over assumptions on the age index, confirming a negative effect. Under assumption $A5'$ in which the age index is determined by the model, high school is estimated to begin at an age index $a = 3$. The associated contemporaneous effect of opposite gender friends is negative, but imprecisely estimated.

This section has provided a cumulative interpretation of the effects of friendship network gender composition. Without assumption $A3$, the contemporaneous and cumulative effects cannot be separated. Empirical results in the paper can therefore be interpreted in two ways. First, this assumption can be discarded, and the original estimated parameter may be interpreted as the cumulative effect of exogenous variation in the share of opposite gender friends induced by an initial dependence of friendship composition on the gender composition of the close neighbourhood. Second, this assumption can be adopted, and data for individuals of different ages identify the parameters that describe the evolution of education production and friendship network gender composition. Furthermore, this allows us to separate the contemporaneous effect of opposite gender friends from the cumulative effect operating through past production. The estimated parameters are consistent with opposite gender friends negatively affecting high school performance under both interpretations.

A.5 Appendix Tables

Table A.1: Descriptive statistics: controls

	Mean		
	All	**Males**	**Females**
Core demographics			
White	0.52	0.52	0.52
Black	0.20	0.20	0.20
Hispanic	0.16	0.16	0.16
Asian	0.09	0.10	0.09
Other	0.03	0.03	0.03
Not born in US	0.09	0.09	0.09
Age (years and months)	16.16	16.24	16.07
Home language			
English spoken at home	0.89	0.89	0.88
Spanish spoken at home	0.08	0.08	0.08
Mother education			
Mother did not graduate high school	0.18	0.17	0.19
Mother graduated high school	0.32	0.34	0.31
Mother attended some college	0.18	0.17	0.19
Mother graduated college	0.25	0.25	0.25
Father education			
Father did not graduate high school	0.14	0.14	0.14
Father graduated high school	0.23	0.24	0.22
Father attended some college	0.13	0.14	0.13
Father graduated college	0.22	0.22	0.21
Parent characteristics			
Interviewed parent not born in US	0.15	0.16	0.15
Not receiving public assistance	0.07	0.05	0.08
Receiving public assistance	0.78	0.80	0.77
Household income			
Household income: <$20k	0.14	0.14	0.15
Household income: $20k-$40k	0.21	0.22	0.21
Household income: $40k-$60k	0.19	0.19	0.18
Household income: >$60k	0.19	0.19	0.19
Household structure			
Mother in household	0.91	0.91	0.91
Father in household	0.71	0.73	0.69
Biological mother in household	0.86	0.86	0.86
Biological father in household	0.60	0.62	0.58
Grade repetition			
Has repeated at least one grade	0.19	0.24	0.15
Observations	8,435	4,124	4,311

Categories for missing such that shares sum to one not reported but included in all analyses.

Table A.2: Instrument balance tests

	(1)	(2)	(3)	(4)
	Share opposite gender			Share white
	All	**Males**	**Females**	**All**
Core demographics				
Female	-0.01***			0.01
Black	0.00	-0.01	0.01	-0.23***
Hispanic	0.00	-0.02	0.02	-0.10***
Asian	0.01	-0.02	0.04***	-0.12***
Other	0.00	-0.00	0.00	-0.05***
Age (years and months)	-0.01	-0.01	-0.01	-0.01***
Not born in US	0.00	0.01	-0.01	0.01*
Home language				
Spanish spoken at home	0.00	-0.01	0.01	-0.04***
Other language spoken at home	-0.02**	-0.06***	0.01	-0.01
At least one ESL course taken	0.01	0.01	0.01	-0.01**
Parent characteristics				
Mother did not graduate high school	0.01	0.00	0.00	-0.01**
Mother attended some college	0.00	-0.00	0.01	0.00
Mother graduated college	0.00	0.01*	-0.01	0.01
Father did not graduate high school	-0.00	-0.00	-0.01	0.00
Father attended some college	-0.00	-0.01	0.00	0.00
Father graduated college	0.01	-0.01	0.01*	0.01
Not born in US	-0.00	0.03***	-0.03***	-0.01
Receiving public assistance	0.01	-0.00	0.01	-0.02***
Annual household income				
Zero	0.01	-0.01	-0.00	-0.01
<$20k	0.00	-0.00	0.01	-0.03***
$20k-$40k	0.00	0.00	0.00	-0.02***
$40k-$60k	0.01	0.00	0.01*	-0.01
Household structure				
Mother in household	0.01	0.02	0.00	0.03**
Father in household	0.01	0.01	0.01	-0.01
Biological mother in household	-0.01	-0.01	-0.01	-0.02***
Biological father in household	-0.00	-0.00	-0.00	0.00
Grade repetition				
Has repeated a grade	-0.01	0.01	-0.01*	0.00
Observations	8,435	4,124	4,311	8,435

