Does Parental Out-migration Benefit Left-behind Children's Schooling Outcomes?

— Evidence from Rural China

Xiaoman Luo *

University of California, Davis

PRELIMINARY VERSION: PLEASE DO NOT CITE OR CIRCULATE WITHOUT THE AUTHOR'S PERMISSION

First draft: March 1, 2019. This draft: July 12, 2019

Abstract

In this paper I investigate how parental out-migration affects the schooling outcomes of children left behind in rural China. In particular, I consider three important and widely-studied mechanisms that migration could affect left-behind children's school performance: direct effect through parental accompaniment, and indirect effect through child's study time, and education spending.The major contribution of this paper is to establish a theoretical framework to clarify different pathways involved in the effect of parental migration on child's schooling performance, and to empirically quantify the importance of these pathways on child schooling in rural China. Applying the model on a household-level data from 9 Provinces, I find that parental migration has significant negative total effect on left-behind children's language scores and math scores, and their language scores are more negatively affected. Further decomposing the total effect into the three channels, I find that the direct effect of migration through parental accompaniment is largely negative, and the indirect effects through study time and income are generally negative as well, but are smaller than the direct effect. Subgroup analysis by

*xmluo@ucdavis.edu

child's gender shows consistent findings, but it calls attention to severe underinvestment in left-behind girls' education in rural China. The results from this paper can help policymakers design and implement education policy in rural China by accounting for the specific barriers to education presented by the high degree of parental migration.

Keywords: Rural-to-urban Migration, Education, Structural Equation Model, Direct Effect, Indirect Effect

1 Introduction

"Left-behind children" in this paper refers to children between 0 and 16 years old with at least one parent moving from rural to other areas, while children do not live with both parents and stay in the rural areas where their *hukou* (household registration) are located.

Parental migration and left-behind children are common phenomena in rural China as a consequence of the *hukou* system.There are two types of *hukou* in China: rural and urban *hukou*, and it has been difficult to transfer from one type to the other. Prior to 1970's, people with rural *hukou* were legally prohibited from migrating to urban areas. Since late 1970's, to meet the huge demand for labor in urban areas generated in the economic reform, the Chinese government gradually relaxed the restriction on *hukou* system and allowed people to migrate from rural to urban areas. However, the transfer of *hukou* status is still highly restrictive, and these migrants and their families with rural *hukou* are generally excluded from the social benefits that urban citizens enjoy. The children of rural migrants have limited access to free public schools, health care benefits, housing support, or social security, etc. If the children migrate with their parents from rural to urban areas, in most cases they can only go to either expensive private schools in cities, or to much cheaper "migrant schools", which are run by local entrepreneurs and the quality of education is commonly unsatisfactory. Therefore, instead of bringing their children to cities, most migrant parents choose to leave children behind with their grandparents or other relatives. According to the 2010 Population Census of China, more than 61 million children have been left behind in rural China by migrant parents, accounting for 37.7% of children in rural areas, and 21.88% of children in China overall. Considering the massive number of left behind children in China, the effect of parental migration on left-behind children's educational outcomes has considerable impact on China's accumulation of human capital in the near future.

Despite the magnitude of this issue, the influence of migrant parents on left-behind children's schooling outcomes in rural China is under-studied. Previous studies have investigated specific mechanisms through which parental migration could influence left-behind children. They typically model a single aspect. Some study the effect of migration through parental accompaniment. Antman (2013) finds that absence of parents could lead to less time investment in children and might have psychological costs for left-behind children. Some study the effect through labor substitution. Chen (2013) and Chang et al. (2011) both use the China Health and Nutrition Survey data to examine study time of left-behind children in China, and conclude that children of migrant households spend more time in household work. Other study the effect through income. Remittances sent home by migrating parents could increase education spending, alleviate household financial burdens, and improve children's living conditions, educational investment, and nutrition status. McKenzie and Rapoport (2011); Bryant et al. (2005); Arguillas and Williams (2010); Edwards and Ureta (2003) have provided evidence in Mexico, Indonesia, Thailand, Philippine, and El Salvador.

However, such studies typically rely on reduced-form empirical strategies which estimate the total effect of parental migration, blending direct effect and indirect effects through different mechanisms. I expand on this by separating direct and indirect effects to provide a deeper understanding and clearer guidance for economists and policy makers. On the other hand, reduced-form estimates are unable to reveal the underlying mechanism of how the parental out-migration decision interacts with the child's incentives and behavior. In principle, the child responds to the parental out-migration decision by substituting between schooling and labor market activities, and in turn the parents anticipate the child's behavior, accounting for the child's well-being when they decide to migrate. Understanding these pathways helps economists and policy makers form a more nuanced view of the problem.

The major contribution of this paper is to establish a theoretical framework to clarify different pathways involved in the short-term effect of parental migration on the child's schooling performance, and to empirically quantify the importance of these pathways on child schooling in rural China. In particular, I consider three specific aspects: (1) Study time effect – children might have to do housework or farm work instead of studying due to the absence of parents; (2) Income effect – remittance sent back by migrant parents could relieve the household's credit constraints in child investment, including paying school fees, hiring after-school teachers, sending children to cram schools, purchasing study materials, or buying nutritious food. (3) Parental accompaniment effect – the absence of migrant parent will lead to less time spent with the children to provide mentoring or reinforce the importance of pursuing education. The theoretical model in this paper considers the migration and education decision as a two-agent model of parent and child.

I fit the both reduced form model and the generalized structural equation model to the data collected by the Rural-Urban Migration in China (RUMiC) project. This is household-level survey data from municipalities that are major senders of rural-to-urban migrants in 9 Provinces in China (Wang, 2010). In summary, I find that both migrant fathers and migrant mothers have significant negative total effect on left-behind children's exam scores. The direct effect is much larger than indirect effect, and both effects are negative.

The rest of this paper is organized as follows. Section 2 starts from the most general form of the theoretical model, which forms the basis for empirical analyses. Section 3 introduces the data and variables in detail, followed by a description on the empirical framework in Section 4. All empirical results are presented in Section 5, including the analysis for all samples and subgroups analysis. Section 6 concludes and remarks on the findings.

2 Theoretical Modeling Framework

2.1 A Two-Agent Model

I consider a simple model with a household of one child and one parent, and there's no borrowing or savings in the model. The model considers two periods. In the first period, the parent is at work age and the child is at school age, but the child could also work at home or outside if he or she wants. In the second period, child has grown up and fully entered the labor market while parent has retired, so the household consumption only rely on child's income in the second period.

Let $u_t(c_t)$ be the utility of parent in period t from consumption c_t, where $t \in \{1,2\}$, $\frac{\partial u_t}{\partial c_t} > 0$ and $\frac{\partial^2 u_t}{\partial c_t^2} < 0$. Let e be the human capital level of child in period 1, and e_0 denote the ability gift. Let d be the willingness for parent to migrate out and leave child behind, and $d \in [0,1]$. Let s be the share of time that the child spends studying, so $(1-s)$ denote the share of time that the child spends working. Let W_p be parent income from work and W_k be child income from work. β_p is the discount factor of parent. The utility of parent is

$$\begin{aligned} \max_{d} \quad & u_1(c_1) + \beta_p u_2(c_2), \\ s.t. \quad & c_1 \leq W_p(d) + W_k(s(d)), \\ & c_2 \leq g(e), \\ & e \leq f(d, s(d), c_1(d), e_0). \end{aligned} \tag{1}$$

I assume that $\frac{\partial f}{\partial s} \geq 0$, $\frac{\partial f}{\partial c_1} \geq 0$. These are standard assumptions that education and consumption could weakly increase the production of human capital, and I further assume their decreasing marginal returns to human capital accumulation, i.e., $\frac{\partial^2 f}{\partial s^2} \leq 0$, $\frac{\partial^2 f}{\partial c_1^2} \leq 0$. Furthermore, I assume $\frac{\partial^2 f}{\partial s \partial c_1} = \frac{\partial^2 f}{\partial c_1 \partial s} \leq 0$. In addition, I assume that higher human capital of the child in period 1 will lead to weakly higher income in period 2, that is, $\frac{\partial g}{\partial e} \geq 0$. As for income in period 1, I assume that $\frac{\partial W_p}{\partial d} \geq 0$ and $\frac{\partial W_k}{\partial s} \leq 0$. Parent maximize utility by choosing the optimal d.

