Journal of Development Economics 111 (2014) 181–195

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Does parental absence reduce cognitive achievements? Evidence from rural China☆

Hongliang Zhang [a,*], Jere R. Behrman [b], C. Simon Fan [c], Xiangdong Wei [c], Junsen Zhang [a]

[a] Department of Economics, The Chinese University of Hong Kong, Shatin, Hong Kong
[b] Department of Economics, University of Pennsylvania, McNeil 160, 3718 Locust Walk, Philadelphia, PA 19104-6297, United States
[c] Department of Economics, Lingnan University, Tuen Mun, Hong Kong

ARTICLE INFO

Article history:
Received 4 September 2013
Received in revised form 16 September 2014
Accepted 20 September 2014
Available online 2 October 2014

Keywords:
Left-behind children
Migrating parents
Cognitive achievement
China

ABSTRACT

Many children worldwide are left-behind by parents migrating for work — over 61 million in rural China alone, almost half of whom are left-behind by both parents. While previous literature considers impacts of one parent absent on educational inputs (e.g., study time, enrollment, schooling attainment), this study directly investigates impacts on children's learning (test scores) and distinguishes impacts of absence of one versus both parents. Dynamic panel methods that control for both unobserved individual heterogeneity and endogeneity in parental absence are used with data collected from rural China. The estimates indicate significant negative impacts of being left-behind by both parents on children's cognitive development, reducing their contemporary achievements by 5.4 percentile points for math and 5.1 percentile points for Chinese, but much smaller insignificant impacts of being left-behind by one parent. Cross-sectional evidence indicates that only absence of both parents is associated with substantially lower family inputs in after-school tutoring.

1. Introduction

One in three children under age 17 in rural China is living without one or both parents who have migrated in search of work in cities. Almost half of these children have been left-behind by both parents.[1] Despite its degree and scale, the "left-behind children" phenomenon in China remains understudied because of both theoretical ambiguities and empirical challenges. The existing literature has long highlighted the various channels through which parental migration can affect the human capital development of children left-behind (e.g., Dustmann and Glitz, 2011; Haveman and Wolfe, 1995; Stark, 1993). On the one hand, parents increase their earnings through migration and remittances of these earnings can ease the household budget constraint and thereby increase household spending on education and reduce child labor.[2] This theoretical prediction has been empirically supported by studies on the effect of remittances from migrants on children left-behind in El Salvador, Guatemala, Mexico, the Philippines, and some Pacific countries (Tonga and Vanuatu).[3] On the other hand, parental migration inherently leads to parental absence from home, which can have negative effects on children left-behind through channels such as the loss of local earnings and labor, the lack of parenting inputs, and the psychological costs associated with family separation.[4] Moreover, parental migration also increases the migration prospects of children and can induce more or less educational investment in children depending on the difference in the rates of return to human capital between the migration destination and the place of origin (Beine et al., 2008). Therefore, the sign of the overall effect of parental migration on the education of children left-behind is a priori unclear and remains an empirical question.

Recently, there is a growing empirical literature that examines the effects of parental migration on the outcomes of children left-behind, focusing mainly on dimensions of time allocation and schooling attainment.

☆ We thank Shuaizhang Feng, John Ham, Chih-Sheng Hsieh, Lars Lefgren, Hyungsik Roger Moon, and seminar participants at the 2013 Asian Conference on Applied Micro-Economics/Econometrics and the 2014 Asian Meeting of the Econometric Society for valuable comments and discussions. We are also grateful to the Co-editor and two anonymous referees for helpful comments and feedback. We acknowledge funding support provided by the Hong Kong Research Grants Council General Research Fund (No. 458910), Chinese University of Hong Kong South China Program, and Grand Challenges Canada Grant 0072-03. We also thank the Educational Bureau of Longhui County in China for administrative support, and Fengjiao Chen, Vincent Chow, and Jia Wu for excellent research assistance.

* Corresponding author. Tel.: +852 3943 8233.
E-mail addresses: hongliang@cuhk.edu.hk (H. Zhang), jbehrman@econ.upenn.edu (J.R. Behrman), fansimon@ln.edu.hk (C.S. Fan), xdwei@ln.edu.hk (X. Wei), jszhang@cuhk.edu.hk (J. Zhang).

[1] See Section 2.1 for the source of these figures.

[2] For a survey of the remittances literature, see Rapoport et al. (2006) and Yang (2011).

[3] See, for example, Edwards and Ureta (2003), Yang (2008), Adams and Cuecuecha (2010), Alcaraz et al. (2012), Clemens and Tiongson (2012), and Gibson and McKenzie (2014).

[4] See Antman (2013) for a survey of the literature on the adverse effects of parental absence.

http://dx.doi.org/10.1016/j.jdeveco.2014.09.004
0304-3878/

Antman (2011), for example, finds that Mexican fathers' migration to the U.S. decreases study hours and increases work hours for children left-behind. Chang et al. (2011) and Chen (2013) both employ the China Health and Nutrition Survey (CHNS) to examine the time allocation of left-behind children in China, and find that children of migrant households spend more time in household work.[5] In contrast to the consistent results on children's time allocation, findings on children's schooling attainment are mixed, even among studies of the same context. For example, in studying the impact of Mexican emigration to the U.S. on children's schooling attainment, McKenzie and Rapoport (2011) find a negative effect of living in a migrant household on schooling of older children left-behind, while Antman (2012) reports a positive effect of paternal migration on the schooling attainment for girls.[6] In a study of the impact of New Zealand's Recognized Seasonal Employer (RSE) program on Pacific countries, Gibson and McKenzie (2014) find that the seasonal migration of some household member(s) has a large positive impact on school attendance for 15–18 year olds in Tonga, but no impact for children at any schooling stage in Vanuatu.[7,8]

This paper examines the immediate net impacts of parental migration on the cognitive achievements of their children left-behind in rural China. It makes two important contributions to the existing literature. First, we are able to distinguish between absences of one versus both parents and estimate their effects separately because of the large numbers of children in both categories. Clearly, the learning implications for children of a family structure with one parent at home can be drastically different from that with neither parent at home. The phenomenon of "left-behind children" elsewhere almost entirely refers to cases in which only one parent (usually the father) is absent from home; in contrast, in rural China, often both parents migrate simultaneously. If paternal and maternal inputs are closer substitutes in the production of child human capital than those of other caregivers, such as grandparents, the absence of both parents has larger impacts on child human capital and may deserve greater policy attention than the commonly considered cases of the absence of one parent. Nonetheless, research on the impacts of migration of both parents on children left-behind is very rare. As for the two previous studies on left-behind children in China, Chen (2013) limits her analysis to households with fathers away from home only, and Chang et al. (2011) restrict the effects to be linear in the number of parents away.

Second, we are among the first to examine the impacts of parental absence on children's learning outcomes measured by test scores. Most previous research is limited to inputs into education such as study time, school enrollment, and schooling attainment. This distinction is important because parental absence may have very different impacts on the time students devote to education than on what they learn if parents provide inputs that are complements to students' own efforts. Therefore, it is possible that parental absence generates positive impacts on the time children spend on education but still negative impacts on their learning outcomes. To the best of our knowledge, the only published paper that examines the effect of parental absence on children's academic achievement is in the context of military deployments (Lyle, 2006), which finds a tenth of a standard deviation decline in the test scores of enlisted soldiers' children.

The data used in this paper were collected by the authors from Longhui County in Hunan Province of China. The county was selected for this study to represent the country's poorest rural areas with a high prevalence of parental absence: per capita GDP is less than a quarter of the national average and over two-thirds of children are left-behind by at least one parent. Working with the county's educational bureau, we randomly selected over 5000 third to fifth graders (9 to 11 year olds) enrolled in the county's primary schools, and collected longitudinal information on their parental absence status and test scores in math and Chinese for every school term since their enrollment in the first term of grade 1. The identification strategy used in this paper follows the same spirit as Andrabi et al. (2011), who apply dynamic panel methods to evaluate the effect of private schooling on student achievement in Pakistan. To address the possibility that the contemporary parental migration status and child outcomes are shaped by common past factors such as genetics and experience, we adopt a value-added specification of human capital accumulation to control for the impacts of all historical schooling inputs and heritable endowments on current child outcomes. We further include child fixed effects in the value-added model to control for unobserved individual heterogeneity in learning. That is, for children whose parents' migration status changes over time, we identify the effect of parental absence by comparing a child's *achievement progress* in periods with parental absence to his/her *achievement progress* in periods without parental absence. Finally, even after controlling for lagged achievement and individual heterogeneity in achievement progress, changes in parental absence status may still be correlated with changes in the time-varying component of the unobserved determinants of learning. To further address this concern, we employ a GMM framework and instrument changes in parental absence with its longer lags, and explore the robustness of our results to a range of persistence parameters.

Our estimates show significant adverse effects of the absence of both parents on the cognitive achievements of children left-behind, reducing their contemporary test scores by 5.4 and 5.1 percentile points in math and Chinese, respectively. However, we find that the effects of the absence of a single parent, though still negative, are much smaller and insignificant, suggesting that there may be a high degree of substitution between fathers and mothers in educating children. That is, when only one parent is away, the remaining parent may assume the roles of both in terms of educating their children, resulting in little reduction in family inputs on children's education. This hypothesis is also consistent with the cross-sectional evidence that only the absence of both parents is associated with substantially lower family inputs in after-school tutoring. As previous research suggests a key role of the persistence parameter in estimating the value-added achievement function, in addition to estimating the persistence parameter empirically, we also allow it to be exogenously assigned, varying from 1.0 to 0.4 with decrements of 0.2. Our conclusion that only the absence of both parents has significant adverse effects on children's cognitive achievement is robust to this variation in the persistence parameter value. Furthermore, we also conduct sensitivity/robustness analysis using alternative classifications of parental absence status, sample selection rules, and achievement measures, and obtain similar results in all these exercises. Our results suggest that the absence of both parents, which is quite common in rural China, is a much more serious problem in shaping the educational outcomes of the next generation than the usually considered cases elsewhere of the absence of a single parent, and therefore deserves greater policy attention.

The remainder of this paper is organized as follows. Section 2 provides the background of the "left-behind children" phenomenon, describes our sample, and provides the "first-look" evidence of the relationship between parental absence status and children's cognitive achievement

[5] Chen (2013) also finds that mothers spend less time in both household and income-generating activities after fathers' migration, and interprets the findings as mothers' non-cooperative behaviors as a result of fathers' imperfect monitoring after migration.

[6] McKenzie and Rapoport (2011) match the observed decrease in schooling to increased housework for girls and migration for boys, and link the latter to the increased migration prospects of boys and lower returns to schooling in the U.S. for Mexican migrants. Antman (2012) interprets the differential effects of paternal migration by gender as the results of increased bargaining power for mothers who spend the marginal dollars on the education of girls.

[7] Gibson and McKenzie (2014) attribute the divergent impacts between Tonga and Vanuatu to their differences in both the nature of selection into migration and schooling fee policies. The RSE households are relatively better off in Vanuatu. In addition, many schools in Vanuatu also allow students to remain enrolled even with unpaid fees from previous years.

