Measuring Educational Peer Effects through Multiple Overlapping Groups*

Carlos Bethencourt[†] Daniel Santos[‡]

June 2024

Abstract

In this paper, we investigate the influence of peer effects on academic performance by presenting an empirical strategy that addresses the reflection problem without relying on friendship nominations. Within each school, we categorize students into two distinct groups: those in the same grade and those participating in the same extracurricular activities. Since extracurricular activities span multiple grades, functioning as a school-wide phenomenon, we use the idiosyncratic characteristics of participants from other grades involved in activities where a specific grade does not participate but has peers who do, serving as instruments for the academic performance of peers in the same school-grade. This approach mitigates the issues of endogenous friendship formation and measurement errors typically found in self-reported friendship nomination networks, providing a more credible estimation of endogenous peer effects. Across various specifications, we identify a positive and statistically significant endogenous peer effect. Specifically, a one-unit increase in the average academic performance of peers (measured on a scale of 1 to 4, from D or lower to A, respectively) is associated with an anticipated increase of 0.66 units in an individual's academic performance. Furthermore, when exploring underlying mechanisms, we observe a "mentorship" effect from grade 12 to lower grades in high school.

Keywords: Overlapping groups, Peer effect, Education.
JEL Classification: I20, D85, D91.

*We gratefully acknowledge the valuable discussions and comments provided by Effrosyni Adamopoulou, Clyde Schechter, Kevin Shih, Francesc Ortega, Suleyman Taspinar, Alexander Smith and Jenifer Ruiz-Valenzuela. We also thank participants in several seminars for useful comments. Daniel Santos is grateful to CUNY Graduate Center and Queens College NY, and University of Barcelona, Institut d'Economía de Barcelona, for their hospitality and stimulating environment. Carlos Bethencourt is a member of CAERP, and he wish to acknowledge its support. Any remaining errors are ours. We thank the Spanish Ministry of Science and Innovation, AEI and FEDER for providing financial support through Grant PID2022-137352NB-C44.

†Corresponding author. Universidad de La Laguna, Tenerife. cbethenc@ull.es.
‡Universidad de La Laguna, Tenerife. dsantost@ull.edu.es.

1 Introduction

This paper investigates the presence of peer effects on academic performance, with a specific focus on estimating the endogenous peer effect. Specifically, it examines how the academic performance of an individual's grade mates within the same grade influences the individual's own academic achievements. As direct estimation is not feasible due to issues such as simultaneity or reflection, we employ an innovative instrumental variables approach to estimate peer effects. Our instruments are derived from analyzing the influence of extracurricular mates from other grades on the academic performance of an individual's grade peers. This method outperforms the traditional "friends of my friends who are not my friends" approach, which relies on self-reported friendship nominations, by avoiding its two main caveats: the need to assume an exogenous friendship network formation, which empirical data does not support; and the measurement error in connections, which jeopardizes the exclusion restriction required for identification, that is, peers of an individual's peers can actually be the individual's own peers (Goldsmith-Pinkham and Imbens 2013; Blume et al. 2015; Bramoullé et al. 2019).

We utilize data from the first wave of Add Health and consider different socialization groups within a school (middle or high). First, we consider peers in the same grade and school as our reference group, aligning with common practices in the literature (Bifulco et al. 2011, 2014; Olivetti et al. 2020; Cools et al. 2022). Consistent with the notation of De Giorgi et al. (2020), we refer to these peers as our 1st distance peers. Secondly, we investigate the extracurricular activities in which an individual's 1st distance peers participate but the individual does not. We define 2nd distance peers as the activity mates of an individual's 1st distance peers who belong to a different grade than the individual. By using more than one group or network instead of a single network, we reinforce the absence of connections required to identify a causal endogenous peer effect (Bramoullé et al. 2019; De Giorgi et al. 2020; Nicoletti et al. 2018).

Extracurricular activities allow us to formulate a robust identification strategy for two main reasons: *(i)* they provide opportunities to acquire values and social skills that complement and enrich the formal education obtained in the classroom. For instance, these activities can enhance school belonging, pro-social behavior, self-esteem, and social status (Coleman 1961a; Feldman and Matjasko 2005), making the influence of team and club mates crucial in understanding an individual's academic achievement. *(ii)* Interactions between students from different grades within a school are rare, with extracurricular activities serving as one of the few means to facilitate such cross-grade interactions (Schaefer et al. 2011; Fredricks and Simpkins 2013). As a result, by ensuring that our instruments are based on activities in which the individual does not participate but their grade mates do, we establish the exclusion restriction necessary for the validity of the instruments.

To illustrate our identification strategy, consider an individual in the 10th grade at a specific school. Suppose she is not a player on the soccer team, but some of her grade mates (1^{st} distance peers) are. Her grade mates regularly attend training and matches, sharing the field and dressing room with students from other grades (2^{nd} distance peers). Through these interactions, they collaborate, invest effort, and demonstrate discipline to achieve their shared goals. Values such as teamwork and the ability to challenge oneself are then brought into the classroom, fostering a more productive environment and inspiring classmates to emulate these behaviors and attitudes. Given that the soccer team is organized at the school level, we can use the influence of teammates from other grades (who are not her peers) on her grade mates as an instrument to estimate how her grade mates ultimately influence her. This approach also applies to activities where students compete individually, such as tennis or chess, where rivals and other participants can stimulate and motivate individuals to improve.

This identification strategy makes a novel contribution to the literature on peer effects in education in two key aspects: Firstly, the challenge of estimating a comprehensive model of peer effects, given the reflection problem highlighted by Manski (1993), has led most studies to focus on estimating how peers' backgrounds influence academic performance (Hoxby 2000b; Sacerdote 2011). Our approach introduces a new perspective to address this reflection problem, enabling us to investigate the impact of peers' behavior on academic performance. Secondly, some studies have addressed the reflection problem by estimating peer effects within friendship networks (Bramoullé et al. 2009; Lin 2010).[1] However, to estimate the causal effect, these studies must rely on a strong and likely unrealistic assumption: that friends are chosen randomly.[2]Additionally, and perhaps more importantly, analyzing a single network where individuals are connected directly or indirectly requires perfect knowledge of the network to leverage variations in exogenous traits of an individual's peers' peers who are not her peers as instruments to solve the reflection problem. This lack of perfect knowledge about the network and all its connections can lead to incorrectly labeling individuals as unconnected when they

[1]De Giorgi et al. (2010) also applies the strategy of Bramoullé et al. (2009) to estimate endogenous peer effect in the choice of college major.

[2]While it is true that the choice of a school and the decision to engage in extracurricular activities may be subject to some selection bias, individuals still interact with others who may differ from themselves. This is particularly evident in settings such as clubs and teams, where interaction arises from a shared commitment to and preference for those activities, rather than being entirely voluntary. It's important to note that friendships operate differently. Friendships are typically formed based on homophily, not merely on shared preferences for a particular activity (McPherson et al. (2001)). Furthermore, interaction within friendships is voluntary and not a requirement for participating in the same club. In practice, while some teammates may become friends, others may not form such connections.

might actually be connected. This issue is present in Add Health friendship network, which has been basis for most studies estimating endogenous peer effect in the literature so far, as pointed out by Blume et al. (2015). In fact, in their attempt of tackle the problem of endogenous network formation, Goldsmith-Pinkham and Imbens (2013) discovered that friends of friends who are supposedly not directly connected in Add Health data show correlations in outcomes. They state that measurement error in the friendship network is biasing the results and these are, at best, correlational, but not causal.

Therefore, our innovation within these two strands of literature takes a twofold approach: presenting the first estimation of both endogenous and contextual peer effects on academic performance, moving beyond the sole focus on contextual effects and; introducing a definition of peer groups that can be tested as more exogenously formed than friendship networks and where absence of connections in the data are more likely to reflect reality.

Therefore, our contribution to these two strands of literature is twofold: first, we present the first estimation of both endogenous and contextual peer effects on academic performance, moving beyond the sole focus on contextual effects; second, we introduce a definition of peer groups that is more exogenously formed than friendship networks, where the absence of connections in the data is more likely to reflect reality.

Furthermore, estimating the full linear-in-means model, as proposed by Manski (1993) and Moffitt (2001), allows us to uncover a new dimension of peer influences, offering valuable insights for policy design. While contextual peer effects can inform policy decisions regarding the allocation of specific types of peers[3], it is crucial to consider endogenous peer effects for assessing and capturing the social multiplier. Carrell et al. (2008) and Nicoletti et al. (2018), among others, point out that the possible amplification of shocks experienced by individuals through endogenous peer effects leads to a multiplying effect that can bring about significant changes in group composition, even when there is only a small response to individual variation. This social multiplier effect is particularly relevant for policies aimed at influencing the dynamics of the distribution and concentration of certain types within the population (Mas and Moretti 2009).

To the best of our knowledge, only two previous papers have employed a similar empirical strategy: Nicoletti et al. (2018) and De Giorgi et al. (2020). However, our approach differs from theirs in the nature of our instruments for two reasons: *(i)* unlike the 2$^{\text{nd}}$ distance peers in Nicoletti et al. (2018) and De Giorgi et al. (2020), who may only be engaged in one exclusive

[3] Carrell et al. (2013) show that sorting high ability with low ability peers can lead to performance improvement for low ability individuals without harming high ability individuals. They find that low ability and high ability individuals do not mix with each other, and therefore sorting peers does not yield the expected outcome.

group, such as a neighborhood or firm, our case involves individuals with the opportunity to participate in multiple extracurricular activities (31 activities, to be precise). Each of these activities may entail distinct contexts of socialization, encompassing different values, social norms, and cultures. This means that 2$^{\text{nd}}$ distance peers can have multiple interactions with 1$^{\text{st}}$ distance peers, making the weight of these 2$^{\text{nd}}$ distance peers endogenous and non-uniform. This introduces a novel approach that should enable us to obtain more accurate F-statistics in the first stage.[4] *(ii)* Due to the inherent nature of our instruments, we do not assume the existence of socialization at the 2$^{\text{nd}}$ distance; rather, it occurs as a consequence of participation in the same activity (they must interact as part of the same club or team). In contrast, in a firm or neighbourhood, it is assumed that the individual interacts with coworkers or neighbours. In conclusion, *(i)* and *(ii)* jointly imply that our understanding of social interactions beyond the first distance of socialization (the grade) is more accurate than in the cases of Nicoletti et al. (2018) and De Giorgi et al. (2020). Therefore, our instrumental variable estimation should produce a more precise endogenous peer effect coefficient.

