The Peer Effects of Persistence on Students' Academic and Non-cognitive Outcomes*

Jian Zou†

December 2019

Abstract

This paper studies the peer effects of persistence, a psychological factor reflects the ability to persevere when facing challenges and setbacks, on students' academic and non-cognitive outcomes. By exploiting student-classroom random assignments in high schools and a nationally representative sample in China, I find positive impacts of peer persistence on students' academic performance. Moreover, in the follow-up wave, results show that having peers with higher persistence leads to enduringly improved academic outcomes and increased mental status. While there is no impact of peer persistence on students' own persistence, I find suggestive evidence of increased self-disciplined behaviors and friendship formation with "good" but not "bad" peers as potential mechanisms. I account for friendship formation evidence by developing a simple model of human capital formation with personality feature in which students' test scores is increasing with peer persistence. The model implies that the friendship sorting, as an equilibrium outcome, can be driven by the fact that students sharing similar persistence are more likely to make friends and is socially efficient in maximizing the total human capital production.

Keywords: peer effects, personality traits, human capital
JEL Code: I21, I24, J24

*I thank faculties at UIUC for their guide and feedback: Dan Bernhardt, Mark Borgschulte, Benjamin Marx, Elizabeth Powers, Julian Reif, and Rebecca Thornton, and helpful comments from Alex Eble, Roberts Brent, Bobby Chung, Nicolás Salamanca, Yiming Xia, Jianfeng Xu, Yuhao Yang, and Xiaoyang Ye, as well as seminar participants at the 2019 CES Annual Conference, the briq/IZA Workshop on Behavioral Economics of Education, Summer Seminar and Applied Micro Research Lunch at University of Illinois, 2019 CIEFR Summer School at Peking University, and Human Capital and Economic Development in China Conference at University of Buffalo. All errors are mine. This paper has previously circulated with the title *Understanding the Peer Effects of Non-Cognitive Ability on Academic Outcomes*.

†Department of Economics, University of Illinois at Urbana-Champaign, 1407 West Gregory Drive, Urbana, IL, 61801. jianzou2@illinois.edu.

1 Introduction

The impacts of non-cognitive abilities on both educational attainment and labor market outcomes have been well documented in the economics literature (Heckman and Rubinstein, 2001; Borghans et al., 2008; Heckman et al., 2019). The predictive power of non-cognitive abilities, such as persistence and self-control, rivals that of cognitive skills in improving many life outcomes, which include student's academic outcomes (Roberts, Wood and Caspi, 2008; Heckman, Pinto and Savelyev, 2013; Kautz et al., 2014). However, evidence related to the impacts of personality through social interaction is scarce. Understanding the peer effects of personality traits will help to unfold the mechanism behind results found in mindset-based intervention and thus future policy design (Alan, Boneva and Ertac, 2019). In this study, I investigate the peer effects of persistence, a personality trait reflects the ability to persevere when facing challenges and setbacks, on academic outcomes and non-cognitive outcomes and its potential mechanism, by exploiting student-classroom random assignments in high schools and a nationally representative sample in China.

Causal inference in estimating the impacts of peer persistence mainly encounters two empirical challenges: selection problem and reflection problem (Manski, 1993). Firstly, I address the selection problem by exploiting the student-classroom random assignment within grade in middle schools in China. Random assignment of newly enrolled students to the classroom at the beginning of Grade 7 becomes more widespread in China's middle high schools, as it is recommended by the Ministry of Education after tracking is prohibited in 2006.[1] The student-classroom random assignment within high schools provides exogenous variations of peer composition, making us able to examine the impacts of peers persistence. Secondly, to address the reflection problem, I use students' prior information of personality when they were at Grade 6 in the primary schools. Moreover, there is an additional challenge on external validity when estimating the impacts of peer persistence. Since most existing datasets do not contain measurements on persistence, study that uses self-collected survey data to overcome this challenge might still need to address the external validity issue. However, by exploiting a national representative sample of middle school students in China from the China Education Panel Survey (CEPS), this study is able to provide the external validity.

The results of this study show positive impacts of peer persistence on students' academic performance.

[1] See "Compulsory Education Law of the People's Republic of China (2006 Amendment)", http://www.gov.cn/flfg/2006-06/30/content_323302.htm.

Moreover, in the follow-up wave, I find impacts of peer persistence on sustainably improving students' academic outcomes and increasing their mental stress. Next, I examine the heterogeneity of impacts of peer persistence on academic outcomes. I find that high-persistence students benefit from persistent peers in achieving higher academic outcomes. However, no evidence is found on showing low and middle-persistence students benefit from having more persistent peers in the classroom.

Three potential mechanisms are examined in this study: students' own persistence, student behavioral respond, and friendship formation. While I find there is no impact of peer persistence on affecting students' own persistence, I find evidence that higher peer persistence is associated with spending less time watching TV and playing video games, and decreased behaviors on both tardy and truancy. Those results suggest increased self-disciplined behaviors as one mechanism because of having higher persistent peers. Next, by examining students' friendship formation, I find that having higher persistent peers in the classroom, students will more likely to make friends with "good" peers who are study hard and avoid making friends with "bad" peers who misbehave.

These results consistently suggest that, while having more persistent peers in the classroom improve student's academic achievement, it also makes them feel more unhappy after a long-term interaction with high persistent peers, probably due to their increased self-disciplined behaviors in response to high persistence peers.

To further understand the friendship formation mechanism and thus the role of peer persistence on students' human capital formation, I develop a simple model, following framework of Almlund et al. (2011), in which student's human capital is increasing in own persistence and peer persistence. Moreover, results in the heterogeneity section suggest that students' own persistence and peer persistence are complements on the human capital production, which means a positive cross-partial. A illustrative example consists two students with different type of persistence, and the positive cross-partial implies that students with high (low) persistence will more likely to make friendship with peers who also have high (low) persistence. The friendship sorting, as an equilibrium outcome of the model, can be driven by the fact that students sharing similar persistence are more likely to make friends and leads to socially efficient in maximizing the total human capital production. The empirical evidence of friendship formation found in this study can be

treated as a proxy evidence for the friendship sorting outcome predicted by this model.

The contribution of this study is threefold. Firstly, this study generalizes the recent study on impacts of peer personality by Golsteyn, Non and Zölitz (2019) beyond developed countries and to the secondary educational settings. Golsteyn, Non and Zölitz (2019) makes use of random assignment of students to teaching sections in a Dutch business school and find positive impacts of peer persistence on students' academic achievement, which are consistent with results found in this study.[2] Moreover, this study: a. extends the outcomes from traditional academic achievement to non-cognitive outcomes; b. provides empirical evidence and theoretical model on understanding the mechanisms; c. strengthens the external validity by exploiting a nationally representative student sample in China. Broadly, this study speaks to the peer effects in the economics of education literature.[3]

Secondly, this study also contributes to a growing body of literature documenting the importance of non-cognitive abilities, especially of personality traits, by cooperating knowledge from developmental psychology into economics (Heckman and Rubinstein, 2001; Heckman, Stixrud and Urzua, 2006; Borghans et al., 2008; Lindqvist and Vestman, 2011; Almlund et al., 2011; Heckman et al., 2019).

Lastly, the spillover pattern of peer personality traits on academic and non-cognitive outcomes found in this study potentially helps researchers to understand the mechanism for recent mindset-based intervention literature (Damgaard and Nielsen, 2018; Bettinger et al., 2018; Alan, Boneva and Ertac, 2019) and policymakers for future policy design.

The rest of the paper is organized as follows. Section 2 introduces the student-classroom random assignment and data. Section 3 and 4 presents the identification strategy and main empirical results, respectively. Section 5 discusses the mechanism and the human capital production model that illustrates the friendship formation. Section 6 concludes the paper.

[2]Another related study is by Shure (2017), in which she found consistent positive impacts of peer persistence on students' academic outcomes, by employing the school fixed-effects framework and focusing on secondary setting in UK. Also, by exploiting the preschool setting and plausible student-classroom random assignment within schools in US, Neidell and Waldfogel (2010) shows there are negative impacts of peer externalizing behaviors on students' academic outcomes.

[3]Examples includes Feld and Zölitz (2017); Carrell, Sacerdote and West (2013); Carrell, Fullerton and West (2009); Kremer, Miguel and Thornton (2009), and Lyle (2007) For a survey, see Sacerdote (2011).

2 Randomization Background and Data

2.1 Randomization in Middle Schools in China

In the pre-university education system in China, a typical student in China goes through 6 years in primary school (Grade 1 to 6), 3 years in junior high school (Grade 7 to 9), and 3 years in senior high school (Grade 10 to 12). The nine-year compulsory education in China requires students to finish junior high school education before they can decide whether to continue their education in senior high school.