For the child, let $\tilde{u}_t$ be the utility of child in period t, β_k be the discounting factor, and other notations are the same as for parent. In particular, I assume $\beta_k < \beta_p$ because child is more myopic compared to parent and cares more about utility in the current period. I assume that child's utility not only comes from consumption c_1, but is also affected by s, and child's utility in period 2 is purely dependent on consumption c_2, where $\frac{\partial \tilde{u}_t}{\partial c_t} > 0$, and $\frac{\partial^2 \tilde{u}_t}{\partial c_t^2} < 0$, $\frac{\partial^2 \tilde{u}_1}{\partial s^2} < 0$. I further assume that $\frac{\partial^2 \tilde{u}_1}{\partial c_1 \partial s} < 0$ and $\frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1} < 0$. Child maximize utility by choosing the optimal level of s. The utility of child is

$$\begin{aligned} \max_{s} \quad & \tilde{u}_1(s, c_1) + \beta_k \tilde{u}_2(c_2), \\ s.t. \quad & c_1 \leq W_p(d) + W_k(s(d)), \\ & c_2 \leq g(e), \\ & e \leq f(d, s(d), c_1(d), e_0). \end{aligned} \tag{2}$$

2.2 Parent Utility Maximization

For parent utility maximization, there is a trade-off between current consumption and future consumption, which guarantees an interior solution. At an interior optimum, from the first order condition,

$$\frac{\partial e}{\partial d} = -\frac{\overbrace{\frac{\partial u_1}{\partial c_1}}^{(+)}[\overbrace{\frac{\partial W_p(d)}{\partial d}}^{(+)} + \overbrace{\frac{\partial W_k(s(d))}{\partial s}}^{(-)}\overbrace{\frac{\partial s(d)}{\partial d}}^{(?)}]}{\underbrace{\beta_p \frac{\partial u_2}{\partial c_2}\frac{\partial g(e)}{\partial e}}_{(+)}}. \tag{3}$$

The sign of each part is marked in Equation (3) [1]. The denominator of $\frac{\partial e}{\partial d}$ is positive, but the sign of its numerator is undetermined because the sign of $\frac{\partial s(d)}{\partial d}$ is undetermined. If $\frac{\partial s(d)}{\partial d} < 0$, then $\frac{\partial e}{\partial d} < 0$, suggesting that the decrease in study time when parent migrate away is correlated with less human capital accumulation. If $\frac{\partial s(d)}{\partial d} > 0$, then the sign of $\frac{\partial e}{\partial d} < 0$ depends on the relative size of $\frac{\partial W_p(d)}{\partial d}$ and $\frac{\partial W_k(s(d))}{\partial s}\frac{\partial s(d)}{\partial d}$, that is, the relative size of change in parent income and child income as migration status changes.

2.3 Child Utility Maximization

For child utility maximization, there is also a trade-off between current consumption and future consumption, which guarantees an interior solution. At an interior equilibrium,

[1] Derivation is in the Appendix

$$\frac{\partial s(d)}{\partial d} = -\frac{\Lambda + \Gamma + \Theta}{\Psi + \Phi}, \tag{4}$$

where

$$\Lambda = \underbrace{\frac{\partial W_k}{\partial s}}_{(-)}[\underbrace{\frac{\partial^2 \tilde{u}_1}{\partial c_1^2}\frac{\partial W_p}{\partial d}}_{(-)} + \underbrace{\beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}}_{(+)}\underbrace{(\frac{\partial^2 f}{\partial c_1 \partial d} + \frac{\partial^2 f}{\partial c_1^2}\frac{\partial W_p}{\partial d})}_{(-)}] > 0,$$

$$\Gamma = \underbrace{\beta_k}_{(+)}\underbrace{(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial W_p}{\partial d})}_{(?)}\underbrace{(\frac{\partial f}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial W_k}{\partial s})}_{(?)}\underbrace{[\frac{\partial^2 \tilde{u}_2}{\partial c_2^2}(\frac{\partial g}{\partial e})^2 + \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial^2 g}{\partial e^2}]}_{(-)},$$

$$\Theta = \underbrace{\frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1}\frac{\partial W_p}{\partial d}}_{(-)} + \underbrace{\beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}}_{(+)}\underbrace{(\frac{\partial^2 f}{\partial s \partial d} + \frac{\partial^2 f}{\partial s \partial c_1}\frac{\partial W_p}{\partial d})}_{(-)} < 0,$$

$$\Psi = 2\underbrace{\frac{\partial^2 \tilde{u}_1}{\partial c_1 \partial s}\frac{\partial W_k}{\partial s}}_{(+)}\underbrace{(1 + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e})}_{(+)} + \underbrace{\frac{\partial^2 f}{\partial c_1^2}\frac{\partial W_k}{\partial s}}_{(+)} > 0,$$

$$\Phi = \underbrace{\frac{\partial^2 \tilde{u}_1}{\partial s^2}}_{(-)} + \underbrace{\frac{\partial \tilde{u}_1}{\partial c_1}\frac{\partial^2 W_k}{\partial s^2}}_{(-)} + \underbrace{\frac{\partial^2 \tilde{u}_1}{\partial c_1^2}(\frac{\partial W_k}{\partial s})^2}_{(-)} + \underbrace{\beta_k (\frac{\partial f}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial W_k}{\partial s})^2}_{(+)}\underbrace{[\frac{\partial^2 \tilde{u}_2}{\partial c_2^2}(\frac{\partial g}{\partial e})^2 + \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial^2 g}{\partial e^2}]}_{(-)} +$$

$$+ \underbrace{\beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}}_{(+)}\underbrace{(\frac{\partial^2 f}{\partial s^2} + \frac{\partial f}{\partial c_1}\frac{\partial^2 W_k}{\partial s^2})}_{(-)} < 0.$$

The sign of each part of $\frac{\partial s(d)}{\partial d}$ is marked in Equation (4)[2]. In Γ, since $\frac{\partial f}{\partial d}$ is unknown, the sign of $(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial W_p}{\partial d})$ is undetermined. Since $\frac{\partial f}{\partial s}$ is positive and $\frac{\partial f}{\partial c_1}\frac{\partial W_k}{\partial s}$ is negative, the sign of $(\frac{\partial f}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial W_k}{\partial s})$ is also undetermined. Therefore, the sign of Γ is unknown. Because neither its numerator or denominator has a determined sign, the sign of $\frac{\partial s(d)}{\partial d}$ is undetermined.

$\frac{\partial s(d)}{\partial d}$ shows how left-behind child's study time changes when parent migrate away, and $\frac{\partial e}{\partial d}$ shows how child human capital changes when parents migrate away. I will estimate $\frac{\partial s(d)}{\partial d}$ and $\frac{\partial e}{\partial d}$ in the empirical section.

[2]Derivation in Appendix.

3 Data

3.1 Data Source

The dataset used in this paper is collected by the Rural-Urban Migration in China (RUMiC) Project, which is an longitudinal survey carried out in China in a five-year time span. This project is a joint effort by the Australian University, University of Queensland, Beijing Normal University, and Institute for the Study of Labor (IZA). Starting in 2008, the project covers 9 provinces or province-level municipalities that are major sending or receiving areas of rural-to-urban migration: Anhui, Chongqing, Guangdong, Hebei, Henan, Hubei, Jiangsu, Sichuan, and Zhejiang. The RUMiC survey includes 8,000 samples in rural household survey (RHS), 5,000 in urban household survey (UHS), and 5,000 in rural-to-urban migrant household survey (MHS), all samples in each category randomly selected in each province.

Since this paper focuses on rural-to-urban migration, data from RHS and MHS can be used for analysis. However, because RHS beats MHS in both sample size and attrition rate (0.4% v.s. 58.4% attrition at the individual level, and 0.1% v.s. 63.6% at the household level, according to Akgüç et al. (2014), this paper restricts the main analysis to rural households. The RHS draws random samples from the annual household income and expenditure surveys carried out in rural villages, and tracks subjects having permanent living addresses.

Survey documents and data for 2008 and 2009 are available. However, since the 2008 dataset does not include important outcome variables such as children's exam scores or study hours, and has no information on migrants' destination or industry information, I only use the cross-sectional data in 2009 survey in this paper. Originally, the dataset has 6899 children in 4843 households. Since the focus of this paper is on school-aged children, the original samples are filtered by children's age, education status, marital status, and parents' age, child history, etc. 2393 children in 1760 households are left in the data after cleaning. The parents in the dataset for analysis come from 81 cities in 9 Provinces, and their migration destinations cover 176 cities in 31 Provinces.

3.2 Descriptive Statistics

In this section, I will use data visualization to briefly show what the data looks like. Figure 1 shows children with different parental migration status. Left-behind children account for roughly 30% of children in rural areas. As introduced in the following section for defining the treatment variable, this is because I use a stricter definition of left-behind children and require parents to migrate away for over 3 months. If I use the same standard as the National Bureau of Statistics in China, then the proportion of left-behind children in my sample is 37.5%, which is quite close to the 37.7% measurement by the National Bureau of Statistics, so the sample I use is quite representative of children in rural areas.