Our main results can be summarized as follows: Firstly, we identify a significant and positive endogenous peer effect on academic performance through our Two-Stage Least Squares estimation. Specifically, when considering the average of all extracurricular activities in which individuals do not participate but at least a 1$^{\text{st}}$ distance peer does (without distinguishing between types of activity), we find that a one-unit increase in the peers' average GPA (measured from 1 to 4, "D or lower" to "A") results in an increase of an individual's GPA by 0.66 units in our preferred specification. Secondly, we refine our instruments by categorizing them based on different types of extracurricular activities: arts, academics, excellence, and athletics.[5] In doing so, we find that only the specification involving athletics activities satisfies the relevance condition of the instruments, although the peer effect in this case turns out not statistically significant. Specifically, we find that a one-unit increase in the peers' average GPA results in an increase in an individual's GPA by 0.51 units. Although not statistically significant, this distinction help us to shed light on the differences between types of activity as a source of instrument relevance. Thirdly, we examine a mentorship effect from 2$^{\text{nd}}$ distance peers in

[4]The drawback here is that the exclusion restriction could be compromised if an individual encounters a 2$^{\text{nd}}$ distance peer in an extracurricular activity who was instrumenting a 1$^{\text{st}}$ distance peer through a different activity. Essentially, we assume that each activity entails a distinct socialization context, potentially resulting in different effects of the 2$^{\text{nd}}$ distance peer on both the individual and her 1$^{\text{st}}$ distance peer. In this scenario, the validity of the exclusion restriction hinges on the premise that we are instrumenting for members of other grades participating in activities where the individual is not involved, but at least one of her 1$^{\text{st}}$ distance peer is.

[5]Please note that Add Health does not provide any classification, and we categorize activities into each type based on their nature.

grade 12 on individuals in grades 9, 10, and 11 as a potential underlying mechanism. We find that the endogenous peer effect in high school is explained by those students with a greater presence of 2$^{\text{nd}}$ distance peers from grade 12.

We compare our results to the two strands of the education literature previously mentioned: First, among the group of papers that focus on contextual peer effect, the most relevant studies analyze the effect of peers' background on individuals' achievement, with mixed results (Sacerdote (2011)). While the majority of studies document a positive and sizable peer effect on academic achievement, such as Hoxby (2000b), who finds that an increase in female cohort composition, translated to an increase in peers' average scores, results in improving reading scores by 0.3 to 0.5 points and math scores by 1.7 to 6.8 points, other studies find modest peer effects. For instance, Angrist and Lang (2004) or Imberman et al. (2012) respectively find that the busing of Metco students into suburban Boston schools does not statistically impact non-Metco students, and that the arrival of Katrina evacuees modestly impact achievement in receiving schools in Louisiana and Houston, Texas, considering the linear-in-means model. Moreover, evidence suggests that peer effects are non-linear (see Imberman et al. (2012) and Lavy et al. (2012)). In this respect, we also find evidence of non-linearities when we distinguish between middle and high school. Secondly, among the studies that have exploited social networks to achieve identification following Lee (2007) and Bramoullé et al. (2009), contributions have been scarce. A notable contribution is Lin (2010), who finds a positive endogenous peer effect of 0.27. We will compare our results to this study because Lin (2010) also uses Add Health data and exactly the same definition of GPA as we do, with the sole difference being the being the definitions of peer groups and the identification strategy, since she uses friendship nominations.

Finally, in relation to the literature that has focused on addressing endogenous peer effects by instrumenting peers with secondary peers (Nicoletti et al. (2018) and De Giorgi et al. (2020)), our main contribution lies in utilizing an identification strategy based on the actual interactions among individuals. Our approach benefits from involving a second group with a deeper understanding of the relevant peers and the actual interactions among its members, in which socialization is guaranteed and weights of peers are determined by the inherent dynamics of their own socialization.

The paper is organized as follows: Subsection 1.1 reviews the literature on peer effects in general and, more specifically, peer effects in education. Section 2 provides a description of the data. In Section 3, we detail the identification strategy. Section 4 presents the results and Section 5 explore underlying mechanisms. Finally, Section 6 offers concluding remarks.

Additional supplementary material are contained in the Appendix.

1.1 Previous Literature

There is abundant evidence indicating that individuals' behavior is not entirely explained by their own characteristics or utility. Rather, their actions are influenced by other relevant individuals, such as peers. Peer effects manifest in various contexts and have diverse impacts, including areas such as education, crime, consumption, and labor (Hoxby 2000a; Sacerdote 2001; Patacchini and Zenou 2012; De Giorgi et al. 2020; Olivetti et al. 2020). Despite their significance for policymakers' decisions, the empirical estimation of these effects has presented considerable challenges (Angrist 2014).[6] The primary reason lies in the inherent nature of peer effects. Specifically, peers can influence us in two ways: through their own behavior and choices, such as their dedication to studies or decision to smoke (referred to as the endogenous peer effect), and through their inherent exogenous characteristics, such as their parents' socioeconomic status (contextual or exogenous peer effect). The existence of these two possible types of effects and their identification generates three empirical challenges (Manski 1993): *(i)* the reflection problem, stemming from the simultaneity between an individual's outcome and her peers' outcomes, prompting the question of who exactly influences whom; *(ii)* the challenge of endogenous network or group formation, implying that individuals selects their peers, resulting in individuals self-selecting into a network or group and, consequently, making it challenging to draw causal estimates from the effects of their peers on themselves; and *(iii)* the correlated effects problem, arising from common unobserved shocks impacting both the individual and her peers.

The existence of correlated effects poses technical challenges, however, the main difficulties stem from the reflection problem and the endogeneity of group formation. In the literature, researchers have primarily focused on estimating the exogenous peer effect and addressing the endogenous group formation, often assuming the absence of an endogenous peer effect. A common approach involves eliminating the unobserved heterogeneity driving group composition and utilizing the remaining variation in idiosyncratic characteristics across groups as quasi-

[6] Angrist (2014) emphasizes the common mistake of drawing causal impact conclusion instead of correlation when estimating peer effects. To elucidate his argument, he focuses especially on the effects of peers' characteristics on individual's outcome as the difference between the OLS estimate between an individual's outcome on his peers' characteristics and the 2SLS estimate of an individual's outcome on his individual characteristics, instrumenting them using group dummies. In the end, the instruments are very unlikely to be orthogonal to the individual's outcome, as plenty of confounding and unobserved variables can be misleading the estimates (an omitted variable bias is violating the exclusion restriction). Therefore, as the regression of an individual's outcome on his peers' mean outcome is tautological, Angrist (2014) points out at the orthogonality condition between group belonging and characteristics as the main goal to achieve in order to estimate credible (contextual) peer effects.

random. This strategy effectively solves the endogenous group formation and correlated effects problems, guaranteeing a robust identification strategy for obtaining causal estimates of peers' background on individuals' own outcomes (Hoxby 2000b,a; Bifulco et al. 2011, 2014; Lavy and Schlosser 2011; Lavy et al. 2012; Carrell et al. 2018; Olivetti et al. 2020; Cools et al. 2022). Another common approach involves using natural events or experiments to guarantee the randomness of peer group composition (Sacerdote 2001; Angrist and Lang 2004; Imberman et al. 2012). Regarding academic achievement, the results found by studies which estimate some peers' exogenous characteristic such as gender, race or ability, range from positive increases of less than one point score to increases that surpass one point score and even come close to two points score, depending on the subject, the level of education (primary or secondary) and if allowing for non-linearities (Sacerdote 2011). However, these results documented by the literature are hard to compare to our results, as they only consider the contextual peer effect, whereas we estimate both the endogenous and contextual peer effects.

While the previous line of research primarily focuses on estimating peer effects in group contexts, another branch has concentrated on estimating peer effects within social network contexts. Pioneering works in this area include Lee (2007), Bramoullé et al. (2009), Calvó-Armengol et al. (2009), Lin (2010), and De Giorgi et al. (2010). Lee (2007) establishes the theoretical conditions necessary to identify both endogenous and exogenous peer effects in linear-in-means models, particularly when there are varying sizes across individuals' networks. If the size of the group differs for each individual, variations in size can be utilized for identification. Meanwhile, Bramoullé et al. (2009) are the first to discuss the use of "friends of friends who are not my friends" for identification, although they assumed that network formation is strictly exogenous. In contrast, Calvó-Armengol et al. (2009) focus on how proximity to the center of the network impacts academic achievement. Building upon the concepts introduced by Lee (2007) and Bramoullé et al. (2009), Lin (2010) uses the variability in individual-level friendship networks and estimates significant and positive peer effects on academic achievement. Nevertheless, much like to Bramoullé et al. (2009), she has to assume exogeneity in the formation of the friendship network. Similarly, De Giorgi et al. (2010) embraced the concept of "peers of my peers who are not my peers" and estimated peer effects on major choice at Bocconi University. By exploiting Bocconi University's policy of randomizing students in mandatory courses, they addressed the issue of endogenous network formation. They discovered a significant and positive endogenous peer effect in major choice, yet their findings provided limited insights into contextual peer effects. Interestingly, the coefficients of contextual peer effects appeared to exert minimal influence on the endogenous peer effect coefficient, being statistically insignificant. The authors argue that the randomness within their peer group effectively nullifies the influence of contextual peer effects. More recently, Goldsmith-Pinkham and Imbens (2013), Blume et al. (2015), Hsieh and Lee (2016), and Hsieh et al. (2020) have expanded upon the concepts presented by Lee (2007), Bramoullé et al. (2009), and Lin (2010) by considering the endogenous formation of friendship networks. Goldsmith-Pinkham and Imbens (2013) model the endogenous network formation as an omitted variable bias, introducing a parameter to account for that. They also examine differences between past friends and current friends influences on individuals. Their results are quite relevant as they find a positive correlation between current friends' academic performance and individual's academic performance, but also a positive correlation between past friends' academic performance and individual's academic performance and between friends of friends who are not friends's academic performance and individual's academic performance. As they mention, these results cast several doubts on the causal effect of friends' outcomes on individual's outcomes, even controlling for the endogenous formation of the network. Whereas Hsieh and Lee (2016) and Hsieh et al. (2020) make great contributions to better modeling the endogenous formation of the network, the absence of causality is still present in both studies as another serious problem pointed out by Blume et al. (2015) arises: when using self-reported friendship nominations, especially Add Health data base, identification mainly stems from knowing who you are not friend with rather that knowing who are your friends. To do so, complete knowledge of the network is pivotal, i.e. knowing the actual links within the network. As Add Health only allows to nominate up to 5 male friends and 5 female friends, failing in identifying two individuals as friends do not necessarily means that there is a 0 between the two individuals in the sociomatrix, i.e. they are in fact friends. This could explain why Goldsmith-Pinkham and Imbens (2013) find correlation between friends of friends who are not friends and individuals, because the friendship network in Add Health is measured with error. In conclusion, because of the error measurement of friendship network, papers that model the endogenous network formation analyzing academic performance report correlations, but no causal effects. On the contrary, our two definitions of network are partitioned groups of a population, representing isolated networks from each other. In this scenario, we improve the two problems which friendship networks suffer from: endogenous formation and measurement error. Regarding academic performance, Lin (2010) estimates the endogenous and contextual peer effects using as peers friends within the same school and grade. Moreover, she uses Add Health data and measures academic performance like we do; using the same GPA variable. Her findings from the maximum likelihood estimation of the Spatial Autoregressive Model reveal a positive endogenous peer effect with a coefficient of 0.274, which corresponds with the 34% of a standard deviation in the average GPA in her sample. One particular case of study is Hanushek et al. (2003), who study endogenous and contextual peer effects using the UTD Texas Schools Project Microdata, a data panel focusing on primary school. Their difference yields in using two years lagged values of academic achievement to avoid simultaneity or reflection, but not the current achievement. In order to be a good proxy of current achievement, correlation between scores two years before and scores currently must be very high. However, unobserved variations over time are not captured leading to, at best, an underestimation of the coefficient of the endogenous peer effect. They estimate an endogenous peer effect coefficient ranging from 0.15 to 0.24, using peers in the same school and grade as the reference group.