While tracking is commonly used in the senior high school in China, it is explicitly prohibited in junior high school by the Ministry of Education since 2006. As recommended by the Ministry of Education for students' equal and fair opportunities purpose, at the beginning of Grade 7 in middle high schools, the random assignments of students to the classroom increases its popularity as an alternative. Two random assignment strategies are commonly used in those middle schools in China: i) purely random assignment, and ii) "balanced assignment" rule. The former uses a computer program, incorporating student's demographics, classrooms, and other information, to carry the randomization process. The latter, "balanced assignment" rule, refers to the process of evenly assigning students into the classroom with one restriction on having different class compositions with similar test score on average.[4] While we cannot differentiate these two types in our analysis, both institutions are considered to be a legitimate randomization process to assess a causal impact in peer effect studies.[5]

2.1 Data and Variables

Not all middle schools adopt the random assignment policy, I choose schools with random assignment in CEPS following similar criteria used by Gong, Lu and Song (2018). Schools that assign students to classrooms

[4] An example of the "balanced assignment": We assume there is a middle school with 5 classes and 200 students at Grade 7. Based on those 200 students' test scores in exams at the end of Grade 6 or diagnostic test scores at the beginning of Grade 7, the school will rank those students from 1 to 200. First, starting with the top 5 students, the school will assign the student with rank 1 to the Class 1, rank 2 student to the Class 2, rank 3 student to the Class 3, and all the way to the student with rank 5 to the Class 5. Next, the student with rank 6 will be assigned to the Class 5, rank 7 student will be in Class 4, and all the way back to assigning rank 10 student to the Class 1. The school will keep continuing the above process until all the students are assigned in the classroom. The "balanced assignment" rule is similar to the "snake draft" rule in the American football game. Detailed examples of using "balanced assignment" for the random assignment process in China's middle schools can be found in He and Ross (2017), Feng and Li (2016), and Carman and Zhang (2012).

[5] This random assignment strategy was firstly used in estimating the gender peer effects by Hu (2015) and effects of teacher-student gender match by Gong, Lu and Song (2018), and is also used in several follow-up peer effects studies. (Eble and Hu, 2018, 2017; Hu, 2018; Gong, Lu and Song, 2019; Xu, Zhang and Zhou, 2019)

randomly should meet the following two conditions: (i) the school principal reports that standards used for arranging new students into classrooms are random assignment; and (ii) all headteachers in the seventh grade report that students are not assigned by test scores.[6] This leads to 59 schools remain, which is 52.6% of the school sample. Since the variation comes from with-grade class averages, I drop 3 schools where there is only one class in Grade 7. Thus, the final sample I analyze in wave 1 consists of 5,105 students across 112 classrooms in 56 schools.

For schools in the follow-up wave, some of those schools will reassign their students to a different class at the beginning of Grade 8, with or without using random assignment policy. Since the information on how those schools reassign their students is not available, I exclude all the schools in the second wave with reassignment policy. I select schools in the second wave by further meeting the following two conditions: (i) the school principal reports that there was no reassignment of students to classrooms at the beginning of Grade 8; and (ii) there is 'w2clsra" indicator in the CEPS second wave to identify which whether there was reassignment at the beginning of Grade 8.[7] Therefore, I further exclude 8 schools, and the final sample used as the estimation sample ending up with 4,321 students across 96 classrooms and 48 schools. See **Table 1** for statistical description.

Seven questions related to student's personality traits at Grade 7 are measured in the CEPS Wave 1, which can be found in **Table 2**.[8] The first three questions measure students' persistence, a lower aspect of Conscientiousness.[9] Persistence reflects people's extent of hard-working and ability to persevere in the face of challenges and setbacks, which is strongly correlated with Grit. People who are persistent tend to be achievement-oriented, self-disciplined, and efficient. The measurement of persistence are similar to items used in Golsteyn, Non and Zölitz (2019), which is originally from the Student Motivation Scale of Martin

[6]Compared to the conditions used in this study, Gong, Lu and Song (2018) have one more condition in their study's criteria: after students have been assigned to classrooms at the beginning of the seventh grade, there is not a rearrangement of their classes for grades 8 and 9. However, in this study I only focus on students from the seventh grade in the first wave of CEPS, instead of students from both seventh and ninth grade, so I do not need to meet this condition in my criteria.

[7]While schools' information on reassignment is almost identical under those two conditions, there is one school where the principal reports they did reassign students at Grade 8 but "w2clsra" indicates there was not. I drop that school as well in the second wave.

[8]One of the main models used in psychology is the Big Five model, which includes five factors of personality traits: Openness to experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

[9]Firstly, persistence is also referred as industriousness in the psychology literature. Secondly, in the psychology literature, a traditional lower-order structure of Conscientiousness consists of two aspects: industriousness and orderliness. However, there are alternatives on how to construct lower-order structure of Conscientiousness in the psychology literature. For example, Jackson and Roberts (2017) argues that recent studies support the four-factor solutions which divide Conscientiousness into orderliness, industriousness, responsibility, and self-control.

(2009). The rest four questions are items related to another two personality traits of the Big Five model, Extraversion and Openness. Since those four questions are not precise measurements neither for any of those two personality traits nor any of their lower facets, I use them as student's baseline personality traits control to obtain more precise estimates.[10] I average those questions to come up with three indexes: one measurement for persistence and two personality controls that are related to Extraversion and Openness. For estimate interpretation later, following Golsteyn, Non and Zölitz (2019), I standardized persistence over the entire estimation sample with mean 0 and standard deviation 1.

Academic outcomes are consist of total scores in midterm tests at Grade 7 and 8. Test scores of three subjects (Chinese, Math, and English) provided by the CEPS were standardized with mean 70 and standard deviation 10. I take the sum of them to come up with the total scores and then standardize the total scores by school-grade level with mean 0 and standard deviation 1, for the purpose of estimates interpretation.

Non-cognitive outcomes include student's mental stress measurement at Grade 7 and 8. For mental stress measurements, I followed Gong, Lu and Song (2018) and selected the following four questions in both waves. The CEPS asked students "Do you have the feelings below in the last seven days?" with a scale from 1 (never) to 5 (always): (1) depressed, (2) blue, (3) unhappy, or (4) that life is meaningless.

I choose one set of students' predetermined variables to implement balancing tests for the random assignment and include them as controls in the later estimations. Those control variables include: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6.

[10] I aim to use the measurements of students' personality traits at Grade 6 to address the reflection problem (Manski, 1993). The issue here is that personality traits at Grade 6 were surveyed when the students are at Grade 7. Therefore, one might think about the measurement error issue. In other words, those personality measurements may be a function of students' status at Grade 7. I address these issues later in the balancing tests discussion in Section 4.

3. Identification

3.1 Balancing Tests

To verify the student-classroom random assignments in middle high schools, I examine the balancing tests using the following equation:

$$\overline{\mathrm{P}}_{\text{-i,j,k}} = \beta_0 + \beta_1 \mathrm{X}_{\mathrm{i,j,k}} + \delta_{\mathrm{k}} + \epsilon_{\mathrm{i}} \tag{1}$$

Where $\overline{\mathrm{P}}_{\text{-i,j,k}}$ is the leave-me-out-mean of peers' persistence in class j school-grade k. $X_{i,j,k}$ is student's time-invariant variables, which include student's age, gender, minority, hukou status, migrant status, whether from one-child family, whether attended kindergarten, age attending primary school, class percentile ranking and personality traits at Grade 6, and father's and mother's years of schooling. δ_k is the school-grade fixed effects, and ϵ_i is the error term. Follows Abadie et al. (2017), I cluster the standard errors at school-grade level, the level of randomization, which allows correlation across students within each school-grade.

We should expect that, without school-grade fixed effects, student's predetermined variables will correlate with her peers' persistence average. By only including an additional school-grade fixed effect, there should be no correlation between student's predetermined variables and peers' persistence average.

3.2 Linear-in-Mean Model

To examine the peer effects of persistence on students' academic outcomes, I use the following linear-in-mean model:

$$\mathrm{Y}_{\mathrm{i,j,k}} = \beta_0 + \beta_1 \mathrm{P}_{\mathrm{i,j,k}} + \beta_2 \overline{\mathrm{P}}_{\text{-i,j,k}} + \beta_3 \mathrm{X}_{\mathrm{i,j,k}} + \delta_{\mathrm{k}} + \epsilon_{\mathrm{i}} \tag{2}$$

Where $Y_{i,j,k}$ are student's academic outcomes at Grade 7 and 8. $\mathrm{P}^{6\mathrm{th}}$ is student i's personality when she was at a student at Grade 6. $\overline{\mathrm{P}}_{\text{-i,j,k}}$ is the leave-me-out-mean of peers' persistence in class j school-grade k, $X_{i,j,k}$ is one set of student's time-invariant variables.

The β_2 is the parameter of interests, which will give us the estimates of impacts of peer's persistence on academic outcomes. In the regressions, I also include leave-me-out average of other peer characteristics – gender, migrant, and percentile ranking at Grade 6 – to obtain more estimates.

4. Main Findings

4.1 Tests for Random Assignment

The results of balancing tests of random assignment can be found in **Table 3**, which shows the estimations of regressing peers' persistence mean on the student's predetermined variables. While results in column (1) are with school-grade fixed effects, results in column (2) are the corresponding balancing test without school-grade fixed effects. We can see that, without school-grade fixed effects, some of student's characteristics are correlated with the peers' persistence. However, conditional on which school-grade to attend, none of the student' time-invariant variables is correlated with her peer mean of persistence at Grade 7, which implies good balancing for student-class random assignment within the school-grade level. Moreover, **Table 3** also provides the F-statistics and corresponding p-value of a joint hypothesis test of student's predetermined variables. The insignificance of F-statistics implies that student's predetermined variables jointly can not predict her peers' persistence mean.