Figure 1

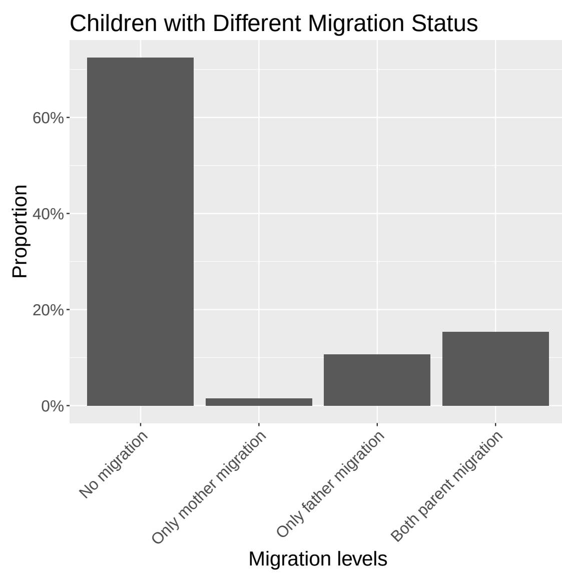
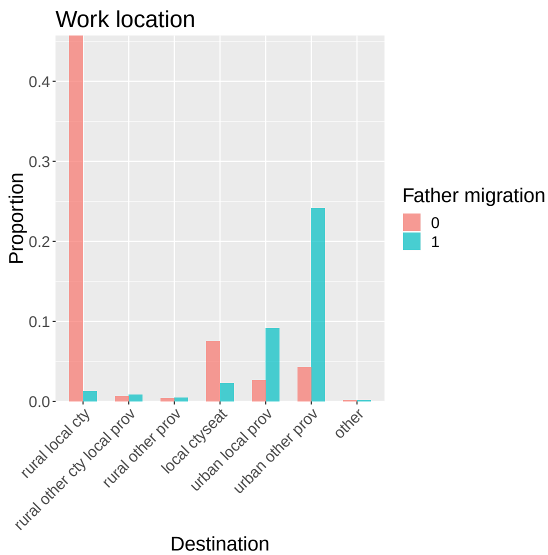


Figure 2 shows the migration destinations for parents. Since father and mother could have different destination, the bar chart is drawn separately for father migration and mother migration. On the x-axis, the first three categories are migration from rural to rural areas, which are rural area in local county, rural area in other county in the same Province, and rural area in other Province. The last two categories are migration from rural to urban areas, which are cities of local Province, and city of other Province. The middle category, local county seat, is in between rural and urban areas, which is less developed than cities but more developed than rural areas. We could see that most people migrate from rural to urban areas, which is the focus of my analysis.

Figure 2: Destination of Work



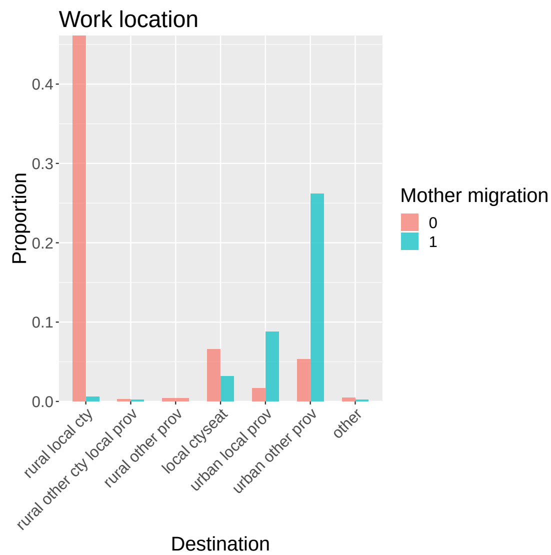


Figure 3 depicts the reasons why parents do not bring children when migrating to work in cities. High living cost and education cost in cities are among the Top 3 reasons. This is partly because of the *hukou* restriction I mentioned in the Introduction. Children with rural *hukou* could hardly benefit from the social benefits such as education and housing, which increases their living cost and education cost if they migrate with their parents. Another important reason is because parents are too busy to take care of children if bringing them along. This is especially true when other family members such as grandparents are unable to migrate together with the parents, so if parents are busy working, they will not have enough time to take care of children.

Figure 4 shows whom the children in rural areas lives with. Usually, we assume it's best for children to live with parents, so the first three categories on the x-axis are the best case scenario, where children live with both parents, or with either father or mother. In the next three categories, children are taken care of by other people, such as grandparents, other relatives, or by teachers at school. In the last case, children live by themselves in off-campus rental room. We could see that when parents migrate away, children are most likely taken care of by

Figure 3

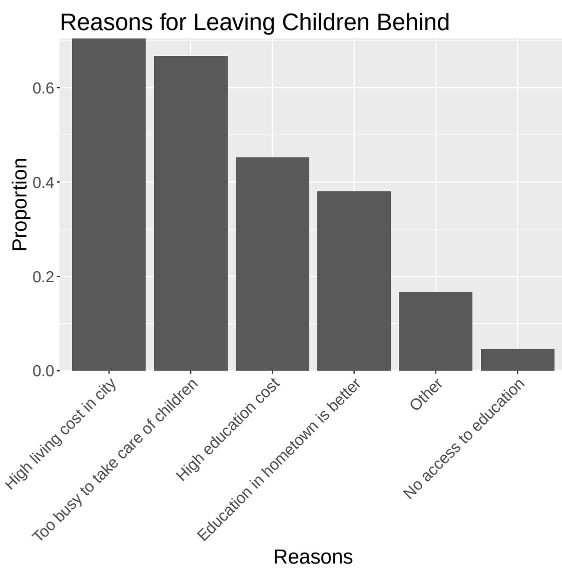
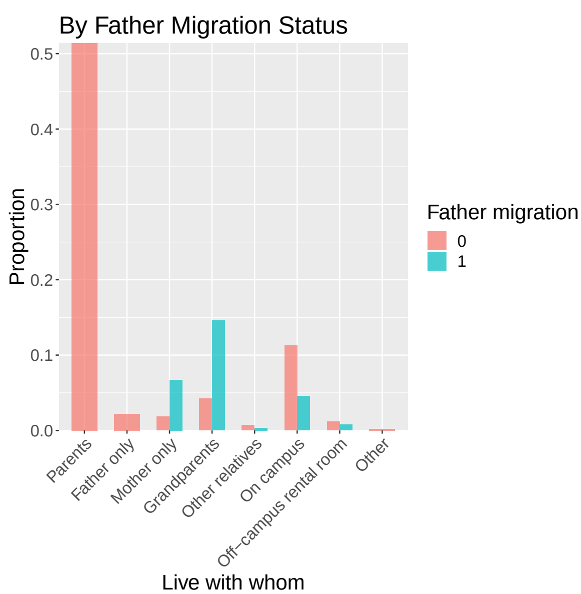
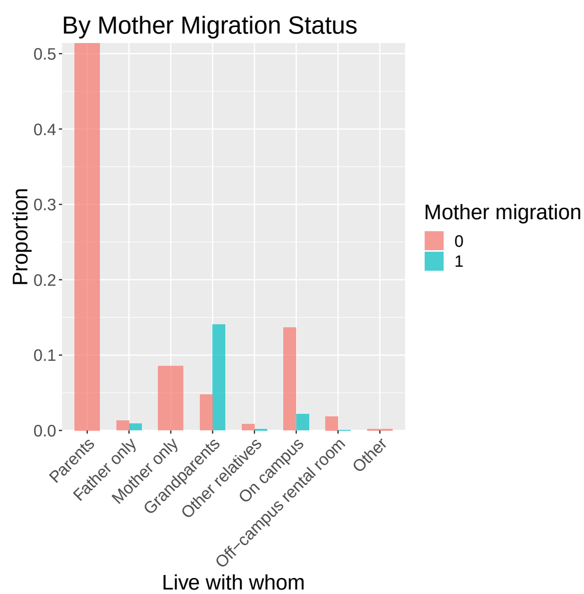


Figure 4





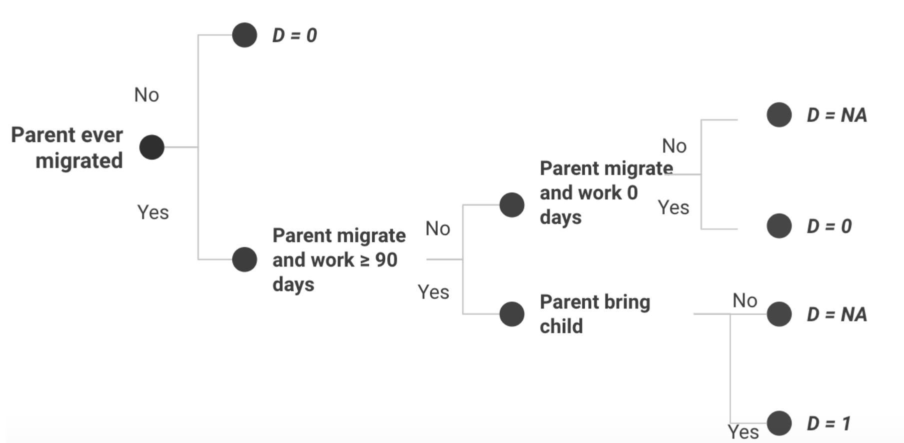
grandparents, who are generally not quite well-educated or have much modern parenting knowledge or skills as children's parents do.