[8] There are also studies on the effects of the migration of family members other than parents on children's education. For example, Kuhn (2007) shows that the migration of brothers was associated with improvements in children's pace of school completion in Bangladesh, while the migration of sisters was not. Gibson et al. (2011) use a migration lottery program to study the effects of the permanent emigration of some household member(s) from Tonga to New Zealand on the remaining household members and find insignificant impact of the migration of some household members, typically uncles and aunts, on children's school enrollment and schooling attainment.

based on a difference-in-differences strategy. Section 3 introduces our empirical framework, which employs dynamic panel methods to evaluate the effects of parental absence on children's cognitive achievements in a value-added achievement function. Section 4 presents our empirical results. Section 5 concludes.

2. Background and data

2.1. China's rural-to-urban migration and left-behind children

China started its household registration system, commonly known as the *Hukou* system, in the mid-1950s. Under this system, rural residents were barred from entering cities to look for jobs from the mid-1950s to the end of 1970s. The main purpose of this system was to stem a flood of rural migrants to cities that was feared would paralyze the infrastructure and cause huge social problems there. However, since China embarked on its economic reform in the late 1970s, there has been a rising demand for cheap labor in cities. The government began to gradually relax its control of the *Hukou* system and allowed rural residents to come to cities to work. Ever since the relaxation of this system in the beginning of the 1980s, hundreds of millions of rural migrant workers have gone to cities to find jobs. According to the latest *2013 Investigational and Monitoring Report of Chinese Migrant Workers* by the National Bureau of Statistics, there were an estimated 161 million[9] rural migrants employed outside their home area for a period of over six months, of which over 130 million were individual migrants who left rural family members behind.[10] Such a huge wave of rural-to-urban migration is unprecedented and has been called the largest peace-time migration in history (Roberts et al., 2004).

Despite the Chinese government's decision to allow rural residents to work in cities, it has not dismantled its *Hukou* system. The rural migrant workers are still being treated as "second-class" citizens and are usually without entitlements for city welfare including free public education for their children (Chen and Feng, 2013). As a consequence, the majority of migrant parents choose to leave their children behind in their home townships/villages, leading to a huge left-behind children phenomenon in the countryside. According to the All-China Women's Federation's (ACWF, 2013) report based on the 2010 Population Census, there were over 61 million children aged 17 years or below left-behind by one or both parents in the countryside, of which 46.7% were left by both parents. These left-behind children accounted for 37.7% of rural children and 21.9% of all children in China.

2.2. Data description

This study uses data collected by the authors in Longhui County of Hunan Province in Central China, which is among the poorest counties in the nation. According to ACWF (2013), Hunan is also among the six provinces in China with more than half of the rural children left-behind by one or both parents.[11] The population of the county was about 1.2 million in 2011 and 90% of the population were rural residents. The county was selected to represent the country's poorest rural areas. It has been designated as a national poverty county since 1994.[12] In 2010, the county's per capita GDP was RMB 6922, less than a quarter of the national average of RMB 29748. Another feature that Longhui shares with many other poor counties in China is the high prevalence of left-behind children. Among its students in primary (grades 1–6) and middle schools (grades 7–9), at any recent time over two-thirds have at least one parent away from home (working elsewhere), of which more than half have both parents absent. According to the statistics for year 2011 from the county's educational bureau, enrollment was universal (100%) for primary school and almost universal (99.2%) for middle school. However, less than three-quarters of students (74.7%) proceeded to secondary school (grades 10–12) after completing the nine-year compulsory schooling. Among those who continued secondary school, 39% attended one of the county's two elite high schools (Nos. 1 and 2 High Schools), 32% attended a regular high school, and the remaining attended a vocational secondary school. In 2011, while the overall college admission rate[13] (i.e., tier 3 colleges or above) was 60.6% among college entrance exam takers,[14] the admission rates for tier 1 colleges and tier 1 and 2 colleges combined were only 7.8% and 30.2%, respectively. However, for the two elite high schools, the admission rate for tier 1 colleges (12.8%) was more than 3.6 times that of the regular high schools (3.5%), and the admission rate for tier 1 and 2 colleges combined (43.1%) was more than 2.8 times that of the regular high schools (18.1%). This suggests that it is important for students to be admitted to the two elite high schools in order to gain access to higher education. Under the Chinese education system, admission to elite high schools is by and large based on students' test scores on the Middle School Exit Exam, and so it matters a lot for students at the margin whether their test scores (or class rank) can be improved.

Working with the Educational Bureau of Longhui County, we randomly selected 23 primary schools in this county, including 20 schools within the jurisdictions of the randomly selected five townships (out of a total of the county's 25 townships excluding the county seat) and three randomly selected schools (out of a total of 11 schools) in the county seat, and collected both administrative and survey information on all third to fifth graders enrolled in April 2011. Fig. 1 shows the locations of the selected townships and schools in this county (in the main map), as well as the county's location in the province and the province's location in the nation (in the overview maps). We choose to focus our study on third to fifth graders in primary school for three reasons. First, primary school students probably are more vulnerable to parental absence than are older students. Second, grade 3–5 students have sufficiently long past academic achievement records for us to build a panel data on student academic achievements. Third, many middle schools in rural areas provide boarding facilities for needy students, therefore making it hard to isolate the impact of parental absence on student learning in middle school.

Our sample consists of over 5000 students, accounting for roughly one-fifth of all students enrolled in grades 3 to 5 in the study county. The data were collected from three sources. First, a student information sheet was filled out by the master teacher of each class, with information from the school's administrative records on each student's final exam scores in math and Chinese for every school term[15] since the student enrolled in the school. These exams were designed by either the township's school board or each individual school to assess students' end-of-term mastery of academic subjects covered in each school term. Although raw scores from different schools are not directly comparable due to differences in the exam contents, they are still informative about students' relative performances within the same school. We therefore utilize a student's percentile rank in a class as a relative measure of his/her cogni-

[9] This figure does not include 103 million engaged in off-farm employment within their hometowns for a period of over six months.

[10] The onset of the global financial crisis in 2008 had some severe adverse impact on off-farm employment of China's rural labor force and is estimated to have resulted in lay-offs of 49 million rural workers (Huang et al., 2011). The impact, however, seems to have been short-lived and did not change the trend of rising off-farm employment of China's rural labor force. For comparison with the numbers in the text for 2013, in 2008, there were an estimated 140 million rural migrants employed outside their home area, of which 112 million were individual migrants.

[11] The other five provinces are Anhui, Chongqing, Jiangsu, Jiangxi, and Sichuan.

[12] It was designated as a provincial poverty county in 1986 when China launched its first poverty alleviation program, and was upgraded to a national poverty county in 1994 in the second wave. In the fourth wave of national poverty alleviation program announced in 2012, there are a total of 592 national poverty counties in 21 provinces. For further discussion of the poverty alleviation program and the designation of poverty counties in China, see Park et al. (2002) and Meng (2013).

[13] This statistic does not include three-year colleges offering associates degrees.

[14] Note that the college entrance exam takers in a year include both high school graduates in that year and repeated takers who finished high school previously. Because the repeated takers are much more likely to enroll in a regular high school than an elite high school, regular high schools are overrepresented in college entrance exam takers relative to their enrollment size at high school admission.

[15] In China, a school year is divided into two school terms: Fall term (September to January) and Spring term (February to June).

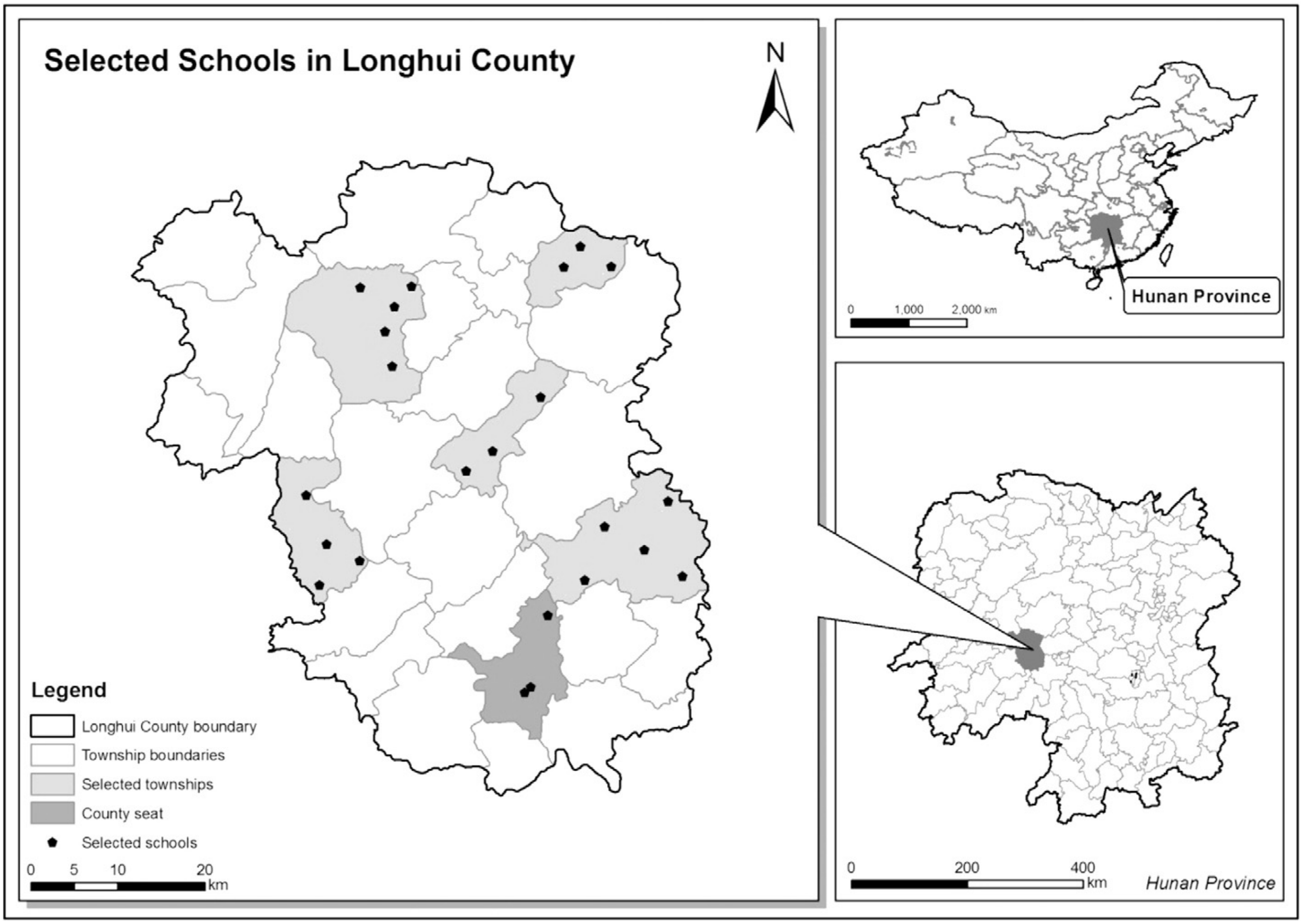

Fig. 1. Sample townships and sample schools in Longhui county.

tive performance within a class.[16] Second, a student survey was conducted in class under the instruction of our surveyors, collecting students' family and personal information, including time allocation after school, whether parents or others helped with study, and self-reported satisfaction levels on their relationships with parents, other family members, teachers, and classmates. Third, a household survey was completed by students' parents or primary caregivers when both parents were absent from home,[17] collecting information on family composition, parents' ages, schooling attainment, migration status, main economic activities, and incomes. Very important for this study, we asked parents or caregivers in the absence of parents to report retrospectively the paternal and maternal migration status since the child started school. Specifically, for every school term since the first term of grade 1, each parent indicated (or was indicated by a proxy) one of the following four options that most closely described his/her residence/migration status: (i) was always present at home; (ii) was present at home for more than half of the time but not always; (iii) was absent from home for more than half of the time but not always; (iv) was always absent from home.