One concern on the personality traits measurements is that those traits at Grade 6 were surveyed when students are at Grade 7, thus recalled personality traits measurements may reflect and be correlated with peers' characteristics at Grade 7. If this is true, then we should expect there are correlations between student's own personality traits and her peer persistence. However, balancing tests of **Table 3** shows that, with school-grade fixed effects, student's own personality traits are not correlated with peer's persistence. Also, as shown later in the results of **Table 8**, having peers with higher persistence will not affect students' own persistence, which makes it unlikely the case that we have retrospective persistence being affected by peers' own persistence.

Another concern is that personality traits measurements in the CEPS survey are poor measures of the potential underlying confounders, which makes the above balancing tests unable to well address student's clustering in the classroom caused by personality characteristics. I use a more powerful test proposed by Pei, Pischke and Schwandt (2019), where controls are used as dependent variables in balancing regressions to detect potential selection. This method not only helps to address the potential poor measure issue of personality traits but also can be used to check the robustness of balancing tests in **Table 3**. Appendix

Table A1 shows that none of the balancing tests for various student's predetermined controls yields a significant correlation with peers' persistence.

4.2 Impacts of Peer Persistence on Academic Outcomes

Before examining the impacts of peer persistence on students' academic and non-cognitive outcomes, **Table 4** gives the correlation of student's own persistence and total scores. From **Table 4**, we can see that persistence is positively correlated with students' academic performance. Taking students at Grade 7 as an example, a one standard deviation increase in persistence is correlated with a 7.5 percent increase in total score. **Figure 1** visualizes the regression results with the same set of controls in **Table 4**, using binned scatter following Chetty, Friedman and Rockoff (2014).

Table 5 shows the peer effects of persistence on students' total scores at Grade 7 and 8. As shown in **Table 5** 's column (1), a one standard deviation increase in peer persistence leads to a 13.2 percent increase in standardized total scores. To obtain more precise estimates, I further control for other peer characteristics: leave-me-out average of peers' gender, migrant status, and percentile ranking at Grade 6.[11] As shown in column (2), the significance of estimate remains unchanged and the magnitude becomes slightly smaller, which implies a robust estimation. In column (3) and (4) of **Table 5**, I examine the impacts of peer personality on students' academic performance in the follow-up wave. From the results of column (3) and (4), we can find that, by continually interacting with the same classmates, impacts of peer persistence are significant and sustaining at Grade 8. One concern from the linear-in-mean framework is that we are imposing a linear relationship between the variance of peer mean and dependent variable. I address this concern by visualizing their relationship and whether if the relationship is linear. **Figure 2** visualizes the regression results in columns (2) and (4) of **Table 5**, where we can see the relationship between total scores and peer persistence in both subfigures are quite linear.

Compared to the impacts of peer persistence in Golsteyn, Non and Zölitz (2019), this study finds consistent but relatively larger estimates. One potential reason could be due to, compared to a standard randomized control trial, we don't have ideal baseline measurements on academic outcomes in this study to

[11]There are studies showing the impacts of female peers (Hu, 2015; Gong, Lu and Song, 2019), migrant peers Hu (2018) and low-ability peers Xu, Zhang and Zhou (2019) on student's academic and non-cognitive outcomes, using the same settings in term of data and identification strategy.

control for. Even though this study has controlled almost all students' characteristics before and at Grade 6 that CEPS has, the estimates here are still not precise enough, because of the lack of baseline information. Another reason could be due to the setting differences– middle school setting in developing counties could be different from higher education setting in the developed county in Golsteyn, Non and Zölitz (2019).[12]

4.3 Impacts of Peer Persistence on Non-cognitive Outcomes

I further examine the impacts of peer persistence on students' non-cognitive outcomes: mental stress at Grade 7 and 8. I first examine the four mental stress items separately, and then calculate an index which takes an average of those four items. Finally, to get the aggregated effects of peer persistence on mental stress, I follow Kling, Liebman and Katz (2007) to compute the average effect size (AES), where the estimation takes a weighted average of those four items as the dependent variable.[13] As shown in Panel A of Table 6, while there is a significant impact on Pessimistic item, in general, there are no impacts of peer persistence on mental stress for students at Grade 7. However, when we look at students in the follow-up wave, the results in Panel B show that having higher persistent peers will increase student's mental stress. Moreover, estimates from both the index and average effect size show the same level of significance and similar magnitude, which help to address the multiple hypothesis testing issues. Results in column (12) of Table 6 give the aggregated impacts of peer persistence on mental stress: a one standard deviation increase in peer persistence leads to a 5.3 percent increase in student's mental stress.

Therefore, those results consistently suggest that, while having higher persistence classmates increases students' academic achievement, it also worsens their mental status, probably due to students' increased self-disciplined behaviors (as shown in the mechanism section) in response to interacting with higher persistent peers.

[12]One related evidence is that, using the same dataset and random assignment identification strategy, Hu (2018) studies the impacts of migrant peers in the classroom and finds relatively larger estimates compared to estimates found in the developed country.

[13]The definition of AES follows Gong, Lu and Song (2018). Specifically, the AES of peer persistence on outcome k in non-cognitive category c is defined as the following: $\text{AES}_c = (1/n_c)\sum_{n=1}^{n_c} e_{kc}/\sigma_{kc}$, where n_c is the number of outcomes in category c, e_{kc} is the estimated effect for outcome k, and σ_{kc} is the standard deviation of the outcome variables. The standard error and p-value of AES is calculated by following the method provided in the Web Appendix B of Kling, Liebman and Katz (2007).

4.4 Heterogeneity

Exogenous peers' variation, under the linear-in-mean framework, runs into challenges from heterogeneous priors, which imply that the same peer contextual effect can generate different impacts among different subset of the sample. In other words, even an effective peers' variation may produce spuriously null average treatment effects, since there are offsetting heterogeneous treatment effects of opposite signs among different subsamples[14]. Thus, it will shed more lights on the main results by looking at the heterogeneity of impacts of peer persistence on subsample with different prior persistence.

For the heterogeneity behind impacts of peer persistence on students' academic outcomes, I use the following equation (3) to look at how peer persistence impacts vary by interacting with student's baseline persistence. I first divide student's baseline persistence into three dummy groups: low, medium, and high persistence. Next I obtain the interaction terms between peer persistence and those three dummy groups. By replacing student's own persistence and include the three interaction terms into the equation (2), we have the following specification:

$$\begin{aligned} \mathrm{Y_{i,j,k}} = \beta_0 + \sum_{\text{level}=1}^{3} \alpha^{\text{level}} \mathrm{P^{level}_{i,j,k}} + \sum_{\text{level}=1}^{3} \gamma^{\text{level}} \mathrm{P^{level}_{i,j,k}} * \overline{\mathrm{P}}_{\text{-i,j,k}} \\ + \psi \overline{\mathrm{P}}_{\text{-i,j,k}} + \lambda \mathrm{X_{i,j,k}} + \delta_k + \epsilon_i \end{aligned} \tag{3}$$

Where $Y_{i,j,k}$ are student's academic outcomes at Grade 7. $\mathrm{P^{level}_{i,j,k}}$ is the dummy group measuring student i's own persistence with level $\in$ [low, medium, high]. γ^{level} is parameter of interest that reflects the heterogeneity of peer effects, as it is indicated by interaction term $\mathrm{P^{level}_{i,j,k}} * \overline{\mathrm{P}}_{-i,j,k}$. The rest are the same as those in equation (2).

Table 7 shows the heterogeneous effects of peer persistence on students' academic outcomes. I find that only high-persistence students benefit from more persistent peers in achieving higher academic outcomes, which implies a complementary relationship between high persistence students and having higher persistent peers.

[14] Ability peer effects in the educational literature

5. Mechanism

The main results in **Table 5** and **6** suggest the To understand the mechanism behind the impacts of peer persistence on students' academic and non-cognitive outcomes, I examine the following mechanisms: direct personality spillover, student behavior changes, and friendship formation.

5.1 Does Peer Persistence Change Students' Own Persistence?

A recent study by Alan, Boneva and Ertac (2019) finds positive impacts of boosted grit on student's academic achievement, and suggests peer effects as a potential mechanism where treated students may lead to beliefs and behaviors changes of other untreated students in the same classroom. Together with the fact that persistence is highly closed with grit, one potential mechanism behind the impacts of peers' persistence could come from the direct spillover of persistence. In other words, having higher persistent peers in the classroom will increase students' own persistence or related behaviors, which will lead to improved academic performance eventually.

In CEPS wave 2, students at Grade 8 were asked about their own persistence at Grade 7 using four questions.[15] To examine the persistence spillover mechanism, I first use the same three questions items as those in wave 1, and then compute an index by averaging those three questions, and finally come up with a weighted average to estimate the AES. The results in column (1) -(4) of **Table 8** shows that there are no impacts of having higher persistent peers on student's own persistence. Also, as we can see in the column (5) of **Table 8**, the estimate for aggregated effect is positive but insignificant.