In the next subsections, I will introduce in more details about how the treatment variables, outcome variables, and covariates are defined.

3.3 Treatment Variable

According to Meng and Yamauchi (2015), a good indicator for parental migration is based on very recent migration experience. Based on our models derived, this paper focuses on the binomial decision of left-behind status. For households where both father and mother migrate away from home, a consideration proportion of them have different migration destinations, so I generate the migration decision D separately for father and mother, which corresponds to d in my model setup. Figure 5 shows my definition of the dummy variable for child's left-behind status. The detailed migration destination is only recorded if migrants work away from home for more than 90 days in the last year, so I restrict migrant parent to those who migrate for over 90 days in the past year. In addition, since the control group in this paper is children in rural areas with non-migrant parents, not children who migrate away with their migrant parents, migrant children are not included in analysis. In addition, to keep the difference between treatment and control groups clear, I do not include children whose parents migrate out to work for more than 0 but less than 90 days in analysis.

Figure 5: Definition of Child Left-behind Status



3.4 Variables for Mechanisms

For the measure of child study time T_S, I use the variable recording child's weekly study hour reported by their guardians, which corresponds to s in my model set up. For the measure of spending on child education W_T, it is calculated by adding up spending on child's tuition at school, supplemental classes inside and outside of school, food and accommodation, and sponsorship fees at school in the year 2008.

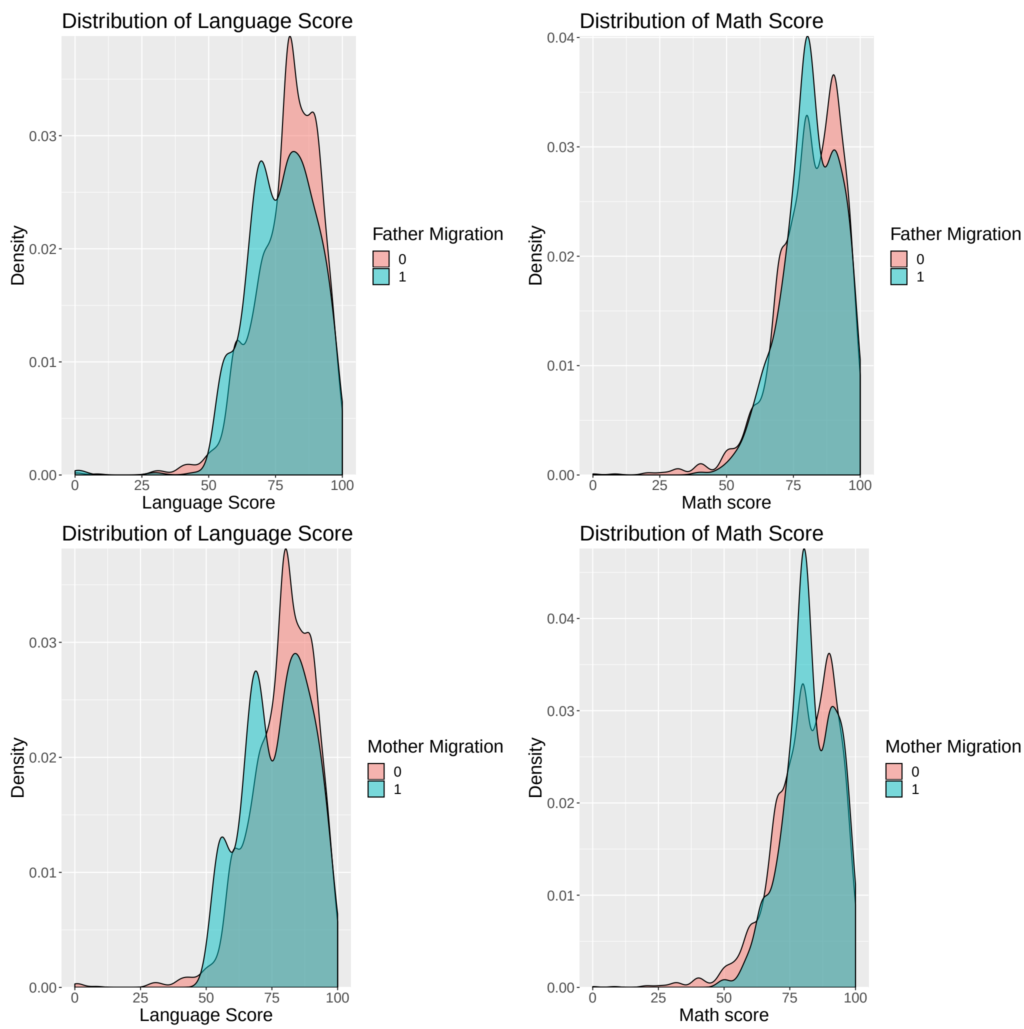
3.5 Dependent Variables

In the model setup, I define child human capital as e. In the RUMiC data, I choose child exam scores P as a proxy for child human capital. The outcome variables used to record children's school performance are final exam scores in the last school term for the subjects of language and mathematics if still at school. Note that since less than 2.5% school-aged children drop out in my sample, the exam scores is not likely biased by the "still at school" requirement.

These outcome variables are reported by parents or guardians, who know children's test scores because they are informed of children's scores during parental meetings at school every semester. In addition, they receive the hard copy of children's score reports from school at the end of every semester. Thus, the reported score is quite reliable. The test scores are also comparable across children in the sample. Since 7 out of 9 provinces use the same version of textbooks, while only a few villages in the remaining 2 provinces use another two versions of textbook. All of the three versions of textbook and exams are designed closely following the Curriculum Standard designed by the Ministry of Education of China. Particularly, the materials are highly consistent for core subjects such as language and mathematics. I normalize test scores by dividing them by the full test score used in the child's school to eliminate the difference caused by grading schemes, which are also reported by parents or guardians. However, the cultural background of different regions might also influence test scores at the province level. For instance, provinces such as Jiangsu and Zhejiang have been famous for culture and education. Therefore, I include provincial dummy variables to account for this factor.

Figure 6 shows the distribution of exam scores. We could see that for left-behind children, the distribution is more right skewed, suggesting that these children perform worse in exams in general. And the difference is more obvious in language scores.

Figure 6



3.6 Covariate Variables

As for other covariates, I first include the personal characteristics of child, such as age, gender, height and weight, birth weight, health status, and whether the child goes to boarding school. I also include the parent characteristics such as the age and years of education. In addition, I include province dummies to account for systematic differences in cultural background and governmental financial support.

Note that some important variables, such as parents' total years of education, have many missing values in the 2009 dataset. Considering that these variables are relatively stable for adults, I replace the missing values in 2009 with variable values from 2008 for people with the same household ID and same household member ID. If the two years records different education years, then the higher one is used for 2009.

Table 1: Summary Statistics

Variable	Migrant Parents	Non-migrant Parents	Difference (P-value)
Dependent Variables			
Language score	77.33	79.45	0.00***
Math score	81.79	81.19	0.27
Covariates: Child			
Male	0.52	0.54	0.35
Age	11.59	12.11	0.00***
Height	138.06	144.91	0.00***
Weight	39.65	42.81	0.00***
Birthweight	3.27	3.23	0.05**
Covariates: Parents			
Mother age	35.96	38.49	0.00***
Father age	37.31	40.21	0.00***
Mother edu year	7.15	7.28	0.19
Father edu year	8.08	8.21	0.19
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 1 shows the summary statistics of some important dependent and independent variables. From the table, left-behind children perform significantly worse than children with non-migrant parents in language exam, but not significantly different in math exam. Left-behind children are also significantly younger, lighter, and shorter than their counterparts. The difference in weight and height is probably due the difference in age, which is then probably due to the difference in parents' age. As shown in the table, migrating parents are significantly younger than non-migrant parents, but the difference in education levels in two groups is not statistically significant. In the empirical analysis, I control for covariates that are significantly different across treatment and control groups, and also include covariates that do not differ significantly to increase estimation efficiency.