We then linked the history of parental migration status from the household survey to the longitudinal test scores from the student information sheet to construct a panel for up to five, seven, and nine school terms for the third, fourth, and fifth graders, respectively. We impose a set of restrictions to confine our analysis to students with complete histories of parental migration status and test scores. First, we exclude 534 students (~10.1%) who started schooling elsewhere and later transferred to the current schools.[18] Second, we exclude 114 students (~2.2%) with missing test score information in one or more school terms. As primary school enrollment is almost universal in China,[19] missing test scores were mainly due to absences on the exam days or temporary transfers to other schools. Third, we exclude 61 students (~1.1%) with incomplete parental information, including migration history and marital status. Finally, we further exclude 70 students (~ 1.4%) who have one or both parents deceased. However, we keep in our sample 123 students (~2.3%) who reported having one parent permanently missing due to divorce because of our concern that selecting the sample based on parents' marital status may lead to biased estimates if parental divorce is an outcome of their separation due to migration (see Section 4.4 for more detailed discussion). Thus, our final data set consists of 4528 students – 1668 third graders, 1536 fourth graders, and 1324 fifth graders – for whom the full histories of parental migration status and child test scores are available since the beginning of primary school.

Table 1 shows the summary statistics for students in our final sample as well as four subsamples defined by students' current parental

[16] Specifically, a student's percentile rank in a particular term is calculated according to the raw scores of all his/her *current* classmates in the survey term, which is also the comparison group for all previous terms.

[17] The primary caregivers were asked to verify the information with students' parents by phone when filling out the family survey.

[18] The majority of these students started schooling in satellite school units in remote villages and later transferred to the primary schools we surveyed. These satellite school units offer junior (usually first and second) grade classes only.

[19] According to China's Ministry of Education, the primary school enrollment rate was 99.5% in 2008. http://www.china.org.cn/government/scio-press-conferences/2009-09/11/content_18508942.htm.

Table 1
Summary statistics by current parental absence status.

	All	Both parents at home	Only mother at home	Only father at home	Both parents absent
Percentage share of the full sample	100.00%	28.9%	27.4%	4.7%	38.9%
Primary caregivers when both parents were absent					
Paternal grandparents	–	–	–	–	66.3%
Maternal grandparents	–	–	–	–	15.8%
Older siblings	–	–	–	–	1.9%
Other close relatives (uncles, aunties, cousins)	–	–	–	–	6.6%
Others (friends, neighbors, etc.)	–	–	–	–	9.4%
Number of children in family	2.09	2.06	2.16	2.05	2.06
Proportion of children w/ no sibling	0.164	0.177	0.117	0.243	0.177
Proportion of children w/ two or more sibling	0.207	0.191	0.233	0.210	0.200
Proportion of children w/ at least one older sibling	0.458	0.503	0.489	0.444	0.404
Paternal years of completed schooling	8.99	9.31	8.98	8.54	8.81
Maternal years of completed schooling	8.19	8.38	8.06	8.08	8.15
Median household income in the previous year (RMB)	30000	26000	30000	29000	40000
Number of students	4528	1310	1243	214	1761

Notes: The number of observations used to calculate the median household income is 4040, smaller than what is reported in the bottom row due to missing data.

absence status. Unless otherwise noted in Section 4.3.2, we define throughout the paper a parent as absent from home in a school term if he/she was away from home for more than half of the time during a school term. For the school term when the survey was conducted, only 29% of the children in our sample had both parents present at home. Of the 71% of the children who were left-behind by at least one parent, 39% had both parents absent from home, 27% had only mothers at home, and the remaining 5% had only fathers at home. For children with only one parent at home, in 85% of the cases the parent was the mother. Therefore in most, but not all, cases having only one parent at home meant having the mother at home and the father absent. For children with both parents away from home, 66% had paternal grandparents as the primary caregivers, 16% had maternal grandparents as the primary caregivers, and the remaining were in the care of other close relatives, friends, or neighbors. It turns out that the One-Child Policy was not very strictly enforced in the county under study. Families in our sample on average have 2.1 children. Less than one-sixth of these families have only one child, whereas in contrast more than one-fifth of them have three or more children. Also, 46% of children in our sample have at least one older sibling, who is a major source of after-school tutoring for their younger siblings as later discussed in Section 4.5. This ratio is highest among students with both parents at home (0.503) and lowest among those with both parents absent (0.404). Fathers and mothers in our sample on average have 9.0 and 8.2 completed grades of schooling, respectively. Among the four subsamples defined by parental absence status, families with both parents at home have the highest paternal (9.3) and maternal completed grades of schooling (8.4), whereas families with only fathers/mothers at home have the lowest paternal/maternal completed grades of school (8.5/8.1). We also report the median household income of each subsample based on households reporting non-zero income for the previous year.[20] Not surprisingly, the median annual household income was highest among families with both parents away as migrant workers (RMB 40000) and lowest among those with both parents staying at home (RMB 26000). Because of the small number of children with only fathers at home, we consolidate maternal absence (only) and paternal absence (only) into a single category of having (only) one parent absent from home for the remaining of the paper, unless otherwise noted in Section 4.3.2.

2.3. Difference-in-differences estimates

In the dynamic panel model used in this paper, identification is based on the 1419 students whose parental absence status changed during the observation period. Table 2 illustrates the working of this identification strategy by employing a difference-in-differences strategy to compare children's test score changes over two periods when parental absence status varied. We subdivide transitions in parental absence status into six categories by changes in the number of parents at home, as illustrated in column (1) of Table 2. For children falling into each category, the difference-in-differences estimation compares the means of their test score changes between the period immediately before and the period immediately after the transition took place. Taking categories (1a) and (1b) as examples, among the 288 cases in which a child had transitioned from having both parents at home to having both parents away, the migration of both parents is associated with 3.0 and 2.6 percentile-point reductions in test score changes for math and Chinese, respectively, whereas among the 252 cases where a child had transitioned in the opposite way from having both parents away to having both parents at home, the return of both parents is associated with 2.2 and 3.1 percentile-point increases in test score changes for math and Chinese, respectively. However, only the difference for Chinese associated with the transition from the absence of both parents to the presence of both parents is significant at the 10% level. In column (6), we pool these two categories together and examine instead the difference in the means of test score changes when both parents were away versus when both parents were present at home. This exercise shows that the absence of both parents is associated with 2.6 and 2.8 percentile reductions in test score changes for math and Chinese, respectively, compared to the presence of both parents. Because of the reductions in standard errors with sample pooling, the coefficient estimate in this case is significant at the 10% level for math and at the 5% level for Chinese. In another exercise comparing the means of test score changes when both parents were away and when only one parent was away, we also find the difference to be negative and significant for both subjects. However, the comparison between the absence of a single parent and the presence of both parents shows smaller and insignificant differences.

While the intuition is clear, the consistency of the difference-in-differences estimates hinges on two critical assumptions, which are rather restrictive and may not necessarily hold in practice. First, it adopts a canonical restricted version of the value-added model that assumes perfect persistence of lagged achievement. However, it has been widely acknowledged in the literature that achievement exhibits mean reversion, which can be more salient in our context as these exams mainly attempted to assess students' understanding of subjects covered in each school term and varied substantially in content over time and across grades. Second, the difference-in-differences estimation requires that changes in a child's parental absence status are exogenous to changes in the unobserved time-varying determinants of achievement progress. However, in reality, parents may change their

[20] 89% of the households in our survey report non-zero income. The proportions reporting zero income are roughly the same different subsamples.

Table 2
Means and differences in test score changes for children whose parental absence status varied.

	Change in number of parents at home from $t-1$ to t	Number of observations	Mean test score change in $t-1$	Mean test score change in t	Difference in test score changes between $t-1$ and t	Difference in test score changes by number of parents at home
	$n_{t-1} \rightarrow n_t$		Δy_{t-1}	Δy_t	$\Delta y_t - \Delta y_{t-1}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Math						
(1a)	$2 \rightarrow 0$	288	1.159	−1.831	−2.990	
			(1.041)	(1.167)	(1.919)	$\Delta y_{(n=0)} - \Delta y_{(n=2)}$
(1b)	$0 \rightarrow 2$	252	−2.402**	−0.247	2.155	−2.572*
			(1.206)	(1.119)	(1.897)	(1.356)
(2a)	$1 \rightarrow 0$	396	0.812	−0.536	−1.348	
			(0.806)	(0.888)	(1.351)	$\Delta y_{(n=0)} - \Delta y_{(n=1)}$
(2b)	$0 \rightarrow 1$	420	−1.302	1.367*	2.669**	−2.008**
			(0.804)	(0.820)	(1.325)	(0.947)
(3a)	$2 \rightarrow 1$	505	0.315	−0.011	−0.326	
			(0.764)	(0.752)	(1.250)	$\Delta y_{(n=1)} - \Delta y_{(n=2)}$
(3b)	$1 \rightarrow 2$	400	0.889	0.981	0.092	−0.209
			(0.781)	(0.795)	(1.240)	(0.893)
Panel B. Chinese						
(1a)	$2 \rightarrow 0$	288	1.159	−2.008*	−3.167**	
			(1.041)	(1.057)	(1.563)	$\Delta y_{(n=0)} - \Delta y_{(n=2)}$
(1b)	$0 \rightarrow 2$	252	−3.259***	−0.181	3.078*	−2.821**
			(1.026)	(1.165)	(1.780)	(1.257)
(2a)	$1 \rightarrow 0$	396	1.017	−0.477	−1.493	
			(0.838)	(0.879)	(1.425)	$\Delta y_{(n=0)} - \Delta y_{(n=1)}$
(2b)	$0 \rightarrow 1$	420	−1.711**	1.998**	3.709***	−2.601***
			(0.824)	(0.853)	(1.384)	(0.993)
(3a)	$2 \rightarrow 1$	505	0.209	−0.935	−1.143	
			(0.714)	(0.758)	(1.201)	$\Delta y_{(n=1)} - \Delta y_{(n=2)}$
(3b)	$1 \rightarrow 2$	400	−0.133	1.062	1.195	−1.169
			(0.732)	(0.798)	(1.298)	(0.888)

Notes: Cells contain the mean test score changes in columns (3) and (4), the differences in the test score changes from t-1 to t in column (5), and the differences in test score changes by the number of parents at home in column (6). Standard errors are in parentheses.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

residence/migration status in response to unobserved family-level shocks, which could also affect their child's learning progress, raising concern over the consistency of the difference-in-differences estimates. To address both threats to the consistency of the difference-in-differences estimates, we adopt in the next section a more generalized empirical framework in which both imperfect persistence of past learning outcomes and endogeneity of changes in parental absence status are allowed.