One may argue that it will need a longer time to observe personality development, which is similar to the fact that, in general, it takes a longer time to observe the belief or norm changes if there is any. In other words, there may be persistence changes if we observe or interact the student with higher persistent peers for a longer time. However, since the dependent variables on persistence were measured one year after the initial random assignment, the results of **Table 8** may indicate that there are indeed do not have impacts of peers' persistence on students' own persistence. Since psychology literature describe persistent individuals

[15]The CEPS Wave 2 asked students "How much do you agree with each of the following statements about your experiences in GRADE 7?" with a scale from 1 (strongly disagree) to 4 (strongly agree), using the same three questions on persistence as those in **Table 2** with an additional question "I would persist in my interests and hobbies".

as people who tend to finish what they start, work long hours or even take on extra work (Jackson et al., 2010), I use students' study hour as secondary measurements for potential persistence changes because of peer persistence. However, as shown in Appendix **Table A2**, there are also no impacts of peer persistence on student's time use on study.

Thus, those findings imply that personality spillover may not work in a similar way as human capital spillover does, where human capital displays a direct and explicit spillover pattern. Therefore, a natural question is raised: If there is no direct spillover of persistence, then what is the causal link behind the impacts of peer persistence on students' human capital formation and also on the increased mental stress.

5.2 Students' Self-Disciplined Behaviors

While there are no student persistence changes due to peer persistence, there may be students' behavioral changes by having higher persistent peers, which leads to changes in student's academic and non-cognitive outcomes.

CEPS surveys various questions from students so that I am able to examine student's behavioral responses in a detailed way. One set of questions CEPS asks students is the time use of students on various activities at last weekday and weekend. I first look at student's time use on study hour and entertainment last weekday and weekend, measured by hour per day.[16] From panel A of **Table 9**, results imply that students with higher persistent peers spend less time on watching TV last weekday and weekend and playing video games last weekday. The decreased time spent on entertainment is mainly driven by the time use decrease during the weekday. Consistently, results show that student's own persistence is highly correlated with decreased time spent on entertainment.

Next, I investigate whether there are some changes in student's school attendance behaviors.[17] Results in Panel B of **Table 9** show consistent evidence on the impacts of peer persistent in decreasing students' tardy and truancy behaviors. Combined with the suggestive evidence on time use on entertainment in Panel A of **Table 9**, I propose that those results consistently indicate students' increased self-disciplined

[16] I look at how many hours per day the student spent last weekday and weekend on: 1. homework assigned by teacher; 2. tutoring homework assigned by parents or tutoring school; 3. watching TV; 4. playing video games. While the first second items are time use on study, the last two are time use on entertainment.

[17] The CEPS 2013-14 asked students "How much do you agree with each of the following statements about your school life?" with a scale from 1 (strongly disagree) to 4 (strongly agree): 1. "I am always late for class." and 2. "I always skip classes.".

behaviors as one mechanism in explaining the impacts of peer persistence on improving students' academic achievement. Moreover, increased self-disciplined behaviors are also more likely to increase student's mental stress as students need to restrict their own behaviors, especially on entertainment time use. Finally, there are also results of increased student's education expectations as a potential mechanism. As shown in panel C of **Table 9**, I find that having higher persistent peers has positive impacts on the students' expectations of obtaining a college degree in the future.

Recall that persistence, in this study, describes the personality trait at which students will be more achievement-oriented and self-disciplined. The impacts of peer persistence on student's increased self-disciplined behaviors are also in line with the psychological root of the persistence concept. Moreover, in contrast to the traditional channel of improved students' effort on human capital formation, the mechanism behind peer persistence impacts yield new understanding for future research that works on the peer effects of other non-cognitive abilities. In contrast to direct persistence spillover, I propose the increased self-disciplined behaviors as a direct but inexplicit spillover of peer persistence, and those spillovers potentially become one channel that helps to improve students' human capital formation.

5.3 Endogenous Friendship Formation

People with similar characteristics are more likely to make friends with each other because of the factors that psychologists call *homophily*.[18] Homophily has been widely studied in friendship formation in both psychology (Selfhout et al., 2010; Wrzus et al., 2017) and economics literature(Jackson, 2010). Therefore, students with personality similarity should be more likely to form friendship with each other. However, while students with similar personality traits are more like to interact with each other, personality similarity not necessarily leads to improvement on students' academic achievement. Recall the results in the previous mechanism part, where we found that interacting with persistent peers will increase student's school-oriented self-disciplined behaviors. Therefore, two students with higher persistence interact with each other in the classroom, we can infer that they will decrease their "tasks-trading" cost mutually, and thus increasing their total human capital production.

[18]Homophily is one among the three factors that shape the primates friendship formation in psychology literature (Laakasuo et al., 2017).

CEPS surveys students' five best friends in the survey and their own description of the behaviors of those five best friends.[19] While the five best friend network data is not publicly available in CEPS, I am able to investigate student's friendship formation using the self-reported description of five best friends' behaviors. For interpretation, I create a dummy variable for each behavior description with value 1 indicating at least one of five best friends is fitting in that behavior and value 0 indicating none of those five best friends in line with that behavior description.

Panel A of **Table 10** shows the results using the linear-in-mean model, which shows the positive impacts of peer persistence on the likelihood of having at least one best friend who is working hard, and positive impacts on likelihood of making friends with peers who are smoking or drinking, going to net cafe often, or in love relationship. Interestingly, I did not find evidence on having higher persistent classmates will affect friendship formation with peers who have an excellent grade. Results in Panel A provides suggestive evidence that, having higher persistent peers in the classroom, will affect student's friendship formation by making friends with "good" peer and avoid making friends with "bad" peers. To further look at the heterogeneous impacts of peer persistence on friendship formation, I use the same specification in the heterogeneity section with dependent variables replaced to friendship outcomes. Results in Panel B of **Table 10** shows the only students with high persistence perform the above friendship formation patterns.

There are two potential scenarios to understand the role of friendship formation. The first scenario comes from the *direct* effects of peer persistence on academic performance, where students benefit from having high persistent classmates directly. After that, those students will have a chance to build friendship networks with "good" peers and not "bad" peers. In this case, friendship formation is the consequence of peer persistence, rather than its mechanism. The second scenario comes from the *indirect* effects of peer persistence. In this case, students benefit from having high persistent classmates, probably because high persistent peers form a better environment to foster academic achievement and high persistent students know how to better make use of this environment by making friends with them. Therefore, in the latter case, friendship formation is treated as the mechanism of impacts of peer persistence on academic achievement.

[19] CEPS asked students "How many of your best friends mentioned above fit in the following descriptions?" with a scale of 0 (none of them), 1 (one or two of them), and 2 (most of them): 1. doing well in academic performance; 2. studying hard; 3. expecting to go to college; 4. skipping classes (truancy); 5. being criticized or punished for violating school rules; 6. always fighting with others; 7. smoking or drinking; 8. always going to net bars or video arcade; 9. has had or is having a romance; 10. dropping out of school.

If the first scenario is true, we should observe that only student with high persistence will have their academic performance improved, which have been shown as true in **Table 10**. However, results found in **Table 10** also support the second scenario.[20] Therefore, we can not rule out any of those two scenarios by looking at the heterogeneous effects of peer persistence. However, while I cannot rule out the possibility of the first scenario, there are several additional pieces of evidence is consistent in suggesting the friendship formation as a mechanism, which is the second scenario. Firstly, if the first scenario is true, students with high persistence will also more likely to make friends with peers with excellent grades. However, this is not supported by the results found in column (3) of **Table 10**. Secondly, over 70% of students in this sample has most of their five best friends come from the same class, which suggests friendship formation happens within instead of outside class.[21] Therefore, **Table 10** provides suggestive on supporting friendship formation mechanism as a *indirect* mechanism behind impacts of peer persistence on student's human capital formation.

Classical spillover theory proposes the idea that human capital spillover is coming from "ideas/skills exchange" behind agents. (Alfred, 1890; Lucas Jr, 1988). However, while most peer effects studies identify student's behavior or beliefs changes, teacher's or parental investment adjustment as potential channels, few of them are able to indeed detect the friendship network changes. The failure of detecting friendship mechanism, partially due to the lack of friendship information or there is no friendship mechanism, eventually forces those studies to reply on the linear-in-mean framework, which assumes that everyone within the micro-environment except you was treated as your friend in equal weight. On the other side, theoretical literature has been developed trying to estimate peer effects with the endogenous friendship formation accounted for. However, that literature in general needs assumptions on the friendship network side (Griffith, 2018; Auerbach, 2019). Although I am not able to directly analyze the five best friends network due to the data limitation, this study able to detect the friendship formation mechanism with suggestive evidence. Moreover, this study sheds light on understanding human capital formation due to the peer effects by cooperating with the peers' non-cognitive characteristics and potentially could help to address the endogenous peer grouping

[20] In Golsteyn, Non and Zölitz (2019), they find both low and high persistent students benefit from having higher persistent peers. If I also find the same pattern in this study, then I can rule out the first scenario.

[21] Since linking the description of five best friends and the demographic information of those five best friends are not possible in CEPS, I cannot provide direct evidence on within and outsides class friendship formation, which will also be helpful to distinguish those two scenarios.

issue in the optimal policy design in Carrell, Sacerdote and West (2013).