To the best of our knowledge, only two studies have effectively tackled both the reflection problem and endogenous network or group formation, utilizing groups distinct from self-reported friendship nominations, thus providing a better knowledge of networks and interactions and minimizing the risk of violating the exclusion restriction at identification. These studies are Nicoletti et al. (2018) and De Giorgi et al. (2020), introducing a novel element to the empirical estimation of peer effects: individuals can participate in multiple groups which do not perfectly overlap. Nicoletti et al. (2018) employ family peers and neighbors, while De Giorgi et al. (2020) use spouses and coworkers. By considering scenarios where a family member resides in a different neighborhood or a spouse works in a different firm, the neighbors of the family member and the coworkers of the spouse can serve as instruments to explain individual outcomes.

2 Data Description

We use data from the National Longitudinal Study of Adolescent to Adult Health (Add Health).[7] The Add Health dataset was designed to evaluate the influence of family and social environments on individuals' health in the United States throughout their adolescent and adult lives. It provides comprehensive information on a wide range of socioeconomic, familial, social, demographic, behavioral, and health aspects of individuals, and to a lesser extent, their parents, at various stages of their lives. The dataset spans five waves, with Wave V having over 20 years of temporal difference from Wave I. The survey commenced by collecting

[7]This research uses data from Add Health, funded by grant P01 HD31921 (Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

information from students in grades 7–12 during the academic year 1994–1995, originating from a nationally representative sample of 130 private and public schools (Wave I). Add Health employs a school-based design, where these 130 schools are selected from a population of schools in the Quality Education Database (QED), with non-uniform probabilities of selection based on school size. Initially, an interview is conducted with all students present in these 130 schools on a designated school day, constituting the in-school survey, yielding information from 90,118 individuals. Subsequently, a sub-sample of 20,745 individuals completes a questionnaire at home. These individuals form the basis for subsequent follow-up interviews conducted over the years: Wave II (1995-1996), Wave III (2001-2002) when individuals are 18-26 years old, Wave IV (2008-2009) covering ages 24-32, and Wave V (2016-2018) encompassing ages 33-44.

We focus on Wave I in-school survey, where we extract pivotal information to form our two groups: the school code and grade to which each individual belongs, and the extracurricular clubs and teams in which they participate within the school. The following steps outline the process to construct our final sample: first, starting with 90,118 observations from Wave I in-school questionnaire, we exclude 177 observations with duplicated questionnaire IDs; second, we drop 95 observations with missing information in the "What grade are you in?" question, as well as those from grade 6 because they represent 0.11% of the total grade participation; third, an additional 205 observations are eliminated for answering "My school doesn't have grade levels of this kind" to the question "What grade are you in?"; forth, we drop 55 observations with multiple responses and 538 with missing observations. At this point, we proceed to eliminate all observations that could compromise the exclusion restriction in our IV strategy. But first to get to that point, we need to recall first how we construct our instruments. We establish 1st distance and 2nd distance peers for each individual. To achieve this, we employ the two distinct socialization groups within a school that individuals engage in, as previously mentioned. Initially, we focus on peers in the same grade and school (middle or high) as our reference group, aligning with the common practice in the literature (Bifulco et al. (2011), Bifulco et al. (2014), Olivetti et al. (2020), Cools et al. (2022)). In accordance with the notation of De Giorgi et al. (2020), we denote these peers as our 1st distance peers. Secondly, we explore the extracurricular activities that involve the 1st distance peers of an individual, where the individual themselves do not participate. We define 2nd distance peers as the activity peers of an individual's 1st distance peers who are in a different grade than the individual and individual's 1st distance peers.

Therefore, given that we are using 2nd distance peers as instruments for individuals' 1st distance peers, it is crucial to ensure the absence of a direct relationship between the individual and their 2[nd] distance peers. Any such connection would violate the exclusion restriction of the instrument. Therefore, in this context, addressing additional connections within the same school is essential to guarantee the validity of the instruments, and personal and family connections may pose a serious threat. To illustrate this, consider an individual, i, who doesn't participate in the football team but has a grade mate, j, who is part of that team (1[st] distance peer). We can then use idiosyncratic variations in exogenous characteristics (such as mother's education) of football team participants who are not in the same grade as i and j (2[nd] distance peers) as instruments for the endogenous peer effect in the academic performance of individual i, but only in the absence of any additional connections between i and football team participants. If other connections emerge, the exclusion restriction of the instrument would not be satisfied. Consider now that i's brother is a member of the school football team. In this case, it is reasonable to assume that there is a likelihood that i interacts with his brother's teammates. This circumstance would compromise the exclusion restriction, as i is directly connected to 2[nd] distance peers. Therefore, it is necessary to identify cases where individuals may be related to 2[nd] distance peers to guarantee the validity of our instruments. To address the issue of additional family connections, we use the in-school questionnaire. Question 28 informs us if there is an individual in the household attending grades 7 through 12.[8] Additionally, question 33 tells us if those individuals attend the same school as the respondent. By combining information from these questions, we identify 19,982 individuals with at least one household member attending the same school, and we exclude them from the sample.[9] Finally, we exclude 3,419 individuals who belong to grades with fewer than 60 individuals, representing the 5th percentile of the school-grade cohort size (Olivetti et al. (2020)).[10] At this point, the sample has been reduced to 65,385 individuals.

Subsequently, we define the necessary variables at the individual, 1[st] distance peers, and 2[nd] distance peers levels. We exploit as much as individual information from Wave I in-school questionnaire. The variables included as exogenous regressors at the individual level are: gender, with a dummy variable equals 1 if individual is female and 0 otherwise; race, with a dummy variable equals 1 if individual reports being black and 0 otherwise; culture or immigrant status, with a dummy variable equals 1 if individuals was born in the US and 0

[8]Question: "In addition to you, how many other people who are in grades 7 through 12 live in your household?"

[9]Despite mitigating concerns about siblings attending the same school, we cannot determine if non-household familial members of similar age, such as cousins, attend the same school.

[10]While we generate endogenous and contextual peer effects for all individuals who can be instrumented (excluding those who do not participate in any activity), this results in a further reduction in the size of the 1[st] distance peers' group. Nevertheless, the average 1[st] distance peers' group size is 155.62 peers, with a minimum of 28 peers and a maximum of 321 peers, which appears to be within an acceptable range.

otherwise; information regarding the parents, with a dummy variable equals 1 if individual live with both parents and 0 otherwise, and education levels, birthplace and labor status of the mother, with dummy variables reflecting if the mother has secondary education, college degree and post-college degree, and other dummies indicating if mother was born in US and if she works for pay, and 0 otherwise. Parental education and race are commonly considered as exogenous regressors in other studies which examine peer effects on academic performance in secondary school (Hoxby (2000a), Lavy and Schlosser (2011)), as well as gender (Angrist and Lang (2004)). We also include the school-grade leave one out mean of each of these variables, which it is the contextual peer effect (Moffitt (2001)). As we have mention, our variable of interest is the individual GPA, and the endogenous peer effect is the school-grade leave one out mean of GPA. The GPA variable is the average of the grades of the following subjects: English, Mathematics, Sciences and History. Each subject is assessed from "D or lower" to "A".[11]

Then, we drop all observations that have missing values in any of the dependent, independent variables, or instruments. Regarding instruments, we exclude individuals who meet the following criteria: *(i)* do not participate in any extracurricular activity (20,946 observations)[12]; *(ii)* participate in all activities (63 observations); and *(iii)* lack 2^{nd} distance peers, meaning all club/teammates of their 1^{st} distance peers are concentrated in the same grade as the instrumented individual (307 individuals). Dropping individuals with missing values in the dependent or independent variables, or with missing values in the instruments due to reason *(iii)*, results in the removal of an additional 14,560 observations, 14,190 observations with missing value in some control, and 307 observations with missing value in the instruments due to reason *(iii)*. Consequently, our final sample consists of 29,879 observations distributed in three types of school: high schools (grades 9 to 12), middle schools (grades 7 and 8), and high + feeder school (schools with grades spanning from 7 to 12).

Table 1 shows the summary statistics of the demographic characteristics for all individuals, 1^{st} distance peers, and 2^{nd} distance peers. The main takeaway from this table is that the means and standard deviations across variables are similar between all individuals and the two levels of peers. This implies that, on average, the distribution of individuals and peers based on demographic characteristics in extracurricular activities is similar to the distribution of 2^{nd} distance peers. Therefore, if the formation of endogenous peers groups were more prevalent in the 2^{nd} distance than in the 1^{st} distance, those moments of the distribution would exhibit

[11] We recode this variable as 1 = "D or lower", 2 = "C", 3 = "B" and 4 = "A".

[12] These individuals can be instrumented but cannot be used to construct the instruments. A concern about selection bias may arise if those who do not participate in any activity exhibit behavior significantly different from those who do participate. In section 3, we examine whether these excluded individuals differ in terms of the exogenous variables.

significant differences. Our final sample is distributed as follows: the average GPA is 2.95, 53% of our sample is female, 17% is black, 21% is Hispanic, Asian or other race, 94% was born in the USA, 80% live with both parents in home, the average household members are 4.14 (including the individual), 87% have a mother who was born in the USA, and 9.2% of mothers have less than high school education, 54.4% have secondary education, 25.7% have a college degree, 10.7% have a post-college degree, and 83% of them work for pay.

Similarly, Table 2 presents the complete set of extracurricular activities individuals may participate in, along with the participation share by type of activity (Arts, Academics, Excellence, and Athletics). Notably, there is a substantial participation in Athletics (47.53%) compared to the other types.[13]

3 Identification Strategy

We estimate peer effects on academic achievement by exploiting information of two non-perfectly overlapping groups within a school. The first group span the entire grade within a school. This approach is supported by the majority of the literature that studies peer effects in educational settings to assess academic performance (Hoxby (2000a), Angrist and Lang (2004), Lavy et al. (2012)), as well as other long-run outcomes (Bifulco et al. (2011), Bifulco et al. (2014), Olivetti et al. (2020), Cools et al. (2022)). Indeed, according to Hoxby (2000a), including peers at the grade level can more effectively mitigate selection bias compared to the class level, as parents and schools may manipulate the assignment of students to classrooms. In addition, there is literature supporting the choice of peers at grade level rather than classroom level in middle and high school as a better reference group, especially when it comes to small schools (Bellmore et al. (2010)). The reason is that, in contrast to earlier grades, students socialize more out of the classroom. Middle and high school students seldom spend the majority of their school time with the same set of classmates. For instance, they might need to select courses that introduce them to peers from different classrooms, or they may have to switch classrooms between classes. The second group consists of club and teammates at the extracurricular activity level. The importance of this reference group is also emphasized in the education literature, as extracurricular clubs and teams offer a distinctive learning environment to cultivate a different form of human capital beyond formal education (Coleman (1961a),

[13] The distribution is calculated using our final sample of 29,879 observations. Nevertheless, to make sure that the final sample is not biased towards a certain composition, we check the distribution of the sample without removing observations in any dependent and independent variables and instruments used in the analyses. This sample consists of 44,439 observations, and its composition per type of activity is as follows: 21.07% in "Arts", 12.82% in "Academics", 16.47% in "Excellence", 48.80% in "Athletics", and 0.84% in "Farmers". The distributions are almost the same.

