# School Restrictions, Migration, and Peer Effects: A Spatial Equilibrium Analysis of Children's Human Capital in China
# Online Appendix

Zibin Huang* Junsen Zhang†

October 17, 2023

## Contents



*College of Business, Shanghai University of Finance and Economics; Shanghai Institute of International Finance and Economics. Email: huangzibin@mail.shufe.edu.cn

†School of Economics, Zhejiang University. Email: jszhang@cuhk.edu.hk

## A Supplements to the Reduced Form Analysis

### A.1 Robustness Checks for the Peer Effect Regressions

We now check the robustness of our empirical results. Three regressions are run in each specification: the regression using the sample from both years, the regression using the sample from the first year, and the regression using the sample from the second year. We use only those observed in both years.

In Table A1, we change the student performance measure from cognitive test scores to school-level standardized Chinese, Math, and English test scores, which are used in Hu (2018) and Wang, Cheng, and Smyth (2018). The first column of each measurement is a regression with observations from both years. The second column of each measurement is a regression with only observations from the first year, which corresponds to Column (1) in Table 4. The third column of each measurement is a regression with only observations from the second year, which corresponds to Column (2) in Table 4. Our main conclusions still hold. Migrant students negatively affect their classmates' Chinese scores but not significantly affect Math or English scores. However, left-behind students harm their classmates across all three subjects and the magnitudes are much larger. Moreover, we can also detect reductions in the point estimates in the second year. In Table A2, we change the independent variables to be the proportion of rural migrant peers and the proportion of rural left-behind peers in the class. The results are similar to the main regression with less precision. In Table A3, we change the definition of left-behind children. In the main setting, we define left-behind children as children with at least absent parent. In this table, we change it to children with both parents absent. The results show that the negative effects of left-behind peers become larger, which makes sense because these children are usually in more disadvantaged positions than children with at least one parent at home.

Furthermore, we detect a small number of changing Hukou locations during the two waves

(265 out of 10,443 observations). These could be real Hukou changes or measurement errors. It is possible that the shrinking peer effects of migrant students in the second year is because some migrant students change their Hukou and become locals in the second wave. To make sure that the time pattern of migrant students' spillovers is not driven by this, we redefine all students' Hukou registration as the ones in the first year. That is, even if a student changes his/her Hukou in the second year, we still consider them a migrant student since, after all, he or she has migrated from somewhere else. The results are shown in Table A4. Column (2) in Table A4 is exactly Column (1) in Table 4 since migration status is fixed at the first year value. In Column (3) we see that, after redefining, we still detect almost zero peer effects from migrant students in the second year. In addition to the Hukou status change, left-behind status could also change across time since parents may migrate out or return home in the second year. Since left-behind status is not a fixed characteristic, it is more reasonable to make this left-behind status time-varying as in our main context. However, it could also be interesting to see whether the results would change if we fix the identity of both migrant and left-behind students based on year one and estimate the same regressions. The results are shown in Table A5, which shows the same pattern: the negative spillovers of left-behind children persist into the second year and the peer effects from the migrant students fall to zero.

In the main context, we find that the negative spillovers from migrant students vanish in the second year. It is possible that parents of students in classes with many migrant peers may be aware of the negative spillovers after the first year and then devote more money and time to their children to compensate for these spillovers. To test whether this is true, we take parents' time spent on children and educational expenditure on out-of-school courses as dependent variables and regress each on the proportions of migrant and left-behind peers in their children's class. Table A6 shows that there is no significant effect of migrant and left-behind peers on parents' investment. In some specifications, the correlation is even negative.

Tables A7 and A8 show the results when we alter the regression sample. The CEPS uses three different sampling frames. In the national core frame, they randomly choose 15 counties from the 2,870 counties in China. In the Shanghai frame, they choose 3 out of the 18 counties in Shanghai in addition to the national core sample. In the complementary frame, they choose 10 out of 120 counties with substantial immigration. In the previous main regressions, we use survey sampling weights to correct for over-sampling. In Table A7, we exclude all data from the Shanghai frame and the complementary frame, keeping only the national core sample. In Table A8, we drop all data from the Shanghai frame. The results are consistent and robust.

Another concern is that the results in private schools may be very different from the results in public schools and that the negative effects are driven by the students in private migrant schools. To alleviate this problem, we run regressions on the sample of public school students in Table A9, which shows no change in the estimates.

Angrist (2014) raises questions about the interpretation of the peer effect coefficient in the traditional linear-in-mean model. He claims that the estimate of the peer effect ($y$ on $\bar{x}$) is equivalent to the difference (or ratio) between the OLS coefficient of $y$ on individual covariate $x$ and the 2SLS coefficient of $y$ on individual covariate $x$ using group dummies $z$ as instruments. However, the discrepancy between them can be attributed to more than peer effects. He suggests *"mak[ing] a clear separation between the subject of a peer effects investigation and the peers who provide the mechanism for causal effects on these subjects."* Thus, in our case, we run the regressions only on ordinary local students who are neither migrant nor left-behind students to test whether the results are robust.[1] Table A10 shows that none of the qualitative results change.

Another specific concern in Angrist (2014) is that the regression coefficient may be a result of measurement error in the migrant/left-behind status of students rather than the real peer effects. We implement a simulation inspired by Carrell, Hoekstra, and Kuka (2018) and Feld and Zölitz (2017) to further test whether this is the case in our data. The basic idea of the

[1]The regressions are similar to the ones in Carrell, Hoekstra, and Kuka (2018).

simulation is to add measurement error to the sample and run the regressions again. If there is attenuation in the estimates rather than amplification, we can rule out this concern. The detailed algorithm is as follows. (1) First, randomly select $p$% of the sample. (2) Second, in the selected sample, randomly assign 22% of observations to be migrant students and 15% to be left-behind students (the proportions of migrant and left-behind students in this random assignment match the proportions in the original whole sample). These observations are now measured with random error. (3) Third, re-calculate the proportions of migrant and left-behind peers in each class using the whole sample. (4) Fourth, run the main regression using the whole sample. We repeat this process while varying $p$ from 0% to 100%, where 0% means the baseline estimates without any added measurement error and 100% means the extreme case when all observations are measured with error. Figure A1 shows the results. For the estimation of left-behind children's peer effects, it is evident that when additional measurement error is introduced the estimate of the peer effects is attenuated towards zero. For the estimation of migrant children's peer effects, the estimates remain noisy with more measurement error. In general, we do not find any evidence suggesting that the estimated negative peer effects are driven by measurement error.

Table A1: Robustness: Using School-Level Performance Measurement

| | Chinese | | | Math | | | English | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) Both Years | (2) First Year | (3) Second Year | (4) Both Years | (5) First Year | (6) Second Year | (7) Both Years | (8) First Year | (9) Second Year |
| Proportion of Migrant Peers | -4.122** | -8.334* | -2.618 | 1.313 | -1.594 | 1.330 | -3.217 | -6.585 | -3.996 |
| | (1.800) | (4.425) | (2.155) | (5.639) | (7.002) | (6.968) | (2.641) | (5.069) | (3.713) |
| Proportion of Left-Behind Peers | -9.052*** | -15.62* | -9.029*** | -13.43*** | -27.85** | -9.957 | -13.59*** | -23.93** | -13.39*** |
| | (3.238) | (8.147) | (2.628) | (4.541) | (11.71) | (6.015) | (3.620) | (10.01) | (3.160) |
| School FE | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| Year Dummy | YES | NO | NO | YES | NO | NO | YES | NO | NO |
| Personal Controls | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| Household Controls | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| Teacher Controls | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 7,200 | 3,600 | 3,600 | 7,200 | 3,600 | 3,600 | 7,200 | 3,600 | 3,600 |
| R-squared | 0.283 | 0.298 | 0.278 | 0.236 | 0.240 | 0.244 | 0.300 | 0.300 | 0.309 |

Notes: The dependent variables for all regressions are the school-level standardized test scores. Three different school-level performance measurements are considered. The first three columns are test scores of Chinese. The next three columns are test scores of Math. The last three columns are test scores of English. For columns (1), (4) and (7), we use data from both years. For columns (2), (5) and (8), we use data from only the first year. For columns (3), (6) and (9), we use data from only the second year. The set of personal controls includes student age, sex, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes parental education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A2: Robustness: Only Rural Migrant or Left-Behind Children

| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Rural Migrant Peers | 0.0296 | -0.444 | 0.544 |
| | (0.375) | (0.391) | (0.559) |
| Proportion of Rural Left-Behind Peers | -0.722*** | -0.734* | -0.742*** |
| | (0.268) | (0.395) | (0.195) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | YES | YES | YES |
| Teacher Controls | YES | YES | YES |
| Observations | 7,200 | 3,600 | 3,600 |
| R-squared | 0.383 | 0.389 | 0.404 |

Notes: The dependent variable for all regressions is the standardized test score. In this table, we consider the proportion of migrant and left-behind peers with rural Hukou. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A3: Robustness: Left-Behind Children with Both Parents Absent

| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Migrant Peers | -0.0271 | -0.345 | 0.179 |
| | (0.239) | (0.232) | (0.372) |
| Proportion of Left-Behind Peers | -1.127** | -1.854** | -0.710** |
| | (0.461) | (0.706) | (0.280) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | YES | YES | YES |
| Observations | 7,200 | 3,600 | 3,600 |
| R-squared | 0.385 | 0.393 | 0.402 |

Notes: The dependent variable for all regressions is the standardized test score. In this table, we alter the definition of left-behind children to children whose parents are both absent. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A4: Robustness: Fixing Hukou Status for All Students

| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Migrant Peers | -0.127 | -0.507* | 0.123 |
| | (0.221) | (0.303) | (0.327) |
| Proportion of Left-Behind Peers | -0.884*** | -1.114* | -0.709*** |
| | (0.284) | (0.611) | (0.204) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | NO | YES | YES |
| Teacher Controls | YES | YES | YES |
| Observations | 7,200 | 3,600 | 3,600 |
| R-squared | 0.327 | 0.342 | 0.333 |

Notes: The dependent variable for all regressions is the standardized test score. In this table, we fix the Hukou status for all students to their first year Hukou. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A5: Robustness: Fixing Hukou and Left-behind Status for All Students

| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Migrant Peers | -0.299 | -0.507* | -0.0981 |
| | (0.210) | (0.303) | (0.259) |
| Proportion of Left-Behind Peers | -1.290** | -1.114* | -1.348** |
| | (0.497) | (0.611) | (0.526) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | NO | YES | YES |
| Observations | 7,200 | 3,600 | 3,600 |
| R-squared | 0.388 | 0.390 | 0.411 |

Notes: The dependent variable for all regressions is the standardized test score. In this table, we fix the Hukou status and left-behind status for all students to their first year. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A6: Robustness: Parents' Investment

| Dependent Variable | Time Spend on Children | | Education Expenditure | |
|---|---|---|---|---|
| | (1) First Years | (2) Second Year | (3) First Year | (4) Second Year |
| Proportion of Migrant Peers | -2.470 | -4.310 | -1,075* | 185.7 |
| | (1.575) | (3.542) | (604.4) | (398.0) |
| Proportion of Left-Behind Peers | 0.549 | 0.997 | -782.0 | -241.7 |
| | (2.489) | (0.884) | (1,212) | (288.0) |
| School FE | YES | YES | YES | YES |
| Year Dummy | NO | NO | NO | NO |
| Personal Controls | YES | YES | YES | YES |
| Household Controls | YES | YES | YES | YES |
| Teacher Controls | YES | YES | YES | YES |
| Observations | 2,990 | 2,990 | 2,990 | 2,990 |
| R-squared | 0.082 | 0.066 | 0.195 | 0.238 |

Notes: The dependent variables are parents' time and educational expenditures on children. For columns (1) and (3), we use data from only the first year. For columns (2) and (4), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A7: Robustness: Using Only National Core Sample

| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Migrant Peers | -0.152 | -0.591* | 0.111 |
| | (0.254) | (0.336) | (0.385) |
| Proportion of Left-Behind Peers | -0.928*** | -1.242* | -0.688*** |
| | (0.306) | (0.668) | (0.208) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | YES | YES | YES |
| Teacher Controls | YES | YES | YES |
| Observations | 3,380 | 1,690 | 1,690 |
| R-squared | 0.384 | 0.394 | 0.397 |

Notes: The dependent variable for all regressions is the standardized test score. In this table, we use only the national core sample. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A8: Robustness: Dropping Shanghai Sample

| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Migrant Peers | -0.123 | -0.518 | 0.136 |
| | (0.249) | (0.313) | (0.369) |
| Proportion of Left-Behind Peers | -0.899*** | -1.134* | -0.703*** |
| | (0.301) | (0.630) | (0.210) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | YES | YES | YES |
| Teacher Controls | YES | YES | YES |
| Observations | 6,394 | 3,197 | 3,197 |
| R-squared | 0.381 | 0.387 | 0.399 |

Notes: The dependent variable for all regressions is the standardized test score. In this table, we drop all observations from the Shanghai sample. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A9: Robustness: Only Public Schools

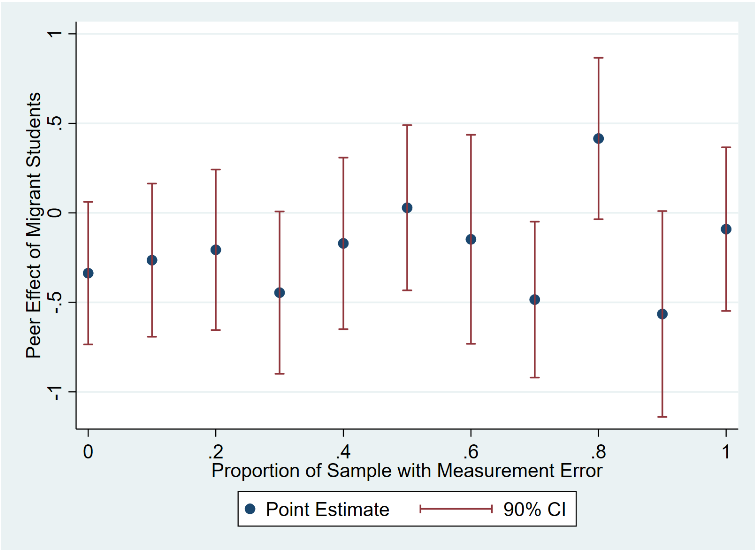
| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Migrant Peers | 0.367 | -0.585 | 1.018* |
| | (0.471) | (0.734) | (0.573) |
| Proportion of Left-Behind Peers | -0.776** | -1.107 | -0.736*** |
| | (0.307) | (0.704) | (0.182) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | NO | YES | YES |
| Teacher Controls | YES | YES | YES |
| Observations | 6,634 | 3,317 | 3,317 |
| R-squared | 0.385 | 0.381 | 0.412 |