5.4 Model

To interpret the friendship formation mechanism, I use a simple model relating to the role of peer persistence on human capital formation in the classroom. The model generates predictions on friendship sorting and thus helps to understand on the role of friendship sorting on human capital formation.

Following the conceptual framework in Almlund et al. (2011), student i's human capital formation function production $h_{i,T}$ in task T depends on her own personality traits p_i and the "effort" she expend on the task $e_{i,T}$: $h_{i,T} = f(p_i, e_{i,T})$. Therefore, while effort is task-specific, personality trait is modeled as public good, which is available in the same amount for all task. Moreover, $e_{i,T} = \phi(p_i)$, which means personality traits are introduced into economic models through constraints that enhance or dampen productivity. [22]

For simplicity and intuitiveness, here I shut down the task-specific subscript T and ignore the potential interacting relationship between students' cognition and personality (and peer personality).[23] Moreover, I model the peer persistence affecting human capital production through effort, and not directly affecting human capital production. Thus, a student i' human capital production function with persistence p_i and peer persistence p_{-i} can be written as:

$$h(p_i, p_{-i}) = f(p_i, \phi(p_i, p_{-i})) \tag{4}$$

where effort is modeling as $e_i = \phi(p_i, p_{-i})$. Given the positive impacts of peer persistence on students' academic outcomes in the **Table 4**, we assume $h(p_i, p_{-i})$ is increasing in each of its arguments. More importantly, empirical results in **Table 7** shows that the cross-partial is positive, i.e. $h_{1,2} = \partial^2 h / \partial p_i \partial p_{-i} > 0$, which means there is complementary relationship between students' own persistence and peer persistence on improving human capital production.

[22] In a more general model, h_i also include the student's cognitive ability and task-specific skills, and e_i depends on three more elements: cognition, task-specific skills, incentives, and preference. See Heckman et al. (2019).

[23] Given that peer persistence has positive impacts on improving students' academic outcomes, it is likely to model a complementary relationship between the student's cognitive ability and peer persistence, which resembles to the dynamic complementarity idea we model relationship between cognitive abilities and non-cognitive abilities on child skill formation (Cunha and Heckman, 2007). Suggestive evidence supporting the complementary relationship between the student's cognitive ability and peer persistence can be found in Appendix **Table A3**.

Suppose a friendship network with two students $i = 1, 2$, and human capital production of those two student are:

$$h_1 = f(p_1, p_2) \text{ and } h_2 = f(p_2, p_1) \tag{5}$$

Given positive cross-partial, we have

$$\frac{\partial^2 h(p_1, p_2)}{\partial p_1 \partial p_2} > 0 \tag{6}$$

where the human capital production function exhibits supermodularity.

Suppose student 1 has higher persistence than student 2, then we write them down as $p_h > p_l$, where student 1 is high type (h) and student 2 is low type (l). Thus, the marginal willingness to pay of a student with high persistence to be in the same friendship network with high persistence student, rather than a low-type student, is $h(p_h, p_h) - h(p_h, p_l)$. Similarly, the marginal willingness to pay of a low persistence student stay in the same friendship network with high persistence student, is $h(p_1, p_h) - h(p_l, p_l)$.

Therefore, by the definition of supermodularity, we have:

$$h(p_h, p_h) + h(p_l, p_l) > h(p_h, p_l) + h(p_l, p_h) \tag{7}$$

The interpretation of equation (7) is that high persistence students will more willing to make friends with the high persistence student, and low-persistence student will also more likely to make friends will low-persistence peers. With human capital production maximizing agents and positive cross-partial, friendship sorting will be the equilibrium outcome. In other words, students sort into specific friendship network to maximize their human capital within those friendship networks. That explains the friendship formation results we had in the **Table 10**. While I cannot have direct evidence on supporting that high persistence student will make friends with high persistence peers, the results of **Table 10** provides suggestive evidence that, with higher peer persistence, student will be sorting into friendship that yields higher return on their academic outcomes (Panel A) and only student with high persistence will perform in that way (Panel B).

6. Conclusion

Exploiting the random assignment of students to classes within middle schools and a nationally representative sample in China, I study the peer effects of persistence on students' academic and non-cognitive outcomes, and its mechanisms. I investigate the impacts of having more persistent peers in the classroom on students' academic performance and their mental stress at two waves. The results show positive impacts of peer persistence on students' academic performance. Moreover, in the follow-up wave, the results find that having peers with higher persistence leads to enduring positive impacts on academic outcomes and negative impacts on mental status. The heterogeneity shows that only high persistent student will be of benefit to having higher persistent peers in the classroom.

Four potential channels are tested in this study: direct persistence spillover, student behavioral respond, and friendship formation. While there is no impact of peer persistence on affecting students' own persistence, I find suggestive evidence of increased self-disciplined behaviors and friendship formation with "good" peers but not "bad" peers as potential mechanisms. Those results consistently suggest that, while having higher persistence classmates will increase students' academic achievement, it will also lower their mental status after a long-term interaction with high persistent peers, probably due to their increased self-disciplined behaviors in response to high persistence peer.

To further understand the friendship formation mechanism and thus the role of peer persistence on students' human capital formation, I develop a simple model, following framework of Almlund et al. (2011), in which student's human capital is increasing in own persistence and peer persistence. Moreover, results in the heterogeneity section suggest that students' own persistence and peer persistence are complements on the human capital production, which means a positive cross-partial. A illustrative example consists two students with different type of persistence, and the positive cross-partial implies that students with high (low) persistence will more likely to make friendship with peers who also have high (low) persistence. The friendship sorting, as an equilibrium outcome of the model, can be driven by the fact that students sharing similar persistence are more likely to make friends and leads to socially efficient in maximizing the total human capital production. The empirical evidence of friendship formation found in this study can be treated as a proxy evidence for the friendship sorting outcome predicted by this model.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge. 2017. "When should you adjust standard errors for clustering?" NBER *Working Paper 24003*.

Alan, Sule, Teodora Boneva, and Seda Ertac. 2019. "Ever failed, try again, succeed better: Results from a randomized educational intervention on grit." *The Quarterly Journal of Economics*, 134(3): 1121–1162.

Alfred, Marshall. 1890. "Principles of Economics, (1890)."

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. 2011. "Personality psychology and economics." In *Handbook of the Economics of Education*. Vol. 4, 1–181. Elsevier.

Auerbach, Eric. 2019. "Identification and Estimation of a Partially Linear Regression Model using Network Data." *arXiv preprint arXiv:1903.09679*.

Bettinger, Eric, Sten Ludvigsen, Mari Rege, Ingeborg F Solli, and David Yeager. 2018. "Increasing perseverance in math: Evidence from a field experiment in Norway." *Journal of Economic Behavior & Organization*, 146: 1–15.

Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel. 2008. "The economics and psychology of personality traits." *Journal of human Resources*, 43(4): 972–1059.

Carman, Katherine Grace, and Lei Zhang. 2012. "Classroom peer effects and academic achievement: Evidence from a Chinese middle school." *China Economic Review*, 23(2): 223–237.

Carrell, Scott E, Bruce I Sacerdote, and James E West. 2013. "From natural variation to optimal policy? The importance of endogenous peer group formation." *Econometrica*, 81(3): 855–882.

Carrell, Scott E, Richard L Fullerton, and James E West. 2009. "Does your cohort matter? Measuring peer effects in college achievement." *Journal of Labor Economics*, 27(3): 439–464.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American economic review*, 104(9): 2633–79.

Cunha, Flavio, and James Heckman. 2007. "The technology of skill formation." *American Economic Review*, 97(2): 31–47.

Damgaard, Mette Trier, and Helena Skyt Nielsen. 2018. "Nudging in education." *Economics of Education Review*, 64: 313–342.

Eble, Alex, and Feng Hu. 2017. "Stereotypes, role models, and the formation of beliefs." *Columbia University, Center for Development Economics and Policy, CDEP-CGEG Working Paper*, , (43).

Eble, Alex, and Feng Hu. 2018. "The sins of the parents: Persistence of gender bias across generations and the gender gap in math performance." *CDEP-CGEG Working Paper No*, 53.

Feld, Jan, and Ulf Zölitz. 2017. "Understanding peer effects: On the nature, estimation, and channels of peer effects." *Journal of Labor Economics*, 35(2): 387–428.

Feng, Han, and Jiayao Li. 2016. "Head teachers, peer effects, and student achievement." *China Economic Review*, 41: 268–283.

Golsteyn, Bart, Arjan Non, and Ulf Zölitz. 2019. "The impact of peer personality on academic achievement."

Gong, Jie, Yi Lu, and Hong Song. 2018. "The Effect of Teacher Gender on Students' Academic and Noncognitive Outcomes." *Journal of Labor Economics*, 36(3): 743–778.

Gong, Jie, Yi Lu, and Hong Song. 2019. "Gender peer effects on students' academic and noncognitive outcomes: Evidence and mechanisms." *Journal of Human Resources*, 0918–9736R2.

Griffith, Alan. 2018. "Random Assignment with Non-Random Peers: A Structural Approach to Counterfactual Treatment Assessment."

Heckman, James J, and Yona Rubinstein. 2001. "The importance of noncognitive skills: Lessons from the GED testing program." *American Economic Review*, 91(2): 145–149.