3. Empirical framework for learning dynamics

We adopt an educational production function approach and consider children's knowledge acquisition as a cumulative process that depends on the histories of family and school inputs as well as on children's inherited endowments. Following Boardman and Murnane (1979) and Todd and Wolpin (2003), we model the cumulative production of cognitive achievement as a dynamic process as follows:

$$y_{it}^{*} = \sum\nolimits_{s=0}^{t} \left(X_{is}\alpha_{t-s} + \theta_{t-s}\varepsilon_{is}\right), \tag{1}$$

where y_{it}^{*} is the *true* cognitive achievement for child i at the end of period t (measured without error), X_{is} is a vector of observed family and school inputs applied to child i in period s, and ε_{is} is the unobserved determinants (including both inputs and endowments) affecting child i's learning in period s. α_{t-s} and θ_{t-s} correspond to, respectively, the impacts of the observed and unobserved factors applied $t-s$ periods prior to the time of assessment on a child's cognitive achievement. In particular, we normalize ε_{it} as the unobserved contemporary component of y_{it}^{*} such that θ_0 always equals unity. Underlying the educational production function in Eq. (1) is the assumption that the input coefficients depend only on the lead time when the inputs are imposed relative to when the achievement is measured (i.e., $t-s$).[21] Further assuming that all input coefficients, both observed and unobserved, decline geometrically at the same rate λ, i.e., $\alpha_{t-s-1} = \lambda\alpha_{t-s}$ and $\theta_{t-s-1} = \lambda\theta_{t-s}\ \forall s \le t-1$, we can obtain a value-added specification that relates a child's current achievement to his/her lagged achievement and the contemporaneous inputs as follows:

$$y_{it}^{*} = \lambda y_{it-1}^{*} + X_{it}\alpha_0 + \varepsilon_{it}. \tag{2}$$

In the value-added specification in Eq. (2), lagged achievement y_{it-1}^{*} is a sufficient statistic for all previous period inputs including heritable endowments, and is linked to current achievement through the persistence parameter λ. However, the estimation of Eq. (2) faces two additional challenges, as discussed in Andrabi et al. (2011). First, lagged achievement only captures individual heterogeneity in learning level at $t-1$, but talented children may also learn faster. Such individual heterogeneity in learning dynamics, if it exists, would result in a common individual-level component in the error term ε_{it}, i.e.,

$$\varepsilon_{it} = \mu_i + \nu_{it},$$

where μ_i reflects the unobserved individual-level heterogeneity in the *average* learning progress and ν_{it} is the time-varying deviation in the unobserved individual-level learning progress that has a zero mean across time for the same child. Second, we do not observe the latent

[21] A more general specification of Eq. (1) denotes the input coefficients as $\alpha_{t,s}$ and $\theta_{t,s}$ to allow them to vary by both the timing of the inputs (s) and the timing of achievement measurement (t). However, such a specification is not empirically estimable with our data.

learning outcome (y^*_{it}) but instead only a noisy measure (y_{it}):

$$y_{it} = y^*_{it} + e_{it},$$

where e_{it} is independently distributed across both individuals and time periods. Specifically, for this paper we define a time period to be the same as a school term and measure a child's cognitive achievement in each period by his/her percentile rank in class in the term-end exam. In addition to circumventing the incomparability of raw scores across schools, the use of within-class percentile score also eliminates the effects of any school inputs that have common additive influences on students within the same class, allowing us to focus attention on family inputs only for the input vector X in Eq. (1).

Replacing latent achievements with their observed measures in Eq. (2) yields:

$$y_{it} = \lambda y_{it-1} + X_{it}\alpha_0 + \mu_i + \varphi_{it}, \tag{3}$$

where $\varphi_{it} = \nu_{it} + e_{it} - \lambda e_{it-1}$. To address individual-level heterogeneity in learning progress, we can difference Eq. (3) as follows

$$\Delta y_{it} = \lambda \Delta y_{it-1} + \Delta X_{it}\alpha_0 + \Delta\varphi_{it}, \tag{4}$$

where $\Delta\varphi_{it} = \nu_{it} - \nu_{it-1} + e_{it} - (1+\lambda)e_{it-1} + \lambda e_{it-2}$. Though the individual fixed effect μ_i is gone in Eq. (4), the lagged dependent variable Δy_{it-1} is endogenous to the error term $\Delta\varphi_{it}$ by construction as they both contain ν_{it-1}, e_{it-1}, and e_{it-2}.[22] In the standard dynamic panel model without measurement errors, Arellano and Bond (1991) propose instrumenting Δy_{it-1} with two or more lags of y_{it} if the unobserved individual-level, time-varying shocks ν_{it} are serially uncorrelated. However, with measurement error in achievement in our context, y_{it-2} is still endogenous to $\Delta\varphi_{it}$ as they both contain e_{it-2}. We thus use three or more lags of y_{it} to instrument for Δy_{it-1} in Eq. (4). As e_{it} is independently distributed across both individuals and time by construction, the validity of these higher order lags as instruments for Δy_{it-1} hinges only on the assumption of no serial correlation in v_{it}, which we will test empirically using the Arellano–Bond test.[23]

We next consider addressing the possible endogeneity of ΔX_{it} in the differenced specification, a problem raised in Section 2.3. In our empirical estimation, X_{it} includes only parental absence status – i.e., dummy indicators for whether both parents are primarily at home, only one parent is primarily at home, and both parents are primarily absent from home – as a proxy for family inputs. Hence, ΔX_{it} measures the change in parental absence status from period $t-1$ to period t. We first exclude the strict exogeneity of X_{it} in our context as parents are likely to adjust their migration status based on observed past achievements, i.e., X_{it} is endogenous to y_{is} for $s < t$. As both v_{is} and e_{is} are contained in y_{is}, this implies $E[X_{it}v_{is}] \neq 0$ and $E[X_{it}e_{is}] \neq 0$ for $s < t$. Second, we rule out parents anticipating and adjusting to future shocks by assuming that past parental absence status (X_{is}) is uncorrelated with the future realizations of v_{it} and e_{it}. That is, $E[X_{is}v_{it}] = 0$ and $E[X_{is}e_{it}] = 0$ for $s < t$, the former of which hinges on the aforementioned assumption of no serial correlation in v_{it}, whereas the latter of which follows immediately from e_{it} being independently distributed across both individuals and time.[24] Third, in terms of the contemporaneous relationship, we allow parents to adjust their migration status (X_{it}) in response to the realization of the unobserved individual-level shocks affecting their children's learning progress (v_{it}), i.e., $E[X_{it}v_{it}] \neq 0$. However, as e_{it} is pure measurement error not realized until the exam day, we assume exogeneity of X_{it} to the contemporary measurement error term e_{it}, i.e., $E[X_{it}e_{it}] = 0$. Now consider the following moment condition:

$$\begin{aligned} E[X_{is}\Delta\varphi_{it}] = {} & E[X_{is}v_{it}] - E[X_{is}v_{it-1}] + E[X_{is}e_{it}] - (1+\lambda)E[X_{is}e_{it-1}] \\ & + \lambda E[X_{is}e_{it-2}]. \end{aligned} \tag{5}$$

Under the above assumptions, all terms on the right-hand side of Eq. (5) will be 0 for $s \le t-2$. That is, two and more lags of X_{it} are exogenous to $\Delta\varphi_{it}$, satisfying the validity requirement as instruments for the potentially endogenous ΔX_{it}. We thus implement our differenced GMM estimation by instrumenting Δy_{it-1} and ΔX_{it} using three or more lags of y_{it} and two or more lags of X_{it}, respectively.

4. Results

4.1. Main results

Table 3 presents our main estimation results. In all specifications, we include class-by-term fixed effects and cluster the standard errors at the class level. In columns 1 and 4, we start with the simple linear estimation of a canonical restricted version of Eq. (4), assuming perfect persistence in achievement (i.e., $\lambda = 1$)[25] and exogeneity of the change in parental absence status. That is, for each subject, we regress the difference in test score changes on the change in parental absence status, which is summarized by the two dummy indicators for the absence of one (s_{it}) and both parents (b_{it}). The coefficient estimates on the dummy indicator for the absence of both parents are negative and significant for both subjects, suggesting that the absence of both parents is associated with a decline in test score of 2.5 and 3.6 percentile points for math and Chinese, respectively, compared to having both parents at home. Though still negative, the coefficient estimates on the dummy indicator for the absence of a single parent are much smaller (−0.4 for math and −0.9 for Chinese) and insignificant, showing little evidence that the absence of a single parent (usually the father) is associated with any salient decline in a child's test score for either subject. These coefficients are qualitatively the same as findings from the difference-in-differences estimates reported in the last column of Table 2: that is, the absence of both parents is associated with significantly lower achievement relative to the presence of either one or both parents, whereas the difference between the latter two cases is small and insignificant.

However, the simple regression coefficients presented in columns 1 and 4 are subject to the potential endogeneity problem discussed in Section 2.3. As a first attempt to address this issue, we still restrict λ to be 1 but employ a GMM estimation using two and higher order lags of parental absence status as instruments for its change in columns 2 and 5. While the estimates of the coefficient on the absence of one parent remain small and insignificant, those on the absence of both parents – i.e., −6.9 for math and −5.3 for Chinese – are consistently larger in magnitude than in the linear regression estimates reported in columns 1 and 4, suggesting that the absence of both parents may be positively correlated with the omitted time-varying determinants of a child's cognitive achievement. There are at least two possible reasons that might explain the existence of such a positive correlation. First, parents may choose to both migrate and leave their child behind when a suitable guardian becomes available, such as the retirement of a grandparent, the return migration of a sibling to the home township/village, etc. Second, the return of one/both parents may be induced by some family-level shocks, which demand a lot of their time and attention after return. Examples for such shocks include the worsening of the health condition of a grandparent, the pregnancy of the mother, the birth of a younger sibling, etc. In such cases, the observed adverse

[22] Note that Δy_{it-1} $(= y_{it-1} - y_{it-2})$ contains v_{it-1} and e_{it-1} through the term y_{it-1}, and contains e_{it-2} through the term y_{it-2}.

[23] For an application of the Arellano–Bond method to estimate path dependency in a schooling context, see also Mani et al. (2012), who find significant but imperfect path dependency in the schooling progress of children in rural Ethiopia.

[24] Note that this is a weaker assumption than requiring X_{it} to be predetermined as it still allows X_{it} to depend on the current realization of v_{it} as discussed in the next point in the text.

[25] This restricted version of the value-added achievement function is also commonly used in the existing literature (e.g., Hanushek, 2003; Hanushek et al., 2003).

Table 3
OLS and GMM estimates of the First-differenced achievement function.