Notes: The dependent variable for all regressions is the standardized test score. In this table, we drop all observations from the Shanghai sample. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

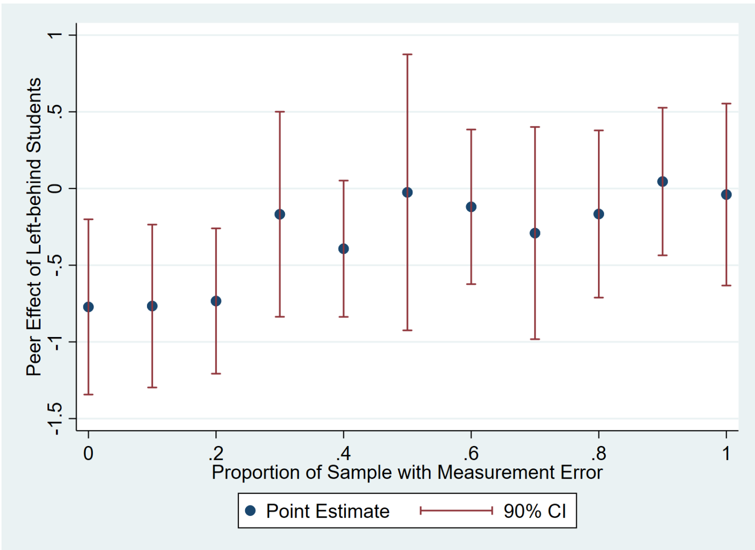
Table A10: Robustness: Only on Ordinary Local Students

| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Migrant Peers | -0.490 | -0.362 | 0.00111 |
| | (0.539) | (0.792) | (0.573) |
| Proportion of Left-Behind Peers | -1.572*** | -1.075 | -0.684*** |
| | (0.305) | (0.838) | (0.165) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | YES | YES | YES |
| Teacher Controls | YES | YES | YES |
| Observations | 4,732 | 2,366 | 2,366 |
| R-squared | 0.378 | 0.386 | 0.418 |

Notes: The dependent variable for all regressions is the standardized test score. In this table, we keep only local students who are neither migrant nor left-behind students. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.




(a) Peer Effects of Migrant Students




(b) Peer Effects of Left-behind Students

Figure A1: Main Regression Coefficients with Simulated Samples

Notes: This figure shows the estimated peer effects of migrant and left-behind peers when we add measurement errors. The x-axis represents the proportion of the sample assigned measurement errors. The blue dots represent the point estimates. The red interval represents the 90% confidence intervals of the peer effect point estimates. Subfigure (a) shows the peer effect for migrant students. Subfigure (b) shows the peer effect for left-behind students.

## A.2 Extending the Balance Check

In this subsection, we implement more balance checks to validate our regression assumption, that is, the assignment of students to different classrooms in the same school is random.

First, in the main context, we implement the balance check by estimating regressions including all covariates. Now we run balance checks separately for each variable rather than combining them in the same regression. The results are shown in Table A11. Even if we scrutinize for any potential non-randomness by checking the covariates one by one, there is still no evidence that randomness is violated. After controlling for school fixed effects, no coefficient is significant either statistically or economically. The joint F-test cannot reject the hypothesis that the covariates are jointly uncorrelated with the proportions of migrant or left-behind peers in the class.

Second, we run the regression separately for the 2013 and 2014 cohorts. Tables A12 and A13 show that in neither wave is there evidence of violation of randomness after we control for school fixed effects.

Third, we investigate the balance check results when we change the dependent variable to the average ability of students in the class. We proxy a child's ability by their sixth grade class ranking. A higher value means a lower ranking. Table A14 shows that all sorting patterns disappear once school fixed effects are included.

Table A11: Balance Check in One-by-One Regressions

| | Proportion of Migrants | | Proportion of Left-Behinds | |
|---|---|---|---|---|
| | Without School FE | With School FE | Without School FE | With School FE |
| Age | -0.0147 | 0.00146 | 0.0520*** | 0.00277 |
| | (0.0115) | (0.00115) | (0.0135) | (0.00349) |
| Sex | 0.0128** | 0.00195 | 0.00874 | 0.000540 |
| | (0.00509) | (0.00140) | (0.00677) | (0.00132) |
| Board | -0.0574 | -0.0111 | 0.126*** | 0.00225 |
| | (0.0360) | (0.00822) | (0.0374) | (0.00255) |
| Hukou Type | -0.0288* | 0.000325 | 0.0851*** | 0.00599 |
| | (0.0148) | (0.00361) | (0.0198) | (0.00733) |
| Whether Migrant Student | 0.258*** | 0.0247 | -0.0659*** | -0.00538 |
| | (0.0507) | (0.0192) | (0.0183) | (0.00390) |
| Whether Left-behind Student | -0.0494*** | -0.00383 | -0.140*** | 0.00536 |
| | (0.0130) | (0.00285) | (0.0204) | (0.00636) |
| Only Child | -0.00381 | 0.00233 | 0.121*** | 0.00510 |
| | (0.0264) | (0.00168) | (0.0204) | (0.00526) |
| Father's Education Years | 0.00263 | -0.000586 | -0.0149*** | -0.00113 |
| | (0.00303) | (0.000667) | (0.00288) | (0.00102) |
| Mother's Education Years | 0.00300 | -0.000314 | -0.0162*** | -0.000890 |
| | (0.00281) | (0.000546) | (0.00228) | (0.000990) |
| Whether Parents Have Conflicts | 0.00336 | -0.000353 | 0.0261** | -0.00108 |
| | (0.0114) | (0.00153) | (0.0107) | (0.00353) |
| Sixth Year Ranking | -0.000633** | 0.0000956 | 0.000929*** | 0.0000219 |
| | (0.000307) | (0.0000711) | (0.000350) | (0.0000866) |
| Teacher Has College Degree | 0.0353 | 0.0151 | -0.0179 | 0.0105 |
| | (0.0260) | (0.0164) | (0.0339) | (0.0155) |
| Teacher Sex | -0.0166 | 0.00593 | 0.0746** | 0.0213 |
| | (0.0252) | (0.0103) | (0.0362) | (0.0162) |
| School Fixed Effect | NO | YES | NO | YES |
| Year Fixed Effect | YES | YES | YES | YES |
| Observations | 10,443 | 10,443 | 10,443 | 10,443 |
| Joint F>Prob | 0.000 | 1.000 | 0.000 | 1.000 |

Notes: In the first two columns, we run regressions for the proportion of migrant children on different variables with/without controlling for school fixed effects. In the third and the fourth columns, we run regressions for the proportion of left-behind children on different variables with/without controlling for school fixed effects. All regressions are run separately for each variable. The coefficients in this table can be interpreted as the correlation between the independent variable and the composition of children in the class. The joint F-test is calculated using a SUR system. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A12: Balance Check in 2013

| | Proportion of Migrants | | Proportion of Left-Behinds | |
|---|---|---|---|---|
| | Without School FE | With School FE | Without School FE | With School FE |
| Age | -0.00189 | 0.000548 | 0.0164* | 0.000588 |
| | (0.00690) | (0.00106) | (0.00879) | (0.00132) |
| Sex | 0.0129** | 0.000959 | 0.00881 | -0.000421 |
| | (0.00533) | (0.00113) | (0.00540) | (0.00111) |
| Board | -0.0241 | -0.00567 | 0.0808** | 0.00164 |
| | (0.0267) | (0.00618) | (0.0308) | (0.00188) |
| Hukou Type | -0.0240** | -0.00144 | 0.0135 | 0.00353 |
| | (0.0104) | (0.00336) | (0.00916) | (0.00381) |
| Whether Migrant Student | 0.245*** | -0.150 | -0.0551*** | -0.00854* |
| | (0.0464) | (0.136) | (0.0199) | (0.00508) |
| Whether Left-behind Student | -0.0213** | -0.00382 | 0.0860*** | 0.00268 |
| | (0.00985) | (0.00276) | (0.0171) | (0.00653) |
| Only Child | 0.0131 | -0.00100 | 0.0724*** | 0.000766 |
| | (0.0168) | (0.00280) | (0.0125) | (0.00154) |
| Father's Education Years | -0.000522 | -0.000274 | 0.000893 | -0.000343 |
| | (0.00189) | (0.000361) | (0.00147) | (0.000307) |
| Mother's Education Years | 0.00142 | 0.000197 | -0.00689*** | -0.000260 |
| | (0.00143) | (0.000293) | (0.00148) | (0.000277) |
| Whether Parents Have Conflicts | 0.00265 | 0.000723 | 0.0131 | 0.00215 |
| | (0.00734) | (0.00348) | (0.00891) | (0.00332) |
| Sixth Year Ranking | -0.000687** | 0.0000739 | 0.000451 | 0.0000802 |
| | (0.000296) | (0.0000656) | (0.000294) | (0.0000756) |
| Teacher Has College Degree | 0.0265 | 0.0295 | 0.00321 | 0.0247 |
| | (0.0263) | (0.0241) | (0.0285) | (0.0223) |
| Teacher Sex | -0.00184 | 0.00523 | 0.0314 | 0.0281 |
| | (0.0204) | (0.0121) | (0.0261) | (0.0193) |
| School Fixed Effect | NO | YES | NO | YES |
| Observations | 6,240 | 6,240 | 6,240 | 6,240 |
| Joint F>Prob | 0.000 | 0.3233 | 0.000 | 0.5194 |

Notes: In this table, we implement the balance check in 2013. In the first two columns, we run regressions for the proportion of migrant children on different variables with/without controlling for school fixed effects. In the third and the fourth columns, we run regressions for the proportion of left-behind children on different variables with/without controlling for school fixed effects. The coefficients in this table can be interpreted as the correlation between the independent variable and the composition of children in the class. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A13: Balance Check in 2014

| | Proportion of Migrants | | Proportion of Left-Behinds | |
|---|---|---|---|---|
| | Without School FE | With School FE | Without School FE | With School FE |
| Age | -0.0204** | -0.000546 | 0.00710 | 0.000991 |
| | (0.00970) | (0.00120) | (0.00927) | (0.00196) |
| Sex | 0.0145* | 0.00326 | -0.000401 | -0.000456 |
| | (0.00783) | (0.00212) | (0.00536) | (0.00248) |
| Board | -0.0173 | -0.00630 | 0.0478 | -0.00170 |
| | (0.0428) | (0.00481) | (0.0298) | (0.00274) |
| Hukou Type | -0.0254 | -0.000444 | 0.00973 | 0.00167 |
| | (0.0167) | (0.00263) | (0.0112) | (0.00415) |
| Whether Migrant Student | 0.266*** | -0.223 | -0.0633*** | -0.00566 |
| | (0.0554) | (0.187) | (0.0156) | (0.00603) |
| Whether Left-behind Student | -0.0456*** | -0.00507 | 0.115*** | 0.0487 |
| | (0.0120) | (0.00487) | (0.0322) | (0.0457) |
| Only Child | 0.00872 | -0.00272 | 0.0269** | 0.00115 |
| | (0.0195) | (0.00322) | (0.0117) | (0.00280) |
| Father's Education Years | 0.000162 | -0.000725 | -0.00194 | -0.000471 |
| | (0.00219) | (0.000632) | (0.00141) | (0.000573) |
| Mother's Education Years | 0.0000151 | -0.0000241 | -0.00654*** | -0.000508 |
| | (0.00217) | (0.000257) | (0.00138) | (0.000780) |
| Whether Parents Have Conflicts | -0.00680 | -0.00221 | 0.0112 | -0.00683 |
| | (0.00861) | (0.00271) | (0.0136) | (0.00873) |
| Sixth Year Ranking | -0.000480 | 0.000134 | 0.000357 | -0.0000713 |
| | (0.000430) | (0.000107) | (0.000254) | (7.89e-05) |
| Teacher Has College Degree | 0.00991 | 0.00736 | 0.0253 | 0.00489 |
| | (0.0304) | (0.0301) | (0.0204) | (0.0206) |
| Teacher Sex | -0.000949 | 0.00454 | 0.0431* | 0.0111 |
| | (0.0240) | (0.0183) | (0.0237) | (0.0163) |
| School Fixed Effect | NO | YES | NO | YES |
| Observations | 4,203 | 4,203 | 4,203 | 4,203 |
| Joint F>Prob | 0.000 | 0.8264 | 0.000 | 0.4976 |

Notes: In this table, we implement the balance check in 2014. In the first two columns, we run regressions for the proportion of migrant children on different variables with/without controlling for school fixed effects. In the third and the fourth columns, we run regressions for the proportion of left-behind children on different variables with/without controlling for school fixed effects. The coefficients in this table can be interpreted as the correlation between the independent variable and the composition of children in the class. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A14: Balance Check for Student Ability

| | Average Sixth Year Ranking | |
|---|---|---|
| | Without School FE | With School FE |
| Age | 0.215 | 0.110* |
| | (0.270) | (0.0627) |
| Sex | -0.126 | 0.0634 |
| | (0.119) | (0.0847) |
| Board | -0.569 | 0.127 |
| | (0.862) | (0.0775) |
| Hukou Type | 0.439 | -0.0176 |
| | (0.284) | (0.109) |
| Whether Migrant Student | -1.247** | 0.0973 |
| | (0.558) | (0.132) |
| Whether Left-behind Student | 0.787*** | 0.0438 |
| | (0.281) | (0.0897) |
| Only Child | 0.111 | 0.165 |
| | (0.448) | (0.158) |
| Father's Education Years | -0.00924 | -0.0185 |
| | (0.0413) | (0.0158) |
| Mother's Education Years | -0.103** | -0.00100 |
| | (0.0466) | (0.0149) |
| Whether Parents Have Conflicts | 0.0226 | 0.169 |
| | (0.228) | (0.108) |
| Sixth Year Ranking | 0.0777*** | 0.0127 |
| | (0.0162) | (0.0105) |
| Teacher Has College Degree | 0.000396 | 0.963 |
| | (0.808) | (1.050) |
| Teacher Sex | -0.112 | 0.652 |
| | (1.089) | (1.087) |
| School Fixed Effect | NO | YES |
| Year Fixed Effect | YES | YES |
| Observations | 10,443 | 10,443 |
| Joint F>Prob | 0.000 | 1.000 |

Notes: In each column, we regress the average sixth year ranking of students in the class on different variables with/without controlling for school fixed effects. The coefficients in this table can be interpreted as the correlation between the independent variable and the average ability of the children in the class. The joint F-test is calculated using a SUR system. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

## A.3 Heterogeneity in Peer Effects

In this section, we implement two heterogeneity analyses.