Heckman, James J, Jora Stixrud, and Sergio Urzua. 2006. "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." *Journal of Labor economics*, 24(3): 411–482.

Heckman, James J, Tomás Jagelka, Tim Kautz, et al. 2019. "Some Contributions of Economics to the Study of Personality." NBER *Working Paper 26459*.

Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes." *American Economic Review*, 103(6): 2052–86.

He, Leshui, and Stephen L Ross. 2017. "Classroom peer effects and teachers: Evidence from quasi-random assignment in a Chinese middle school." *Human Capital and Economic Opportunity Global Working Group Working Paper*, 14: 2017.

Hu, Feng. 2015. "Do girl peers improve your academic performance?" *Economics Letters*, 137: 54–58.

Hu, Feng. 2018. "Migrant peers in the classroom: Is the academic performance of local students negatively affected?" *Journal of Comparative Economics*, 46(2): 582–597.

Jackson, Joshua J, and Brent W Roberts. 2017. "Conscientiousness." In *The Oxford handbook of the five factor model*.

Jackson, Joshua J, Dustin Wood, Tim Bogg, Kate E Walton, Peter D Harms, and Brent W Roberts. 2010. "What do conscientious people do? Development and validation of the Behavioral Indicators of Conscientiousness (BIC)." *Journal of Research in Personality*, 44(4): 501–511.

Jackson, Matthew O. 2010. *Social and economic networks.* Princeton university press.

Kautz, Tim, James J Heckman, Ron Diris, Bas Ter Weel, and Lex Borghans. 2014. "Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success." National Bureau of Economic Research.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica*, 75(1): 83–119.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to learn." *The Review of Economics and Statistics*, 91(3): 437–456.

Laakasuo, Michael, Anna Rotkirch, Venla Berg, and Markus Jokela. 2017. "The company you keep: Personality and friendship characteristics." *Social Psychological and Personality Science*, 8(1): 66–73.

Lindqvist, Erik, and Roine Vestman. 2011. "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment." *American Economic Journal: Applied Economics*, 3(1): 101–28.

Lucas Jr, Robert E. 1988. "On the mechanics of economic development." *Journal of monetary economics*, 22(1): 3–42.

Lyle, David S. 2007. "Estimating and interpreting peer and role model effects from randomly assigned social groups at West Point." *The Review of Economics and Statistics*, 89(2): 289–299.

Manski, Charles F. 1993. "Identification of endogenous social effects: The reflection problem." *The Review of Economic Studies*, 60(3): 531–542.

Martin, Andrew J. 2009. "Motivation and engagement across the academic life span: A developmental construct validity study of elementary school, high school, and university/college students." *Educational and psychological measurement*, 69(5): 794–824.

Neidell, Matthew, and Jane Waldfogel. 2010. "Cognitive and noncognitive peer effects in early education." *The Review of Economics and Statistics*, 92(3): 562–576.

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt. 2019. "Poorly measured confounders are more useful on the left than on the right." *Journal of Business & Economic Statistics*, 37(2): 205–216.

Roberts, Brent W, Dustin Wood, and Avshalom Caspi. 2008. "The development of personality traits in adulthood." *Handbook of personality: Theory and research*, 3: 375–398.

Sacerdote, Bruce. 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*. Vol. 3, 249–277. Elsevier.

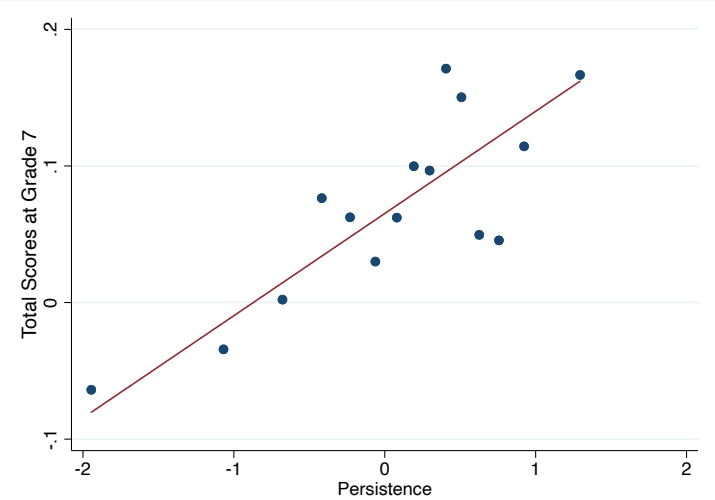
Selfhout, Maarten, William Burk, Susan Branje, Jaap Denissen, Marcel Van Aken, and Wim Meeus. 2010. "Emerging late adolescent friendship networks and Big Five personality traits: A social network approach." *Journal of personality*, 78(2): 509–538.

Shure, Nikki. 2017. "Non-cognitive peer effects in secondary education." Unpublished manuscript.

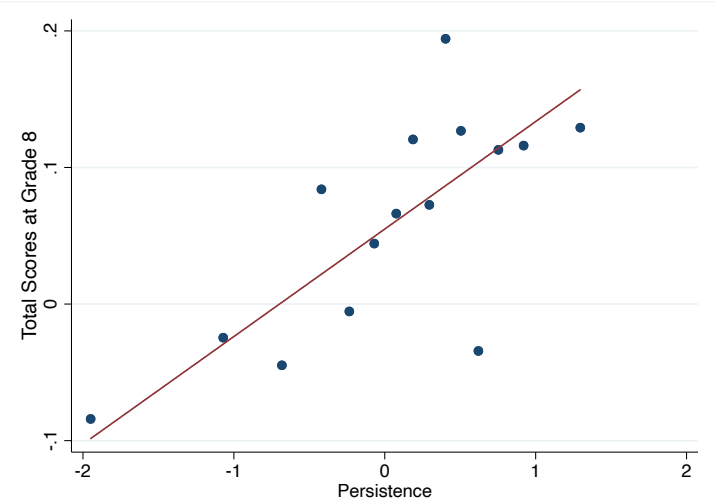
Wrzus, Cornelia, Julia Zimmermann, Marcus Mund, and Franz J Neyer. 2017. "Friendships in young and middle adulthood: Normative patterns and personality differences." *The psychology of friendship*, 21–38.

Xu, Di, Qing Zhang, and Xuehan Zhou. 2019. "The Impact of Low-Ability Peers on Cognitive and Non-Cognitive Outcomes: Random Assignment Evidence on the Effects and Operating Channels."

Figure 1: Correlation between Peer Personality on Academic Outcome



(A) Grade 7

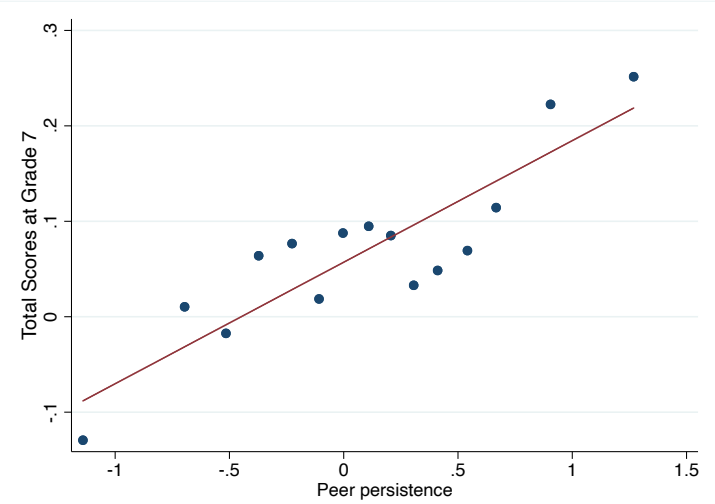


(B) Grade 8

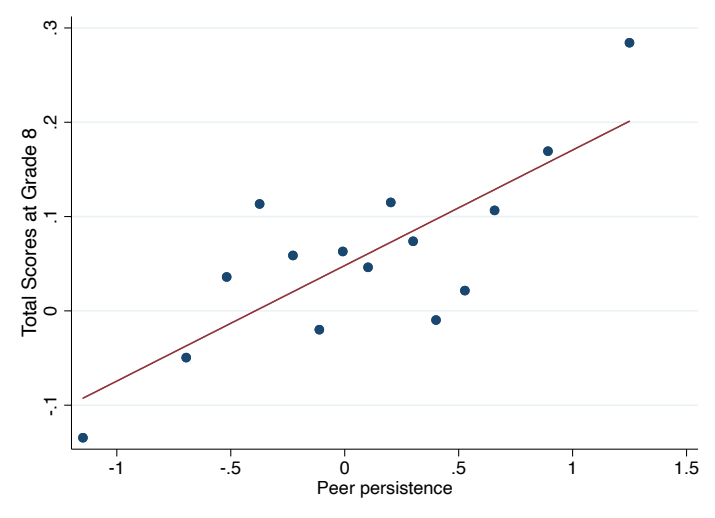
Note. Subfigure (A) and (B) visualize the regression results with the same set of controls in column (1) and (2) of Table 4, respectively. All regressions include student's age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as school-grade fixed effects.

Source. China Education Panel Survey 2013-14.