	Math			Chinese		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged achievement	–	–	0.539***	–	–	0.437***
			(0.051)			(0.070)
Only one parent absent	−0.360	−0.546	−1.988	−0.934	−2.018	−1.986
	(0.806)	(2.583)	(2.150)	(0.912)	(2.448)	(2.160)
Both parents absent	−2.547***	−6.895***	−5.490**	−3.359***	−5.275**	−5.070**
	(0.947)	(2.486)	(2.314)	(1.045)	(2.234)	(2.024)
p-Value of the Arellano–Bond test for AR(3) in the first-differenced equation	–	0.188	0.207	–	0.609	0.332
p-Value of the Sargan test of overidentification	–	0.950	0.201	–	0.513	0.029
Number of observations	21,952	21,952	21,952	21,952	21,952	21,952
Number of students	4528	4528	4528	4528	4528	4528

Notes: Columns (1) and (4) report the OLS estimates of a restricted version of the first-difference achievement function, Eq. (4), assuming perfect persistence in achievement. Columns (2) and (5) report the GMM estimates of a restricted version of Eq. (4) assuming perfect persistence in achievement, using parental absence status lagged two and more periods as instruments. Column (3) and (6) report the GMM estimates of an unrestricted version of Eq. (4) with empirically estimated persistence parameter, using both parental absence status lagged two and more periods and achievement lagged three and more periods as instruments. All regressions include class-by-term fixed effects. Robust standard errors clustered at the class level are reported in parentheses.
*** $p < 0.01$.
** $p < 0.05$.

relationship between a child's contemporaneous achievement progress and his/her parental absence status is mitigated because of the existence of unobserved confounding factors that affect both parental absence status and achievement progress. The advantage of the dynamic panel approach adopted here is that it can address the contamination of these unobserved confounding factors through the employment of lagged parental absence status as instruments.

The GMM estimates in columns 2 and 5 provide a good benchmark to compare to the simple linear regression estimates as they both assume perfect persistence in a child's achievement. However, perfect persistence of lagged achievement may be a too restrictive assumption in our context, particularly because these exams, which aimed to assess students' learning outcomes in a particular school term, differed substantially in test content across terms. Moreover, a large number of empirical studies also suggest that the test score effects of educational inputs fade out fairly rapidly.[26] We thus further estimate a generalized version of the value-added achievement function, allowing the persistence parameter λ to be determined empirically. Specifically, we conduct a GMM estimation of Eq. (4), employing the three and higher order lags of achievement and two and higher order lags of parental absence status as instruments in this first-differenced specification. The results, as reported in columns 3 and 6, are qualitatively the same as those in columns 2 and 5 under the assumption of perfect persistence in achievement. The point estimates of the coefficients on the absence of both parents drop only slightly to 5.5 percentile points for math and 5.1 percentile points for Chinese, both of which are significant at the 5% level. While the estimated coefficients on the absence of one parent rise to around 2.0 for both subjects, they remain insignificant. The magnitude of the size of the immediate effect of the absence of both parents we find here is comparable to what Krueger (1999) finds for the project STAR: students' test scores increase by about four percentile points for the first year they attend smaller classes (with a reduction of class size from 22 to 15 students).

Finally, we consider our empirical estimates of the persistence parameter (λ): 0.54 for math and 0.44 for Chinese. They are both significantly and substantially lower than 1, showing that persistence in achievement is indeed far from perfect in our context. Note that the parameters on the parental absence dummies we estimate above only correspond to the immediate impact of parental absence on a child's current achievement. With our specification of the value-added achievement function, the effect of the permanent absence of both parents on a child's long-run achievement equals the product of its immediate impact on current achievement and the inverse of one minus the persistence parameter. That is, given our point estimates in columns 3 and 6 of Table 3, the permanent absence of both parents can lower a child's within-class percentile score by 11.9 points ($=\frac{5.5}{1-0.54}$) for math and 9.0 points ($=\frac{5.1}{1-0.44}$) for Chinese. With an average class size of 40, these estimates indicate that the permanent absence of both parents lowers a student's within-class rank by approximately four positions, which are notable though still modest long-run effects. As mentioned before, the class rank can at the margin significantly affect a student's chances of continuing secondary education, getting into an elite high school, and admission to a reputable college.

4.2. Regression diagnosis

The consistency of our estimates hinges upon the assumption of no serial correlation in the unobserved individual-level, time-varying determinants v_{it}. Arellano and Bond (1991) propose a test for autocorrelation in v_{it} based on the residuals from the first-differenced equation. In the standard dynamic panel estimation *without* measurement errors, the Arellano–Bond test is performed by examining the existence of second-order serial correlation in the error term $\Delta\varphi_{it}$ in the first-differenced transformation. However, because $\Delta\varphi_{it}$ in Eq. (4) further contains e_{it}, e_{it-1}, and e_{it-2} due to measurement error in test scores in our context, it is serially correlated in both the first and second orders even if v_{it} itself is serially uncorrelated. Therefore, we implement the Arellano–Bond test by performing a test of third-order serial correlation in $\Delta\varphi_{it}$. The p-values of the Arellano–Bond AR(3) test, reported in the bottom of each column in Table 3, are larger than 0.18 for math and 0.33 for Chinese in all of the specifications, showing little evidence for the existence of serial correlation in v_{it}.

We next check the strength of our instruments to investigate whether the IV estimates in Table 3 suffer from the weak instrument problem.[27] While Stock and Yogo (2005) provide formal tests and tabulate the set of critical values to assess the strength of instruments for the two-stage least squares (TSLS) and limited information maximum likelihood (LIML) estimators, there is no such formal procedure (as far as we know) for evaluating the strength of the instrument set for the dynamic panel GMM estimator. Nonetheless, in a recent paper studying the effect of board structure on firm performance, Wintoki et al. (2012) adapt the

[26] For example, Jacob et al. (2010) and Rothstein (2010) find low persistence of teacher effects over time, and Currie and Thomas (1995) and Banerjee et al. (2007) show that the impacts of the Head Start program in the United States and of a remedial education program in India fade out rapidly after the intervention.

[27] For more detailed discussion of the weak instrument problem, see Bound et al. (1995), Staiger and Stock (1997), and Stock et al. (2002).

Table 4
First-stage regression statistics.

	Angrist–Pischke multivariate first-stage F statistics	Kleibergen–Paap rk Wald F statistics
	(1)	(2)
Panel A First-stage regressions with $\lambda = 1$, math and Chinese		
Δs_{it}	24.03	18.92
Δb_{it}	24.68	
Panel B First-stage regressions with empirically estimated λ, math		
Δy_{it-1}	9.18	
Δs_{it}	17.82	8.02
Δb_{it}	20.94	
Panel C First-stage regressions with empirically estimated λ, Chinese		
Δy_{it-1}	13.41	
Δs_{it}	18.24	11.64
Δb_{it}	18.27	

Notes: The table reports the Angrist–Pischke multivariate first-stage F statistics and Kleibergen–Paap rk Wald F statistics of the first-stage regressions of the first-differenced endogenous variables on their lagged levels. All specifications include class-by-term fixed effects and cluster standard errors at the class level.
Panel A corresponds to the first-stage regressions for columns (2) and (5) of Table 3. The endogenous regressors are the first differences of one parent absent dummy (Δs_{it}) and both parents absent dummy (Δb_{it}), and the instruments are two and higher order (up to eight) lags of s_{it} and b_{it}. Stock–Yogo critical values for 2 endogenous variables and 14 instruments: 5% maximal IV relative bias (19.83), 10% maximal IV relative bias (10.89).
Panels B and C correspond to the first-stage regressions for columns (3) and (6) of Table 3, respectively. The endogenous regressors are the first differences of lagged achievement (Δy_{it-1}), one parent absent dummy (Δs_{it}) and both parents absent dummy (Δb_{it}), and the instruments are two and more (up to eight) lags of s_{it} and b_{it} and three and higher order (up to eight) lags of y_{it}. Stock–Yogo critical values for 3 endogenous variables and 20 instruments: 5% maximal IV relative bias (19.56), 10% maximal IV relative bias (10.60), 20% maximal IV relative bias (5.93).

procedure outlined by Stock and Yogo (2005) for the TSLS estimator to assess the strength of the instruments in dynamic panel GMM estimations. Following the same spirit as Wintoki et al. (2012), we also carry out the first-stage regressions of our endogenous variables on the instruments corresponding to the dynamic panel GMM estimations in Table 3, and present in Table 4 both the Angrist–Pischke multivariate first-stage F-statistics and the Kleibergen–Paap Wald rk F-statistics.[28] Specifically, Panel A of Table 4 corresponds to the first-stage regressions for columns (2) and (5) of Table 3, which have the same first-stage regressions as they only differ in the dependent variable, and Panels B and C correspond to the first-stage regressions for columns (3) and (6) of Table 3, respectively. With the single exception for the first-stage regression with lagged achievement in math as the dependent variable in which the Angirst–Pischke multivariate first-stage F-statistic is 9.18, this statistics for all the other eight first-stage regressions exceeds the rule-of-thumb threshold of 10 suggested by Staiger and Stock (1997) for assessing instrument strength. Note that the Kleibergen–Paap Wald rk F-statistics is reported here instead of the Cragg–Donald Wald statistics because the latter is not valid with clustered standard errors. However, no corresponding critical values are available for testing weak identification using the Kleibergen–Paap Wald rk F-statistics. As suggested by Baum et al. (2007), we compare with caution the Kleibergen–Paap Wald rk F-statistics to the critical values compiled for Cragg–Donald Wald statistics under the assumption of i.i.d. errors. The Kleibergen–Paap Wald rk F-statistics for both Panel A (18.92) and Panel C (11.64) exceed the corresponding critical values for 10% maximal IV relative bias in Table 5.1 of Stock and Yogo (2005), i.e., 10.89 (for 2 endogenous regressors and 14 instruments) and 10.60 (for 3 endogenous regressors and 20 instruments) respectively, implying that any bias from the TSLS estimates using the instruments is less than 10% of the bias from the OLS regression with 5% significance level. This statistic for Panel B (8.02) also exceeds the critical value for 20% maximal IV relative bias, i.e., 5.93 (for 3 endogenous regressors and 20 instruments). Overall, the results of these tests show little evidence that our estimates suffer from the weak instrument problem.

Finally, we consider the results of the Sargan test of overidentification, reported in the bottom of each column in Table 3. When only lagged values of parental absence status are used as instruments for its change in estimating the restricted model of learning dynamics assuming perfect persistence in columns 2 and 5, the Sargan statistics pass the overidentification test with p-values of 0.95 and 0.51 for math and Chinese, respectively. However, when we further employ the lagged dependent variables as instruments to empirically estimate the persistence parameter, the p-values of the Sargan statistics fall substantially to 0.20 for math (column 3) and 0.03 for Chinese (column 6). The failure of the overidentifying restrictions in the latter case suggests that some of the local parameters identified by different instruments are significantly different from each other for Chinese. Since we have employed lagged variables as instruments to identify both the persistence parameter and the coefficients on parental absence dummies, the failure of the overidentifying restrictions can come from deviations in the local estimates of either of them. As the main purpose of this paper is to identify the impact of parental absence, we would be particularly concerned if the rejection of the overidentifying restrictions were caused by differences in the local estimates of the coefficients on parental absence status. However, the fact that the Sargan tests pass for both subjects in columns 2 and 5 suggests that this is not the case when the persistence parameter is set to 1. Moreover, when the persistence parameter λ is set to other constant values rather than being estimated empirically in Table 5 (see Section 4.3.1 below for more detailed discussion), the p-values of the Sargan statistics exceed 0.55 in all cases, further suggesting that the failure of the overidentifying restrictions for Chinese in column 6 of Table 3 is caused by differences in the local estimates of the persistence parameter rather than in the coefficients on parental absence status.