4 Empirical Framework

4.1 Reduced Form Model

If we are interested in the total effect of parental migration on the educational outcomes of left-behind children, then we will refer to the reduced form model. The model for estimation is:

$$
\begin{aligned}
P_{ij} &= \delta_0 + \delta \cdot D_{ij} + \xi \cdot X_{ij} + \omega_j + \epsilon_{ij}, \qquad (5)\\
D_{ij} &= \mathbb{1}(a_D + \xi \cdot X_{ij} + \omega_j + \eta_{ij} \geq 0)
\end{aligned}
$$

where P_{ij} is the schooling performance of child i in province j, such as final exam scores in language and mathematics. D_{ij} is the measure for parental migration. To account for individual heterogeneity, other covariates and error terms are included. X_{ij} is the set of control variables, including characteristics of child (study hours, gender, age, birth weight, current weight, current height, health) and parents (education, age). ϵ_{ij} and η_{ij} are random errors. ω_j is province fixed-effect. Note that *exam score* is normalized to a continuous variable ranging from 0 to 100.

Since there are many unobserved factors that correlate with both parental migration decisions and children's school performance, ordinary least squares (OLS) estimates tend to have omitted variable bias. For instance, parents who highly value children's education and development might be less likely to migrate away, and children might study harder and perform better at school because of parents' values and attitudes toward education. Such variables of attitude and values are hard to observe, so the omission of such variables might lead to omitted variable bias, and I need at least one instrumental variable for identification of the total effect. I will introduce the instrumental variables I choose in details in Section 4.3.

4.2 Structural Form Model

If we are interested in the direct and indirect effect of parental migration on the educational outcomes of left-behind children, then we need to refer to the structural form model

$$
\begin{aligned}
P_{ij} &= \gamma_0 + \gamma_T \cdot T_{S_{ij}} + \gamma_W \cdot W_{T_{ij}} + \gamma_D \cdot D_{ij} + \xi \cdot X_{ij} + \omega_j + \phi_{ij} \qquad (6)\\
T_{S_{ij}} &= a_T + b_T \cdot D_{ij} + \xi \cdot X_{ij} + \omega_j + u_{ij}\\
W_{T_{ij}} &= a_W + b_W \cdot D_{ij} + \xi \cdot X_{ij} + \omega_j + v_{ij}\\
D_{ij} &= \mathbb{1}(a_D + \xi \cdot X_{ij} + \omega_j + \zeta_{ij} \geq 0)
\end{aligned}
$$

In this model, D_{ij} has direct effect on performance denoted by γ_D, and the indirect effect on performance through study time $T_{S_{ij}}$ is denoted by $b_T \cdot \gamma_T$, and the indirect effect through income $W_{T_{ij}}$ is denoted by $b_W \cdot \gamma_W$. These paths depict the mechanisms of interest. The covariates X_{ij} and province fixed-effects are defined in the same way as in the reduced form model. The error terms ϕ_{ij}, u_{ij}, v_{ij}, and ζ_{ij} are random errors. In this structural equation model, not only D_{ij} is an endogenous variable, as discussed in Section 4.1, but $T_{S_{ij}}$ and $W_{T_{ij}}$ are also endogenous. To identify the coefficients in the structural equation model, I need at least three instrumental variables for identification of the direct and indirect effects, and I will explain in Section 4.3.

4.3 Choice of Instrumental Variables

Some popular choices of instrumental variables for migration decision include: religious preference uncommon in urban locations, dummy variable indicating whether the householder's first occupation was as a farmer, distance from home village to provincial capital, and the average migration rate in the village (Fisher, 2005; Xiang et al., 2016; Meng and Yamauchi, 2015). However, these are not excellent choices for the scope of this research. First, religion in China is not widespread, and all religions are common ones, so the uncommon religious preference variable is not quite feasible. Second, the householder's first occupation as a farmer is also not quite feasible since the data of this research is in rural China, where farming is the fundamental industry and the coverage of farmers is predominantly high, and this instrument still suffers from endogeneity issue. Third, the distance from home village to provincial capital also suffers from endogeneity because parents from villages closer to the capital have lower migration cost and thus are more likely to migrate, and the general education facilities in these regions are possibly better, leading to better schooling outcomes in children. This concern could be relieved if school size, school rank, the number and quality of teachers, class size, or per capita educational investments in each village are taken into consideration. But unfortunately, these variables are not controlled for in the paper mentioned above. Last, the average migration rate would not only influence the migration decision of each household, but also influence tax revenues and educational investment in the region, thereby influencing the schooling outcomes of children.

For the reduced form model, this paper follows the method of Bartik (1991) and uses a Bartik-like instrument that combines migrants' destination-industry information with changes in employment rate at destination by industry. The migration information is generated based on migrant's origin city, destination city, and the industry they work for using data from China 1% National Population Sample Survey 2005. The employment information is extracted from Urban Statistical Yearbook of China. The change in employment rate is generated using 2007 and 2008 employment data of each industry in all cities in China. These years are chosen such that there is sufficient time for migration flow to change as employment changes, but not too early so that the correlation between migration and employment would fade away. The Bartik instrument is generated as below:

$$Z_{B\,o,2008} = \frac{\sum_{d=1}^{D}\sum_{k=1}^{K}(Mig_{o,d,k,2005} \cdot \triangle Employment_{d,k,2007-2008})}{\sum_{d=1}^{D}\sum_{k=1}^{K} Mig_{o,d,k,2005}},$$

where o stands for origin city of migrants, d stands for their destination city, and k represents the industry that migrants work for. $Mig_{o,d,k,2005}$ is the total number of migrant workers from city o to city d that work in industry k in 2005. $\triangle Employment_{d,k,2007-2008}$ is the estimator of the industry growth rate of industry k in destination d during 2007 and 2008. Since the migration in my analysis is composed of both inter-city migration and within-city migration, I generate two Bartik instruments Z_{IB} and Z_{WB} respectively. Bartik instrument is widely used in migration literature. It is correlated with migration decision, but is arguably exogenous in the equations of performance, study time, and income, which makes it a valid instrument.

For the structural form model, in addition to Bartik instruments, I included two other instruments. One instrument is the size of farmable land in the household, Z_W. The other is adult male share Z_T in the household, defined as

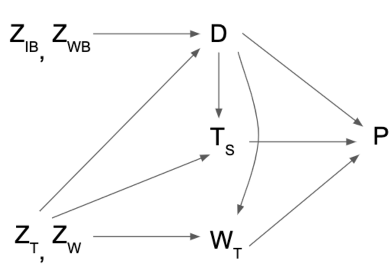
$$Z_T = \frac{Number\ of\ adult\ males\ in\ household}{Household\ size}.$$

These two instrumental variables are correlated with migration decision, and arguably, they are exogenous to performance. But unlike Bartik instruments, they may not be exogenous to study time or education spending.

4.4 Identification of Coefficients

With Bartik instruments, the coefficient on migration decision could be identified in the reduced form model. As for the structural model, one necessary and sufficient condition for identification is order condition, that is, for each equation, the number of excluded exogenous variables should be larger or equal to the number of included endogenous variables minus one. Figure 7 is a diagram showing the the paths of effect on child performance after including all the exogenous instrumental variables. Path diagram is an alternative representation of structural equation model, where each edge represents the inclusion of a variable into a certain equation. For instance, in the path diagram below, at the performance node P, there are three edges pointing to it: migration decision D, child's study hours T_S, and total education spending W_T, and it is equivalent to Equation 1 in the structural equation model to its right. By the same reasoning, the four equations to the right of the path diagram is equivalent to the diagram.

Figure 7: Path diagram and Structural Equation Model



$$Eq1 : P \sim D + T_S + W_T$$
$$Eq2 : T_S \sim D + Z_T + Z_W$$
$$Eq3 : W_T \sim D + Z_T + Z_W$$
$$Eq4 : D \sim Z_{IB} + Z_{WB} + Z_T + Z_W$$

Table 2: Order Condition of Structural Equation Model

	# Excluded Exogenous	# Included Endogenous - 1
Eq 1	4	3
Eq 2	2	1
Eq 3	2	1
Eq 4	0	0

For the entire system of equations, P, T_S, and W_T are endogenous variables, and instrumental variables Z_{IB}, Z_{WB}, Z_T, and Z_W are exogenous variables. As shown in Table 2, the order conditions are satisfied for all equations, and thus all the coefficients in the structural equation model in Equation (6) are identifiable. Although the covariates X are left out of the equations and the path diagram for simplicity, it will not change the result of order conditions. With the identified structural model, if we define δ to be the total effect of migration on children's schooling outcomes, then the total effect can be decomposed into the following three part:

$$\delta = \gamma_D(parental\ accompaniment) + \gamma_T b_T(time\ allocation) + \gamma_W b_W(income), \quad (7)$$

where γ_D captures the direct effect of migration, $\gamma_T b_T$ captures the indirect effect of migration through child's study time, and $\gamma_W b_W$ captures the indirect effect of migration through total education spending.