In the first regression, we interact the main independent variables with an indicator of whether the student is an ordinary local student or not. The results are shown in Table A15. We find that ordinary local students are more likely to be negatively affected by their migrant and left-behind peers. Additionally, migrant peers have a smaller negative effect than left-behind peers on both ordinary local and non-local students.

One peer effect concern is possible heterogeneity across students with different abilities. Are high or low ability students affected more? We first categorize all observations into four quartile groups according to their test scores in 2013 and then run the main regressions using only 2014 data while interacting the peer proportions with the quartile group indicators. The baseline group is the fourth quartile group, which is the highest ability group of students. In each regression we control for prior test scores, which leads to a conventional value-added style model.

The results of the regressions are shown in Table A16. For convenience, we calculate all the marginal effects of the proportions of migrant peers and left-behind peers on each quartile group of students by adding up the corresponding regression coefficients in Table A17. A Wald test is implemented for each of the marginal effects. We find that the proportion of left-behind peers negatively affects students in the first and the third quartile groups. For instance, in column (2), an increase of ten percentage points in the proportion of left-behind peers results in a 0.743 (0.275-1.018) point decrease in the standardized test score for students with the lowest previous test scores. Meanwhile, the proportion of migrant peers does not have significant negative effects on any quartile.

A general conclusion is that in the second year students in the lowest quartile group of ability are most negatively affected by the left-behind peers in the class, but the effect falls to zero

Table A15: Peer Effects of Migrant and Left-Behind Children by Type

| | (1) | (2) | (3) |
|---|---|---|---|
| Proportion of Migrant Peers | -0.391 | -0.384 | -0.584** |
| | (0.262) | (0.260) | (0.277) |
| Proportion of Left-Behind Peers | -0.955** | -0.945** | -0.812** |
| | (0.415) | (0.411) | (0.326) |
| Proportion of Migrant Peers × Local | -0.308* | -0.297* | -0.218 |
| | (0.162) | (0.160) | (0.156) |
| Proportion of Left-Behind Peers × Local | -0.0749 | -0.0628 | -0.0516 |
| | (0.112) | (0.114) | (0.114) |
| School FE | YES | YES | YES |
| Year Dummy | YES | YES | YES |
| Personal Controls | YES | YES | YES |
| Household Controls | NO | YES | YES |
| Teacher Controls | NO | NO | YES |
| Observations | 10,267 | 10,267 | 10,267 |
| R-squared | 0.355 | 0.356 | 0.364 |

Notes: The dependent variable for all regressions is the standardized test score. In these regressions, we investigate the heterogeneity of the peer effects on local and non-local students. For column (1), we do not control for personal characteristics and household characteristics. For column (2), we do not control for household characteristics. For column (3), we control for all independent variables. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

for students with the highest ability. Migrant students do not negatively affect students in any quartile group of ability.

Table A18 shows a heterogeneity analysis to investigate how the peer effects differ across children with different parental skills. We define high-skill families as those with at least one college-educated parent. We interact the proportion of migrant and left-behind peers with whether a student comes from a high-skill family. We find that, in general, the negative impact of migrant and left-behind peers is smaller on children from high-skill families.

Table A16: Heterogeneous Peer Effects of Migrant and Left-behind Children by Ability

| | (1) | (2) |
|---|---|---|
| Proportion of Migrant Peers | 0.284 | 0.0196 |
| | (0.288) | (0.411) |
| Proportion of Left-Behind Peers | 0.209 | 0.275 |
| | (0.334) | (0.280) |
| Proportion of Migrant Peers × Cognitive Skill Q1 | -0.287 | -0.263 |
| | (0.258) | (0.254) |
| Proportion of Left-Behind Peers × Cognitive Skill Q1 | -0.968*** | -1.018*** |
| | (0.278) | (0.271) |
| Proportion of Migrant Peers × Cognitive Skill Q2 | 0.104 | 0.0685 |
| | (0.273) | (0.261) |
| Proportion of Left-Behind Peers × Cognitive Skill Q2 | -0.454* | -0.472** |
| | (0.232) | (0.225) |
| Proportion of Migrant Peers × Cognitive Skill Q3 | 0.113 | 0.106 |
| | (0.213) | (0.207) |
| Proportion of Left-Behind Peers × Cognitive Skill Q3 | -0.730*** | -0.730*** |
| | (0.213) | (0.215) |
| Test Score in 2013 | 0.356*** | 0.346*** |
| | (0.0439) | (0.0413) |
| School-Grade FE | YES | YES |
| Personal Controls | YES | YES |
| Household Controls | NO | YES |
| Teacher Controls | NO | YES |
| Observations | 3,472 | 3,472 |
| R-squared | 0.512 | 0.518 |

Notes: The dependent variable for all regressions is the standardized test score. These two regressions use only data from the second wave since previous test scores are available only for the second wave. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A17: Marginal Peer Effects of Migrant and Left-behind Children by Ability

| Marginal Effects | (1) | (2) |
|---|---|---|
| Proportion of Migrant Peers on Q1 | -0.003 | -0.243 |
| Proportion of Migrant Peers on Q2 | 0.388* | 0.088 |
| Proportion of Migrant Peers on Q3 | 0.397* | 0.126 |
| Proportion of Migrant Peers on Q4 | 0.284 | 0.020 |
| Proportion of Left-Behind Peers on Q1 | -0.759** | -0.743*** |
| Proportion of Left-Behind Peers on Q2 | -0.245 | -0.197 |
| Proportion of Left-Behind Peers on Q3 | -0.521** | -0.455** |
| Proportion of Left-Behind Peers on Q4 | 0.209 | 0.275 |
| School FE | YES | YES |
| Personal Controls | YES | YES |
| Household Controls | NO | YES |
| Teacher Controls | NO | YES |
| Test Score in 2013 | YES | YES |

Notes: The marginal effects are calculated using the regressions from Table A16. Wald tests are implemented and the null hypothesis is whether the marginal effect is zero. We control for student's prior test score in 2013 in both columns. For column (1), we do not control for household and teacher characteristics. For column (2), we control for all independent variables. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents always have conflict with each other. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A18: Heterogeneity: Peer Effects by Parents' Skills

| | (1) Both Years | (2) First Year | (3) Second Year |
|---|---|---|---|
| Proportion of Migrant Peers | -0.439 | -0.740** | -0.348 |
| | (0.314) | (0.336) | (0.405) |
| Proportion of Migrant Peers × High-skill Family | 0.0921 | -0.327 | 0.291 |
| | (0.197) | (0.348) | (0.195) |
| Proportion of Rural Left-Behind Peers | -0.928*** | -1.070** | -0.567** |
| | (0.315) | (0.482) | (0.233) |
| Proportion of Rural Left-Behind Peers × High-skill Family | 0.650 | 1.121* | 0.191 |
| | (0.401) | (0.606) | (0.312) |
| School FE | YES | YES | YES |
| Year Dummy | YES | NO | NO |
| Personal Controls | YES | YES | YES |
| Household Controls | YES | YES | YES |
| Teacher Controls | YES | YES | YES |
| Observations | 6,944 | 3,472 | 3,472 |
| R-squared | 0.387 | 0.399 | 0.406 |

Notes: The dependent variable for all regressions is the standardized test score. For column (1), we use data from both years. For column (2), we use data from only the first year. For column (3), we use data from only the second year. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

## A.4 Mechanism Analysis: Student Misbehavior

In this section, we consider an important mechanism through which migrant students and left-behind students can affect their classmates that rationalizes why only left-behind students have persistent negative effects on their classmates. As some previous studies have found (Case and Katz, 1991; Gaviria and Raphael, 2001; Li, Zang, and An, 2013), student misbehavior can "contaminate" classmates. It is likely that students will mimic the behavior of their friends who smoke, fight or drop out of school.

In the 2014 wave, the survey asks students whether they are often involved in fights, whether they often bully others, whether they often skip classes, whether they often cheat on exams, whether they often smoke, and whether they often go to video gaming bars. In Table A19, for each of the indicator variables created by the answers to these questions and using only the 2014 data, we regress the misbehaving indicator on the proportion of migrant peers in his or her class, the proportion of left-behind peers in his or her class, the indicator of whether the student is a migrant, the indicator of whether the student is left-behind, and other controls. Then we check for two effects: first, the causal peer effects of migrant and left-behind classmates on student misbehavior; second, the correlation between being either migrant, left-behind or local and student misbehavior.

Table A19 displays the results. In several cases, we find statistically significant positive peer effects on misbehavior. For instance, an increase of ten percentage points in the proportion of left-behind peers in the class leads to a 1.1 percentage point increase in the probability of a student being often involved in fights. The average probability of being often involved in fights is 5.9 percent, which means a 1 percentage point increase is equivalent to a 18.6 percent increase relative to the average rate. In addition, left-behind students are 3.1 percentage points (52.5%) more likely to be often involved in fights. Combining these two facts, we can infer that left-behind students may often get into trouble with others and trigger more classroom fights. Similar peer effects also exist for cheating on exams and going to video gaming bars. An increase of ten percentage points in the proportions of migrant peers and left-behind peers will respectively lead to 2.5 percentage point (32.9%) and 1.2 percentage point (15.8%) increases in the probability of a student cheating on exams. An increase of ten percentage points in the proportions of migrant peers and left-behind peers will respectively lead to increases of 1.2 percentage points (30.8%) and 1.8 percentage points (46.2%) in the probability of the student often going to video gaming bars. Moreover, left-behind students are 2.0 percentage points (97.6%) more likely to often smoke. In the last two columns of the table, we combine the six measures of misbehavior together into one index and rerun the regressions. First, we use a simple mean score of the previous six variables to measure student misbehavior. Second, we use the first principal component of the previous six variables as a new combined index. The results show that, generally, more migrant and left-behind classroom peers increase student misbehavior, and the effect is more severe from left-behind peers. All the estimates suggest that, even in the second year, more left-behind classmates still leads to more misbehavior. Consistently with this story, in Appendix A.3 we also find that students with the lowest abilities (usually students who are most likely to misbehave) are affected the most by their left-behind peers.

An anecdotal story claims that the misbehavior of left-behind students is due to the lack of supervision from their parents. To test this story, we run more regressions to discover the correlation between parents' care and being left behind. The dependent variables are whether parents care much about their child's exam scores, general school performance, time spent on the Internet, time spent watching TV, and whether the relationships between the student and his or her mother and his or her father are good. These questions are answered by the children. Tables A20 and A21 reveal an obvious negative correlation between being left behind and parental supervision in both the first and the second year, which implies persistent damage to the child-parents relationship for left-behind students. Parents of left-behind students care less about their children's exams, school performance, time spent online or watching TV, and the relationships between parents and left-behind children are worse. The child-parents relationship worsens with time. The last two columns of Tables A20 and A21 are similar to the last two columns of Table A19, where we use two combined indexes as the dependent variables.

Generally, the misbehavior of students is an important channel through which migrant and left-behind students negatively affect their classmates, especially for left-behind students. The misbehavior of left-behind students may result from the absence of their parents. Furthermore, the detrimental effect of left-behind students on their classmates' misbehavior and the damage to the child-parents relationship of left-behind students persists across time, which may explain the persistence of the negative peer effect of left-behind children on the test scores of their classmates. This could suggest a policy role for government in providing better support for left-behind children.