Figure 2: Impact of Peer Personality on Academic Outcome



(A) Grade 7



(B) Grade 8

Note. Subfigure (A) and (B) visualize the regression results with the same set of controls in column (2) and (4) of Table 5, respectively. All regressions include peer characteristic controls: classroom proportion of female, migrant, and percentile ranking at Grade 6, and student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as school-grade fixed effects.

Source. China Education Panel Survey 2013-14.

Table 1: Statistical Description

	N	Mean	SD	Min	Max
A. Outcome variables					
Academic outcomes:					
Total score	4,195	210.0	25.94	89.19	293.9
Total score (8th)	4,057	210.1	26.43	49.33	285.7
Noncognitive outcomes:					
Depressed	4,212	2.14	0.98	1	5
Blue	4,196	1.85	1.02	1	5
Unhappy	4,210	2.14	1.03	1	5
Pessimistic	4,193	1.66	1.03	1	5
Depressed (8th)	4,018	2.28	1.08	1	5
Blue (8th)	4,009	2.17	1.12	1	5
Unhappy (8th)	4,011	2.26	1.10	1	5
Pessimistic (8th)	4,006	1.90	1.11	1	5
B. Regressor of interests:					
Personality variables:					
Persistence at Grade 6	4,145	3.44	0.64	1	4
Peer variables:					
Peer persistence	4,897	3.30	0.24	2.48	3.90
C. Predetermined variables:					
Student age	4,228	13.48	0.67	12	17
Female student	4,321	0.49	0.50	0	1
Minority	4,313	0.10	0.28	0	1
Agricultural Hukou	4,321	0.43	0.5	0	1
Nonlocal residence	4,283	0.24	0.43	0	1
Only child in family	4,269	0.55	0.50	0	1
Attend kindergarten	4,284	0.85	0.36	0	1
Age attending primary school	4,274	6.66	1.00	0	9
Percentile ranking at Grade 6	3,685	0.66	0.24	0	1
Mother's education	4,311	10.19	3.52	0	18
Father's education	4,311	10.77	3.32	0	18
Extraversion at Grade 6	4,105	3.16	0.67	1	4
Openness at Grade 6	4,084	3.31	0.64	1	4

Note. Variables from second wave of China Education Panel Survey are indicated by "8th" in parentheses.
Source. China Education Panel Survey 2013-14 and 2014-15.

Table 2: Measurement of Student Personality Traits

Trait	Definition	Measurement
Persistence	1. "I would try my best to go to school even if I was not feeling very well or I had other reasons to stay at home." 2. "I would try my best to finish even the homework I dislike." 3. "I would try my best to finish my homework, even if it would take me quite a long time."	All questions are measured on a scale of 1-4: 1 for strongly disagree 2 for somewhat disagree 3 for somewhat agree 4 for strongly agree
Extraversion	1. "I was able to express myself clearly." 2. "I was able to give quick responses."	
Openness	1. "I was a fast learner." 2. "I was a curious about new stuff."	

Note. CEPS asked students at Grade 7 about their perception on own personality traits when they were students at Grade 6, by asking them "How much do you agree with each of the following statements about your experiences in GRADE 6? " While the first three items measure students' persistence, a lower-order of Conscientiousness, the second and third two items are measurements related to Extraversion and Openness to Experience in the Big Five model, respectively. The measurement of persistence are similar to items used in Golsteyn et al. (2019), which is originally from the Student Motivation Scale of Martin (2009).

Source. China Education Panel Survey 2013-14.

Table 3: Balancing Test for Random Assignment

	Dependent variable: Peers' Personality	
	(1) Peer Persistence	(2) Peer Persistence
Student age	0.012	−0.089*
	(0.020)	(0.045)
Female student	−0.003	0.001
	(0.016)	(0.032)
Minority	0.030	−0.353
	(0.038)	(0.217)
Agricultural Hukou	−0.042	0.123
	(0.039)	(0.123)
Non-local residence	0.002	−0.207*
	(0.028)	(0.115)
Only child in family	0.012	0.095
	(0.044)	(0.091)
Attend kindergarten	−0.018	0.019
	(0.050)	(0.082)
Age attending primary school	0.016	0.116**
	(0.018)	(0.045)
Percentile ranking at Grade 6	−0.025	0.054
	(0.095)	(0.128)
Mother's education	−0.005	0.003
	(0.005)	(0.011)
Father's education	0.005	−0.014
	(0.004)	(0.011)
Extraversion at Grade 6	0.031	0.148***
	(0.022)	(0.041)
Openness at Grade 6	−0.016	−0.064*
	(0.020)	(0.034)
Persistence at Grade 6	0.009	0.106***
	(0.022)	(0.034)
School-Grade FE	yes	no
Joint test F-statistic	1.54	3.67
Joint test p-value	0.136	0.000
Adjusted R-squared	0.598	0.074
Observations	3,292	3,292

Note. The dependent variable is leave-me-out average of peers' persistence. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as corresponding school-grade fixed effects. Robust standard errors clustered at the school-grade level are in parentheses. Joint test F-statistics and p-value are for joint testings of all predetermined variables. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.

Table 4: Correlation between Persistence and Academic Outcome

	Dependent variable: Standardized total test score	
	(1)	(2)
	Total score at Grade 7	Total score at Grade 8
Std. Persistence at Grade 6	0.075***	0.079***
	(0.018)	(0.016)
School-Grade fixed effects	yes	yes
Control variables	yes	yes
Adjusted R-squared	0.358	0.370
Observations	3,224	3,252

Note. The dependent variables are standardized total scores at Grade 7 and 8, while the independent variable, student's own persistence, is standardized over the estimation sample. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as corresponding school-grade fixed effects. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.

Table 5: The Impacts of Peer Persistence on Academic Outcome

	Dependent variable: Standardized total test score			
	(1) Total at Grade 7	(2) Total at Grade 7	(3) Total at Grade 8	(4) Total at Grade 8
Std. Peer Persistence	0.132***	0.127***	0.127***	0.122***
	(0.041)	(0.039)	(0.039)	(0.037)
Std. Persistence at Grade 6	0.073***	0.073***	0.078***	0.077***
	(0.017)	(0.018)	(0.016)	(0.016)
School-Grade FE	yes	yes	yes	yes
Control variables	yes	yes	yes	yes
Controlling peer characteristics	no	yes	no	yes
Adjusted R-squared	0.366	0.368	0.377	0.377
Observations	3,224	3,224	3,252	3,252

Note. The dependent variables are standardized total scores at Grade 7 and 8, while the independent variable, peer persistence, is standardized over the estimation sample. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as corresponding school-grade fixed effects. Peer characteristic controls include classroom proportion of female, migrant, and percentile ranking at Grade 6. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.

Table 6: The Impacts of Peer Persistence on Student's Mental Stress

	Panel A: Mental Stress at 7th Grade					
	Depressed (1)	Blue (2)	Unhappy (3)	Pessimistic (4)	Average (5)	AES (6)
Std. peer persistence	0.022	0.020	0.028	0.061*	0.040	0.033
	(0.028)	(0.028)	(0.030)	(0.034)	(0.032)	(0.026)
Std. persistence at Grade 6	-0.066**	-0.071***	-0.054**	-0.141***	-0.102***	-0.083***
	(0.025)	(0.026)	(0.020)	(0.026)	(0.026)	(0.021)
Adjusted R-squared	0.050	0.048	0.049	0.068	0.069	-
Observations	3,226	3,226	3,226	3,226	3,226	3,226
	Panel B: Mental Stress at 8th Grade					
	Depressed (7)	Blue (8)	Unhappy (9)	Pessimistic (10)	Average (11)	AES (12)
Std. peer persistence	0.062**	0.054*	0.036	0.061**	0.061*	0.053*
	(0.029)	(0.030)	(0.040)	(0.030)	(0.033)	(0.029)
Std. persistence at Grade 6	-0.037*	-0.054**	-0.013	-0.067***	-0.049**	-0.043**
	(0.020)	(0.023)	(0.024)	(0.023)	(0.023)	(0.020)
Adjusted R-squared	0.042	0.053	0.044	0.047	0.054	-
Observations	3,265	3,265	3,265	3,265	3,265	3,265
School-Grade FE	yes	yes	yes	yes	yes	yes
Control variables	yes	yes	yes	yes	yes	yes
Controlling peer characteristics	yes	yes	yes	yes	yes	yes

Note. The dependent variables are items on mental stress measurement at Grade 7 and 8, which are from the questions "Do you have the feelings below in the last seven days?" with a scale from 1 (never) to 5 (always): (1) depressed, (2) blue, (3) unhappy, or (4) that life is meaningless. Average is an index averaging the four mental stress items, while AES = average effect size from Kling et al. (2007). All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as corresponding school-grade fixed effects. Peer characteristic controls include classroom proportion of female, migrant, and percentile ranking at Grade 6. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14 and 2014-15.