4.3. Sensitivity and treatment heterogeneity analysis

In this subsection, we investigate the sensitivity of our estimated parental absence effects to variation in the persistence parameter value (Section 4.3.1) and alternative classifications of parental absence status (Section 4.3.2), and also explore the potential heterogeneity in treatment effects by child gender and sibship structure (Section 4.3.3).

4.3.1. Variation in the persistent parameter value

Some previous research demonstrates that the values of the persistence parameter could play a central role in treatment effect evaluations employing value-added models of learning and sometimes even lead to qualitatively different conclusions. For example, Chay et al. (2005) show that the conventional difference-in-differences assessment of a school intervention program in Chile targeting low-performing schools substantially overstates the true program effect because of ignoring the mean reversion in test scores. Andrabi et al. (2011) also find that the

[28] Both statistics are obtained using the ivreg2 module in Stata 12.

Table 5
GMM estimates with other exogenously assigned persistence parameter value.

	Math			Chinese		
	(1)	(2)	(3)	(4)	(5)	(6)
Persistence parameter (exogenously assigned)	0.8	0.6	0.4	0.8	0.6	0.4
Only one parent absent	−0.841	−1.136	−1.432	−1.719	−1.420	−1.120
	(2.356)	(2.235)	(2.237)	(2.192)	(2.074)	(2.118)
Both parents absent	−6.246***	−5.596**	−4.946**	−5.085**	−4.894**	−4.703**
	(2.347)	(2.322)	(2.416)	(2.046)	(1.980)	(2.047)
p-Value of the Arellano–Bond test for AR(3) in the first-differenced equation	0.191	0.210	0.262	0.552	0.452	0.304
p-Value of the Sargan test of overidentification	0.935	0.913	0.881	0.553	0.614	0.694
Number of observations	21,952	21,952	21,952	21,952	21,952	21,952
Number of students	4528	4528	4528	4528	4528	4528

Notes: The table reports the GMM estimates of the first-differenced achievement function, Eq. (4), with exogenously assigned persistence parameter as indicated in the first row. All regressions use parental absence status lagged two and more periods as instruments and include class-by-term fixed effects. Robust standard errors clustered at the class level are reported in parentheses.
*** $p < 0.01$.
** $p < 0.05$.

estimated effects of private schooling in Pakistan are highly sensitive to the persistence parameter value: while estimates of the generalized value-added model (allowing λ to be determined empirically) yield large, positive, and significant impacts of private schooling on achievement, those of the restricted model (i.e., $\lambda = 1$) suggest no advantage of private schools over public schools.

To investigate the sensitivity of our estimated parental absence effects to variation in the persistence parameter value, we conduct in Table 5 an exercise in which we assign λ to a set of alternative constant values, varying from 0.8 to 0.4 with decrements of 0.2. The coefficient estimates on parental absence dummies of all specifications in this exercise unanimously point to the same conclusion: the absence of both parents has significant adverse effects on children's test scores, whereas the impacts of the absence of a single parent are much smaller and insignificant. The results in Tables 3 and 5, taken together, show that the persistence parameter value makes little difference in estimating the parental absence effects in our context.

4.3.2. Alternative classifications of parental absence status

While the results from Tables 4 and 5 suggest that only the absence of both parents leads to significant declines in student achievement, estimates of the coefficient on the absence of a single parent are driven mainly by the effect of paternal absence (only), which accounts for 85% of the cases of one parent absent in our sample. As an attempt to disentangle the impact of the absence of both parents from that of the absence of the mother (only), we separate the cases of maternal absence (only) from paternal absence (only), and rerun the GMM estimations in columns 3 and 6 of Table 4. The point estimates of the coefficients on the separate indicators for maternal and paternal absence (only), reported in columns 1 and 2 of Table 6, suggest that for both subjects, maternal absence generates no larger adverse effect than paternal absence. Actually, the coefficient on maternal absence is even estimated to be positive for math (0.521), although we probably should not take this positive sign too seriously given the size of its standard error (4.973). The results of this sensitivity analysis seem to suggest that the large and significant adverse effect of having both parents migrating is indeed due to the absence of both of them rather than the mother alone.[29]

Given our retrospective survey, the residence status of a parent during a school term is characterized by an ordinal indicator, I_f for father and I_m for mother, that equals 0 if he/she was always present at home, 1 if he/she was present at home for more than half of the time but not always, 2 if he/she was absent from home for more than half of the time but not always, and 3 if he/she was always absent from home. So far, we have defined a parent as absent from home if this ordinal indicator takes a value of 2 or 3. To investigate the sensitivity of our results to alternative definitions of parental absence, we further separate $I_f/I_m = 2$ from $I_f/I_m = 3$, and employ four dummy indicators (instead of two) to classify parental absence status: (i) only one parent often absent. i.e., ($I_f = 2, I_m \leq 1$) or ($I_f \leq 1, I_m = 2$); (ii) only one parent always absent, i.e., ($I_f = 3, I_m \leq 1$) or ($I_f \leq 1, I_m = 3$); (iii) both parents often absent, i.e., ($I_f = 2, I_m = 2$) or ($I_f = 2, I_m = 3$) or ($I_f = 3, I_m = 2$); and (iv) both parents always absent, i.e., ($I_f = 3, I_m = 3$). We report estimates employing this alternative classification of parental absence status in columns 3–4 of Table 6. While the coefficients on all the four dummy indicators for varying degrees of parental absence are negative for both subjects, only that on the dummy indicator for having both parents always absent is significant at the 10% level for math and at the 5% level for Chinese. As a further sensitivity analysis, we combine the two ordinal indicators, I_f and I_m, to construct a proxy index for the fraction of the time of parental absence (L) such that $L = (I_f + I_m)/6$. The estimates of the coefficient on this proxy index, as reported in columns 5–6 of Table 6, are significant at the 5% level for both subjects. The point estimates suggest that complete parental absence (i.e., $L = 1$) lowers a student's percentile score by 6.0 points for math and 4.8 points for Chinese, both of which are very close to estimates of the coefficient on absence of both parents in various other specifications. It is important to note that given the relatively large standard errors of these coefficients (2.7 for math and 2.5 for Chinese), only rather large adverse effects can be detected as significant. Therefore, it is not surprising that our estimates for the effects of partial parental absence are always insignificant, regardless of the definition and specification used. While we find no statistical evidence that partial parental absence has adverse effects on student achievement, we also cannot rule out the linearity of the effects to the length of parental absence.

4.3.3. Treatment heterogeneity

To examine the possible heterogeneity in parental absence effects by child gender, we conduct separate estimates for boys (columns 1–2) and girls (columns 3–4) in Table A1. The results for boys are the same as those for the combined sample: the absence of both parents has significant and adverse effects on their son's achievements in both subjects while the absence of a single parent does not. However, the results for girls are different by subject: for math, the absence of a single parent is found to have almost the same adverse effect as the absence of both parents (−5.9 vs. −6.0), both of which are significant at the 10% level, whereas, for Chinese, both coefficients are much smaller and neither of them is significant. However, given the relatively large standard errors of these coefficients (around 3.0), we cannot

[29] Nonetheless, we could not provide support to this statement with statistical significance because of the very imprecise estimates of the coefficient on maternal absence (only) due to its rare occurrences.

Table 6
Sensitivity analysis to alternative classifications of parental absence status.

	Math	Chinese	Math	Chinese	Math	Chinese
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged achievement	0.504***	0.415***	0.446***	0.378***	0.553***	0.480***
	(0.049)	(0.068)	(0.048)	(0.067)	(0.054)	(0.075)
Only father absent	−3.074	−2.416	–	–	–	–
	(2.294)	(2.325)				
Only mother absent	0.521	−2.197	–	–	–	–
	(4.973)	(4.225)				
Both parents absent	−5.606**	−5.036**	–	–	–	–
	(2.282)	(1.977)				
Only one parent often absent	–	–	−0.442	−2.654	–	–
			(3.200)	(2.848)		
Only one parent always absent	–	–	−2.631	−1.252	–	–
			(2.416)	(2.542)		
Both parents often absent	–	–	−3.619	−0.366	–	–
			(2.489)	(2.934)		
Both parents always absent	–	–	−5.029*	−5.901**	–	–
			(2.568)	(2.427)		
Fraction of the time of parental absence	–	–	–	–	−6.043**	−4.801*
					(2.661)	(2.521)
p-Value of the Arellano–Bond test for AR(3) in the first-differenced equation	0.219	0.317	0.220	0.273	0.191	0.367
p-Value of the Sagan test of overidentification	0.193	0.060	0.201	0.081	0.010	0.007
Number of observations	21,952	21,952	21,952	21,952	21,952	21,952
Number of students	4528	4528	4528	4528	4528	4528

Notes: The table reports the GMM estimates of the first-differenced achievement function, Eq. (4), with alternative classifications of parental absence as defined in the row headings. All regressions use parental absence status lagged two and more periods and achievement lagged three and more periods as instruments, and include class-by-term fixed effects. Robust standard errors clustered at the class level are reported in parentheses.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

reject the equality of any coefficient by gender for either subject. As an attempt to investigate the possible heterogeneity in the impacts of parental absence by sibship structure, we subdivide our sample into three subsamples by the presence of any sibling, and if so, any older sibling, and conduct separate estimations for each subsample in columns 5–10 of Table A1. The coefficients for the two parental absence dummies turn out to be even more imprecisely estimated. Out of the 12 coefficients estimated, only two are significant at the 10% level, whereas the coefficient on one parent absence dummy is positive – but small in magnitude (1.02) and statistically insignificant – for math for students with only younger sibling(s). Taken together, the results in Table A1 show little evidence for heterogeneous effects of parental absence by child gender or sibship structure, although it is in part because of the limitations of the statistical power of our data.

4.4. Robustness checks

As mentioned in Section 2.2, there are 123 students in our sample who reported having one parent permanently missing due to divorce.[30] We do not exclude them from our baseline analysis because of the concern that selecting the sample based on parents' current marital status, which could be affected by their migration history, may lead to biased estimates of the impacts of parental absence. For example, if parental divorce, which results in the permanent missing of one parent, has a larger adverse impact on the affected child than temporary parental absence due to migration, excluding students with divorced parents may underestimate the impact of parental migration. We thus conduct robustness checks in columns 1 and 2 of Table A2 to further exclude students with divorced parents, i.e., restricting only to students whose (biological) parents maintain their marriage relationship. Moreover, as further robustness checks, we expand our sample to also include students with deceased parent(s) – but treat them as absent – in columns 3–4 of Table A2. Both exercises yield estimates similar to our baseline analysis, showing that our results are robust to these alternative sample selection rules.

To investigate the stability of our results to measures of learning outcomes used, we also conduct robustness analysis employing two alternative achievement measures. In Panel A of Table 7, we replicate the same estimations as in Table 4 but replace the dependent variable with the within-class z-scores. The estimation results are almost exactly the same as those in Table 4. For all the six specifications, the coefficients are negative and significant for the absence of both parents, but are smaller and insignificant (though still negative) for the absence of a single parent. The point estimates in columns 3 and 6 suggest that the absence of both parents lowers a student's test scores in math and Chinese by 0.21 and 0.14 within-class standard deviations, respectively, both of which are significant at the 5% level. In Panel B of Table 7, we use as the dependent variable students' raw test scores in the term-end exams (ranging from 0 to 100). The coefficients are largely consistent with our previous estimates using relative performance measures, but are somewhat less precisely estimated. For example, in columns 3 and 6, the coefficient on the absence of both parents is estimated to be −1.55 for math and −1.60 for Chinese, but only the latter is significant at the 10% level. We believe that the relative imprecision of the coefficient estimates for the absolute performance is to a large extent due to the incomparability of raw test scores over time (even within the same class), which could introduce a lot of noise in the estimates when the persistent parameter is restricted to be time-invariant.[31]

[30] In such cases, the missing parent due to divorce is also treated as absent from home in our analysis.