Table A19: Students' Misbehaviors and the Peer Effects: Second Year

| | Often Fight | Often Bully | Often Skip Class | Often Cheat | Often Smoke | Often Gaming | Average Index | First Principal Component |
|---|---|---|---|---|---|---|---|---|
| Proportion of Migrant Peers | -0.105 | 0.0427 | -0.0374* | 0.247*** | 0.0246 | 0.123*** | 0.0491 | 0.485 |
| | (0.0888) | (0.0682) | (0.0188) | (0.0423) | (0.0401) | (0.0430) | (0.0351) | (0.490) |
| Proportion of Left-Behind Peers | 0.103* | -0.00787 | 0.0107 | 0.116** | 0.0123 | 0.179*** | 0.0689 | 0.765 |
| | (0.0552) | (0.0506) | (0.0453) | (0.0535) | (0.0537) | (0.0593) | (0.0427) | (0.648) |
| Whether Is a Migrant | 0.00594 | 0.00197 | -0.00531 | -0.0220 | 0.0140 | 0.0385** | 0.00553 | 0.0940 |
| | (0.0188) | (0.0109) | (0.0103) | (0.0175) | (0.0120) | (0.0156) | (0.00866) | (0.126) |
| Whether Is a Left-Behind | 0.0309* | 0.000388 | -0.00555 | 0.000806 | 0.0196** | 0.0116 | 0.00962 | 0.119 |
| | (0.0178) | (0.0102) | (0.00573) | (0.0212) | (0.00919) | (0.0159) | (0.00749) | (0.0969) |
| School FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Personal Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Household Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Teacher Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 4,203 | 4,203 | 4,203 | 4,203 | 4,203 | 4,203 | 4,203 | 4,203 |
| R-squared | 0.064 | 0.028 | 0.059 | 0.090 | 0.045 | 0.082 | 0.095 | 0.084 |

Notes: The dependent variables of the first six columns are respectively the indicators of whether the student is often involved in fights at school, whether the student often bullies his or her classmates, whether the student often skips classes, whether the student often cheats in exams, whether the student often smokes and whether the student often goes to video gaming bars. In the seventh column, we use the mean of the first six variables as an average index of students' misbehavior to be the dependent variable. In the last column, we use the first principal component of the first six variables to be the dependent variable. The set of personal controls includes student age, gender, Hukou type, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes parental education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A20: Relation With Parents: First Year

| | On Exam | On School Performance | On Internet | On TV | Relation with Mother | Relation with Father | Average Index | First Principal Component |
|---|---|---|---|---|---|---|---|---|
| Whether Is a Migrant | 0.0414 | 0.00131 | 0.0632* | 0.00133 | 0.00313 | -0.0250 | 0.0142 | 0.0897 |
| | (0.0391) | (0.0425) | (0.0322) | (0.0378) | (0.0288) | (0.0376) | (0.0256) | (0.133) |
| Whether Is a Left-Behind | -0.110*** | -0.0719** | -0.0236 | -0.0534 | -0.0448* | -0.0424 | -0.0577*** | -0.295*** |
| | (0.0333) | (0.0292) | (0.0271) | (0.0333) | (0.0242) | (0.0321) | (0.0133) | (0.0700) |
| School FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year Dummy | YES | YES | YES | YES | YES | YES | YES | YES |
| Personal Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Household Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Teacher Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 3,514 | 3,514 | 3,514 | 3,514 | 3,514 | 3,514 | 3,514 | 3,514 |
| R-squared | 0.056 | 0.054 | 0.068 | 0.040 | 0.095 | 0.109 | 0.102 | 0.094 |

Notes: All the regression are run on the sample of the first wave (2013). The dependent variables of the first six columns are respectively the indicators of whether parents care much about student's exam, whether parents care about the student's general school performance, whether parents pay attention to the time the student spends on the Internet, whether parents pay attention to the time the student spends on TV, whether the relation between the student and his or her mother is good and whether the relation between the student and his or her father is good. In the seventh column, we use the mean of the first six variables as an average index of students' relations with their parents to be the dependent variable. In the last column, we use the first principal component of the first six variables to be the dependent variable. The set of personal controls includes student age, gender, Hukou type, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes parental education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A21: Relation With Parents: Second Year

| | On Exam | On School Performance | On Internet | On TV | Relation with Mother | Relation with Father | Average Index | First Principal Component |
|---|---|---|---|---|---|---|---|---|
| Whether Is a Migrant | 0.0104 | -0.0649** | 0.00105 | 0.0330 | -0.00382 | 0.0102 | -0.00234 | -0.0146 |
| | (0.0294) | (0.0319) | (0.0365) | (0.0419) | (0.0298) | (0.0327) | (0.0190) | (0.0981) |
| Whether Is a Left-Behind | -0.0123 | -0.0215 | 0.000125 | -0.00232 | -0.0467* | -0.105*** | -0.0313 | -0.115 |
| | (0.0523) | (0.0233) | (0.0402) | (0.0265) | (0.0272) | (0.0311) | (0.0205) | (0.109) |
| School FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year Dummy | YES | YES | YES | YES | YES | YES | YES | YES |
| Personal Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Household Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Teacher Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 3,514 | 3,514 | 3,514 | 3,514 | 3,514 | 3,514 | 3,514 | 3,514 |
| R-squared | 0.077 | 0.065 | 0.065 | 0.043 | 0.097 | 0.129 | 0.133 | 0.113 |

Notes: All the regression are run on the sample of the second wave (2014). The dependent variables of the first six columns are respectively the indicators of whether parents care much about student's exam, whether parents care about the student's general school performance, whether parents pay attention to the time the student spends on the Internet, whether parents pay attention to the time the student spends on TV, whether the relation between the student and his or her mother is good and whether the relation between the student and his or her father is good. In the seventh column, we use the mean of the first six variables as an average index of students' relations with their parents to be the dependent variable. In the last column, we use the first principal component of the first six variables to be the dependent variable. The set of personal controls includes student age, gender, Hukou type, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes parental education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

## A.5 Family Background as a Channel

One potential policy implication is that we can reduce aggregate negative spillovers by relaxing the enrollment restriction on migrant students and encouraging parents to take their children with them when they migrate. However, peer effects may result from the selection into migration. If the differences in peer effects between left-behind and migrant students are fully attributed to left-behind students having lower "ability" and being from families with lower socioeconomic status, but not from the consequences of being left behind, then the external validity of the peer effects estimates could be limited. After a policy change, the peer effects would also change according to the changing composition of migrants. If selection is the whole story, then the policy recommendation of relaxing the public school enrollment restriction on migrant students to reduce overall negative spillovers is not cogent. In this section, we try to test whether selection is the whole story. We alter the regression specifications in different ways and the main conclusions do not change.

In the main regression, we regress the test score of a student on the proportion of migrant and left-behind peers in his or her class. To investigate the effect resulting from peers' pre-determined family background, we additionally control for the average family background of a student's classmates. The control variables include the average self-reported socioeconomic condition of the classmates and the average years of schooling completed by each parent of the classmates. Since the assignment of students into classes is random, these averaged variables capturing classmate family background should also be exogenous. If we detect zeros for the coefficients of the proportions of migrant and left-behind peers, this means family background is likely to be the only channel for the peer effects.

Results are exhibited in Table A22. Column (1) displays the result from the main regression without classmates' family background controls as a reference (same as column (3) in Table 3). We do find evidence that after controlling for these pre-determined family backgrounds the point estimates of the peer effects shrink a little, though they are still negative. Comparing column (5) with column (1), we find that more than 65%-80% of the negative spillovers from left-behind/migrant children cannot be explained by the inclusion of pre-determined factors. The results reveal that migrant and left-behind students affect their classmates because they migrate or are left behind, rather than just because they are from disadvantaged families. Meanwhile, in all specifications, when we net out the pre-determined family background, the negative peer effects of left-behind peers are still more significant than the negative peer effects of migrant peers, both economically and statistically.

Table A22: Peer Effects Netting Out Average Family Background

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Proportion of Migrant Peers | -0.337 | -0.379 | -0.223 | -0.244 | -0.271 |
| | (0.242) | (0.244) | (0.251) | (0.241) | (0.239) |
| Proportion of Left-Behind Peers | -0.772** | -0.625 | -0.604* | -0.587* | -0.503 |
| | (0.347) | (0.396) | (0.314) | (0.320) | (0.351) |
| Average Socioeconomic Condition of Classmates | | 0.379 | | | 0.265 |
| | | (0.301) | | | (0.317) |
| Average Father Education of Classmates | | | 0.0875** | | 0.0357 |
| | | | (0.0358) | | (0.0603) |
| Average Mother Education of Classmates | | | | 0.0811** | 0.0425 |
| | | | | (0.0316) | (0.0542) |
| School FE | YES | YES | YES | YES | YES |
| Year Dummy | YES | YES | YES | YES | YES |
| Personal Controls | YES | YES | YES | YES | YES |
| Household Controls | YES | YES | YES | YES | YES |
| Teacher Controls | YES | YES | YES | YES | YES |
| Observations | 10,443 | 10,443 | 10,443 | 10,443 | 10,443 |
| R-squared | 0.361 | 0.363 | 0.363 | 0.364 | 0.365 |

Notes: The dependent variable for all regressions is the standardized test score. The first column is the baseline estimate from the main regression. In columns (2), (3), and (4), we run regressions adding average family background characteristics of the classmates one by one. In column (5), we run regression with the whole set of family background characteristics. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

## A.6 Class Size Effect

Another important part of the peer effect is the class size effect. It is widely known that smaller classes can positively affect children's learning outcomes (Krueger, 1999). Thus, if class sizes matter a lot in China, this should be considered in the cost-benefit analysis when we calculate the resources required to keep school quality unchanged as more migrant children move to big cities. Table A23 shows the main regression results when class size effects are also considered. We do not find any significant changes in the peer effects estimates for either migrant or left-behind children. Additionally, the effect of class size is very small and statistically insignificant.

Table A23: Peer Effects of Migrant Children and Left-Behind Children with Class Size Effects

| | (1) | (2) | (3) |
|---|---|---|---|
| Proportion of Migrant Peers | -0.521* | -0.504* | -0.395 |
| | (0.298) | (0.289) | (0.255) |
| Proportion of Left-Behind Peers | -1.031** | -0.994** | -0.800** |
| | (0.454) | (0.443) | (0.352) |
| Class Size | -0.000602 | -0.000802 | -0.00124 |
| | (0.00278) | (0.00276) | (0.00300) |
| School FE | YES | YES | YES |
| Year Dummy | YES | YES | YES |
| Personal Controls | YES | YES | YES |
| Household Controls | NO | YES | YES |
| Teacher Controls | NO | NO | YES |
| Observations | 10,443 | 10,443 | 10,443 |
| R-squared | 0.353 | 0.354 | 0.361 |

Notes: The dependent variable for all regressions is the standardized test score. Compared with the main regression, we additionally control for class size in this set of regressions. For column (1), we do not control for household characteristics or teacher characteristics. For column (2), we do not control for teacher characteristics. For column (3), we control for all independent variables. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently conflict with each other. The set of teacher characteristics includes whether the head teacher has a college degree and head teacher's sex. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

## B Supplements to the Structural Model Analysis

### B.1 Human Capital Equation By Type

As explained in the main paper, there are three types of families: stayers, migrant families, and left-behind families. Migrant families may have children in public or migrant schools, leading to four types of children. Here we explain the human capital production function in each of these four cases.

For stayers (both parents and children stay), we have:

$$v_{iio}^{s} = \chi_0^s + \mathbf{\Theta^s} Peer_{i,pub} + \phi^s + \kappa_{r_i}^s \tag{1}$$

Stayers study in public schools in their hometown $i$, so their peer composition is $Peer_{i,pub}$. They can enjoy the public school premium $\phi^s$ since $Pub_o = 1$.

For left-behind families, we have:

$$v_{ijo}^{s} = \chi_0^s + \mathbf{\Theta^s} Peer_{i,pub} + \phi^s + \nu^s + \kappa_{r_i}^s \tag{2}$$

Left-behind children study in public schools in their hometown $i$, so their peer composition is $Peer_{i,pub}$. They can enjoy the public school premium $\phi^s$ since $Pub_o = 1$. The only difference between left-behind families and stayers is the former bear a left-behind cost $\nu^s$ since $lb_o = 1$.

For migrant families who successfully enroll their children in public schools, we have:

$$v_{ijo}^{s} = \chi_0^s + \mathbf{\Theta^s} Peer_{j,pub} + \phi^s + \eta^s + \kappa_{r_j}^s \tag{3}$$

Migrant children who successfully enroll in public schools in their migration destination $j$ have peer composition $Peer_{j,pub}$. They can enjoy the public school premium $\phi^s$ since $Pub_o = 1$. They have to bear the migration cost on human capital $\eta^s$.

For migrant families who enroll their children in private migrant schools, we have:

$$v^s_{ijo} = \chi^s_0 + \mathbf{\Theta^s} Peer_{j,pri} + \eta^s + \kappa^s_{r_j} \tag{4}$$

Migrant children who enroll in private migrant schools in their migration destination $j$ have peer composition $Peer_{j,pri}$. They cannot enjoy the public school premium $\phi^s_{r_j}$ since $Pub_o = 0$. They have to bear the migration cost on human capital $\eta^s$.

### B.2 Deriving the Gravity Equation

Given a worker endowed with skill $s$ and home city $i$, the cumulative density function (CDF) of his or her distribution of utility across working cities $j$ is:

$$G^s_{ij}(u) = Pr[U \leq \tilde{U}] = F(\frac{\tilde{U}\tau^s_{ij}}{w^s_j}(u^s_{ij})^{-\beta}) = e^{-\Phi^s_{ij}\tilde{U}^{-\epsilon}} \tag{5}$$

The probability density function (PDF) of utility is $g^s_{ij}(u) = \Phi^s_{ij}\epsilon u^{-(\epsilon+1)}e^{-\Phi^s_{ij}u^{-\epsilon}}$. We denote $G^s_i(u)$ to be the CDF of the optimized utility for the worker:

$$G^s_i(u) = \prod_r G^s_{ir}(u)$$

where the left-hand side refers to the probability of having a utility less than $u$, and the right-hand side refers to the probability of having utility less than $u$ for all possible choices of $j$. As a type of extreme value distribution, the maximum of several random variables with Fréchet distributions is also Fréchet distributed. In this case, we have:

$$G^s_i(u) = e^{-u^{-\epsilon}\sum_r \Phi^s_{ir}} = e^{-u^{-\epsilon}\Phi^s_i} \tag{6}$$

where $\sum_r \Phi_{ir}^s = \Phi_i^s$. Then for workers endowed with skill $s$ and hometown $i$, the proportion of them working in city $j$ is:

$$
\begin{aligned}
\pi_{ij}^s &= Pr[u_{ij}^s \geq \max_r \{u_{ir}^s\}] \\
&= \int_0^\infty \prod_{r \neq j} G_{ir}^s(u) g_{ij}^s(u) du \\
&= \int_0^\infty \Phi_{ij}^s \epsilon u^{-(\epsilon+1)} \cdot e^{-\Phi_{ij}^s u^{-\epsilon}} \cdot e^{-\sum_{r \neq j} \Phi_{ir}^s u^{-\epsilon}} du \\
&= \int_0^\infty \Phi_{ij}^s \epsilon u^{-(\epsilon+1)} \cdot e^{-\Phi_i^s u^{-\epsilon}} du \\
&= \int_0^\infty \Phi_{ij}^s d[\frac{1}{\Phi_i^s} e^{-\Phi_i^s u^{-\epsilon}}] \\
&= \frac{\Phi_{ij}^s}{\Phi_i^s}
\end{aligned} \tag{7}
$$

As a result, we can express $\pi_{ij}^s$ as a gravity equation:

$$
\pi_{ij}^s = \frac{\Phi_{ij}^s}{\Phi_i^s} = \frac{(w_j^s (k_{ij}^{s'})^\beta)^\epsilon (\tau_{ij}^s)^{-\epsilon}}{\sum_r (w_r^s (k_{ir}^{s'})^\beta)^\epsilon (\tau_{ir}^s)^{-\epsilon}} \tag{8}
$$

### B.3 Multiple Equilibria

As in many other urban economics models with spillover effects, this model may have multiple equilibria. Fortunately, with all the parameters and observed data, we can still identify unobserved city characteristics $\mathbf{A}$ and $\tau$ uniquely. Hence, although there may be other equilibria of this model, the equilibrium reflecting the real world can be recovered from the data.