Table 7: Effect of Peer Personality on Academic Outcome: Heterogeneous Effects

	Dependent variable: Standardized test score			
	Total	Chinese	Math	English
	(1)	(2)	(3)	(4)
Low persistence * Std. peer persistence	0.100	0.082	0.122	0.046
	(0.112)	(0.091)	(0.115)	(0.089)
Medium persistence * Std. peer persistence	0.060	-0.001	0.079	0.077
	(0.058)	(0.061)	(0.052)	(0.070)
High persistence * Std. peer persistence	0.143***	0.094**	0.152***	0.138***
	(0.041)	(0.040)	(0.045)	(0.046)
School-Grade FE	yes	yes	yes	yes
Control Variables	yes	yes	yes	yes
Controlling peer characteristics	yes	yes	yes	yes
Adjusted R-squared	0.382	0.319	0.261	0.337
Observations	3,224	3,226	3,228	3,227

Note. The dependent variables are standardized total scores and test scores from three main subjects – Chinese, Math, and English – at Grade 7. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as corresponding school-grade fixed effects. Peer characteristic controls include classroom proportion of female, migrant, and percentile ranking at Grade 6. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.

Table 8: Mechanism I: Direct Persistence Spillover

	Dependent variable: Persistence at Grade 7				
	School Attendance	Disliked Homework	Challenging Homework	Average	AES
	(1)	(2)	(3)	(4)	(5)
Std. peer persistence	0.019	0.020	0.021	0.023	0.020
	(0.028)	(0.027)	(0.031)	(0.029)	(0.024)
Std. persistence at Grade 6	0.159***	0.194***	0.201***	0.216***	-
	(0.025)	(0.023)	(0.019)	(0.023)	-
School-Grade FE	yes	yes	yes	yes	yes
Control variables	yes	yes	yes	yes	yes
Controlling peer characteristics	yes	yes	yes	yes	yes
Adjusted R-squared	0.049	0.088	0.089	0.097	-
Observations	3,286	3,286	3,286	3,286	3,286

Note. The dependent variables are items on persistence measurement at Grade 7. Average is an index averaging the three persistence items, while AES = average effect size. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as corresponding school-grade fixed effects. Peer characteristic controls include classroom proportion of female, migrant, and percentile ranking at Grade 6. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14 and 2014-15.

Table 9: Mechanism II: Student Response

	Panel A: Time Use on Entertainment			
	(1) Watching TV weekday	(2) Watching TV weekend	(3) Playing game weekday	(4) Playing game weekend
Std. peer persistence	-0.118**	-0.090*	-0.108**	-0.074
	(0.046)	(0.050)	(0.045)	(0.073)
Std. persistence at Grade 6	-0.147***	-0.129***	-0.131***	-0.164***
	(0.038)	(0.038)	(0.030)	(0.045)
Adjusted R-squared	0.125	0.111	0.086	0.116
Observations	3,217	3,236	3,225	3,240
	Panel B: School Attendance		**Panel C: Education Expectation**	
	(5) Tardy: "I am always late for class"	(6) Truancy: "I always skip classes."	(7) Years of schooling	(8) Going to college
Std. peer persistence	-0.039**	-0.026*	0.297	0.049***
	(0.017)	(0.013)	(0.187)	(0.018)
Std. persistence at Grade 6	-0.041***	-0.006	0.008	0.005
	(0.014)	(0.010)	(0.070)	(0.009)
Adjusted R-squared	0.064	0.060	0.264	0.212
Observations	3,286	3,285	3,198	3,198
School-Grade FE	yes	yes	yes	yes
Control variables	yes	yes	yes	yes
Controlling peer characteristics	yes	yes	yes	yes

Note. The dependent variables in Panel A are hours per day used on entertainment at weekday/weekend last week, while dependent variables in Panel B are dummy variables indicating student's self-reported school attendance behaviors. The dependent variables in Panel C is student's education expectation. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as corresponding school-grade fixed effects. Peer characteristic controls include classroom proportion of female, migrant, and percentile ranking at Grade 6. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.

Table 10: Mechanism III: Friendship Formation

	Panel A: "Good" peer			Panel B: "Bad" peer		
"Any of your five best friends fit in the following descriptions: "	Hard Working	Aspiration to College	Excellent Grade	Smoking or Drinking	Going to Net Cafe Often	In Love Relationship
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Linear-in-Mean Model						
Std. peer persistence	0.015**	0.010	0.009	-0.013**	-0.017**	-0.016**
	(0.006)	(0.006)	(0.008)	(0.006)	(0.007)	(0.008)
Std. persistence at Grade 6	0.018***	0.006	0.013**	-0.019***	-0.035***	-0.047***
	(0.006)	(0.004)	(0.005)	(0.006)	(0.007)	(0.009)
R-squared	0.074	0.051	0.074	0.059	0.121	0.083
Observations	3,249	3,236	3,247	3,244	3,245	3,243
Panel B: Heterogeneous Effects	(7)	(8)	(9)	(10)	(11)	(12)
Low persistence * Std. peer persistence	0.025	-0.000	0.006	-0.016	-0.014	-0.015
	(0.027)	(0.028)	(0.033)	(0.033)	(0.037)	(0.038)
Medium persistence * Std. peer persistence	0.007	0.009	0.008	-0.009	-0.018	-0.014
	(0.014)	(0.015)	(0.017)	(0.015)	(0.014)	(0.016)
High persistence * Std. peer persistence	0.016***	0.010*	0.010	-0.014***	-0.017**	-0.016*
	(0.006)	(0.005)	(0.007)	(0.005)	(0.007)	(0.009)
R-squared	0.074	0.053	0.074	0.059	0.123	0.083
Observations	3,249	3,236	3,247	3,244	3,245	3,243
School-Grade FE	yes	yes	yes	yes	yes	yes
Control Variables	yes	yes	yes	yes	yes	yes
Controlling peer characteristics	yes	yes	yes	yes	yes	yes

Note. The dependent variables are dummy variables indicating whether the student has any friend among her five best friends fit in those descriptions. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as corresponding school-grade fixed effects. Peer characteristic controls include classroom proportion of female, migrant, and percentile ranking at Grade 6. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.

Appendix to: "The Peer Effects of Personality on Students' Academic and Non-cognitive Outcomes"

Jian Zou

November 2019

Table A1: Balancing tests

	Age	Female	Minority	Agricultural Hukou	Nonlocal residence	Only child in family	Attend kindergarten
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Peer Persistence	0.010	-0.002	0.004	-0.015	0.001	0.005	-0.005
	(0.016)	(0.010)	(0.005)	(0.014)	(0.009)	(0.020)	(0.014)
	Age attending primary school	Percentile ranking	Mother's education	Father's education	Extraversion at Grade 6	Openness at Grade 6	Persistence at Grade 6
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Peer Persistence	0.029	-0.003	-0.065	0.063	0.043	-0.023	0.016
	(0.029)	(0.012)	(0.072)	(0.057)	(0.026)	(0.025)	(0.035)
School-Grade FE	yes	yes	yes	yes	yes	yes	yes
Control Variables (expect the dep. var.)	yes	yes	yes	yes	yes	yes	yes

Note. The coefficient on peer persistence in regressions with individual controls as dependent variable is displayed. All regressions follow the same set of student's predetermined variables in Table 3 expect the dependent variable, as well as school-grade fixed effects. Robust standard errors clustered at school-grade level are in the parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.

Table A2: Mechanism: Time Use on Study

	Dependent Variable: Study Hour			
	Homework at weekday	Homework at weekend	Tutoring HW at weekday	Tutoring HW at weekend
	(1)	(2)	(3)	(4)
Std. peer persistence	-0.048	-0.019	-0.008	0.035
	(0.089)	(0.072)	(0.048)	(0.038)
Std. persistence at Grade 6	0.0131	-0.000	-0.052	0.049*
	(0.051)	(0.057)	(0.040)	(0.027)
School-Grade FE	yes	yes	yes	yes
Control Variables	yes	yes	yes	yes
Controlling peer characteristics	yes	yes	yes	yes
R-squared	0.077	0.109	0.055	0.081
Observations	3,228	3,246	3,234	3,247

Note. The dependent variables are hours per day used on homework at weekday/weekend last week. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as school-grade fixed effects. Peer characteristic controls include classroom proportion of female, migrant, and percentile ranking at Grade 6. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.

Table A3: Model: Complements of Rank and Persistence

	Dependent variable: Standardized total score at Grade 7				
	(1)	(2)	(3)	(4)	(5)
	0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1
Std. Peer Persistence	−0.088	0.082	0.114	0.174***	0.107**
	(0.198)	(0.072)	(0.070)	(0.053)	(0.040)
Std. Persistence at Grade 6	0.033	0.080	0.059	0.062**	0.100***
	(0.142)	(0.054)	(0.041)	(0.028)	(0.034)
School-Grade FE	yes	yes	yes	yes	yes
Control variables	yes	yes	yes	yes	yes
Controlling peer characteristics	yes	yes	yes	yes	yes
Adjusted R-squared	0.320	0.133	0.152	0.171	0.139
Observations	124	301	631	941	1,176

*Note.*The dependent variables are standardized toal test scores at Grade 7. Sample are divided by percentile ranking $\in [0, 1]$ at Grade 6 in each column. All regressions include student's predetermined variables: age, gender, ethnicity, rural status, local residency status, age attending primary school, dummy variables indicating only-child family, attended kindergarten, and academic ranking and personality measures at Grade 6, as well as school-grade fixed effects. Peer characteristic controls include classroom proportion of female, migrant, and percentile ranking at Grade 6. Robust standard errors clustered at the school-grade level are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

Source. China Education Panel Survey 2013-14.