4.5 Nonrandom Missing Patterns

Previous studies simply remove observations with missing values in empirical analysis without accounting for nonrandom missing patterns. However, in my samples, I find that children with missing values in study time and education spending perform much worse than those with non-missing values, and these two variables are particularly important in studying the indirect effects of parental migration. Simply removing the observations with missing values in these variables will lead to wrong estimation of the effect of migration. Instead, I use the Heckman model to impute for the missing values in these two variables. Comparison of results with and without imputation are included in Section 5.

5 Empirical Results

5.1 Results on All Samples

5.1.1 Reduced Form

The reduced form model is estimated using maximum likelihood with a non-linear first stage. Results in Table 3 are based on samples with imputed study time and education spending measures. Panel A of Table 3 shows the effect of parental migration on left-behind children's exam scores based on the reduced form model. Since no mechanisms or paths are specified, the reported effect can be regarded as the total effect that blends direct and indirect effects from different pathways. Out of a 100 point scale, on average, language scores of children whose father out-migrates is about 9 points lower than that of children whose father does not, and this difference is significant at the 1% level. Language scores of left-behind by migrant mothers is roughly 6 points lower than that of children whose mother does not, and this difference is also significant at 1% level. The math scores of children whose fathers out-migrate is about 6 points lower than those of children whose fathers do not migrant, whereas that of children whose mother out-migrates tends to be 5 points lower than that of children whose mother does not, and both differences are significant at 1% level.

In summary, parental migration has significant negative total effect on left-behind children's language and math scores, and children's language scores are affected more negatively. Mother migration tends to yield a slightly smaller negative total effect than father migration.

5.1.2 Structural Form

The structural form is estimated with maximum likelihood with the four equations derived in Section 4.2, which focuses on the direct effect of parental migration and leaving children behind through parental accompany (γ_D), and the indirect effects through children's study time($\gamma_T b_T$) and household's spending on children's education ($\gamma_W b_W$).

Panel B of Table 3 shows the direct effect, indirect effect, and total effect of parental migration on left-behind children's exam scores based on the structural equation model. For the direct effect on language scores, children whose father out-migrates achieve almost 6 points lower than children whose father does not, and this difference is significant at the 1% level; children whose mother out-migrates achieve roughly 7 points lower than children whose mother does not, and this difference is also significant at 1% level. For direct effect on math score, children whose father out-migrates achieve 3 points lower than children whose fathers do not migrate, whereas children whose mother out-migrates achieve slightly less than 3 points lower than children whose mothers do not migrate. The direct effect on math scores are both significant at the 5% level.

As for indirect effect, parental migration has significant negative indirect effect on left-behind children's language scores through study time, but the effect on math score through this mechanism is small in size and insignificant, and it has generally negative effect through education spending. Mother migration has a larger negative effect on language scores through child's reduced study time, and father migration has a larger negative effect on language scores through reduced educational spendings.

5.2 Results without Imputation

Table 4 shows results using exactly the same methods as in Table 3, and the only difference is that observations with missing study time or education spending are simply removed in Table 4. Comparing with Table 3, the number of observations immediately shrink from 2199 to 1148, and the total effect in reduced form analysis and the direct and indirect effects in structural form analysis generally shrink in size. Again, this shows that children with missing values in these measures are those who are more negatively affected by parental outmigration, so simply removing observations with missing values will underestimate the negative effect of being left behind. This confirms the necessity to impute for the missing values.

5.3 Exploring Heterogeneous Treatment Effects

I also investigate the heterogeneous effects in different subgroups partitioned by gender. For each subgroup, I repeat the process of estimating the structural

Table 3: Effect of Parental Migration on Child Schooling Outcomes (All Samples, Imputed)

Panel A: Reduced Form				
	Language Score		Math Score	
	(1) Father	(2) Mother	(3) Father	(4) Mother
Effect	-8.967***	-6.234***	-6.205***	-4.665***
	(0.001)	(0.000)	(0.004)	(0.002)
First Stage				
Inter Bartik	1.587**	2.444***	1.718**	2.613***
	(0.017)	(0.002)	(0.026)	(0.002)
Within Bartik	-1.179***	-1.585***	-1.335***	-1.667***
	(0.000)	(0.000)	(0.000)	(0.000)
Adult male share	0.187	0.390	0.228	0.462
	(0.328)	(0.124)	(0.319)	(0.106)
Farm size	0.019***	0.029***	0.016**	0.026***
	(0.005)	(0.001)	(0.033)	(0.002)
Province FE	Yes	Yes	Yes	Yes
N	2199	2199	2199	2199
Panel B: Structural Form				
Direct Effect				
Parental Accompany	-5.685***	-7.290***	-2.936**	-2.786**
	(0.000)	(0.000)	(0.020)	(0.011)
Indirect Effect				
Study Time	-0.173**	-2.424**	-0.037	-0.140
	(0.038)	(0.023)	(0.613)	(0.211)
Education Spending	-2.150**	-0.201	-1.353**	-1.631***
	(0.013)	(0.295)	(0.010)	(0.001)
First Stage				
Inter Bartik	1.678***	1.865***	1.688**	2.286 ***
	(0.007)	(0.005)	(0.019)	(0.002)
Within Bartik	-1.397***	-0.577**	-1.656***	-1.759***
	(0.000)	(0.030)	(0.000)	(0.000)
Adult male share	0.276	0.235	0.328	0.451
	(0.120)	(0.325)	(0.116)	(0.055)
Farm size	0.019***	0.023***	0.009***	0.018***
	(0.009)	(0.004)	(0.236)	(0.026)
Province FE	Yes	Yes	Yes	Yes
N	2199	2199	2199	2199

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 4: Effect of Parental Migration on Child Schooling Outcomes (All Samples, Not Imputed)

	Language Score		Math Score	
Panel A: Reduced Form				
	(1) Father	(2) Mother	(3) Father	(4) Mother
Effect	-4.639***	-4.164***	-3.026**	-2.779**
	(0.005)	(0.005)	(0.044)	(0.036)
First Stage				
Inter Bartik	2.731**	3.356***	2.980**	3.654***
	(0.029)	(0.010)	(0.024)	(0.006)
Within Bartik	-1.795***	-1.986***	-1.822***	-2.011***
	(0.000)	(0.000)	(0.000)	(0.001)
Adult male share	0.735**	0.786*	0.789**	0.848*
	(0.042)	(0.075)	(0.040)	(0.076)
Farm size	0.034***	0.038***	0.031***	0.034**
	(0.004)	(0.004)	(0.010)	(0.013)
Province FE	Yes	Yes	Yes	Yes
N	1148	1148	1148	1148
Panel B: Structural Form				
Direct Effect				
Parental Accompany	-3.453***	-3.231***	-1.669	-1.375
	(0.005)	(0.008)	(0.146)	(0.228)
Indirect Effect				
Study Time	-0.077 **	-0.088	0.079	0.075
	(0.019)	(0.431)	(0.506)	(0.484)
Education Spending	-2.232	-2.394**	-1.898**	-1.570**
	(0.505)	(0.014)	(0.013)	(0.011)
First Stage				
Inter Bartik	2.280**	2.597***	2.455**	2.480**
	(0.017)	(0.009)	(0.015)	(0.010)
Within Bartik	-1.582***	-1.592**	-1.656***	-1.853***
	(0.000)	(0.000)	(0.000)	(0.000)
Adult male share	0.469*	0.431	0.488*	0.487
	(0.084)	(0.144)	(0.091)	(0.144)
Farm size	0.025**	0.026**	0.019*	0.019
	(0.028)	(0.035)	(0.087)	(0.120)
Province FE	Yes	Yes	Yes	Yes
N	1148	1148	1148	2199

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

equation models using data with imputed study time and education spending, and the results are presented in Table 5.