**Proposition 1** *Given the parameter vector* $\mathbf{\Gamma} = \{\beta, \epsilon, \alpha, \chi, \Theta, \phi, \nu, \eta, \kappa, \zeta, \sigma\}$, *data on the endogenous variable vector* $\mathbf{\Delta} = \{\mathbf{w}, \mathbf{L}, \mathbf{Peer}\}$ *and data on the observed city characteristics* $\mathbf{\Omega}^* = \{\mathbf{\Xi}, \mathbf{p}\}$, *the unobserved city characteristics* $\mathbf{\Omega}' = \{\mathbf{A}, \tau\}$ *can be uniquely identified.*

The proposition is proved in section 4.7 and Appendix B.4. Another multiple equilibria concern is the equilibrium selection in the counterfactual analysis. We choose the equilibrium that is the closest to the one recovered by the data in the real world. The detailed algorithm is shown in Appendix B.10.

## B.4 Solving the Model

### B.4.1 Solving for Local Productivity

Equations (18) and (19) represent the first order conditions of firms with respect to high-skill and low-skill workers. By solving these two equations, we can derive the productivities $A_j^h$ and $A_j^l$ given wages and the numbers of workers in the data:

$$A_j^h = \frac{w_j^h}{[1 + (\frac{w_j^l L_j^l}{w_j^h L_j^h})]^{\frac{1-\rho}{\rho}}} \tag{9}$$

$$A_j^l = \frac{w_j^l}{[1 + (\frac{w_j^h L_j^h}{w_j^l L_j^l})]^{\frac{1-\rho}{\rho}}} \tag{10}$$

### B.4.2 Solving for Migration Costs

As there is no migration cost (iceberg cost is equal to one) when an individual chooses to stay in his or her home city $i$, the commuting probability equation (16) with $i = j$ can be written as:

$$\pi_{ii}^s = \frac{(w_i^s (u_{ii}^s)^\beta)^\epsilon}{\Phi_i^s} \tag{11}$$

Then since $u_{ij}^s$ can be calculated using the given parameters and the data, we can solve for $\Phi_i^s$ as:

$$\Phi_i^s = \frac{(w_i^s (u_{ii}^s)^\beta)^\epsilon}{\pi_{ii}^s} \tag{12}$$

With $\Phi_i^s$ obtained, the migration cost $\tau_{ij}^s$ is determined by the original commuting probability equation:

$$\begin{aligned} \pi_{ij}^s &= \frac{(w_j^s (u_{ij}^s)^\beta)^\epsilon (\tau_{ij}^s)^{-\epsilon}}{\Phi_i^s} \\ \Rightarrow \tau_{ij}^s &= \frac{w_j^s (u_{ij}^s)^\beta}{(\pi_{ij}^s \Phi_i^s)^{\frac{1}{\epsilon}}} \end{aligned} \tag{13}$$

### B.5 Algorithm to Solve City Fixed Effects

Since we have 276 cities and 2 types of skills, there will be $276 \times 2 = 552$ fixed effect terms. It is difficult to estimate all of them using a traditional optimization algorithm. Rather, we utilize the contraction algorithm recommended in Berry, Levinsohn, and Pakes (1995). We estimate the fixed effects terms by equalizing the empirical share and the predicted share of migrant/left-behind children for each destination/Hukou city.

Denoting $m$ as migrant children and $l$ as left-behind children, we have:

$$\begin{aligned} S_{jm} &= \frac{N_{jm}}{N_j}, \quad S_{il} = \frac{N_{il}}{N_i}, \quad S_s = \frac{N_{sm}}{N_s} \\ \hat{S}_{jm}(\zeta) &= \frac{\sum_o P_{oijm}}{N_j}, \quad \hat{S}_{il}(\zeta) = \frac{\sum_o P_{oijl}}{N_i}, \quad \hat{S}_s(\zeta) = \frac{\sum_o P_{os}}{N_s} \end{aligned}$$

$S_{jm}$ is the empirical share of migrant workers who migrate to destination city $j$ with their children. $N_{jm}$ is the total number of migrant children in destination city $j$. $N_j$ is the total number of migrant workers in destination city $j$. Similarly, $S_{il}$ is the empirical share of migrant workers who migrate from home city $i$ and leave their children behind. $N_{il}$ is the total number of left-behind children in home city $i$. $N_i$ is the total number of migrant workers from home city $i$. $S_s$ is the empirical share of migrant workers with skill $s$. $N_{sm}$ is the total number of migrant workers with skill $s$. $N_s$ is the total number of migrant workers with skill $s$. $\hat{S}_{jm}(\zeta)$, $\hat{S}_{il}(\zeta)$, and $\hat{S}_s(\zeta)$ are the corresponding predicted shares using the model for each trial of the parameters.

We equate them to solve for all the fixed effects in a system of linear equations as follows:

$$S_{jm} = \hat{S}_{jm}(\zeta), \quad S_{il} = \hat{S}_{il}(\zeta), \quad S_s = \hat{S}_s(\zeta)$$

To solve this system, we first guess a set of $\zeta^0$ and then update them using the following rule:

$$\begin{aligned}
\zeta_{jm}^{t+1} &= \zeta_{jm}^{t} + ln(\frac{S_{jm}}{\hat{S}_{jm}(\zeta^t)}) \\
\zeta_{il}^{t+1} &= \zeta_{il}^{t} + ln(\frac{S_{il}}{\hat{S}_{il}(\zeta^t)}) \\
\zeta_{s}^{t+1} &= \zeta_{s}^{t} + ln(\frac{S_{s}}{\hat{S}_{s}(\zeta^t)})
\end{aligned}$$

The procedure is completed once we achieve convergence.

### B.6 Willingness to Pay for Public School

We denote $c$ as consumption. Since $z_{ijo}$ and $\tau_{ij}^{s}$ will not be affected by agents' choices, the indifference curve of the utility function can be expressed as:

$$\bar{U} = c_{ij}(u_{ij}^{s})^{0.64} \tag{14}$$

For a low-skill worker with a public school premium of 0.222 (column (2) in Table 8), we have:

$$\frac{c_1}{c_2} = e^{0.222\times0.64} \approx 1.15 \tag{15}$$

$$\Rightarrow c_2 \approx \frac{17}{20}c_1 \tag{16}$$

This means that average low-skill Chinese parents are indifferent to having only $\frac{20}{23}$ of their current consumption if they can enroll their children in public schools. In other words, they are willing to pay about $\frac{3}{23}$, or 13% of their annual wages, to enroll their children in public schools.

This translates to about 1,645 RMB or 243 US dollars in 2010.[2]

## B.7 Validation of the Calibration of $\beta$

We employ two methods to validate the calibration of the weight parents put on children's human capital in their utility. Furthermore, we investigate the robustness of our main results by re-calibrating the model using different $\beta$.

### B.7.1 By Using Housing Premia

The first validation method is to match the revealed willingness to pay in an external market. However, we must clarify that this is only a simple back-of-the-envelope calculation.

One of the traditional topics in urban economics is the housing premium of good schools, which reflects parents' willingness to pay for a better education. There are many studies examining how the academic performance of schools impacts local housing prices (Black and Machin, 2011). However, very few rigorous studies have been done in China due to the lack of school performance data. The only study on this topic is Chan et al. (2020). In this paper, the authors estimate the housing premium of school quality in Shanghai in 2015 and 2016. They find that a one-standard-deviation increase in tournament (academic competitions, for instance, Math Olympiad) performance raises housing prices by 2.4%. They also show that the average housing price in Shanghai in 2015 was 3,189,100 RMB (about 470,000 USD). Thus, a one-standard-deviation increase in tournament performance raises housing prices by $3,189,100 \times 2.4\% = 76538.4$ RMB. Assuming that agent pay the housing cost in 10 years (the housing price-annual wage ratio in China is about 10 to 20), we obtain a yearly school premium of $76538.4 \div 10 = 7653.84$ RMB. Since the standard deviation of the cognitive test score used in this study is 0.886, the yearly school premium expense per point would be $7653.84 \div 0.886 \approx 8638.65$ RMB. Based on the China Statistical Yearbook, average income in

[2]The average annual wage for low-skill workers in China in 2010 was 12,655 RMB.

Shanghai in 2015 was 49867.2 RMB. Thus, the share of annual income that parents are willing to pay for one point is $8638.65 \div 49867.2 \approx 0.17$. Similar to Appendix B.6, we solve for the parameter $\beta$ that rationalizes this school premium using the indifference curve of the utility function:

$$\frac{c_1}{c_2} = e^{\hat{\beta}^{cali1}} \approx \frac{1}{1 - 0.17} \tag{17}$$

$$\Rightarrow \hat{\beta}^{cali1} \approx 0.19 \tag{18}$$

This yields a smaller estimate of the parameter $\hat{\beta}$ than in the main calibration. There are two possible reasons. First, housing prices are very high in Shanghai and many people with relatively low wealth cannot afford them. Consequently, they are not responsive to the impact of good schools on housing prices. In other words, the housing market in good school districts would have been bid up if most people were not priced out of the market. Second, Math Olympiad is different from normal education, focusing only on elite students. Math Olympiad success may not be as attractive as success on the High School Entrance Exam to common families. Thus, the housing premium may underestimate parents' willingness to pay for education quality in China. We check the robustness of our results in Appendix B.8 by recalculating the counterfactual results using $\beta = 0.19$ to make sure the main conclusion is not changed.

#### B.7.2 Survey

The second validation method is to give a short survey to parents of migrant children and directly ask their willingness to pay for public schools. We implemented this survey during June and July 2022 in Shanghai. The survey asks two main questions. First, how much would you be willing to pay for a public school seat if it was not free? Second, what is the average annual income for you and your spouse? Due to the Covid-19 lockdown in Shanghai, implementing the survey was extremely difficult. It was distributed to two groups of parents and taken online using the Tencent Survey App.

The first group is parents of migrant children who are studying with a non-profit, NGO"$S$". "$S$" focuses on providing migrant children with free out-of-school education. Their children are aged between 5 and 15, covering both primary and middle school. The NGO offers many courses including piano, guitar, choir, cooking, programming, biology, psychology, and economics. They also help children with school homework. All teachers are volunteers from universities in Shanghai. In total, there are 385 students studying with this NGO. We randomly selected 177 children to have their parents take this survey. After dropping extreme values, there were 150 effective responses.

The second group is parents of migrant children from a public school "$Q$". "$Q$" is located in the south suburban region of Shanghai with a high proportion of migrant students. There are 399 migrant students in this school and we randomly selected 213 children to have their parents take this survey. After dropping extreme values, there were 182 effective responses.

Table B1 shows the average WTP for all surveyed parents and the two groups separately. We find that the average WTP for public school is 15.2% of income. As over 70% of parents do not hold any college degree in our survey, the results are very close to Appendix B.6, suggesting the estimate of $\beta$ in the model is valid.

Table B1: Survey Results of WTP for Public Schools

| | Total | NGO S | Public School Q |
|---|---|---|---|
| Number of Parents | 332 | 150 | 182 |
| Average WTP | 15.2% | 12.6% | 17.4% |

Notes: This table displays the average WTP from the survey. The first column shows the results for both groups. The second column shows the results for parents from NGO "S". The third column shows the results for parents from School "Q".

## B.8 Sensitivity of the Main Results

In this section, we check the robustness of our main results by choosing alternative values of $\beta$ that are larger and smaller that used in the main text. We set the larger value to be 1, which is equivalent to low-skill families being willing to spend 20% of their annual income to enroll their children in public schools. We set the smaller value to be 0.19, which is derived from Appendix B.7.1. We re-calibrate the model in these two settings and calculate the main counterfactual results. Tables B2 and B3 show that when $\beta$ is larger we have a larger human capital gain. Meanwhile, reducing $\beta$ decreases the human capital gain. Overall, within the reasonable range of this parameter, the changes are small compared with the results in the main setting in Table 10.

Table B2: Changes in Human Capital: High $\beta = 1$

| Variables | Changes (Test Score s.d.) |
|---|---|
| Average HC | 0.0079 |
| Average HC of High-skill from Big | -0.0096 |
| Average HC of Low-skill from Big | -0.030 |
| Average HC of High-skill from Small | 0.026 |
| Average HC of Low-skill from Small | 0.0078 |

Notes: HC stands for Human Capital. "Big" means big cities and "Small" means small cities. In this setting, we assume $\beta = 1$. In the first column, we show the changes in human capital when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

Table B3: Changes in Human Capital: Low $\beta = 0.19$

| Variables | Changes (Test Score s.d.) |
|---|---|
| Average HC | 0.0075 |
| Average HC of High-skill from Big | -0.018 |
| Average HC of Low-skill from Big | -0.028 |
| Average HC of High-skill from Small | 0.012 |
| Average HC of Low-skill from Small | 0.0097 |

Notes: HC stands for Human Capital. "Big" means big cities and "Small" means small cities. In this setting, we assume $\beta = 0.19$. In the first column, we show the changes in human capital when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

### B.9 Model Fit

To check the fit of the estimation, we calculate all the endogenous variables in equilibrium using the model with parameters we have estimated and compare them with real data.