Table 1: Sample description

All	Obs	Mean	Std
GPA	29,879	2.95	0.77
Female	29,879	0.53	0.50
Black	29,879	0.17	0.38
Other	29,879	0.21	0.41
Born US	29,879	0.94	0.24
Both parents in home	29,879	0.80	0.40
Household members	29,879	4.14	1.08
Mother born US	29,879	0.87	0.34
Mother with less than HS	29,879	0.092	0.29
Mother with secondary edu	29,879	0.544	0.50
Mother with college edu	29,879	0.257	0.44
Mother with post-college edu	29,879	0.107	0.31
Mother working for pay	29,879	0.83	0.37
1st distance peers	Obs	Mean	Std
GPA	29,879	2.90	0.26
Female	29,879	0.50	0.10
Black	29,879	0.18	0.23
Other	29,879	0.23	0.20
Born US	29,879	0.92	0.10
Both parents in home	29,879	0.74	0.11
Household members	29,879	4.14	0.22
Mother born US	29,879	0.85	0.17
Mother with less than HS	29,879	0.098	0.08
Mother with secondary edu	29,879	0.545	0.13
Mother with college edu	29,879	0.253	0.09
Mother with post-college edu	29,879	0.104	0.08
Mother working for pay	29,879	0.83	0.06
1st distance peers	29,879	155.62	72.35
2nd distance peers	Obs	Mean	Std
Female	29,879	0.49	0.11
Black	29,879	0.18	0.22
Other	29,879	0.27	0.20
Born US	29,879	0.90	0.10
Both parents in home	29,879	0.73	0.12
Household members	29,879	4.08	0.29
Mother born US	29,879	0.83	0.18
Mother with less than HS	29,879	0.106	0.08
Mother with secondary edu	29,879	0.502	0.14
Mother with college edu	29,879	0.259	0.10
Mother with post-college edu	29,879	0.132	0.09
Mother working for pay	29,879	0.82	0.07
2nd distance peers	29,879	31.4	17.0
No. schools	112		

Notes: The table reports descriptive statistics for the main variables used in the analysis. There is a detailed definition of each variable in Appendix A. The sample includes students in grades 7 through 12 with at least 59 peers. Source: Add Health.

Table 2: Sample description - Extracurricular activities (Share (%))

Arts	**Academics**	**Excellence**	**Athletics**	**Farmers**
Book	French	Debate	Baseball	Future Farmers
Band	German	Newspaper	Basketball	
Drama	Latin	Honor Society	Field hockey	
Cheerleader	Spanish	Student Council	Football	
Chorus	Computer	Yearbook	Ice Hockey	
Orchestra	History		Soccer	
	Math		Swimming	
	Science		Tennis	
			Track	
			Volleyball	
			Wrestling	
21.55%	12.67%	17.51%	47.53%	0.75%

Notes: The table reports the extracurricular activities we use to build our instruments, and the share of participation by type of activity. The share of participation is calculated using the observations before dropping those with missing value in one of the main variables used in the regressions. Source: Add Health.

Feldman and Matjasko (2005), Fredricks and Eccles (2005), Fredricks and Eccles (2006)).

Extracurricular activities play a crucial role in our analysis in terms of socialization. The organization of extracurricular clubs and teams at the school level in the U.S. facilitates interaction among individuals from different grades. Literature highlights that outside of these activities, there is limited socialization and contact across grades and ages (Schaefer et al. (2011), Patrick et al. (1999), Fredricks and Simpkins (2013)). Consequently, participation in extracurricular activities becomes the primary channel of contact between students from different grades within secondary schools (Shrum et al. (1988)). This implies the existence of non-perfectly overlapping groups within schools: while all individuals belong to a particular grade (group), some of them may simultaneously belong to different extracurricular activities (groups) with students from the same or different grades. This feature allows us to exploit the intransitivity condition outlined by Bramoullé et al. (2009) and De Giorgi et al. (2010). In fact, our paper is similar in spirit with the works of Nicoletti et al. (2018) and De Giorgi et al. (2020), because each individual can participate in more than one group, whereas in the cases of Bramoullé et al. (2009) and De Giorgi et al. (2010), each individual only belongs to one group, but every group is different by individual. To see it more clearly, look at 1 and let's suppose that an individual within a particular school, denoted by S1, who is in grade 9, do not participate in an extracurricular activity, denoted by Act 1, but there is a grade mate, individual S4, who does. Therefore, S1, from grade 9, can participate with S7, who belong to grade 10, in Act 1. Then, as long as S1 and S7 are only indirectly connected through S4, S7 can serve as an instrument to estimate the peer effect from S4 to S1.

Figure 1 represents the sharpest case between groups, because they non-overlap perfectly, looking alike to the cases of Nicoletti et al. (2018) and De Giorgi et al. (2020). However, we have 31 extracurricular activities, and there is no exclusivity across groups (that is, I can be in the basketball team, and in the math club, for example). This provokes that our second definition of groups overlaps but not perfectly. To understand why, look at **??**. In this case, for S1, we can use the information of S2, who participates in Act 2 and Act 3, whereas S1 only participates in Act 1. Nonetheless, S6, from grade 10, participates in both Act 1 and Act 2, so he is con contact with both S1 and S2. The question here is: Is still S6 valid for instrumenting S1 through S2? We argue that, given the fact that the overlapping is not perfect, we can exploit the variation in idiosyncratic characteristics of activity mates from other grades of S2. Moreover, the activities are different, implying different schedules and/or facilities, so it is unlikely that S6 is affecting in the same way S1 and S2. As a conclusion, the non-perfectly overlapping of the second group should still satisfy the exclusion restriction in our instrumental variables setting.

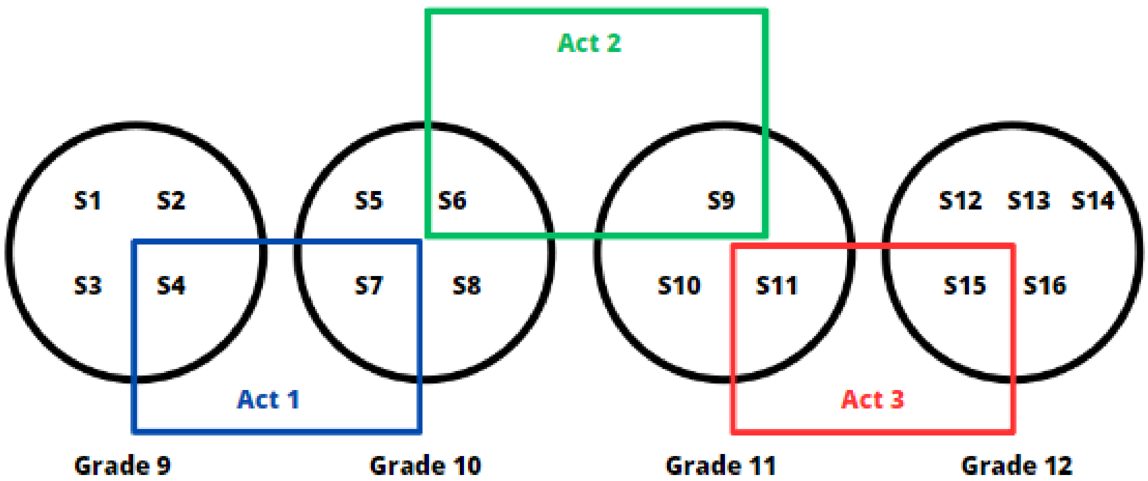

Figure 1: A simple example of non-overlapping groups within a school

Once we have clarified our identification strategy, we present the following equation to be estimated using both Ordinary Least Squares (hereafter OLS) and Two-Stage Least Squares (henceforth 2SLS):

$$y_{igs} = \alpha_g + \gamma_s + \delta_s \tilde{g} + \varphi y_{-igs} + \beta x_{-igs} + \theta x_{igs} + \epsilon_{igs}, \tag{1}$$

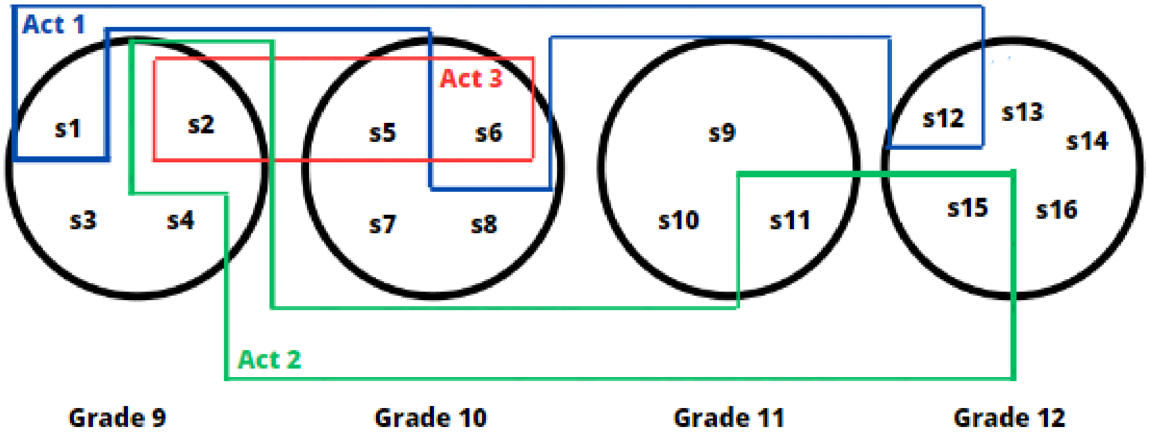

Figure 2: A simple example of non-perfectly overlapping groups within a school

where students are denoted by i, grade or cohorts are denoted by g and schools are denoted by s. Variable y_{-igs} denotes the average GPA of individual i peers (within the same school and grade), excluding i from the distribution. Similarly, x_{-igs} represents the average observable traits of i peers after excluding i from the distribution, while x_{igs} denotes the individual's observable traits. The scalar parameters φ and β quantify the endogenous and the exogenous peer effects, respectively. The vector α_g comprises dummy variables controlling for grade fixed effects, and γ_s comprises dummy variables controlling for school fixed effects. The necessity of incorporating school fixed effects becomes obvious when considering the availability of extracurricular activities. Despite the wide and diverse range of activities, as shown in Table 2, the existence of 31 activities does not imply that every school offers all of them. The provision of activities is likely tied to factors such as the financial status, available resources, or educational policies of each school. Unfortunately, our data only indicates whether individuals participate or not in each activity. We lack information on whether every school offers all activities. Consequently, if we observe that no one in a particular school is participating in swimming, we cannot discern whether this is due to the activity being offered but garnering no interest or if the school does not provide the activity. So, to account for these unobserved differences between schools—such as variations in income, resources, or educational policies influencing the availability of extracurricular activities—we introduce this set of school fixed effects. The variable $\delta_s\tilde{g}$ represents a school-specific linear time trend, where $\tilde{g}$ measures the distance between the grade that the individual attends a reference grade. Here we, consider the lowest grade, i.e., grade 7, as the reference, so $\tilde{g} = g - 7$ for $g = \{7, 8, 9, 10, 11, 12\}$. This variable aims to capture whether the selection bias varies across students within the same school depending on the grade. Finally, ϵ_{igs} is the i.i.d error term.[14]

As previously highlighted in the introduction, the estimation of endogenous and exogenous (contextual) peer effects faces challenges such as the reflection problem, endogenous group formation, and correlated effects Manski (1993). First, to address the reflection problem, we employ instrumental variables — specifically, "peers of my peers who are not my peers". More precisely, we use as instruments the variations in observable traits among the team and club mates of the individual's peers who belong to different grades and participate in activities she does not take part in. In other words, we use variations in observable traits of 2$^{\text{nd}}$ distance peers as instruments to estimate how her 1$^{\text{st}}$ distance peers influence her.