As shown in Table 5, parental migration has a much larger negative direct effect on left-behind boys than on girls, and children are more negatively affected in language scores, which is consistent with our finding in Table 4. As for indirect effects, parental migration has a large indirect negative effect on boy's exam scores through reduced study time, and this effect is particularly significant when mothers migrate away. This might be partially explained by the role that mother plays in child's education, and by the difference in time management skills and study habits between boys and girls. For girls, parental migration has a large negative indirect effect on their exam scores through reduced educational spendings, and this is significant no matter it is father or mother who migrates, although the effect size due to father migration is much larger. This finding might be partially explained by the unfair treatment of girls and underinvestment in girl's education in rural China, especially when the girl's parents migrate away.

6 Conclusion and Remarks

In this paper I established a theoretical framework to unify different pathways including parental accompaniment, children's study time, and education spending. The empirical analysis uses the household-level data from 9 Provinces that are major sending areas of rural-to-urban migration. Both the reduced and the structural models show significant negative direct effect of father and mother migration on left-behind children's language scores and math scores, and language scores tend to be more negatively affected. The different indirect effect patterns for father and mother migration might be explained by different roles that father and mother play in child's study time management and education investment, which is worth further exploring and has significant migration policy implications.

Structural form results by subgroups complements the reduced form subgroup results, and reveals how parental migration affect left-behind children differently in different pathways. Results from subgroup analysis by gender draws attention to time management issues of left-behind boys and severe underinvestment in education for left-behind girls in rural China. Understanding these pathways helps

Table 5: Effect of Parental Migration on Child Schooling Outcomes (Structural Form, Imputed, Subgroup by Gender)

	Language Score		Math Score	
Panel A: Boys				
	(1) Father	(2) Mother	(3) Father	(4) Mother
Direct Effect				
Parental Accompany	-8.856**	-6.751**	-2.868*	-5.388**
	(0.014)	(0.012)	(0.083)	(0.050)
Indirect Effect				
Study Time	-3.714	-3.598*	-0.134	-6.485**
	(0.107)	(0.077)	(0.574)	(0.057)
Education Spending	-0.737	-0.339	-1.366**	-0.111
	(0.245)	(0.300)	(0.018)	(0.403)
First stage				
Inter Bartik	1.889**	2.478***	1.675*	2.042**
	(0.015)	(0.005)	(0.076)	(0.014)
Within Bartik	-0.447	-0.262	-1.706***	0.093
	(0.109)	(0.411)	(0.000)	(0.716)
Adult male share	-0.064	-0.119	0.043	-0.072
	(0.808)	(0.727)	(0.875)	(0.848)
Farm	0.016*	0.021**	0.020**	0.022**
	(0.061)	(0.025)	(0.020)	(0.015)
Province FE	Yes	Yes	Yes	Yes
N	1202	1202	1202	1202
Panel B: Girls				
Direct Effect				
Parental Accompany	-6.848***	-5.089***	-3.866**	-3.215**
	(0.003)	(0.000)	(0.012)	(0.012)
Indirect Effect				
Study Time	-0.020	-0.117	-0.016	-0.123
	(0.720)	(0.119)	(0.486)	(0.153)
Education Spending	-6.952*	-1.659*	-3.405*	-1.470**
	(0.089)	(0.089)	(0.053)	(0.022)
First stage				
Inter Bartik	0.417	0.938	0.989	1.541
	(0.479)	(0.346)	(0.228)	(0.172)
Within Bartik	-0.881*	-1.713***	-1.354***	-2.075***
	(0.055)	(0.001)	(0.008)	(0.000)
Adult male share	0.441*	0.971**	0.534*	1.062**
	(0.097)	(0.012)	(0.068)	(0.012)
Farm	0.014	0.032**	-0.008	0.007
	(0.186)	(0.011)	(0.522)	(0.602)
Province FE	Yes	Yes	Yes	Yes
N	997	997	997	997

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

economists and policy makers form a more nuanced view of the problem, and separating direct and indirect effects could provide a clearer guidance for policy makers to make policies addressing specific influence mechanism for specified subgroups.

Although I consider a particular specification, our model is not limited to this setting. In principle, for any utility functions and any functional relationship between the children schooling performance and other variables, one can derive the general equilibrium. The only technical difficulty lies in the econometric tools to handle the complicated nonlinear structural form models. I leave it to future works. On the other hand, it is straightforward to add other pathways into this theoretical framework. For instance, If I expect an interaction effect among children and collect the data that provides such information, I can build this into the utility maximization part by incorporating the interference. This complicates the model into a multi-agents setting and the general equilibrium can be derived in principle. I also leave it as future research. In addition, the empirical results show different patterns indirect effects through study time and income for different subgroups, which can be further explored in the future.

The results from this paper can help policymakers design and implement education policy in rural China by accounting for the specific barriers to education presented by the high degree of parental migration. In addition, the methodology can be used in other settings to evaluate the effect of parental migration on the education of left-behind children.

Appendix

Appendix A. Parental Utility Maximization

The utility of parent is

$$\begin{aligned}
\max_{d} \quad & u_1(c_1) + \beta_p u_2(c_2), \\
s.t. \quad & c_1 \leq W_p(d) + W_k(s(d)), \\
& c_2 \leq g(e), \\
& e \leq f(d, s(d), c_1(d), e_0).
\end{aligned}$$

Plugging constraints to the utility function

$$L = u_1(W_p(d) + W_k(s(d)) + \beta_p u_2(e).$$

Taking derivative with respect to d and obtain the first order condition

$$\begin{aligned}
& \frac{\partial L}{\partial d} = \frac{\partial u_1}{\partial c_1}(\frac{\partial W_p(d)}{\partial d} + \frac{\partial W_k(s)}{\partial s}\frac{\partial s(d)}{\partial d}) + \beta_p \frac{\partial u_2}{\partial c_2}\frac{\partial g}{\partial e}\frac{\partial e}{\partial d} = 0, \\
\Rightarrow & \frac{\partial e}{\partial d} = -\frac{\frac{\partial u_1}{\partial c_1}[\frac{\partial W_p(d)}{\partial d} + \frac{\partial W_k(s(d))}{\partial s}\frac{\partial s(d)}{\partial d}]}{\beta_p \frac{\partial u_2}{\partial c_2}\frac{\partial g(e)}{\partial e}}.
\end{aligned}$$

Appendix B. Child Utility Maximization

The utility of child is

$$\begin{aligned}
\max_{s} \quad & \tilde{u}_1(s, c_1) + \beta_k \tilde{u}_2(c_2), \\
s.t. \quad & c_1 \leq W_p(d) + W_k(s(d)), \\
& c_2 \leq g(e), \\
& e \leq f(d, s(d), c_1(d), e_0).
\end{aligned}$$

Plugging constraints to utility function

$$\tilde{L} = \tilde{u}_1(s, W_p + W_k(s)) + \beta_k \tilde{u}_2(g(f(s, c_1(s))))$$

Taking derivative with respect to s and obtain the first order condition

$$\frac{\partial \tilde{L}}{\partial s} = \frac{\partial \tilde{u}_1}{\partial s} + \frac{\partial \tilde{u}_1}{\partial c_1}\frac{\partial W_k}{\partial s} + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}(\frac{\partial f}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial s}) = 0.$$

This is a simplified version of first order condition. Writing out arguments of each part, we have

$$\frac{\partial \tilde{L}}{\partial s} = A + B + C \cdot (D + E) = 0,$$

where

$$\begin{aligned}
A &= \frac{\partial \tilde{u}_1}{\partial s}(s(d), W_p(d) + W_k(s(d))), \\
B &= \frac{\partial \tilde{u}_1}{\partial c_1}(s(d), W_p(d) + W_k(s(d))) \cdot \frac{\partial W_k}{\partial s}(s(d)), \\
C &= \beta_k \cdot \frac{\partial \tilde{u}_2}{\partial c_2}(c_2(d)) \cdot \frac{\partial g}{\partial e}(e(d)), \\
D &= \frac{\partial f}{\partial s}(d, s(d), W_p(d) + W_k(s(d))), \\
E &= \frac{\partial f}{\partial c_1}(d, s(d), W_p(d) + W_k(s(d))) \cdot \frac{\partial W_k}{\partial s}(s(d))
\end{aligned}$$