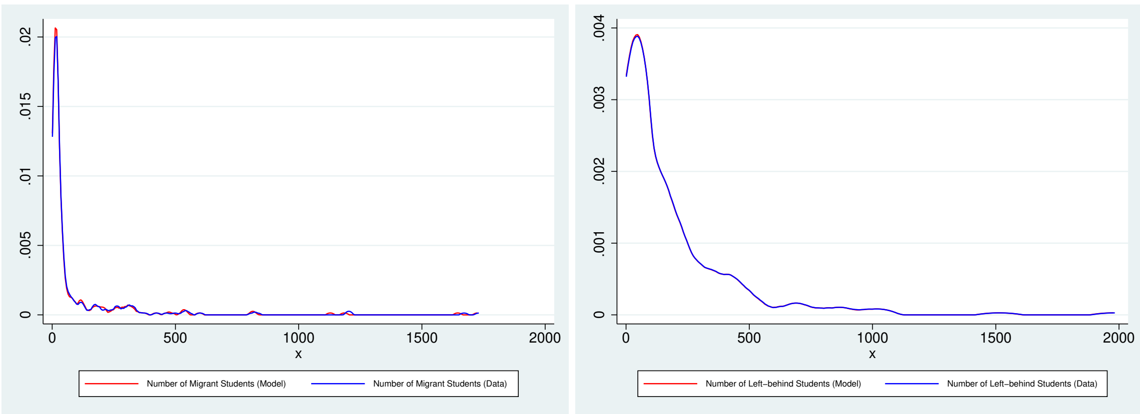
Table B4 shows the fit with aggregate moments. The first column shows the equilibrium results calculated from the model. The second column shows the real data. The third column shows the percentage differences. The difference at the aggregate level is very small. Figures B1 and B2 show the distribution of the model simulated values and the real data across cities. The red curve represents the density at equilibrium from the model and the blue curve represents the density in the data. In Figure B1, we show the fit of the number of migrant and left-behind children across cities. In Figure B2, we show the fit of the wages of high and low-skill workers across cities. In both figures, the red line matches the blue line very well and we can hardly distinguish them from each other. In general, the model fits the data very well in terms of wages and migration of workers and students.

We further compare relative output in different cities generated from the model to relative GDP in different cities in the real world. This is an untargeted moment in the model estimation. We take Beijing as the baseline city. The real-world GDP data come from the *China City Statistical Yearbook*. Table B5 shows the aggregate moments. The first row shows relative output in each city generated from the model. The second row shows relative GDP in the real data. The third row shows the difference between the model and the data. Figure B3 shows the detailed kernel-smoothed density distribution. We find that the model-generated distribution of relative output closely matches the relative GDP data. This out-of-sample investigation further validates our model setting.

Table B4: Model Fit

| Variables | Model | Data | Difference |
|---|---|---|---|
| Total Migrants | 72098 | 72034 | 0.09% |
| Net Migrant Inflow from Small to Big | 38829 | 38803 | 0.06% |
| Total High-skill Migrants | 4618 | 4616 | 0.03% |
| Total Low-skill Migrants | 67480 | 67418 | 0.09% |
| Total Migrant Students | 24195 | 24716 | -2.1% |
| Total Migrant Students to Big | 11252 | 11439 | -1.6% |
| Total Migrant Students to Small | 12943 | 13277 | -2.5% |
| Total Left-behind Students | 47903 | 48048 | -0.3% |
| Total Students in Public in Big | 32630 | 32777 | 0.5% |
| Total Students in Public in Small | 302519 | 302935 | -0.14% |
| Mean Wages of High-skill from Big | 56225 | 56385 | -0.28% |
| Mean Wages of High-skill from Small | 35326 | 35427 | -0.29% |
| Mean Wages of Low-skill from Big | 37237 | 37298 | -0.16% |
| Mean Wages of Low-skill from Small | 25655 | 25693 | -0.15% |

Notes: "Big" means big cities and "Small" means small cities. The first column shows the results for the equilibrium solved in the model. The second column shows the results from the data. The third column shows the difference between the first and the second columns. Sources: Census 2010, China Education Panel Survey 2013 and 2014.




(a) The Density of Migrant Children

(b) The Density of Left-behind Children

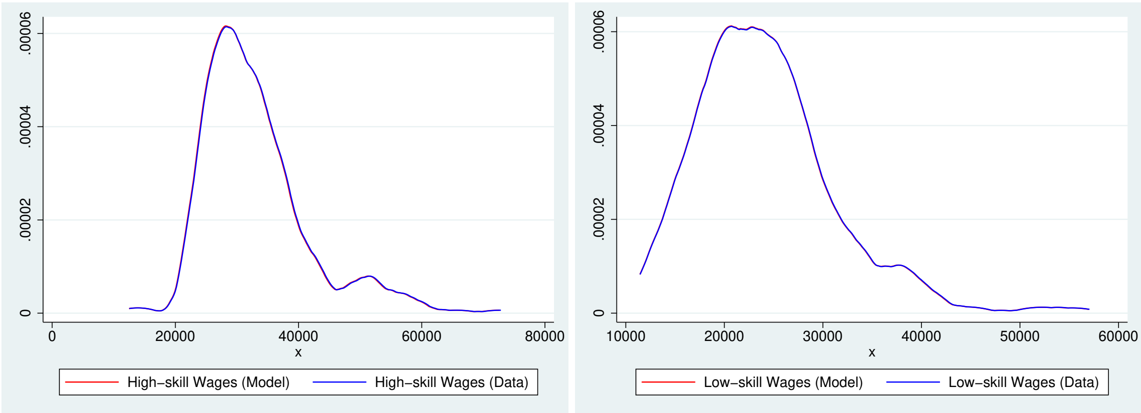
Figure B1: Model Fit of Child Migration

Notes: This figure shows the densities of the number of migrant students in different cities. The red curve represents the density at equilibrium from the model. The blue curve represents the density in the data. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

Table B5: Model Fit of City GDP (Relative to Beijing)

| Variables | Mean | Standard Deviation |
|---|---|---|
| Relative Output (Model) | 0.098 | 0.134 |
| Relative GDP (Data) | 0.109 | 0.143 |
| Difference (Model-Data) | 0.011 | 0.042 |

Notes: This table shows the model fit of the relative city-level output. We take Beijing as the baseline city. The first row shows relative output in each city generated from the model. The second row shows relative GDP in the real data. The third row shows the difference between the model and the data. Sources: Census 2010, China Education Panel Survey 2013 and 2014, China City Statistical Yearbook 2010.

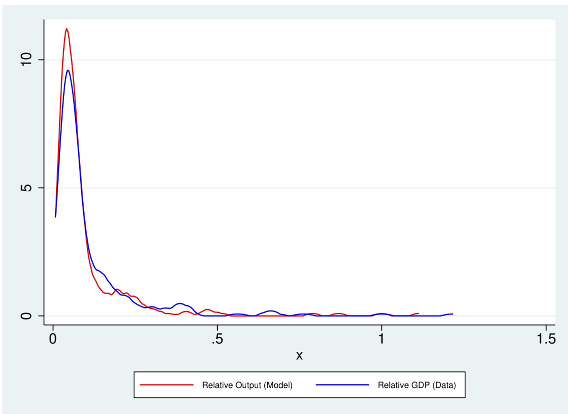



(a) The Density of High-skill Wages

(b) The Density of Low-skill Wages

Figure B2: Model Fit of City-Skill Wages

Notes: This figure shows the densities of wages in different cities for high and low skill workers. The red curve represents the density at equilibrium from the model. The blue curve represents the density in the data. Sources: Census 2010, China Education Panel Survey 2013 and 2014.




(a) The Density of Relative Output and GDP

Figure B3: Fitness of City-level Relative Output

Notes: This figure shows the densities of city-level relative output/GDP in different cities. The red curve represents the density at equilibrium from the model. The blue curve represents the density in the data. The baseline city is Beijing. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

### B.10 Algorithm to Solve the Counterfactuals

Given exogenous variables and parameters, we need to calculate how endogenous variables respond to policy changes. As we have mentioned, we select the equilibrium that is the closest to the one in the real world. Thus, the initial values of the variables are set to be equal to the data. Starting from the initial values of the endogenous variables $\mathbf{\Delta_0} = \{\mathbf{w_0}, \mathbf{L_0}, \mathbf{Peer_0}\}$, we use the following algorithm to find the new equilibrium.

Let $\xi_i^s$ be the number of workers with skill $s$ and Hukou city $i$. In the t-th iteration, let $\hat{V}_t^s$ and $\hat{u}_{ijt}^{s'}$ be the predicted value of children's human capital when the peer composition is $\mathbf{Peer_t}$. For each iteration with $\mathbf{\Delta_t} = \{\mathbf{w_t}, \mathbf{L_t}, \mathbf{Peer_t}\}$, we calculate the predicted values $\hat{\mathbf{\Delta}}_\mathbf{t} = \{\hat{\mathbf{w}}_\mathbf{t}, \hat{\mathbf{L}}_\mathbf{t}, \widehat{\mathbf{Peer}}_\mathbf{t}\}$ as:

$$\hat{L}_{ijt}^s = \frac{(w_{jt}^s (\hat{u}_{ijt}^{s'})^\beta)^\epsilon (\tau_{ij}^{s'})^{-\epsilon}}{\sum_r (w_{rt}^s (\hat{u}_{irt}^{s'})^\beta)^\epsilon (\tau_{ir}^{s'})^{-\epsilon}} \cdot \xi_i^s \tag{19}$$

$$\hat{L}_{jt}^s = \sum_i \hat{L}_{ijt}^s \tag{20}$$

$$\hat{w}_{jt}^h = \frac{1}{\rho}[(A_j^h \hat{L}_{jt}^h)^\rho + (A_j^l \hat{L}_{jt}^l)^\rho]^{\frac{1}{\rho}-1} \rho (A_j^h \hat{L}_{jt}^h)^{\rho-1} A_j^h \tag{21}$$

$$\hat{w}_{jt}^l = \frac{1}{\rho}[(A_j^h \hat{L}_{jt}^h)^\rho + (A_j^l \hat{L}_{jt}^l)^\rho]^{\frac{1}{\rho}-1} \rho (A_j^l \hat{L}_{jt}^l)^{\rho-1} A_j^l \tag{22}$$

$$mi\hat{g\_stu}_{jt} = \sum_s \sum_{i \in \{i:i \neq j\}} \frac{exp(\hat{V}_{ijt}^s(Migchi))}{exp(\hat{V}_{ijt}^s(Migchi)) + exp(\hat{V}_{ijt}^s(Left))} \cdot \hat{L}_{ijt}^s \tag{23}$$

$$l\hat{b\_stu}_{it} = \sum_s \sum_{j \in \{j:j \neq i\}} \frac{exp(\hat{V}_{ijt}^s(Left))}{exp(\hat{V}_{ijt}^s(Migchi)) + exp(\hat{V}_{ijt}^s(Left))} \cdot \hat{L}_{ijt}^s \tag{24}$$

$$loca\hat{l\_stu}_{it} = \hat{L}_{iit}^h + \hat{L}_{iit}^l \tag{25}$$

In equation (19), we multiply the commuting probability from (16) in Section 4.4 with the number of workers holding Hukou in city $i$ to calculate the migration flow from city $i$ to city $j$. Then, we sum over $i$ to get the total number of workers in city $j$ by equation (20). This is the labor market clearing condition. The predicted wage levels in equation (21) and equation (22) are directly from firms' first order conditions (18) and (19) in Section 4.5. The number of migrant children is calculated by multiplying the probability that migrant workers take their children to migrate, by the number of migrant workers. Similarly, the number of left-behind children is calculated by multiplying the probability that migrant workers leave their children behind by the number of migrant workers. The number of local students with parents not migrating is calculated by equation (25). Then the predicted proportions of migrant and left-behind students in each city can be calculated by dividing the number of migrant and left-behind students by the total number of students.

Having these predicted values of the endogenous variables, we use the following updating rule to obtain the values of all variables for the next iteration:

$$\mathbf{\Delta_{t+1}} = \iota \mathbf{\Delta_t} + (1 - \iota)\hat{\mathbf{\Delta}}_{\mathbf{t}} \tag{26}$$

where $0 < \iota < 1$. We iterate until convergence is achieved, that is, $\frac{|\mathbf{\Delta_{t+1}} - \mathbf{\Delta_t}|}{|\mathbf{\Delta_t}|} < \delta$, where the numerator is the L-1 norm of the difference of the endogenous vectors at $t + 1$ and $t$. Then the last value of $\mathbf{\Delta}$ is the new equilibrium.[3]

[3]We set $\iota = 0.5$ and the convergence condition $\delta$ to be 0.01% in the main setting. We also check the robustness of the results when the convergence condition is set at different levels. There is almost no change at all.

### B.11 Counterfactual with Peer Effects Netting Out Family Background

In this section, we consider the counterfactual results if classmates' family backgrounds are considered in the peer effect regression to alleviate the external validity problem. The basic idea is to additionally include the proportion of classmates from high-skill families in the first stage schooling regression and then run the whole model with these parameters. Accordingly, we also need to update the proportions of migrant students from high-skill families in different cities in the calculation of the new equilibrium.

Table B6 shows the results of the first stage human capital equation estimation. The basic conclusions are similar to those in the main setting. Both migrant students and left-behind students have negative spillovers on their classmates. The negative effect from left-behind students is larger. The point estimates are reduced slightly when the proportion of peers from high-skill families is added. It is also obvious that students with better family backgrounds can positively affect their classmates.

Then, we re-calibrate the model and implement the main policy counterfactual to totally remove the enrollment restriction with these new parameters. The responses of migration, wages and human capital are shown in Table B7 and B8. All the conclusions are unchanged.