Secondly, by incorporating an extensive set of fixed effects at the grade and school levels, our goal is to eliminate unobserved heterogeneity that may drive cohort assignment. The reason is that individuals in the same school and grade might exhibit similar behavior due to shared traits, implying self-selection into groups. In this respect, contextual peer effects literature often relies on variation in exogenous characteristics across cohorts within schools as a quasi-random shock (Hoxby (2000a), Lavy and Schlosser (2011), Bifulco et al. (2011)). Moreover, as highlighted by Hoxby (2000a), parents and teachers are more likely to manipulate the assignment of children at the classroom level. Thus, fixed effects should absorb bias arising from endogenous school choice. Note that, while our analysis encompasses two groups of socialization —grade-school and teams and clubs from extracurricular activities— we have introduced fixed effects solely to address the potential endogeneity problem arising from the first group of socialization. The reason is that, unlike commonly used self-reported friendship nominations driven by homophily, participation in extracurricular activities is less prone to being entirely influenced by self-selection. To verify this statement, we assess selection bias in these activities by conducting an OLS regression with the share of each observable trait at the extracurricular level as a dependent variable against each individual-level control. And we compare this with the analogous OLS regression used to assess selection bias in the grade-school group. These balancing tests are a common practice in the literature for testing selection bias (Lavy and Schlosser (2011), Olivetti et al. (2020), Cools et al. (2022)). Tables 3 and 4 show the balancing tests for the idiosyncratic characteristics' composition of both groups, namely school-grade (grades) and school-activity (extracurricular activities). The dependent variables in each table, for every trait, represent the leave-one-out mean[15]. The main objective

[14] Although we will cluster the errors at school level in ours analysis.

[15] We observe a difference in the number of reported observations compared to our main analysis, attributed to two factors: *(i)* in the school-grade balancing test, the disparity arises from missing values in the variable GPA and; *(ii)* in the school-activity balancing test, the results represent the average across all activities. Given that our unit of observation is individual-activity-school, and each individual can participate in up to 31 activities,

is to examine whether the composition of observable traits, such as gender, race or parents' education, within each grade and activity in the school is significantly influenced by possessing a specific exogenous trait. For instance, we aim to determine if being female is correlated with the percentage of females her my grade or in her extracurricular activity. In both tables, we observe that, despite the statistical significance of many coefficients, their magnitudes are consistently very small, tending to approach to zero in almost every case, indicating limited impact on group composition. The exception arises with gender and race concerning participation in extracurricular activities. For being female, it accounts for a 20.5% variation in the percentage of females in the activity, representing almost half of the mean of female participation in extracurricular activities. Nevertheless, unexplained variation persists, even when taking into account the influence of other observable characteristics on the female cohort composition, with a visibly limited impact. Similarly, being black explains approximately 6% of the variation in cohort composition, constituting 30% of the mean of being black in extracurricular activities. Once again, it appears that there is room for unexplained variation in the cohort composition. These results are promising as they indicate that there is unexplained variation in the cohort composition of grades and activities, even after controlling for observable exogenous traits and unobservable heterogeneity through fixed effects.

Another potential concern relates to individuals who do not partake in extracurricular activities. In essence, we need to assess whether our sample is biased due to the distinct characteristics of individuals participating in extracurricular activities. The question at hand is whether participation is endogenous or if it is randomly distributed across various variables such as the socioeconomic backgrounds, race, or gender. It's crucial to emphasize that we are addressing here the fundamental act of participating in at least one activity, regardless of the specific type of activity. Certainly, participation in a specific type of activity may be associated, for instance, with income or gender. However, on average, across the 31 activities, we have observed in the balancing test that participation is not purely driven by endogenous factors. As a precautionary measure, we conduct a simple exercise by looking at the main statistics of observable characteristics for individuals who do not participate in any activity, thereby excluding them from our sample. Out of the initial sample of 90,118 observations, we focus on the 20,946 individuals who do not participate in any activity. We exclude observations with missing values in any of the main individual variables, resulting in a sample size of 11,145 individuals (compared to 29,879 in our final sample). The key statistics are reported in Table 5. We observe that these individuals exhibit slightly lower academic performance compared

we obtain a significantly higher number of observations.

Table 3: Balancing test: school-grade composition

	% Female	% Black	% Other	% Both	% Born US	% Mom S	% Mom C	% Mom P	% Mom born US	% Mom W	Peers' HM
Female	-0.002***	-0.000	-0.000	-0.000	-0.001*	0.001*	-0.000	-0.000	0.000	-0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Both	-0.000	0.001*	0.000	-0.002***	0.000	0.001	-0.000	0.000	0.000	-0.000	0.002
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Born US	-0.003***	-0.001	0.001	0.002***	0.001	0.001	-0.001	0.000	0.000	0.000	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Black	-0.001	-0.000	-0.001*	0.000	0.000	0.002**	-0.001	-0.001	0.001	0.000	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)
Other	-0.001	-0.001	-0.001*	0.001*	0.000	-0.001	0.000	0.000	-0.001	0.000	0.001
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Mom S	0.001**	-0.000	-0.000	0.002***	-0.000	-0.005***	0.001	0.001*	-0.001*	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)
Mom C	0.002**	-0.001	0.000	0.002***	-0.001	-0.002**	-0.002***	0.001**	-0.001**	-0.000	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Mom P	0.001	-0.001	0.000	0.002***	-0.000	-0.002**	0.001	-0.001	-0.002*	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Mom born US	0.000	0.002*	-0.002**	-0.000	0.001**	0.001	-0.001	-0.000	0.000	0.000	-0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Mom W	-0.001	0.000	-0.000	-0.001	-0.000	0.000	-0.000	0.000	0.000	-0.003***	-0.001
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
HM	0.000	0.000	-0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Constant	0.533***	-0.009	0.183***	0.629***	0.982***	0.544***	0.196***	0.023*	0.945***	0.801***	4.137***
	(0.017)	(0.023)	(0.047)	(0.084)	(0.009)	(0.021)	(0.031)	(0.012)	(0.028)	(0.020)	(0.059)
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School T	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	32,956	32,956	32,956	32,956	32,956	32,956	32,956	32,956	32,956	32,956	32,956
R Squared	0.843	0.986	0.973	0.907	0.938	0.895	0.836	0.901	0.970	0.797	0.840

Notes: Standard errors clustered at school-grade level in parentheses. School and grade fixed effects and school trend are included in all specifications. Due to space constraints, we put some variables in acronyms. Mom S stands for Mom secondary, Mom C stands for Mom college, Mom P stands for Mom post, Mom W stands for Mom W and HM stands for Household members. FE stands for fixed effects and School T stands for school trend. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: Add Health.

Table 4: Balancing test: extracurricular activities composition

	% Female	% Black	% Other	% Both	% Born US	% Mom S	% Mom C	% Mom P	% Mom born US	% Mom W	Peers' HM
Female	0.205***	-0.010***	-0.004***	0.013***	-0.000	-0.002	0.003**	-0.000	-0.003***	-0.005***	-0.013***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Both	0.006***	-0.005***	-0.002**	0.002**	0.000	-0.004***	0.003***	0.003***	0.001	-0.000	-0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Born US	-0.001	0.001	-0.009***	0.002	0.018***	0.007***	-0.005**	-0.002	0.012***	0.003*	-0.007*
	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.004)
Black	-0.020***	0.059***	-0.008***	-0.023***	0.005***	0.011***	-0.004***	-0.008***	0.013***	0.005***	-0.000
	(0.003)	(0.004)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.004)
Other	-0.006**	-0.000	0.012***	-0.003***	-0.005***	-0.003***	-0.001	-0.001	-0.010***	-0.001	0.006**
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Mom S	0.006**	-0.003**	-0.006***	0.006***	-0.001	-0.006***	0.008***	0.003***	0.000	0.001	-0.010***
	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)
Mom C	0.013***	-0.007***	-0.005***	0.012***	-0.004***	-0.011***	0.012***	0.008***	-0.003*	0.001	-0.011***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.004)
Mom P	0.018***	-0.012***	-0.005**	0.015***	-0.004***	-0.017***	0.018***	0.008***	-0.003**	-0.000	-0.016***
	(0.004)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.004)
Mom born US	-0.009***	0.009***	-0.017***	-0.005***	0.010***	0.005***	-0.002*	-0.000	0.018***	0.005***	-0.012***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.004)
Mom W	-0.009***	0.002*	-0.001	-0.002**	0.002**	0.002	-0.000	-0.002***	0.003***	-0.001	0.003
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
HM	-0.004***	0.001**	0.001**	-0.001***	-0.000	0.001**	-0.001**	-0.001***	-0.000	0.000	-0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Constant	0.366***	-0.001	0.286***	0.622***	0.966***	0.556***	0.207***	0.037***	0.940***	0.799***	4.024***
	(0.009)	(0.004)	(0.009)	(0.009)	(0.004)	(0.009)	(0.011)	(0.003)	(0.005)	(0.004)	(0.010)
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School T	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	86,228	86,228	86,228	86,228	86,228	86,228	86,228	86,228	86,228	86,228	86,228
Mean	0.52	0.20	0.24	0.74	0.92	0.52	0.26	0.12	0.85	0.84	4.17
R Squared	0.307	0.880	0.821	0.643	0.673	0.601	0.462	0.638	0.822	0.401	0.452

Notes: Standard errors clustered at school-grade level in parentheses. School and grade fixed effects and school trend are included in all specifications. The mean of each variable is provided to better interpret the results. Due to space constraints, we put some variables in acronyms. Mom S stands for Mom secondary, Mom C stands for Mom college, Mom P stands for Mom post, Mom W stands for Mom W and HM stands for Household members. FE stands for fixed effects and School T stands for school trend. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: Add Health.

Table 5: Description of individuals excluded from the final sample

All	Obs	Mean	Std
GPA	11,145	2.57	0.81
Female	11,145	0.50	0.50
Black	11,145	0.16	0.37
Other	11,145	0.27	0.44
Born US	11,145	0.91	0.28
Both parents in home	11,145	0.75	0.43
Household members	11,145	4.15	1.13
Mother born US	11,145	0.82	0.38
Mother with less than HS	11,145	0.17	0.38
Mother with secondary edu	11,145	0.60	0.49
Mother with college edu	11,145	0.18	0.38
Mother with post-college edu	11,145	0.05	0.22
Mother working for pay	11,145	0.80	0.40

Notes: The table reports descriptive statistics for the main individual variables for those individuals who do not participate in any extracurricular activity. There is a detailed definition of each variable in Appendix A. Source: Add Health.

to our final sample (average GPA of 2.57 and 2.95, respectively); and the distribution of their mothers' education is more skewed towards the left tail (17% and 10% of mothers with less than a high school education and 5% and 11% of mothers with post-college education, respectively). Despite these differences, they appear to be similar in all other aspects to individuals who participate in at least one activity, dissipating concerns about selection bias in our final sample.