School and grade fixed effects included. Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Sensitivity analysis - instrument specification

	(1)	(2)	(3)	(4)
	Overall GPA (A=4, D or lower=1)			
Instrument specification				
20 nearest schoolmates	x	x		
Schoolmates within 2km			x	x
Weighted	x		x	
Unweighted		x		x
School friends				
Share opposite gender	-1.05**	-1.02**	-0.79	-1.38
	(0.53)	(0.47)	(0.68)	(1.06)
Controls				
Female	0.18***	0.18***	0.19***	0.19***
	(0.02)	(0.02)	(0.02)	(0.02)
Other controls	x	x	x	x
School and grade fixed effects	x	x	x	x
First-stage coefficients				
Share opposite gender in close neighbourhood	0.13***	0.16***	0.08**	0.06**
	(0.03)	(0.03)	(0.03)	(0.03)
Diagnostics				
F-statistic on excluded instrument	15.32	14.46	6.37	4.46
Observations	8,435	8,435	8,160	8,160

Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Sensitivity analysis - network density and friendship definitions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overall GPA (A=4, D or lower=1)							
Sample restriction								
None	x					x	x	x
Single friend nomination		x						
Five friend nomination			x					
At least two friends				x				
At least 75% nominations matched					x			
Friendship definition								
Any nomination	x	x	x	x	x			
Nominating friendships (out)						x		
Nominated friendships (in)							x	
Reciprocated nominations								x
School friends								
Share opposite gender	-1.05**	-0.84	-1.19	-0.83*	-1.08*	-0.59	-1.08*	-0.45
	(0.53)	(0.66)	(0.79)	(0.49)	(0.55)	(0.56)	(0.55)	(1.32)
Controls								
Female	0.18***	0.18***	0.19***	0.17***	0.20***	0.16***	0.21***	0.17***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.06)
Other controls	x	x	x	x	x	x	x	x
School and grade fixed effects	x	x	x	x	x	x	x	x
Observations	8,435	4,559	3,876	4,110	4,283	6,174	6,017	2,834

Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.5: Sensitivity analysis - school urbanicity

	(1)	(2)	(3)	(4)
	Overall GPA (A=4, D or lower=1)			
School urbanicity				
Urban	x			
Suburban		x		x
Rural			x	x
School friends				
Share opposite gender	0.01	-0.96	-2.44	-1.31***
	(1.24)	(0.64)	(1.67)	(0.60)
Controls				
Female	0.20***	0.20***	0.17***	0.19***
	(0.06)	(0.03)	(0.05)	(0.02)
Other controls	x	x	x	x
School and grade fixed effects	x	x	x	x
First-stage coefficients[a]				
Share opposite gender in close neighbourhood	0.08	0.14***	0.11*	0.13***
	(0.07)	(0.04)	(0.06)	(0.03)
Diagnostics				
F-statistic on excluded instrument	1.39	10.04	3.70	14.30
Observations	1,892	4,456	2,087	6,543

Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Placebo test - effect of share even birth month

	(1)	(2)	(3)	(4)
	First stage		*OLS*	*IV*
	School friends			
	Share opposite gender	Share even birth month	GPA (A=4, D=1)	
Nearest 20 schoolmates				
Share even birth month	0.03	0.21***		
	(0.04)	(0.04)		
School friends				
Share even birth month			0.01	-0.38
			(0.02)	(0.32)
Controls				
Female	-0.01	0.01	0.20***	0.20***
	(0.01)	(0.01)	(0.02)	(0.01)
Other controls	x	x	x	x
School and grade fixed effects	x	x	x	x
Observations	8,435	8,435	8,430	8,430

Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.7: Measurement error from self-reporting bias

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Measurement error proxy (calculated for subset of sample)							
	Share opposite gender				GPA (A=4, D or lower=1)			
	All nominations - reciprocated nominations				*Self-reported - transcript*			
Nearest 20 schoolmates: share opposite gender	-0.00			-0.01	-0.06			-0.03
	(0.04)			(0.04)	(0.07)			(0.07)
Female		0.03***		0.03***		-0.10***		-0.12***
		(0.01)		(0.01)		(0.02)		(0.02)
GPA (A=4, D or lower=1; self-reported)			-0.01	-0.01			0.09***	0.10***
			(0.01)	(0.01)			(0.01)	(0.01)
Observations	2,834	2,834	2,834	2,834	3,811	3,811	3,811	3,811
R-squared	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03

Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: Sensitivity of first stage to distance from community origin

	School friends:		
	Share opposite gender		
	Closest third	*Middle third*	*Furthest third*
	(1)	(2)	(3)
Nearest 20 schoolmates			
Share opposite gender	0.00	0.18***	0.15***
	(0.06)	(0.05)	(0.06)
Controls			
Other controls	x	x	x
School and grade fixed effects	x	x	x
Diagnostics			
F-statistic of excluded instrument	0.01	10.46	5.74
Observations	2,961	2,755	2,719

Indicator variables for school in saturated sample and period of interview included. Robust standard errors clustered by school in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.9: Cumulative effect estimates - math and science GPA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Identifying assumption (A5 or A5′)						
	$\frac{\lambda}{\kappa} \neq 1$ (*A*5)						$\frac{\lambda}{\kappa} = 1$ (*A*5′)
Cumulative effect parameters							
Age index at high school	1	2	3	4	5	6	3.20 (10)
κ: friendship	0.67	0.67*	0.67*	0.67**	0.67	0.67*	0.67**
	(0.53)	(0.35)	(0.38)	(0.31)	(0.44)	(0.38)	(0.32)
γ_0: instrument	0.14***	0.21	0.31	0.47	0.70	1.05	0.34
	(0.04)	(0.34)	(1.21)	(253)	(2313)	(334)	(126)
λ: GPA	0.21	0.49	0.64	0.73	0.78	0.81	0.67*
	(1.23)	(5.70)	(5.63)	(0.70)	(1.53)	(3.04)	(0.32)
γ: opp gender	-1.39*	-0.80***	-0.48**	-0.30	-0.20	-0.14*	-0.43
	(0.82)	(0.34)	(0.22)	(0.20)	(0.34)	(0.08)	(1.06)
Observations	8169	8169	8169	8169	8169	8169	8169
Replications	50	50	50	50	50	50	50

Bootstrapped standard errors in parenthesis (school-grade-gender strata. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B

Negative Externalities in High School Course Repetition

The appendix describes results from a variety of robustness and sensitivity checks.

B.1 Appendix Tables

Table B.1: Descriptive demographic statistics - Pooled (student-years)

	All	First-time course-takers	Failed and repeating course-takers	Course-takers who pass
Math GPA score (transcript)	2.17	2.29	1.26	2.61
Gender and race:				
Female	0.50	0.51	0.41	0.52
White	0.45	0.45	0.32	0.51
Black	0.13	0.12	0.22	0.11
Hispanic	0.21	0.21	0.22	0.18
Asian	0.18	0.21	0.19	0.18
Other	0.02	0.02	0.04	0.02
Age (years and months)	16.80	17.03	17.12	16.65
Immigrant status and home language:				
Not born in US	0.13	0.15	0.11	0.13
Home language: English	0.84	0.83	0.85	0.85
Home language: Spanish	0.10	0.10	0.13	0.09
Home language: Other	0.06	0.07	0.02	0.06
Parent characteristics:				
Mother ed: Less than high school	0.19	0.18	0.21	0.16
Mother ed: High school	0.32	0.31	0.26	0.33
Mother ed: Some college	0.18	0.18	0.19	0.19
Mother ed: College	0.24	0.26	0.25	0.25
Father ed: Less than high school	0.16	0.15	0.18	0.14
Father ed: High school	0.24	0.23	0.21	0.24
Father ed: Some college	0.16	0.17	0.15	0.17
Father ed: College	0.22	0.25	0.14	0.25
Parent not born in US	0.23	0.26	0.26	0.22
Household income:				
Household income: <$20k	0.09	0.08	0.11	0.09
Household income: $20k-$40k	0.23	0.23	0.24	0.23
Household income: $40k-$60k	0.20	0.21	0.17	0.21
Household income: >$60k	0.18	0.19	0.16	0.19
Observations	6341	3379	310	3937
Share	1	0.53	0.05	0.62

Table B.2: Effect of course repeaters on academic performance of first-time course-takers

Dependent variable:	Sample: First-time course-takers						
Math GPA score	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Course-mates:							
Number of students failed and repeating:							
Natural log	-0.58**						
	(0.23)						
Linear		-0.02***	-0.09**	-0.04***	-0.13***		
		(0.004)	(0.03)	(0.01)	(0.04)		
Quadratic				0.001	0.004		
				(0.001)	(0.003)		
Share of students						-0.20	-0.11
failed and repeating						(0.82)	(1.46)
Course size (number of students):							
Linear		-0.001***		0.003			
		(0.0003)		(0.003)			
Quadratic				0.000			
				(0.000)			
Non-parametric	x		x		x		x
Fixed effects[a]	x	x	x	x	x	x	x
Observations (student-years)	3379	3379	3379	3379	3379	3379	3379
Number of students	1810	1810	1810	1810	1810	1810	1810