Table B6: Estimation of the Human Capital Equation: First Stage (External Validity Check)

| | (1) High-skill | (2) Low-skill |
|---|---|---|
| Proportion of Migrant Peers $\theta_1$ | -0.103 | -0.262 |
| | (0.701) | (0.265) |
| Proportion of Left-Behind Peers $\theta_2$ | -0.213 | -0.804** |
| | (0.609) | (0.352) |
| Proportion of High-Skill Peers $\theta_3$ | 0.427 | 0.837*** |
| | (0.312) | (0.304) |
| Whether is migrant student $\eta$ | -0.0848 | 0.0160 |
| | (0.0766) | (0.0353) |
| Whether is left-behind student $\nu$ | -0.0847 | -0.0565** |
| | (0.0640) | (0.0281) |
| School FE | YES | YES |
| Year Dummy | YES | YES |
| Personal Controls | YES | YES |
| Household Controls | YES | YES |
| Observations | 2,716 | 7,775 |
| R-squared | 0.373 | 0.317 |

Notes: The dependent variable is the standard test score. For column (1), we run the regression on students from high-skill families. For column (2), we run the regression on students from low-skill families. The set of personal controls includes student age, gender, Hukou type, whether he or she is a migrant student, whether he or she is a left-behind child, whether he or she is the only child, whether he or she lives at school, and his or her class rank in grade six. The set of household controls includes mother's education, father's education, and whether parents consistently have conflicts. All standard errors are clustered at the school level. Sources: China Education Panel Survey 2013 and 2014. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, and $^{\dagger} < 0.2$.

Table B7: Counterfactual Changes: External Validity Check

| Variables | Changes |
|---|---|
| **Panel. A** | |
| Total Migrants | 2.2% |
| Total High-skill Migrants | 6.6% |
| Total Low-skill Migrants | 1.8% |
| **Panel. B** | |
| Total Migrant Students | 8.9% |
| Total Students in Big | 3.4% |
| Total Students in Small | -0.4% |
| Total Students in Public in Big | 13.7% |
| Total Students in Public in Small | 0.3% |
| Ratio of Left-behind Students/Migrant | -9.3% |

Notes: "Big" means big cities and "Small" means small cities. In this model setting, we include the effect of the proportion of classmates from high-skill families in the human capital equation. We show the changes in migration when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

Table B8: Changes in Human Capital: External Validity Check

| Variables | Changes (Test Score s.d.) |
|---|---|
| Average HC | 0.0072 |
| Average HC of High-skill from Big | -0.012 |
| Average HC of Low-skill from Big | -0.036 |
| Average HC of High-skill from Small | 0.016 |
| Average HC of Low-skill from Small | 0.0089 |

Notes: HC stands for Human Capital. "Big" means big cities and "Small" means small cities. In this model setting, we include the effect of the proportion of classmates from high-skill families in the human capital equation. In the first column, we show the changes in migration when the public school enrollment probability of migrant children in each city is required to be at least 88%. In the second column, we show the changes in migration when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

### B.12 Equating Spillovers

By encouraging formerly left-behind students to migrate with their parents, we can reduce their spillovers. An important concern is that the smaller negative spillover from migrant children compared with left-behind children is due to selection, as we have explained in Section 6.1.4. To check this external validity issue, we set the peer effects of migrant students equal to those of left-behind children and then relax the enrollment restriction. By amplifying the peer effects of migrant children to equal those of left-behind children, we essentially consider the worst case where left-behind students cannot reduce their negative spillovers by reuniting with their parents.

The resulting human capital changes are shown in Table B9. In this setting of equalized peer effects, average human capital increases by 0.0065 (0.0035) standard deviations if we fully relax the enrollment restriction (alternatively, increase the probability floor to 88%). Compared with the increase of 0.0077 (0.0038) standard deviations when the non-spillover channel is allowed, most of the increase is retained. Thus, 85-90% of the policy's positive effects are from the non-spillover channel. The spillover channel is not negligible, but accounts for a relatively small part of the whole effect. As a result, our main conclusion is robust to the possible selection issue since there remains a large human capital gain even in the worst case if left-behind children do not create smaller negative spillovers when they become migrant children.

Table B9: Same Spillover from Migrant and Left-behind Children

| Variables | Human Capital Changes (Test Score s.d.) | |
|---|---|---|
| | 88% Floor | Total Removal |
| Average HC (Original) | 0.0038 | 0.0077 |
| Average HC (Spillover channel muted) | 0.0035 | 0.0065 |

Notes: HC stands for Human Capital. The first row shows the results from the main counterfactual. In the second row, we assume that migrant students' peer effect is equal to that of left-behind students. In this case, we mute the indirect channel because we cannot reduce the overall negative spillovers by simply reuniting left-behind children with their parents. In the first column, we show the changes in human capital when the public school enrollment probability of migrant children in each city is required to be at least 88%. In the second column, we show the changes in human capital when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

### B.13 Sensitivity of $\nu$ and $\eta$

In this section, we investigate the sensitivity of our main counterfactual results when the left-behind cost $\nu$ and the migrant cost $\eta$ in the human capital equation are changed. We check two extreme cases. The first case is when the left-behind cost is zero and the second case is when the migrant cost is zero. We re-calibrate the model in both cases and run the main counterfactual of removing the public school enrollment restriction on migrant children. The results are shown in Tables B10 and B11. The changes are small, which shows that our main results are not sensitive to changes in $\nu$ and $\eta$.

Table B10: Changes in Human Capital: Zero Left-behind Cost

| Variables | Changes (Test Score s.d.) |
|---|---|
| Average HC | 0.0081 |
| Average HC of High-skill from Big | -0.013 |
| Average HC of Low-skill from Big | -0.029 |
| Average HC of High-skill from Small | 0.021 |
| Average HC of Low-skill from Small | 0.0091 |

Notes: HC stands for Human Capital. "Big" means big cities and "Small" means small cities. In this setting, we assume a zero left-behind cost in the human capital equation. We show the results when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

Table B11: Changes in Human Capital: Zero Migrant Cost

| Variables | Changes (Test Score s.d.) |
|---|---|
| Average HC | 0.0077 |
| Average HC of High-skill from Big | -0.014 |
| Average HC of Low-skill from Big | -0.029 |
| Average HC of High-skill from Small | 0.019 |
| Average HC of Low-skill from Small | 0.0087 |

Notes: HC stands for Human Capital. "Big" means big cities and "Small" means small cities. In this setting, we assume a zero migrant cost in the human capital equation. In the first column, we show the results when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

### B.14 Changing the Timing of the Model

In this section, we change the timing of the model to check for robustness. We construct and estimate a model where the uncertainty in accessing public schools is revealed before parents decide on children's migration. That is, in this alternative setting, parents first choose where to migrate and then observe both their children's human capital shock and whether they can be enrolled in public schools in the migration destination. In this case, the expected value of children's human capital for workers who migrate changes to:

$$u_{ij}^{s} = exp[p_{j}^{s} E[max\{V_{ij}^{s}(public) + \upsilon_{oj}, V_{ij}^{s}(Left) + \upsilon_{oi}\}] +$$

$$(1 - p_{j}^{s}) E[max\{V_{ij}^{s}(private_{m}ig) + \upsilon_{oj}, V_{ij}^{s}(Left) + \upsilon_{oi}\}]]$$

We re-calibrate the model in this new setting and implement the main counterfactual. Table B12 shows that the results are not qualitatively changed.

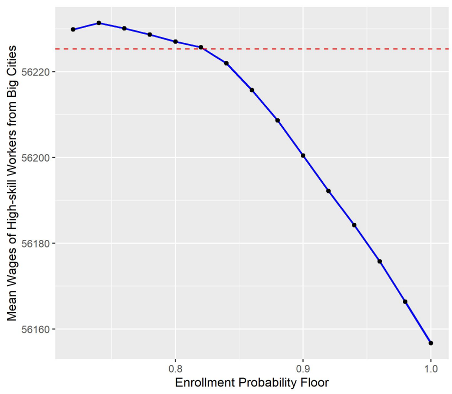
Table B12: Changes in Human Capital: Only Middle School Aged Children

| Variables | Changes (Test Score s.d.) |
|---|---|
| Average HC | 0.0050 |
| Average HC of High-skill from Big | -0.017 |
| Average HC of Low-skill from Big | -0.030 |
| Average HC of High-skill from Small | 0.014 |
| Average HC of Low-skill from Small | 0.0063 |

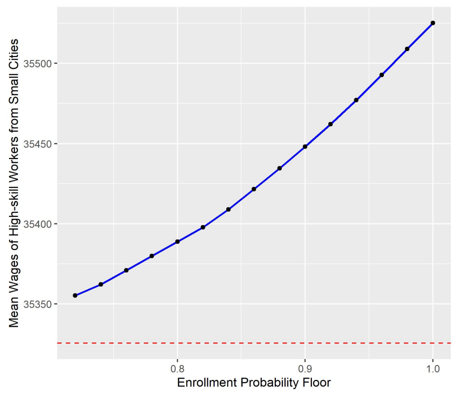
Notes: HC stands for Human Capital. "Big" means big cities and "Small" means small cities. In this setting, we assume that the uncertainty in public school enrollment is revealed before parents make their decision of whether to migrate with their children. We show the results when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

### B.15 Wages Changes in the Main Counterfactual

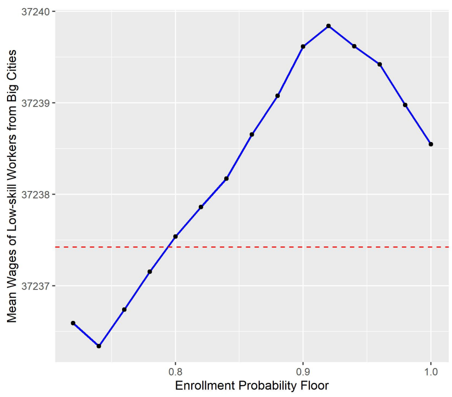
We also find some interesting patterns in how wages change in Figure B4. We look at average wages of people defined by their hometown (Hukou registered) cities. First, only high-skill workers from big cities are harmed. Both low-skill and high-skill workers from small cities enjoy higher average wages since more of them migrate to big cities where they can earn more. In addition, although low-skill families from big cities lose the most in terms of their children's human capital, they are partially compensated by the increase of their wages, which results from the inflow of high-skill migrants from small cities. In general, the magnitudes of the changes are small.



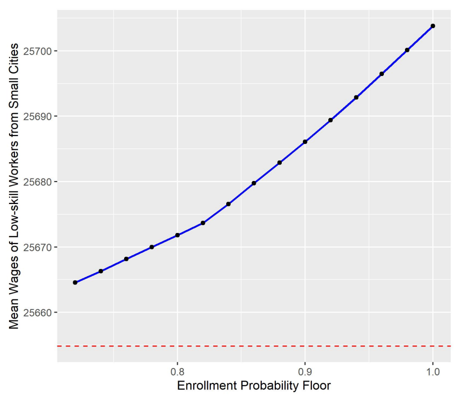

(a) Wages of High-skill from Big Cities




(b) Wages of High-skill from Small Cities




(c) Wages of Low-skill from Big Cities




(d) Wages of Low-skill from Small Cities

Figure B4: Wage Changes with Enrollment Probability Increase

Notes: The x-axis represents the enrollment probability floor the government sets for each city. The blue dots with a solid line show the changes in average wages for different groups of workers as enrollment restrictions are relaxed. The red dashed line represents the baseline level with no relaxation of the enrollment restriction. Subfigure (a) shows the average wages for high-skill workers from big cities. Subfigure (b) shows the average wages for high-skill workers from small cities. Subfigure (c) shows the average wages for low-skill workers from big cities. Subfigure (d) shows the average wages for low-skill workers from small cities. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

### B.16 Calculation Methods for Costs and Benefits

The public financial cost is calculated as follows. We assume that the government keeps the quality of public education unchanged and spends the same amount of money for each newly enrolled migrant child as for current local children. Financial data on Chinese compulsory education are only available at the province-level. We obtain the province-level education expenditure data in 2010 from the website of the Department of Education (MOE, 2011). Then in each equilibrium, we calculate total government expenditure by multiplying the number of students in public schools in each province by the expenditure per student in that province.

The calculation of pecuniary benefit is more difficult. There are two parts to the benefit. First, parents' utility gain. This can be recovered using a method similar to the one described in Appendix B.6. In brief, we can convert the increase in parents' utility to a willingness to pay measure. Second, children's gain in terms of future income. This part is complicated and we employ a method similar to Krueger (1999). First, we link test scores with each child's total years of education when they enter the labor market. Unfortunately, we do not observe the final year of education since our sampled students are in grade 7 and 8. Thus, we use parents' expected education for their children to proxy for total years of education and regress this on test scores. Second, we multiply the regression coefficient by the wage return to an additional year of schooling in China of 17.94% (Churchill and Mishra, 2018), finding that a one-standard-deviation increase in test score is associated with a 9.61% increase in annual income. This is very similar to the results in Krueger (1999). Third, given that average annual income for employed workers in China was 36,539 RMB in 2010, the annual income gain can be calculated by multiplying this figure by the 9.61% increase. Finally, we assume that workers work for 40 years and discount future income at a rate of 3 percent. This is a conservative calculation since wages have been growing very fast in China during the last 30 years.

### B.17 Additional Counterfactual Analysis

#### B.17.1 Separate But Equal

The second counterfactual policy we implement is similar to the notorious “separate but equal” policy of American history. In this case, we assume that the government takes over all private migrant schools and improves the quality of these schools to match public schools. However, children in private migrant schools will still be segregated from local children and the baseline public school enrollment probabilities are maintained. We assume that the cost of taking over private migrant schools and raising their quality to the level of public schools is identical to placing these children in public schools.

The impact of this policy on the directly treated group of families is simple. Since this only increases the quality of their schools without breaking segregation, their human capital gain is the public premium, which is 0.579 for high skill and 0.222 for low skill. The changes in national aggregate/average human capital are more complicated. We show the results in Table B13. The first column shows the changes in human capital for the "separate but equal" policy. The second column repeats the effects of removing all public school enrollment restrictions. The comparison shows that the average gain in human capital is about 20% smaller in the "separate but equal" case. However, the implementation of “separate but equal” policy may encounter less local political resistance since families from big cities are much less affected as they avoid part of the short-term negative spillovers due to segregation. We also implement a cost-benefit analysis as in the last section and find that the average gain and the average cost of this policy are 588 and 289 RMB, which gives a net average gain of 299 RMB per student. This is about 18% smaller than the net gain from the main counterfactual. As a result, we conclude that removing the enrollment restriction is a better policy, even when neglecting the deep moral failures of the "separate but equal" counterfactual.