4 Results

In this section, we report the results for the linear-in-means model (see equation (1)). We report the estimated endogenous and the contextual peer effects stemming from the 1^{st} distance peers, using as instruments the idiosyncratic characteristics of the 2^{nd} distance peers. We present the 2SLS results and compare them with the OLS estimation results. We will exclude from our baseline analyses those individuals belonging to high + feeder schools due to the fact that, even having grades from 7 to 12 in the same buildings, it seems unlikely that a student from grades 7 or 8 share activities with students in high school. Therefore, assuming these connections as real would lead to a measurement error of our groups and their connections, probably exacerbating the peer effects estimated by instrumental variables. Table A.1 in Appendix A illustrates the share of 2^{nd} distance peers by the three different types of school present in our data (high school, middle school and high + feeder school).

Regarding the 2SLS estimation, it's important to note that instruments for each individual have been computed as the average of all activities in which she does not participate but in which at least one of her 1st distance peers does participate. Given the presence of numerous exogenous characteristics (11, to be precise), multiple combinations among them can be created for instruments. However, the inclusion of more instruments increases the likelihood of not satisfying the overidentifying test. To address this, we opted for a maximum of two instruments, experimenting with various combinations and assessing the F first-stage statistic and the Hansen test. The best set of instruments, that is, the most relevant ones for explaining variations in 1st distance peers' GPA and the most likely to influence individuals through their 1st distance peers, consists of "born US" and "Mom work for pay". We report school clustered errors although our treatment variable varies at different levels in the first and the second stage. In the first stage the demographic variables of the 2nd distance peers vary at the school level, since extracurricular activities span different grades within a school, but the endogenous peer effect in the second stage varies at the school-grade level. Given that difference in the variation level of the treatment we opt for taking the more conservative approach and cluster the standard errors at the school level (Abadie et al. (2023)). However, we cannot provide an interpretation of the Stock-Yogo weak ID test critical values under violations of the homoscedastic error assumption, nor can we rely on the "rule of thumb" suggesting that an F statistic equal to or greater than 10 is sufficient to ensure instrument strength (Staiger and Stock (1997)), we are unable to assess the strength or weakness of the instruments. Our only option is to examine the p-value of the F statistic to determine whether the instruments, collectively, have zero coefficients, as illustrated in Nicoletti et al. (2018), without gaining too much insights into their strength.

After these considerations, we then proceed to present the estimation results. Table 6, *OLS* column, shows the OLS results, whereas *2SLS* column shows the 2SLS estimation results. *2SLS Individual IVs* column tackles an important concern in our identification strategy, following the approach of Nicoletti et al. (2018). Our 2SLS might still be suffering from omitted relevant variables. The most important omission comes from the possibility of sorting in the same kind of extracurricular activities. In other words, if individuals are exposed to similar exogenous traits variation in the activities in which they do participate in, the instruments will fail to identify the true peer effect. Therefore, we include in column (3) the instruments at individual level, i.e. the average exogenous traits of their own teammates who belong to other grades. This will be crucial in our identification scheme. First, in the OLS estimation, we observe a significant and positive endogenous peer effect with a coefficient of approximately 0.46. However, as mentioned earlier, this coefficient is susceptible to endogeneity bias due to the reflection problem. To address this, we proceed to estimate it through 2SLS. In this case, we still observe a positive and significant endogenous peer effect of 0.79. This implies that a one-unit increase in peers' GPA is associated with an increase in the individual's GPA of 0.79 units. For example, if peers' GPA, on average, moves from D or lower to C, from C to B, or from B to A, the individual's GPA will significantly increase his mark but he would not reach the next category. In other words, an increase in the variable *peers' GPA* of 0.26, corresponding to one standard deviation (see Table 1), is associated with an increase in individual's GPA of approximately 0.20 units. The F first-stage p-value rejects the null hypothesis of instruments have zero coefficients (thus they are relevant) and the overidentifying test satisfy the relevance and the exclusion condition of the instruments. We compare this result with Lin (2010)'s findings. Despite conducting a similar exercise to ours, using the same dependent variable and the endogenous peer effect variable (GPA from Add Health), we find a larger effect of 0.79 compared to her 0.27. This difference could be attributed to the definition of the reference group. While in our study, it is the school-grade level, in Lin (2010), the reference group is the nominated friends from the same grade as the individual. This definition of the group allows her to use spatial econometrics to leverage the variation across groups, as each group varies at the individual level. However, we address the reflection problem differently by exploiting the fact that each individual can belong to more than one group. Our approach allows us not to assume that the groups are formed randomly and can alleviate the measurement error present in self-reported friendship nominations from Add Health, thus achieving a cleaner identification.

One noteworthy finding is that the 2SLS coefficient is larger than the OLS coefficient, with values of 0.79 versus 0.46. This is unexpected, as the endogeneity bias typically bias OLS estimates upwards. However, previous studies addressing the reflection problem using instrumental variables also report similar observations (De Giorgi et al. (2010), Nicoletti et al. (2018), and De Giorgi et al. (2020)).[16] The common explanation is that measurement errors in the variable of interest bias the coefficient downwards. In our case, there could be at least two sources of measurement error in the GPA variable: *(i)* we only have information about four subjects (see Section 2), and while they are arguably the most important, we are missing others such as foreign language, and; *(ii)* we only observe A, B, C, and D or lower, but we do not observe A+ or B-, and not accounting for those nuances when estimating GPA can result in measurement errors.

This explanation has been complemented by the exclusion bias explained by Caeyers and

[16]This phenomenon is observed in De Giorgi et al. (2010) only when they do not estimate contextual peer effects.

Fafchamps (2016). The exclusion bias automatically generates a negative correlation between a variable at the individual level and at the level of its group when the group mean is calculated excluding the individual. This negative correlation might cause the magnitude of the peer effect coefficient to decrease. Caeyers and Fafchamps (2016) show how the bias is exacerbated when both the sample and the group size go to infinite. Given that our first-distance groups are grades within schools, we can guarantee having a large group size. Indeed, the average school-grade size is 155 individuals, whereas the number of observations is 26,203, so we can expect a non-negligible exclusion bias, biasing the coefficient downwards. Therefore, it seems plausible that the measurement error in variable GPA, along with the magnitude of the exclusion bias in our data, is biasing the OLS endogenous peer effect coefficient downwards. But still our 2SLS estimates might be biased upwards due to the omission of the own extracurricular activity mates. As pointed out by Nicoletti et al. (2018), this omission of relevant variables leads to an overestimation of the 2SLS coefficient. To mitigate this bias, we include the instruments at individual level, i.e. the % of own extracurricular mates from other grades who were born in the U.S. and whose mothers work in the specification showed in *2SLS Individual IVs* column. In such estimation, we obtain a peer effect of 0.66 still satisfying both relevance and exclusion restriction. This coefficient is greater than OLS coefficient but lower than our previous 2SLS. Nicoletti et al. (2018) even find a smaller coefficient in their 2SLS Individual IVs. Whereas they have a negligible exclusion bias due to their sample and peers group sizes, we expect to have a greater exclusion bias which is biasing downwards the OLS estimate jointly with the measurement error. Hausman p-value does not reject in any case significant differences in size between OLS coefficient and 2SLS and 2SLS Individual IVs coefficients. This reflects that, as we mitigate the endogeneity bias stemming from reflection problem and omission of relevant variables (own extracurricular mates), the difference in magnitude between our OLS coefficient (0.46) and our 2SLS coefficient (0.66) is probably due to the exclusion bias.

5 Understanding School Peer Dynamics

In this section, we delve into the validity of our instruments by analyzing various aspects related to socialization in high school and, consequently, the channels that drive the influence of one individual on others. Specifically, we first explore whether there are differences in the influence of peers depending on the type of extracurricular activity they are involved in. Secondly, we analyze possible variations in the peer effect derived from the type of school, such as middle, high, and high+feeder schools. Finally, in relation to this, we explore the existence of a mentorship effect influencing the peer effect.

Table 6: Effect of *Peers' GPA* - All activities

	GPA		
	OLS	2SLS	2SLS Individual IVs
Peers' GPA	0.457***	0.793***	0.658***
	(0.059)	(0.161)	(0.182)
% peers female	-0.067	-0.158**	-0.133
	(0.102)	(0.078)	(0.102)
% peers black	0.035	0.070	0.054
	(0.138)	(0.102)	(0.122)
% peers other	0.106	0.136	0.094
	(0.116)	(0.093)	(0.105)
% peers born US	-0.053	0.061	-0.192
	(0.161)	(0.146)	(0.181)
% peers both parents in home	-0.096	-0.161*	-0.156
	(0.116)	(0.093)	(0.115)
% peers mom secondary	0.231	-0.012	0.096
	(0.150)	(0.182)	(0.201)
% peers mom college	0.205	-0.136	0.027
	(0.159)	(0.230)	(0.248)
% peers mom post	0.264	-0.209	-0.071
	(0.187)	(0.291)	(0.312)
% peers mom born US	0.057	0.042	0.098
	(0.154)	(0.122)	(0.142)
% peers mom work	-0.132	-0.037	-0.264**
	(0.122)	(0.104)	(0.129)
Peers' household members	-0.009	0.013	0.017
	(0.037)	(0.029)	(0.033)
Individual controls and constant	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes
School trends	Yes	Yes	Yes
Instruments (born US, mom work)	No	Yes	Yes
Instruments at individual level	No	No	Yes
R^2	0.176	0.083	0.087
N	26,203	26,203	26,191
F test		5.49	5.48
F test p-value		0.005	0.006
Hausman p-value		0.123	0.338
Hansen p-value		0.152	0.101

Notes: Standard errors clustered at school level in parentheses. School and grade fixed effects and school trend included. Individual controls are: female, black, other, born US, both parents in home, secondary, mom college, mom post, mom born US, mom work, and household members. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: Add Health.

5.1 Differential impact of extracurricular activities on peer influence

In our main specification, we have used 2^{nd} distance peers, averaged across 31 extracurricular activities as an instrument of the 1^{st} distance peers. The large and diverse set of extracurricular activities (refer to Table 2) prompts us to question whether the impact of individuals through activities is uniform or varies depending on the nature of the activity. For example, we wonder whether the influence of 2^{nd} distance peers on 1^{st} distance peers within the French club differs from that in the Student Council, drama club or baseball team. Note that this is specially relevant to the validity of the instruments. The validity of the instruments can be compromised if self-selection depends on the type of activity, because endogenous group formation bias the impact of exogenous demographic characteristics on GPA. Furthermore, the exclusion restriction might be jeopardized if certain activities are more susceptible to self-selection than others. To illustrate this, revisit Figure 2. Let's assume a robust self-selection process in activity 2, where S2, S11, and S15 share highly similar traits, including race, gender, parents' education, cultural background, etc. Now consider S1 in grade 9 and her 1^{st} distance peer S2. In such a scenario, it could seem as though S11 and S15 (both 2^{nd} distance peers) are directly influencing S1, not indirectly through S2, because in terms of observable characteristics, S2, S11 and S15 are essentially alike.