[a]Year, school-cohort, school-course, school-course trends and individual fixed effects included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.3: Robustness check - excluding selected subjects and schools

	Sample: First-time course-takers		
Dependent variable: Math GPA score	(1)	(2)	(3)
Exclusions:			
Algebra I	x		x
Selected schools		x	x
Course-mates:			
Log number of students failed	-0.06	-0.13**	-0.04
and repeating	(0.04)	(0.06)	(0.06)
Fixed effects[a]	x	x	x
Observations (student-years)	2828	2414	2023
Number of students	1565	1291	1130

[a]Year, school-cohort, school-course, school-course trends and individual fixed effects, as well as log number of students in course included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.4: Correlation between course failure rate and subsequent GPA

Dependent variable:	Sample: Course-takers who pass			
Subsequent year math GPA score	(1)	(2)	(3)	(4)
Course-mates:				
Log number of students who	-0.09			-0.13
fail and repeat current course	(0.11)			(0.09)
Log number of students who		-0.10*		-0.03
fail current course		(0.05)		(0.07)
Log number of students who			0.02	0.09
repeat current course			(0.07)	(0.06)
Fixed effects[a]	x	x	x	x
Observations (student-years)	3276	3276	3276	3276
Number of students	1860	1860	1860	1860

[a]Year, school-cohort, school-course, leading school-course and individual fixed effects, as well as log number of students in course included. Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix C

The Effects of Part-Time Work During High School

The appendix describes results from OLS regressions of the various dependent variables on part-time hours worked during high school. These estimates do not have a causal interpretation.

C.1 Appendix Tables

Table C.1: OLS results - educational outcomes

	All grades	*8-10th grades*	*11-12th grades*
	Mean GPA score (Wave 2)		
All hours	-0.002*	0.0004	-0.005***
Hours worked<5	0.004	0.01	-0.01
Hours worked≥5	-0.004***	-0.001	-0.007***
	Graduated high school		
All hours	-0.0008*	-0.001*	-0.0004
Hours worked<5	-0.002	-0.004	-0.0003
Hours worked≥5	-0.001	-0.001	-0.0005
	Years of education		
All hours	-0.012***	-0.013***	-0.012***
Hours worked<5	0.02	0.03	-0.04
Hours worked≥5	-0.016***	-0.01**	-0.017***
	Desire to attend college (Wave 2)		
All hours	-0.005***	-0.004	-0.006**
Hours worked<5	0.01	-0.003	0.04
Hours worked≥5	-0.008***	-0.006*	-0.012***
	Expect to attend college (Wave 2)		
All hours	-0.003*	-0.001	-0.006**
Hours worked<5	0.01	0.02	-0.03
Hours worked≥5	-0.008***	-0.004	-0.012***
	Attended college		
All hours	-0.001	-0.002**	0.0002
Hours worked<5	0.016**	0.02**	0.01
Hours worked≥5	-0.001	-0.001	-0.001

Robust standard errors clustered by school-grade in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

Table C.2: OLS results - labour market outcomes

	All grades	*8-10th grades*	*11-12th grades*
		Age at first full-time job	
All hours	-0.024***	-0.031***	-0.02***
Hours worked<5	-0.01	0.00	-0.09
Hours worked≥5	-0.017***	-0.02***	-0.02**
		Log(number of jobs)	
All hours	-0.001	-0.002	0.000
Hours worked<5	0.01	0.01	0.004
Hours worked≥5	-0.002	0.001	-0.002
		Log(income)	
All hours	0.004***	0.003	0.006**
Hours worked<5	0.01	0.005	0.02
Hours worked≥5	0.003	0.002	0.003
		Log(hours worked per week)	
All hours	0.0013***	0.001	0.002***
Hours worked<5	0.001	0.003	-0.004
Hours worked≥5	0.002**	0.001	0.003***
		Do light or hard physical work	
All hours	0.0002	0.001	0.000
Hours worked<5	0.03***	0.03***	0.02
Hours worked≥5	0.000	0.001	-0.001
		Satisfied with job	
All hours	0.001	0.0002	0.001
Hours worked<5	0.01	0.01	-0.004
Hours worked≥5	0.001	0.001	0.0002

Robust standard errors clustered by school-grade in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.