[31] Specifically, the time-invariant persistence parameter cannot account for the scale difference (as reflected by the variance) of the raw test scores of two consecutive school terms for the same class, which is not an issue for relative achievement measures such as within-class percentile scores and z-scores used previously.

Table 7
Robustness analysis to alternative achievement measures.

	Math			Chinese		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Dependent variable: within-class z-scores						
Lagged achievement	–	–	0.555***	–	–	0.480***
			(0.072)			(0.074)
Only one parent absent	−0.034	−0.091	−0.105	−0.034	−0.024	−0.055
	(0.025)	(0.088)	(0.075)	(0.025)	(0.081)	(0.068)
Both parents absent	−0.088***	−0.284***	−0.205***	−0.079**	−0.122*	−0.141**
	(0.030)	(0.083)	(0.072)	(0.033)	(0.070)	(0.058)
Observations	21,952	21,952	21,952	21,952	21,952	21,952
Number of sid	4528	4528	4528	4528	4528	4528
Panel B. Dependent variable: raw test scores						
Lagged achievement	–	–	0.416***	–	–	0.591***
			(0.116)			(0.084)
Only one parent absent	−0.402	−0.763	−1.406	−0.666*	−0.147	−0.815
	(0.363)	(1.209)	(1.095)	(0.369)	(1.196)	(0.956)
Both parents absent	−1.280***	−2.639**	−1.547	−1.219***	−1.176	−1.595*
	0.412	(1.070)	(0.986)	(0.446)	(0.974)	(0.854)
Observations	21,952	21,952	21,952	21,952	21,952	21,952
Number of sid	4528	4528	4528	4528	4528	4528

Notes: Each column of this table replicates the same estimation in the corresponding column of Table 4, but employs the within-class z — scores as the dependent variable in Panel A and raw test scores as the dependent variable in Panel B. Robust standard errors clustered at the class level are reported in parentheses.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

4.5. Survey evidence on time allocation, family inputs and relationship to others

As mentioned in Section 2.2, we conduct student surveys to collect information on time allocation after school, tutoring help from parents and others, and satisfaction levels on relationships to parents, other adult family members, and teachers and classmates. However, because all these measures are contemporary rather than longitudinal, we cannot apply the same dynamic panel methods to identify the effects of parental absence on them. Nonetheless, we still present the cross-sectional relations between parental absence status and each of these variables in Table 8. To the extent that these cross-sectional relationships partly reflect the impacts of parental absence, they may be informative and shed light on the possible channels through which parental absence affects student achievement. Panel A shows little difference in the average after-school study hours across students with different parental absence status. Students of all three groups spent on average about 70 min each day in study after school, of which about 37 min were for homework. However, Panel B indicates large disadvantages of students left-behind by both parents on family inputs in after-

Table 8
Survey evidence on after-school study hours, family inputs and relationship to others.

	Both parents at home	One parent absent		Both parents absent	
	Mean	Mean	Difference	Mean	Difference
	(1)	(2)	(3)	(4)	(5)
Panel A. After-school study hours					
Minutes spent in learning after school per day	70.523	69.667	−0.856	70.997	0.474
Minutes spent in home work after school per day	38.347	36.867	−1.480	37.683	−0.664
Minutes spent in after-school learning activities other than homework per day	32.176	32.800	0.625	33.313	1.138
Panel B. Family inputs in tutoring					
Ever received tutoring help from someone in the past week	0.767	0.747	−0.020	0.479	−0.288***
Ever received tutoring help from parents in the past week	0.670	0.655	−0.015	0.189	−0.482***
Ever received tutoring help from someone other than parents in the past week	0.420	0.394	−0.026	0.411	−0.009
Ever received tutoring help from an older sibling in the past week	0.266	0.233	−0.032**	0.162	−0.104***
Ever received tutoring help from any adult family member (other than parents) in the past week	0.111	0.117	0.007	0.214	0.103***
Total tutoring minutes per day	25.803	25.403	−0.400	15.234	−10.569***
Tutoring minutes from parents per day	14.450	14.248	−0.202	4.195	−10.255***
Tutoring minutes from others per day	11.353	11.155	−0.198	11.039	−0.314
Tutoring minutes from the older sibling per day	7.053	6.007	−1.046*	4.497	−2.556***
Tutoring minutes from the adult member per day	2.427	2.975	0.548	5.319	2.892***
Panel C. Relationship with others					
Satisfied w/ relationship to parents	0.853	0.827	−0.026*	0.808	−0.045***
Satisfied w/ relationship to other adults in family	0.639	0.574	−0.065***	0.576	−0.063***
Satisfied w/ relationship to teachers and classmates in school	0.840	0.848	0.008	0.848	0.007
Number of students	1310	1457	2767	1761	3071

Notes: For each variable indicated by the row heading, columns (1), (2) and (4) report the means for the sub sample with both parents at home, the subsample with only one parent absent from home, and the subsample with both parents absent from home, respectively; columns (3) and (5) report the difference in the means between columns (2) and (1) and between columns (4) and (1), respectively.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

school tutoring relative to those with one and both parents at home, with the latter two groups having similar statistics in most measures of tutoring inputs. Because we define parental absence status as a parent/parents being away for more than half of the time during a school term, some students classified as being left-behind by both parents happened to have their parent(s) at home in the week prior to our survey and thus 19.0% of them still reported tutoring help from parents, although the number is substantially lower than that of those having one (65.5%) and both parents (67.0%) at home (for more than half of the time). In terms of parental tutoring time, students with both parents absent (for more than half of the time) on average received 4.2 min of tutoring help from their parents per day in the past week, compared to 14.2 and 14.5 min for those with one and both parents at home, respectively. Nonetheless, despite the large discrepancies in parental tutoring inputs between students with and without parent(s) at home, the proportion of students who reported having tutoring help from all others besides parents (about 41%) and the average tutoring minutes from all others (about 11 min per day) are almost the same across students with different parental absence status. However, this equality in overall tutoring inputs from others has masked cross-group differences in tutoring inputs from older siblings and other adult family members. For example, students with both parents absent are 10.4 percentage points less likely to receive tutoring help from an older sibling, but are 10.3 percentage points more likely to receive tutoring help from an adult family member other than parents (e.g., grandparents, uncles, aunties). Both differences are significant at the 1% level. The former difference is at least in part due to differences in the proportion of students with at least an older sibling between these two groups (0.404 vs. 0.503) as shown in Table 1, whereas the latter difference reflects the extent of substitution of tutoring inputs from other adult family members for the loss of parental inputs. While students with both parents absent on average received 2.9 more minutes of tutoring per day from other adult family members compared with those with both parents at home (5.3 vs. 2.4 min), it can only compensate for their disadvantage in tutoring inputs from older siblings (4.5 vs. 7.1 min), and is far from their loss in parental inputs in tutoring due to the absence of both parents, estimated to be 10.3 min per day (4.2 vs. 14.5 min). If after-school tutoring is highly complementary to both school inputs and students' efforts, the lack of parental tutoring inputs may be a main reason for the adverse effects of the absence of both parents on student achievements. Our results in Panels A and B of Table 8 seem to suggest that students with parent(s) at home are able to make more effective progress in learning holding constant the school inputs and their after-school study hours because of the crucial role of family tutoring inputs, which is reduced substantially for those with no parent at home. However, the absence of a single parent seems to have little effect on parental tutoring inputs, implying a very high degree of substitution between paternal and maternal inputs in tutoring, which can also explain our finding of lack of evidence for adverse effects of the absence of one parent.

Panel C presents students' self-reported satisfaction levels on their relationships to parents, other adult family members, and teachers and classmates. We classify a student as satisfied with a relationship if he/she gave it a rating of 4 or more out of a scale of 5. The results suggest that the absence of either a single or both parents has significant adverse effects on their family relationships. The absence of one and both parents increases the probability that a student reported being unsatisfied with his/her relationships to parents by 2.6 percentage points (a 17.7% increase) and 4.5 percentage points (a 30.6% increase), respectively, from a baseline probability of 14.7% for those with both parents at home. Students with one or both parents absent from home were also more likely to be unhappy with their relationships to other adult family members. However, we find no difference in students' satisfaction with their relationships to teachers and classmates by parental absence status. These findings suggest that the migration of parents, even that of only one parent, is likely to have other adverse effects on child development beyond test scores, such as their psychological well-beings. Unfortunately, the lack of longitudinal data on these aspects precludes us from substantiating the suggestive evidence presented here, which leaves room for future investigations.

5. Conclusion

Parental migration with children left-behind is a worldwide phenomenon, but with the largest absolute numbers of children affected in rural China — an estimated 61 million plus such children. The implications of this phenomenon have created considerable interest and concern. But the systematic evidence of the implications of being left-behind for children has been limited and has focused on impacts on time use inputs into learning such as school attendance and time studying. The previous literature generally has not considered the impacts of parental migration on learning itself as measured by children's cognitive achievements. A priori the impacts on learning may be positive or negative, depending on whether, for example, the income effect due to increased parental income outweighs the impacts of reduced parental time inputs into child learning. Therefore it is not surprising that some previous studies have found positive and others have found negative or no significant effects on children's time use related to learning. And, of course, even if children's time spent on learning increases, learning may decrease or be unaffected if parental presence has large effects on child learning.

We investigate how primary-school-aged children's cognitive achievements are affected by the absence of their parents from home. Based on a cumulative educational production function, we derive a value-added specification that relates a child's contemporary learning to his/her lagged achievement, family inputs, and individual heterogeneity in learning. We then estimate this model using a retrospective panel data set that we collected for third to fifth graders in a very poor rural county in China with a considerable proportion of left-behind children, and proxy for family inputs by children's parental absence status. Our GMM estimates of the dynamic panel model show significant adverse effects of the absence of both parents on the cognitive achievements of their children left-behind. In contrast, estimates of the effect of the absence of a single parent are much smaller in magnitude and not significant.

Compared with other contexts in developing countries in which primary school enrollment is not almost universal, the immediate impacts of parental absence on primary-school-aged children in China are largely restricted to the achievement margin; this means that the impacts are likely to be less severe (at least in the short run) since children do not drop out because of parental absence. Both for China and for other contexts, however, there may be longer-run impacts on the enrollment margin through reductions in subsequent entrance exam scores for secondary school and university admissions. Moreover, if, as would seem to be the case based on other estimates for other contexts, primary-school level cognitive achievements have cumulative impacts on upper-school level cognitive achievements and post-school cognitive achievements, which in turn affect lifetime productivities, there may be considerable long-run costs of the reduced primary-school cognitive achievements that we estimate resulting from absences of both parents. Given that a large proportion of children are left-behind by both parents in rural China, our findings thus may have major policy implications. There is a definite value in investigating possible compensatory programs for ameliorating these negative learning effects of absences of both parents or for lessening or eliminating the restrictions that cause family separation due to parental migration.