Therefore, we calculate the instruments per type of activity, specifically arts, academics, excellence, and athletics. Instead of averaging across all activities, we compute the average per type of activity. For simplicity, we exclude from this analysis the activity Future Farmers of America, as it does not belong to any type.[17]

Table 7 provides the 2SLS estimation results. We observe that the size and the significance of the endogenous peer effect varies across the four types. However, the relevance of the instruments is only satisfied in the athletics specification. But these instruments, without considering the other activities, seem to not be enough powerful to achieve identification of the endogenous peer effect. This fact is being exacerbated in the 2SLS Individual IVs, where the coefficient is greater and significant, when should be, at least, close in size to the 2SLS. Nevertheless, the split of the instruments in different types is informative in the sense that tells us that interaction in athletic activities might be the most influential. Of course, it seems that we are not capturing relevance coming from other types by isolating each type, but when all activities are considered regardless any distinction we get enough variation to identify the endogenous peer effect. So, why athletic activities could be more influential than other types?

[17] The activities included in each type can be seen in Table 2.

One plausible hypothesis is that an individual may be attracted to engage in athletics activities for various reasons beyond vocational interests. Participation in athletic activities has been linked to high peer status (Coleman (1961b), Morgan and Alwin (1980)), a greater sense of belonging to school (Eccles et al. (2003)), higher rates of college attendance, or even higher alcohol consumption (Eccles et al. (2003)). Therefore, an individual might be drawn to athletic activities not only out of a genuine interest in athletics but also to socialize more, enhance their peer status, become more popular, or even to participate in social events. In contrast, it is challenging to imagine reasons other than vocational interests for participating in artistic, academic, or excellence activities. Except for the possibility of increasing the probability of college acceptance, which could be a common feature shared among all types of activities, the additional benefits from athletics related to socialization or peer status are less present in artistic, academic, or excellence activities, which are more focused on direct utility. In fact, athletic activities not only have the highest share of participation compared to other types, as seen in Table 2, but also the highest share of contact between 1^{st} distance and 2^{nd} distance peers. Figure 3 shows that approximately 40% of 1^{st} distance peers who participate in an activity in which an individual does not, participate in athletic activities. All in all, the more heterogeneous participation in athletic might foster interactions and values and skill acquisition such as discipline, the will to improve, to excel, etc., to a greater extent than other types.

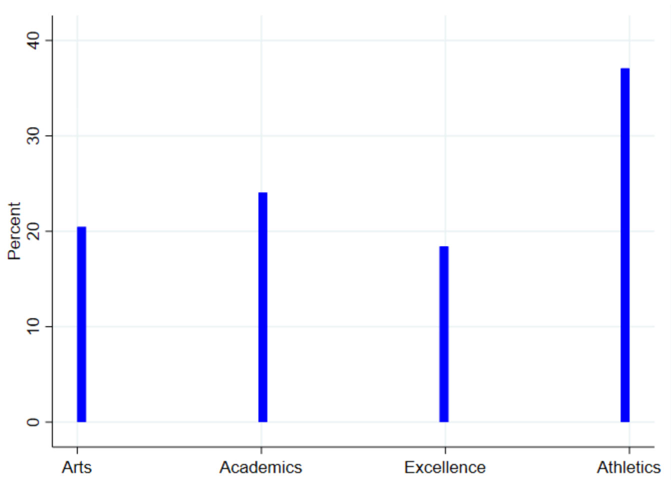

Figure 3: Distribution of type of activities where connection between 1^{st} distance and 2^{nd} distance peers happens.

5.2 Mentorship effect

When contemplating interactions among students in a school setting, it's crucial to acknowledge potential differences in socialization within high school. Given that the dynamics of

Table 7: Effect of *Peers' GPA* - Per type of activity

	GPA				
	Arts	Academics	Excellence	Athletics	Athletics Individual IVs
Peers' GPA	1.324***	0.577	1.801**	0.510	0.832**
	(0.509)	(0.404)	(0.807)	(0.389)	(0.371)
% peers female	-0.302*	-0.107	-0.432	-0.085	-0.146
	(0.164)	(0.153)	(0.299)	(0.156)	(0.197)
% peers black	0.126	0.065	0.168	0.035	0.159
	(0.114)	(0.153)	(0.215)	(0.153)	(0.197)
% peers both parents in home	-0.270	-0.131	-0.349	-0.100	-0.070
	(0.167)	(0.148)	(0.247)	(0.151)	(0.159)
% peers other	0.185	0.112	0.231	0.110	0.270*
	(0.119)	(0.133)	(0.213)	(0.145)	(0.157)
% peers born US	0.244	-0.028	0.402	-0.032	-0.023
	(0.224)	(0.220)	(0.351)	(0.248)	(0.236)
% peers mom secondary	-0.382	0.133	-0.728	0.191	0.109
	(0.390)	(0.332)	(0.653)	(0.340)	(0.357)
% peers mom college	-0.672	0.059	-1.161	0.143	-0.075
	(0.532)	(0.451)	(0.928)	(0.453)	(0.448)
% peers mom post	-0.952	0.094	-1.619	0.181	-0.378
	(0.742)	(0.588)	(1.227)	(0.605)	(0.574)
% peers mom born US	0.008	0.053	-0.026	0.060	0.036
	(0.139)	(0.156)	(0.233)	(0.170)	(0.188)
% peers mom work	0.116	-0.106	0.247	-0.122	0.060
	(0.169)	(0.164)	(0.308)	(0.171)	(0.158)
Peers' household size	0.052	0.006	0.080	-0.006	0.056
	(0.051)	(0.046)	(0.080)	(0.047)	(0.060)
Individual controls and constant	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes	Yes
School trends	Yes	Yes	Yes	Yes	Yes
Instruments (born US, mom work)	Yes	Yes	Yes	Yes	Yes
Instruments at individual level	No	No	No	No	Yes
R^2	0.074	0.085	0.058	0.085	0.085
N	26185	26014	26131	26186	17782
F test	0.93	1.59	2.39	3.26	3.18
F test p-value	0.398	0.209	0.097	0.043	0.046
Hausman p-value	0.163	0.737	0.070	0.930	0.410
Hansen p-value	0.963	0.530	0.242	0.612	0.067

Notes: Standard errors clustered at school level in parentheses. School and grade fixed effects and school trend included. Individual controls are: female, black, other, born US, both parents in home, secondary, mom college, mom post, mom born US, mom work, and household members. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: Add Health.

socialization are influenced by age, there might be variations in socialization depending on school grades.

For instance, seniors might be more influential on sophomores than viceversa, so those students in high school exposed to a larger fraction of grade 12 team and club mates can intensify their learning process as they see seniors as role models. Therefore, the presence of a "mentorship" effect from 2$^{\text{nd}}$ distance peers to 1$^{\text{st}}$ can enhance the endogenous peer effect through 1$^{\text{st}}$ distance peers.

To address this, we distinguish between those individuals with a relative low weight of 2$^{\text{nd}}$ distance peers from those students with a relative high weight of 2$^{\text{nd}}$ distance peers. We define weight of 2$^{\text{nd}}$ distance peers as the share of them coming from grade 12 relatively the total amount of 2$^{\text{nd}}$ distance peers. Therefore, we explore differences in the peer effect between students from grades 9, 10 and 11 depending on their exposure to 2$^{\text{nd}}$ distance peers from grade 12.

We test for differences including interaction terms (Ringdal and Sjursen (2021)). We calculate the distribution of such weight for each grade 9, 10 and 11, and we define a dummy variable equal to one for those students who are above the percentile 75 of the distribution of 2$^{\text{nd}}$ distance peers relatively to their grade and equal to 0 otherwise.[18] Then, we interact the dummy variable with the endogenous peer effect and we estimate the model. The sum of the coefficients of the endogenous peer effect and the interaction variable reflects the peer effect for those individuals who have a high weight, in terms of being above of the percentile 75 (*Peers' GPA (high weight)*), whereas the traditional peer effect coefficient reflects such effect for those with low weight (*Peers' GPA (low weight)*). For simplicity, we only present these two variables in the table. Results are shown in Table 8. Notably, we observe that the endogenous peer effect for grades 9, 10 and 11 is driven by individuals who have a high share of 2$^{\text{nd}}$ distance peers. For instance, in column (3), an increase of one unit in peers' GPA is associated with an increase of approximately 0.45 in students' GPA in grades 9, 10 and 11. In conclusion, we find evidence of a mentorship effect exerted from seniors to the rest of high school students.

6 Conclusions

This paper studies peer effects on academic performance by estimating not only how peers' backgrounds affect achievement, which has been the main focus for most of the literature in education, but also how peers' behavior affects achievement. To overcome the reflection

[18]The distribution of weight or share of 2$^{\text{nd}}$ distance peers varies across grades 9, 10 and 11. In other words, in average, the share of team and club mates from grade 12 interacting with other grades is different for a student from grade 9 than for a student from grade 10 or 11, and we take into account such differences.

Table 8: Effects of *Peers' GPA* - Mentorship effect

	GPA	
	2SLS	2SLS Individual IVs
Peers' GPA (low weight)	0.321	0.298
	(0.321)	(0.313)
Peers' GPA (high weight)	0.461^{**}	0.446^{**}
	(0.267)	(0.215)
Contextual peer effects	Yes	Yes
Individual controls and constant	Yes	Yes
School fixed effects	Yes	Yes
Grade fixed effects	Yes	Yes
School trends	Yes	Yes
Instruments (born US, mom work)	Yes	Yes
Instruments at individual level	No	Yes
R^2	0.081	0.084
N	16,088	16,088
F test	4.39	4.39
F test p-value	0.012	0.012
Hausman p-value	0.977	0.971
Hansen p-value	0.348	0.279

Notes: Standard errors clustered at school level in parentheses. School fixed effects included. Contextual peer effects are: % peers female, % peers black, % peers other, % peers born US, % peers both parents in home, % peers secondary, % peers mom college, % peers mom post, % peers mom born US, % peers mom work and peers' household members. Individual controls are: female, black, other, born US, both parents in home, secondary, mom college, mom post, mom born US, mom work, and household members. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: Add Health.

problem, we exploit non-perfectly overlapping groups within the same school: grade and extracurricular activities. Taking advantage of the fact that extracurricular teams and clubs are organized at the school level, we leverage the variation in idiosyncratic characteristics of the team and club mates of my peers as an instrument, as long as those mates belong to a different grade than mine and my peers, and I do not participate in that activity. We also innovate from the literature that exploits the friendship network structure by utilizing two types of groups that may well be subject to less self-selection than friendship itself and where absence of connections are more likely to hold in reality than in a self-reported friendship nominations network which is usually measured with error. These features allows us to provide with a more credible and causal endogenous peer effect than previous studies in education.

We observe that the instruments satisfy the relevance and the exclusion restriction: extracurricular activities provide an opportunity to learn positive values such as discipline, creativity, hard work, or teamwork that can be taken to the class environment. Without the environment of the extracurricular clubs and teams, connections between pupils from different grades within a school are unlikely to happen. We find a positive and significant endogenous peer effect whose size is 0.66, larger than the literature of peer effects in education. Moreover, not all types of activities exert the same influence, so we explore which type is the most relevant in explaining the relevance of the instruments that allow us to identify the peer effect, with athletics being the one with the strongest instruments. Finally, we carry out an analysis to test whether there is a mentorship effect from pupils who attend grade 12 to team and club mates from grades 9, 10 and 11 in high school. We find that the peer effect in high school is mainly driven by those students with greater exposure to team and club mates from grade 12.