An alternative interpretation of this policy counterfactual exists. A limitation of this study is our inability to differentiate between public schools within the same city due to data availability and the tractability of our model. Should we relax the enrollment restriction, most migrant children would likely enroll in public schools of lower quality with fewer local children. This could lead to 'segregation' within public schools as a result of school sorting. Therefore, this counterfactual analysis can also be considered a simulation of such a scenario, rather than an exact representation of a 'separate but equal' policy.

Table B13: Changes in Human Capital: Two Policies

| Variables | Changes (Test Score s.d.) | |
|---|---|---|
| | Separate But Equal | Total Removal |
| Average HC | 0.0065 | 0.0077 |
| Average HC of High-skill from Big | 0.0064 | -0.014 |
| Average HC of Low-skill from Big | -0.0014 | -0.029 |
| Average HC of High-skill from Small | 0.017 | 0.020 |
| Average HC of Low-skill from Small | 0.0052 | 0.0087 |

Notes: HC stands for Human Capital. "Big" means big cities and "Small" means small cities. In this table, we compare two different policies. In the first column, we show the national average human capital changes from the "separate but equal" policy. In the second column, we show the national average human capital changes when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

#### B.17.2 Long-run Peer Effects

In the main model, we use the peer effects estimates from the pooled regression in Table 7. However, the peer effects are different for students in their first and second years in middle school. Specifically, we find the negative peer effects from migrant students disappear in the second year. If this is the case, then we may underestimate the gains from the relaxation policy. Thus, we re-calculate the main counterfactual when the peer effects of migrant students are reduced to zero.

Table B14 shows that when there are no negative spillovers from migrant children as in the long-run, the average human capital gain is 0.0082 standard deviations, larger than in the main experiment. Furthermore, it is a Pareto improvement in terms of human capital. All students now benefit from the relaxation of the enrollment restriction. This result clearly shows the importance of helping migrant students adapt to their new lives as soon as possible. It reduces their negative spillovers, which benefits both migrant students and their classmates. This also implies that our estimation of the human capital gain in this static model is a very conservative one since we only use the average negative spillovers across the first two years and do not consider the decay of peer effects through time.

Table B14: Changes in Human Capital: Long-run Peer Effects

| Variables | Changes (Test Score s.d.) | |
|---|---|---|
| | 88% Floor | Total Removal |
| Average HC | 0.0041 | 0.0082 |
| Average HC of High-skill from Big | 0.0064 | 0.011 |
| Average HC of Low-skill from Big | 0.0035 | 0.0055 |
| Average HC of High-skill from Small | 0.0095 | 0.020 |
| Average HC of Low-skill from Small | 0.0032 | 0.0064 |

Notes: HC stands for Human Capital. "Big" means big cities and "Small" means small cities. In this counterfactual, we assume that migrant students' peer effect is zero in the long-run and increase the enrollment probability for migrant students in each city. In the first column, we show the changes in human capital when the public school enrollment probability of migrant children in each city is required to be at least 88%. In the second column, we show the changes in human capital when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

#### B.17.3 Decomposition of the Effect on Human Capital

There are three key channels through which relaxing enrollment restrictions can elevate children's human capital. First, more current migrant children gain access to public schools with higher quality. Second, this relaxation enables many left-behind children to migrate to developed regions, reunite with their parents, and reduce their overall negative spillovers. Third, it motivates parents who have remained in their hometowns to migrate to developed regions along with their children. In this section, we perform a simple decomposition to examine these channels.

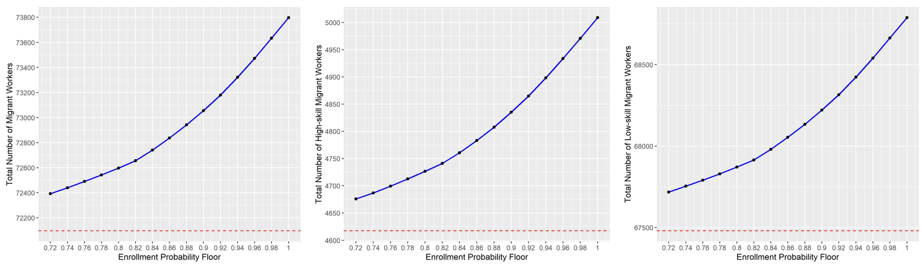
The first row of Table B15 shows the results from the main counterfactual, when the enrollment restriction is totally removed. Additionally, we analyze a scenario where parents remaining in hometowns are prohibited from migrating as a response to the policy relaxation, essentially disabling the third channel. Results, depicted in the second row, reveal a 26% reduction in policy effect when the third channel is incapacitated. Furthermore, we forbid any new migration, both of parents or children. Here, even existing migrant parents could not bring their left-behind children to their migration destinations in response to the policy relaxation, essentially obstructing the second and third channels. The results in the third row indicate a further 26% reduction in policy effect. Conclusively, the three channels contribute to the increase in human capital by 48%, 26%, and 26% respectively.

Table B15: Changes in Human Capital: Channel Analysis (Total Relaxation)

| Variables | Changes (Test Score s.d.) | |
|---|---|---|
| | Average HC | Relative Proportion |
| Main Setting | 0.0077 | 100% |
| No Parents Migration | 0.0057 | 74% |
| No Parents and Children Migration | 0.0037 | 48% |

Notes: HC stands for Human Capital. We consider the case of totally removing the enrollment restriction. In the first row, we show the changes in human capital when both parents and children are allowed to migrate as a response to the policy. In the second row, we show the changes in human capital when parents are not allowed to migrate as a response to the policy. In the third row, we show the changes in human capital when neither parents nor children are allowed to migrate as a response to the policy Sources: Census 2010, China Education Panel Survey 2013 and 2014.

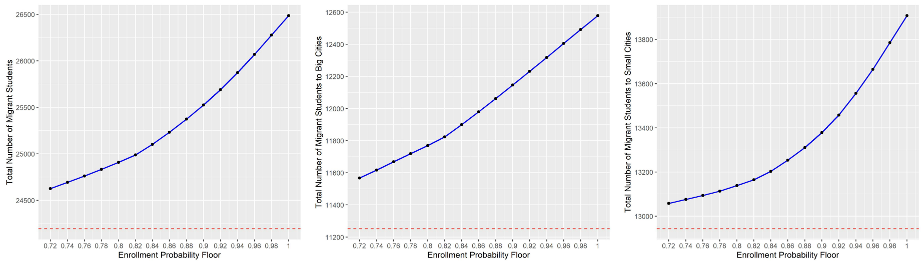
# C Additional Figures and Tables




(a) Total Migration (b) High-skill Migration (c) Low-skill Migration

Figure C1: Worker Migration Changes with Enrollment Probability Increase

Notes: The x-axis represents the enrollment probability floor the government sets for each city. The blue dots with a solid line show the changes in parents' (workers') migration as enrollment restrictions are relaxed. The red dashed line represents the baseline level with no relaxation of the enrollment restriction. Subfigure (a) shows the total migration of workers. Subfigure (b) shows the migration of high-skill workers. Subfigure (c) shows the migration of low-skill workers. Sources: Census 2010, China Education Panel Survey 2013 and 2014.




(a) Total Migration of Students (b) Migrant Students in Big Cities (c) Migrant Students in Small Cities

Figure C2: Student Migration Changes with Enrollment Probability Increase

Notes: The x-axis represents the enrollment probability floor the government sets for each city. The blue dots with a solid line show the changes in students' migration as enrollment restrictions are relaxed. The red dashed line represents the baseline level with no relaxation of the enrollment restriction. Subfigure (a) shows the total migration of students. Subfigure (b) shows the number of migrant students in big cities. Subfigure (c) shows the number of migrant students in small cities. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

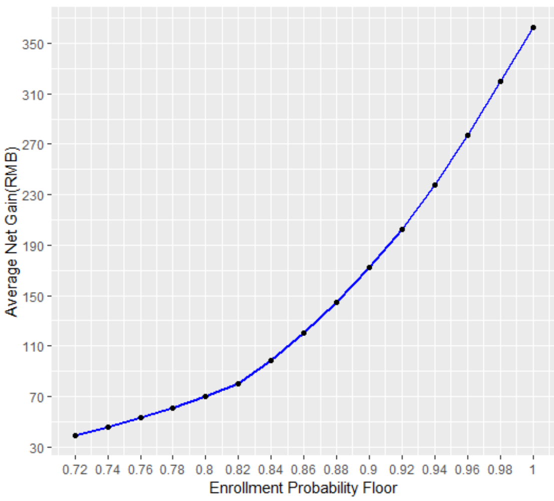



Figure C3: Average Net Gains with Enrollment Probability Increase

Notes: The x-axis represents the enrollment probability floor the government sets for each city. The dots with a solid blue line show the changes in average net gains as enrollment restrictions are relaxed. The net gains are calculated by summing parents' and children's gains, subtracting government's costs. It is measured by RMB. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

Table C1: Summary Statistics of Schools With/Without Random Assignment

| Variable | With Random | Without Random | Differences |
|---|---|---|---|
| Urban School | 0.632 | 0.615 | 0.0169 |
| | (0.484) | (0.490) | (0.0718) |
| Public School | 0.929 | 0.938 | -0.00943 |
| | (0.258) | (0.242) | (0.0374) |
| School Ranking | 3.819 | 3.969 | -0.149 |
| | (0.825) | (0.0.925) | (0.127) |
| Proportion of Migrant Students | 0.219 | 0.179 | 0.0405 |
| | (0.218) | (0.190) | (0.0310) |
| Proportion of Left-behind Students | 0.190 | 0.143 | 0.0467** |
| | (0.167) | (0.120) | (0.0229) |

Notes: In the first column, we show the mean values and standard deviations of different variables for schools with random assignment of students. In the second column, we show the mean values and standard deviations of different variables for schools without random assignment of students, which are dropped in the regression analysis. In the third column, we calculate the differences of the mean values for schools with/without random assignment and their corresponding standard errors. T-tests are also implemented in the third column. Sources: China Education Panel Survey 2013 and 2014.

Table C2: Calibrated Parameters

| Parameter | Description | Value |
|---|---|---|
| $\sigma$ | High/Low-skill labor elasticity of substitution | 1.4 |
| $\alpha$ | Family income effect on human capital | $8 \times 10^{-6}$ |
| $\epsilon$ | Dispersion of the Fréchet distribution | 1.5 |

Notes: This table displays the summary of calibrated parameters. The high/low-skill labor elasticity of substitution is calibrated using Katz and Murphy (1992). The family income effect on human capital is calibrated using Dahl and Lochner (2012). The dispersion of the Fréchet distribution in the unobserved location preference is calibrated using Tombe and Zhu (2019).

Table C3: Counterfactual Changes: Increasing Enrollment Probability for Migrant Students

| Variables | Changes | |
|---|---|---|
| | 88% Floor | Total Removal |
| **Panel. A** | | |
| Total Migrants | 1.2% | 2.4% |
| Total High-skill Migrants | 4.1% | 8.5% |
| Total Low-skill Migrants | 1.0% | 1.9% |
| **Panel. B** | | |
| Total Migrant Students | 4.9% | 9.5% |
| Total Students in Big | 2.1% | 3.4% |
| Total Students in Small | -0.24% | -0.4% |
| Total Students in Public in Big | 7.8% | 13.6% |
| Total Students in Public in Small | 0.0% | 0.3% |
| Ratio of Left-behind Students/Migrant | -5.3% | -9.8% |

Notes: "Big" means big cities and "Small" means small cities. In this counterfactual, we increase the enrollment probability for migrant students in each city. In the first column, we show the changes in migration when the public school enrollment probability of migrant children in each city is required to be at least 88%. In the second column, we show the changes in migration when the enrollment restriction is totally removed and all migrant students can enroll in public schools. Sources: Census 2010, China Education Panel Survey 2013 and 2014.

## References


Angrist, Joshua D. 2014. "The Perils of Peer Effects." *Labour Economics* 30:98–108.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* :841–890.

Black, Sandra E and Stephen Machin. 2011. "Housing Valuations of School Performance." In *Handbook of the Economics of Education*, vol. 3. Elsevier, 485–519.

Carrell, Scott E, Mark Hoekstra, and Elira Kuka. 2018. "The Long-run Effects of Disruptive Peers." *American Economic Review* 108 (11):3377–3415.

Case, Anne C and Lawrence F Katz. 1991. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." Tech. rep., National Bureau of Economic Research.

Chan, Jimmy, Xian Fang, Zhi Wang, Xianhua Zai, and Qinghua Zhang. 2020. "Valuing Primary Schools in Urban China." *Journal of Urban Economics* 115:103183.

Churchill, Sefa Awaworyi and Vinod Mishra. 2018. "Returns to Education in China: A Meta-Analysis." *Applied Economics* 50 (54):5903–5919.

Dahl, Gordon B and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *American Economic Review* 102 (5):1927–56.

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.

Gaviria, Alejandro and Steven Raphael. 2001. "School-based Peer Effects and Juvenile Behavior." *Review of Economics and Statistics* 83 (2):257–268.

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.

Katz, Lawrence F and Kevin M Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *The Quarterly Journal of Economics* 107 (1):35–78.

Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *The Quarterly Journal of Economics* 114 (2):497–532.

Li, Qiang, Wenbin Zang, and Lian An. 2013. "Peer Effects and School Dropout in Rural China." *China Economic Review* 27:238–248.

MOE. 2011. "2010 National Education Expenditure Implementation Statistical Announcement." http://www.moe.gov.cn/srcsite/A05/s3040/201112/t20111223_128871.html. Accessed: 2023-09-09.

Tombe, Trevor and Xiaodong Zhu. 2019. "Trade, Migration, and Productivity: A Quantitative Analysis of China." *American Economic Review* 109 (5):1843–72.

Wang, Haining, Zhiming Cheng, and Russell Smyth. 2018. "Do Migrant Students Affect Local Students' Academic Achievements in Urban China?" *Economics of Education Review* 63:64–77.