Appendix A

Table A1
Treatment heterogeneity analysis by gender and sibship structure.

	Boys		Girls		Students w/ no siblings		Students w/ only younger sibling(s)		Students w/ older sibling(s)	
	Math	Chinese	Math	Chinese	Math	Chinese	Math	Chinese	Math	Chinese
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Only one parent absent	−1.044	−1.848	−5.869*	−1.482	−3.490	−6.430	1.018	−1.174	−4.386	−1.519
	(3.033)	(3.052)	(3.262)	(3.465)	(5.984)	(6.021)	(3.164)	(3.378)	(3.447)	(3.241)
Both parents absent	−5.733*	−5.974**	−5.960*	−2.898	−7.249	−7.097	−2.621	−6.408**	−6.665*	−3.951
	(2.988)	(2.866)	(3.190)	(3.096)	(5.955)	(8.332)	(2.948)	(2.845)	(3.537)	(3.187)
Number of observations	12,709	12,709	9243	9243	3579	3579	8580	8580	9793	9793
Number of students	2609	2609	1919	1919	741	741	1714	1714	2073	2073

Notes: The table reports the GMM estimates of the first-differenced achievement function, Eq. (4), for subsamples defined by gender and sibship structure as noted in the column headings. All regressions use parental absence status lagged two and more periods and achievement lagged three and more periods as instruments and include class-by-term fixed effects. Robust standard errors clustered at the class level are reported in parentheses.
** $p < 0.05$.
* $p < 0.1$.

Table A2
Robustness analysis to alternative sample selections.

	Students w/ no permanent missing parent		All students	
	Math	Chinese	Math	Chinese
	(1)	(2)	(3)	(4)
Panel A. Alternative sample selections				
Only one parent absent from home	−2.140	−2.156	−1.802	−1.961
	(2.114)	(2.124)	(2.156)	(2.149)
Both parents absent from home	−5.802**	−4.391**	−5.281**	−4.807**
	(2.246)	(2.010)	(2.414)	(2.099)
Lagged achievement	0.546***	0.439***	0.550***	0.442***
	(0.050)	(0.070)	(0.051)	(0.068)
p-Value of the Arellano–Bond test for AR(3) in the first-differenced equation	0.299	0.427	0.185	0.450
p-Value of the Sagan test of overidentification	0.432	0.028	0.276	0.036
Number of observations	21,359	21,359	22,312	22,312
Number of students	4405	4405	4598	4598

Notes: The table reports the GMM estimates of the first-differenced achievement function, Eq. (4), using both parental absence status lagged two and more periods and achievement lagged three and more periods as instruments. The sample used in columns 1–2 restricts only to students whose (biological) parents remain in their marriage relationship, whereas the sample used in columns 3–4 contains all students, including those with permanent missing parent(s) due to either divorce or decease. All regressions include class-by-term fixed effects. Robust standard errors clustered at the class level are reported in parentheses.
*** $p < 0.01$.
** $p < 0.05$.

References

Adams Jr., Richard H., Cuecuecha, Alfredo, 2010. Remittances, household expenditure and investment in Guatemala. World Dev. 38 (11), 1626–1641.

Alcaraz, Carlo, Chiquiar, Daniel, Salcedo, Alejandrina, 2012. Remittances, schooling, and child labor in Mexico. J. Dev. Econ. 97 (1), 156–165.

All-China Women's Federation, 2013. A study of the situation of left-behind children in rural China.

Andrabi, Tahir, Das, Jishnu, Khwaja, Asim Ijaz, Zajonc, Tristan, 2011. Do value-added estimates add value? Accounting for learning dynamics. Am. Econ. J. Appl. Econ. 3 (3), 29–54.

Antman, Francisca, 2011. The intergenerational effects of paternal migration on schooling and work: what can we learn from children's time allocations? J. Dev. Econ. 96 (2), 200–208.

Antman, Francisca, 2012. Gender, educational attainment, and the impact of parental migration on children left behind. J. Popul. Econ. 25 (4), 1187–1214.

Antman, Francisca, 2013. The impact of migration on family left behind. In: Zimmerman, Klaus F., Constant, Amelie F. (Eds.), International handbook on the economics of migration. Edward Elgar, United Kingdom.

Arellano, Manuel, Bond, Stephen, 1991. Some tests of specifications of panel data: Monte Carlo evidence and an application to employment equations. Rev. Econ. Stud. 58 (2), 277–297.

Banerjee, Abhijit V., Cole, Shawn, Duflo, Esther, Linden, Leigh, 2007. Remedying education: Evidence from two randomized experiments in India. Q. J. Econ. 122 (3), 1235–1264.

Baum, Christopher F., Schaffer, Mark E., Stillman, Steven, 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. Stata J. 7 (4), 465–506.

Beine, Michel, Docquier, Frederic, Rapoport, Hillel, 2008. Brain drain and human capital formation in developing countries: Winners and losers. Econ. J. 118 (528), 631–652.

Boardman, Anthony E., Murnane, Richard J., 1979. Using panel data to improve estimates of the determinants of educational achievement. Sociol. Educ. 52 (2), 113–121.

Bound, John, Jaeger, David A., Baker, Regina M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. J. Am. Stat. Assoc. 90 (430), 443–450.

Chang, Hongqin, Dong, Xiao-Yuan, Macphail, Fiona, 2011. Labor migration and time use patterns of the left-behind children and elderly in rural China. World Dev. 39 (12), 2199–2210.

Chay, Kenneth Y., McEwan, Patrick J., Urquiola, Miguel, 2005. The central role of noise in evaluating interventions that use test scores to rank schools. Am. Econ. Rev. 95 (4), 1237–1258.

Chen, Joyce J., 2013. Identifying non-cooperative behavior among spouses: Child outcomes in migrant-sending households. J. Dev. Econ. 100 (1), 1–18.

Chen, Yuanyuan, Feng, Shuaizhang, 2013. Access to public schools and the education of migrant children in China. China Econ. Rev. 26 (75), 75–88.

Clemens, Michael, Tiongson, Erwin, 2012. Split decisions: Family finance when a policy discontinuity allocates overseas work. Discussion Paper No. 34/12, Centre for Research and Analysis of Migration. University College London.

Currie, Janet, Thomas, Duncan, 1995. Does head start make a difference? Am. Econ. Rev. 85 (3), 341–364.

Dustmann, Christian, Glitz, Albrecht, 2011. Migration and education. In: Hanushek, Eric A., Machin, Stephen, Woessmann, Ludger (Eds.), Handbook of the economics of education vol. 4. Elsevier, New York, pp. 327–441.

Edwards, Alejandra Cox, Ureta, Manuelita, 2003. International migration, remittances, and schooling: Evidence from El salvador. J. Dev. Econ. 72 (2), 429–461.

Gibson, John, McKenzie, David, 2014. The development impact of a best practice seasonal worker policy. Rev. Econ. Stat. 96 (2), 229–243.

Gibson, John, McKenzie, David, Stillman, Steven, 2011. The impacts of international migration on remaining household members: Omnibus results from a migration lottery program. Rev. Econ. Stat. 93 (4), 1297–1318.

Hanushek, Eric A., 2003. The falure of input-based schooling policies. Econ. J. 113 (485), F64–F98.

Hanushek, Eric A., Kain, John F., Markman, Jacob M., Rivkin, Steven G., 2003. Does peer ability affect student achievement? J. Appl. Econ. 18 (5), 527–544.

Haveman, Robert, Wolfe, Barbara, 1995. The determinants of children's attainments: A review of methods and findings. J. Econ. Lit. 33 (4), 1829–1878.

Huang, Jikun, Zhi, Huayong, Huang, Zhurong, Rozelle, Scott, Giles, John, 2011. The impact of the global financial crisis on off-farm employment and earnings in rural China. World Dev. 39 (5), 797–807.

Jacob, Brian A., Lefgren, Lars, Sims, David P., 2010. The persistence of teacher-induced learning. J. Hum. Resour. 45 (4), 915–943.

Krueger, Alan, 1999. Experimental estimates of the education production function. Q. J. Econ. 114 (2), 497–532.

Kuhn, Randll, 2007. The effect of fathers' and siblings' migration on children's pace of schooling in rural Bangladesh. Asian Popul. Stud. 2 (1), 69–92.

Lyle, David S., 2006. Using military deployments and job assignments to estimate the effect of parental absences and household relocations on children's academic achievement. J. Labor Econ. 24 (2), 319–350.

Mani, Subha, Hoddinott, John, Strauss, John, 2012. Long-term impact of investment in early schooling — Empirical evidence from rural Ethiopia. J. Dev. Econ. 99 (2), 292–299.

Mckenzie, David, Rapoport, Hillel, 2011. Can migration reduce educational attainment? Evidence from Mexico. J. Popul. Econ. 24 (4), 1331–1358.

Meng, Lingshen, 2013. Evaluating China's poverty allevation program: A regression discontinuity approach. J. Public Econ. 101 (1), 1–11.

National Bureau of Statistics, 2014. The 2013 investigational and monitoring report of Chinese migrant workers.

Park, Albert, Wang, Sangui, Wu, Guobao, 2002. Regional poverty targeting in China. J. Public Econ. 86 (1), 123–153.

Rapoport, Hillel, Docquier, Frédéirc, 2006. The economics of migrants' remittance. In: Kolm, Serge-Christophe, Ythier, Mercier, Jean (Eds.), Handbook of the Economics of Giving, Altruism and Reciprocity 2. Elsevier, Amsterdam, pp. 1135–1198.

Roberts, Kenneth, Connelley, Rachel, Xie, Zheming, Zheng, Zhenzhen, 2004. Patterns of temporary labor migration of rural women from Anhui and Sichuan. Chin. J. 52, 49–70.

Rothstein, Jesse, 2010. Teacher quality in educational production: Tracking, decay, and student achievement. Q. J. Econ. 125 (1), 175–214.

Staiger, Douglas, Stock, James H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557–586.

Stark, Oded, 1993. The migration of labor. B. Blackwell, Cambridge, MA.

Stock, James H., Yogo, Motohiro, 2005. Testing for weak instruments in linear IV regression. In: Andrews, D.W.K. (Ed.), Identification and inference for econometric models, chapter 5. Cambridge University Press, New York, pp. 80–108.

Stock, James H., Wright, Jonathan H., Yogo, Motohiro, 2002. A survey of weak instruments and weak identification in generalized method of moments. J. Bus. Econ. Stat. 20 (4), 519–529.

Todd, Petra E., Wolpin, Kenneth I., 2003. On the specification and estimation of the production function for cognitive achievement. Econ. J. 113, F3–F33.

Wintoki, M. Babajide, Linck, James S., Netter, Jeffry M., 2012. Endogeneity and the dynamics of internal corporate governance. J. Financ. Econ. 105 (3), 581–606.

Yang, Dean, 2008. International migration, remittances and household investment: Evidence from Philippine migrants' exchange rate shocks. Econ. J. 118 (528), 591–630.

Yang, Dean, 2011. Migrant remittances. J. Econ. Perspect. 25 (3), 129–152.