Appendix

A Appendix: Additional tables

Appendix Table A.1: 2^{nd} distance peers' weight by grade per type of school

All schools	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Grade 7		82.9%	8.0%	3.4%	3.2%	2.5%
Grade 8	85.5%		7.4%	2.4%	2.5%	2.2%
Grade 9	3.5%	2.7%		37.6%	30.9%	25.1%
Grade 10	1.7%	0.1%	39.3%		32.1%	25.9%
Grade 11	1.7%	1.0%	36.5%	36.6%		24.2%
Grade 12	1.5%	0.9%	34.6%	34.5%	28.4%	
High + feeder schools	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Grade 7		31.3%	32.4%	13.6%	12.8%	9.9%
Grade 8	37.3%		32.2%	10.3%	10.7%	9.5%
Grade 9	22.4%	17.4%		22.5%	20.0%	17.7%
Grade 10	13.7%	7.5%	33.7%		24.2%	20.9%
Grade 11	13.1%	8.0%	32.0%	25.9%		20.9%
Grade 12	11.5%	7.1%	31.8%	25.7%	23.9%	
High schools	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Grade 7						
Grade 8						
Grade 9				40.4%	33.0%	26.5%
Grade 10			40.1%		33.2%	26.7%
Grade 11			37.2%	38.1%		24.7%
Grade 12			35.0%	35.9%	29.1%	
Middle schools	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Grade 7		100%				
Grade 8	100%					
Grade 9						
Grade 10						
Grade 11						
Grade 12						

Notes: The table reports the share in percentage of the 2^{nd} distance peers by grade, for each grade, per type of schools: all (120 schools), high + feeder schools (14 schools), high schools (60 schools), and middle schools (46 schools). Source: Add Health.

Appendix Table A.2: Data description and definitions

Variables	Description
	Wave I - In school questionnaire
GPA	Grade point average across 4 subjects: English, Mathematics, History and Sciences. The variable's values are 1 = "D or lower", 2 = "C", 3 = "B", and 4 = "A". Questions: S10A, S10B, S10C, and S10D.
Female	Dummy variable equal to one if the respondent reported being female. Question: S2.
Black	Dummy variable equal to one if the respondent reported being black. Question: S6B.
Other	Dummy variable equal to one if the respondent reported being Hispanic, Asian or Native American. Questions: S4, S6C, and S6D.
Born US	Dummy variable equal to one if the respondent reported having been born in the U.S. Question: S8.
Both parents in home	Dummy variable equal to one if the respondent reported being living with his two residents parents (parents can be biological, step, foster or adoptive). Questions: S11 and S17.
Household members	Total number of members living in the individual's house, individual included. Question: S27.
Mother/Mom born US	Dummy variable equal to one if the respondent reported his mother was born in the U.S. Question: S13.
Mother/Mom less than HS	Dummy variable equal to one if the respondent reported his mother has less than high school diploma. Question: S12.
Mother/Mom secondary education	Dummy variable equal to one if the respondent reported his mother has completed secondary education. Question: S12.
Mother/Mom college education	Dummy variable equal to one if the respondent reported his mother has graduated from a college or a university. Question: S12.
Mother/Mom post-college education	Dummy variable equal to one if the respondent reported his mother has completed professional training beyond a four-year college. Question: S12.
1^{st} distance peers	Peers who belong to the same school and grade as the individual i. Each variable is calculated at 1^{st} distance peers level by calculating the leave-one-out mean, i.e., excluding individual i from the distribution.
2^{nd} distance peers	Extracurricular peers of individual i's 1^{st} peers who belong to the same school but different grades as the individual i's 1^{st} peers, in those activities in which individual i does not participate in but has at least one 1^{st} distance peer who does.

Source: Add Health and authors' calculations.

References

Abadie, A., Athey, S., Imbens, G.W., Wooldridge, J.M., 2023. When should you adjust standard errors for clustering? The Quarterly Journal of Economics 138, 1–35.

Angrist, J.D., 2014. The perils of peer effects. Labour Economics 30, 98–108.

Angrist, J.D., Lang, K., 2004. Does school integration generate peer effects? Evidence from Boston's Metco Program. American Economic Review 94, 1613–1634.

Bellmore, A., Jiang, X.L., Juvonen, J., 2010. Utilizing peer nominations in middle school: a longitudinal comparison between complete classroom-based and random list methods. Journal of Research on Adolescence 20, 538–550.

Bifulco, R., Fletcher, J.M., Oh, S.J., Ross, S.L., 2014. Do high school peers have persistent effects on college attainment and other life outcomes? Labour Economics 29, 83–90.

Bifulco, R., Fletcher, J.M., Ross, S.L., 2011. The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health. American Economic Journal: Economic Policy 3, 25–53.

Blume, L.E., Brock, W.A., Durlauf, S.N., Jayaraman, R., 2015. Linear social interactions models. Journal of Political Economy 123, 444–496.

Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. Journal of Econometrics 150, 41–55.

Bramoullé, Y., Djebbari, H., Fortin, B., 2019. Peer effects in networks: A survey. Technical Report. IZA Discussion Papers, No. 12947.

Caeyers, B., Fafchamps, M., 2016. Exclusion bias in the estimation of peer effects. Technical Report. National Bureau of Economic Research.

Calvó-Armengol, A., Patacchini, E., Zenou, Y., 2009. Peer effects and social networks in education. The Review of Economic Studies 76, 1239–1267.

Carrell, S.E., Hoekstra, M., Kuka, E., 2018. The long-run effects of disruptive peers. American Economic Review 108, 3377–3415.

Carrell, S.E., Malmstrom, F.V., West, J.E., 2008. Peer effects in academic cheating. Journal of human resources 43, 173–207.

Carrell, S.E., Sacerdote, B.I., West, J.E., 2013. From natural variation to optimal policy? the importance of endogenous peer group formation. Econometrica 81, 855–882.

Coleman, J.S., 1961a. The Adolescent Society: The Social Life of the Teenager and its Impact on Education. Free Press on Glencoe.

Coleman, J.S., 1961b. Athletics in high school. The Annals of the American Academy of Political and Social Science 338, 33–43.

Cools, A., Fernández, R., Patacchini, E., 2022. The asymmetric gender effects of high flyers. Labour Economics 79, 102287.

De Giorgi, G., Frederiksen, A., Pistaferri, L., 2020. Consumption network effects. The Review of Economic Studies 87, 130–163.

De Giorgi, G., Pellizzari, M., Redaelli, S., 2010. Identification of social interactions through partially overlapping peer groups. American Economic Journal: Applied Economics 2, 241–275.

Eccles, J.S., Barber, B.L., Stone, M., Hunt, J., 2003. Extracurricular activities and adolescent development. Journal of social issues 59, 865–889.

Feldman, A.F., Matjasko, J.L., 2005. The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. Review of educational research 75, 159–210.

Fredricks, J.A., Eccles, J.S., 2005. Developmental benefits of extracurricular involvement: Do peer characteristics mediate the link between activities and youth outcomes? Journal of Youth and Adolescence 34, 507–520.

Fredricks, J.A., Eccles, J.S., 2006. Is extracurricular participation associated with beneficial outcomes? concurrent and longitudinal relations. Developmental psychology 42, 698.

Fredricks, J.A., Simpkins, S.D., 2013. Organized out-of-school activities and peer relationships: Theoretical perspectives and previous research. New directions for child and adolescent development 2013, 1–17.

Goldsmith-Pinkham, P., Imbens, G.W., 2013. Social networks and the identification of peer effects. Journal of Business & Economic Statistics 31, 253–264.

Hanushek, E.A., Kain, J.F., Markman, J.M., Rivkin, S.G., 2003. Does peer ability affect student achievement? Journal of applied econometrics 18, 527–544.

Hoxby, C.M., 2000a. The effects of class size on student achievement: New evidence from population variation. The Quarterly Journal of Economics 115, 1239–1285.

Hoxby, C.M., 2000b. Peer effects in the classroom: Learning from gender and race variation. Technical Report. NBER Working Paper Series, No. 7867.

Hsieh, C.S., Lee, L.F., 2016. A social interactions model with endogenous friendship formation and selectivity. Journal of Applied Econometrics 31, 301–319.

Hsieh, C.S., Lee, L.F., Boucher, V., 2020. Specification and estimation of network formation and network interaction models with the exponential probability distribution. Quantitative economics 11, 1349–1390.

Imberman, S.A., Kugler, A.D., Sacerdote, B.I., 2012. Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. American Economic Review 102, 2048–2082.

Lavy, V., Paserman, M.D., Schlosser, A., 2012. 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.

Lavy, V., Schlosser, A., 2011. Mechanisms and impacts of gender peer effects at school. American Economic Journal: Applied Economics 3, 1–33.

Lee, L.F., 2007. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. Journal of Econometrics 140, 333–374.

Lin, X., 2010. Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. Journal of Labor Economics 28, 825–860.

Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. The Review of Economic Studies 60, 531–542.

Mas, A., Moretti, E., 2009. Peers at work. American Economic Review 99, 112–145.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: Homophily in social networks. Annual review of sociology 27, 415–444.

Moffitt, R.A., 2001. Policy interventions, low-level equilibria, and social interactions. Social dynamics 4, 6–17.

Morgan, D.L., Alwin, D.F., 1980. When less is more: School size and student social participation. Social Psychology Quarterly , 241–252.

Nicoletti, C., Salvanes, K.G., Tominey, E., 2018. The family peer effect on mothers' labor supply. American Economic Journal: Applied Economics 10, 206–234.

Olivetti, C., Patacchini, E., Zenou, Y., 2020. Mothers, peers, and gender-role identity. Journal of the European Economic Association 18, 266–301.

Patacchini, E., Zenou, Y., 2012. Juvenile delinquency and conformism. The Journal of Law, Economics, & Organization 28, 1–31.

Patrick, H., Ryan, A.M., Alfeld-Liro, C., Fredricks, J.A., Hruda, L.Z., Eccles, J.S., 1999. Adolescents' commitment to developing talent: The role of peers in continuing motivation for sports and the arts. Journal of youth and adolescence 28, 741–763.

Ringdal, C., Sjursen, I.H., 2021. Household bargaining and spending on children: Experimental evidence from Tanzania. Economica 88, 430–455.

Sacerdote, B., 2001. Peer effects with random assignment: Results for dartmouth roommates. The Quarterly journal of economics 116, 681–704.

Sacerdote, B., 2011. Peer effects in education: How might they work, how big are they and how much do we know thus far?, in: Handbook of the Economics of Education. Elsevier. volume 3, pp. 249–277.

Schaefer, D.R., Simpkins, S.D., Vest, A.E., Price, C.D., 2011. The contribution of extracurricular activities to adolescent friendships: new insights through social network analysis. Developmental psychology 47, 1141.

Shrum, W., Cheek Jr, N.H., MacD, S., 1988. Friendship in school: Gender and racial homophily. Sociology of Education , 227–239.

Staiger, D., Stock, J., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557–586.