The goal is to study the effect of d on s^*, so taking the derivative of $\frac{\partial \tilde{L}}{\partial s}$ with respect to d,

$$\frac{\partial^2 \tilde{L}}{\partial s \partial d} = A' + B' + C' \cdot (D + E) + C \cdot (D' + E')$$

$$A' = \frac{\partial^2 \tilde{u}_1}{\partial s^2}\frac{\partial s(d)}{\partial d} + \frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1}(\frac{\partial W_p}{\partial d} + \frac{\partial W_k}{\partial s}\frac{\partial s(d)}{\partial d}) = \frac{\partial s(d)}{\partial d} \cdot A_1 + A_2,$$

where

$$\begin{aligned}
A_1 &= \frac{\partial^2 \tilde{u}_1}{\partial s^2} + \frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1}\frac{\partial W_k}{\partial s}, \\
A_2 &= \frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1}\frac{\partial W_p}{\partial d}. \\
B' &= \frac{\partial \tilde{u}_1}{\partial c_1}\frac{\partial^2 W_k}{\partial s^2}\frac{\partial s(d)}{\partial d} + \frac{\partial W_k}{\partial s}[\frac{\partial^2 \tilde{u}_1}{\partial c_1 \partial s}\frac{\partial s(d)}{\partial d} + \frac{\partial^2 \tilde{u}_1}{\partial c_1^2}(\frac{\partial W_p}{\partial d} + \frac{\partial W_k}{\partial s})\frac{\partial s(d)}{\partial d}] = \frac{\partial s(d)}{\partial d} \cdot B_1 + B_2,
\end{aligned}$$

where

$$\begin{aligned}
B_1 &= \frac{\partial \tilde{u}_1}{\partial c_1}\frac{\partial^2 W_k}{\partial s^2} + \frac{\partial^2 \tilde{u}_1}{\partial c_1 \partial s}\frac{\partial W_k}{\partial s} + \frac{\partial^2 \tilde{u}_1}{\partial c_1^2}(\frac{\partial W_k}{\partial s})^2, \\
B_2 &= \frac{\partial^2 \tilde{u}_1}{\partial c_1^2}\frac{\partial W_k}{\partial s}\frac{\partial W_p}{\partial d}.
\end{aligned}$$

$$
\begin{aligned}
C' &= \beta_k \frac{\partial^2 \tilde{u}_2}{\partial c_2^2} \cdot (\frac{\partial g}{\partial e})^2 \cdot [\frac{\partial f}{\partial d} + \frac{\partial f}{\partial s}\frac{\partial s(d)}{\partial d} + \frac{\partial f}{\partial c_1}(\frac{\partial W_p}{\partial d} + \frac{\partial W_k}{\partial s}\frac{\partial s(d)}{\partial d})]+ \\
&\quad \beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial^2 g}{\partial e^2}[\frac{\partial f}{\partial d} + \frac{\partial f}{\partial s}\frac{\partial s(d)}{\partial d} + \frac{\partial f}{\partial c_1}(\frac{\partial W_p}{\partial d} + \frac{\partial W_k}{\partial s}\frac{\partial s(d)}{\partial d})] \\
&= \frac{\partial s(d)}{\partial d} \cdot C_1 + C_2,
\end{aligned}
$$

where

$$
\begin{aligned}
C_1 &= \beta_k(\frac{\partial f}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial W_k}{\partial s})[\frac{\partial^2 \tilde{u}_2}{\partial c_2^2}(\frac{\partial g}{\partial e})^2 + \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial^2 g}{\partial e^2}], \\
C_2 &= \beta_k(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial W_p}{\partial d})[\frac{\partial^2 \tilde{u}_2}{\partial c_2^2}(\frac{\partial g}{\partial e})^2 + \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial^2 g}{\partial e^2}]. \\
D' &= \frac{\partial^2 f}{\partial s \partial d} + \frac{\partial^2 f}{\partial s^2}\frac{\partial s(d)}{\partial d} + \frac{\partial^2 f}{\partial s \partial c_1}(\frac{\partial W_p}{\partial d} + \frac{\partial W_k}{\partial s}\frac{\partial s(d)}{\partial d}) = \frac{\partial s(d)}{\partial d} \cdot D_1 + D_2,
\end{aligned}
$$

where

$$
\begin{aligned}
D_1 &= \frac{\partial^2 f}{\partial s^2} + \frac{\partial^2 f}{\partial s \partial c_1}\frac{\partial W_k}{\partial s}, \\
D_2 &= \frac{\partial^2 f}{\partial s \partial d} + \frac{\partial^2 f}{\partial s \partial c_1}\frac{\partial W_p}{\partial d} \\
E' &= \frac{\partial W_k}{\partial s}[\frac{\partial^2 f}{\partial c_1 \partial d} + \frac{\partial^2 f}{\partial c_1 \partial s}\frac{\partial s(d)}{\partial d} + \frac{\partial^2 f}{\partial c_1^2}(\frac{\partial W_p}{\partial d} + \frac{\partial W_k}{\partial s}\frac{\partial s(d)}{\partial d})] + \frac{\partial f}{\partial c_1}\frac{\partial^2 W_k}{\partial s^2}\frac{\partial s(d)}{\partial d} \\
&= \frac{\partial s(d)}{\partial d} \cdot E_1 + E_2,
\end{aligned}
$$

where

$$
\begin{aligned}
E_1 &= \frac{\partial W_k}{\partial s}\frac{\partial^2 f}{\partial c_1 \partial s} + \frac{\partial^2 f}{\partial c_1^2}\frac{\partial W_k}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial^2 W_k}{\partial s^2}, \\
E_2 &= \frac{\partial W_k}{\partial s}(\frac{\partial^2 f}{\partial c_1 \partial d} + \frac{\partial^2 f}{\partial c_1^2}\frac{\partial W_p}{\partial d}).
\end{aligned}
$$

Therefore,

$$\begin{aligned}\frac{\partial s(d)}{\partial d} &= -\frac{A_2 + B_2 + C_2 \cdot (D+E) + C \cdot (D_2 + E_2)}{A_1 + B_1 + C_1 \cdot (D+E) + C \cdot (D_1 + E_1)} \\ &= -\frac{\Lambda + \Gamma + \Theta}{\Psi + \Phi},\end{aligned}$$

where

$$\begin{aligned}
\Lambda &= \frac{\partial W_k}{\partial s}[\frac{\partial^2 \tilde{u}_1}{\partial c_1^2}\frac{\partial W_p}{\partial d} + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}(\frac{\partial^2 f}{\partial c_1 \partial d} + \frac{\partial^2 f}{\partial c_1^2}\frac{\partial W_p}{\partial d})], \\
\Gamma &= \beta_k(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial W_p}{\partial d})(\frac{\partial f}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial W_k}{\partial s})[\frac{\partial^2 \tilde{u}_2}{\partial c_2^2}(\frac{\partial g}{\partial e})^2 + \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial^2 g}{\partial e^2}], \\
\Theta &= \frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1}\frac{\partial W_p}{\partial d} + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}(\frac{\partial^2 f}{\partial s \partial d} + \frac{\partial^2 f}{\partial s \partial c_1}\frac{\partial W_p}{\partial d}), \\
\Psi &= 2\frac{\partial^2 \tilde{u}_1}{\partial c_1 \partial s}\frac{\partial W_k}{\partial s}(1 + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}) \\
\Phi &= \frac{\partial^2 \tilde{u}_1}{\partial s^2} + \frac{\partial \tilde{u}_1}{\partial c_1}\frac{\partial^2 W_k}{\partial s^2} + \frac{\partial^2 \tilde{u}_1}{\partial c_1^2}(\frac{\partial W_k}{\partial s})^2 + \beta_k(\frac{\partial f}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial W_k}{\partial s})^2[\frac{\partial^2 \tilde{u}_2}{\partial c_2^2}(\frac{\partial g}{\partial e})^2 + \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial^2 g}{\partial e^2}] + \\
&\quad + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}(\frac{\partial^2 f}{\partial s^2} + \frac{\partial^2 f}{\partial c_1^2}\frac{\partial W_k}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial^2 W_k}{\partial s^2}).
\end{aligned}$$

References

Akgüç, M., Giulietti, C., and Zimmermann, K. F. (2014). The rumic longitudinal survey: Fostering research on labor markets in china. *IZA Journal of Labor & Development*, 3(1):5.

Antman, F. M. (2013). 16 the impact of migration on family left behind. *International handbook on the economics of migration*, page 293.

Arguillas, M. J. B. and Williams, L. (2010). The impact of parents' overseas employment on educational outcomes of filipino children. *International Migration Review*, 44(2):300–319.

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Bryant, J. et al. (2005). Children of international migrants in indonesia, thailand, and the philippines: A review of evidence and policies.

Chang, H., Dong, X.-y., and MacPhail, F. (2011). Labor migration and time use patterns of the left-behind children and elderly in rural china. *World Development*, 39(12):2199–2210.

Chen, J. J. (2013). Identifying non-cooperative behavior among spouses: child outcomes in migrant-sending households. *Journal of Development Economics*, 100(1):1–18.

Edwards, A. C. and Ureta, M. (2003). International migration, remittances, and schooling: evidence from el salvador. *Journal of development economics*, 72(2):429–461.

Fisher, M. (2005). On the empirical finding of a higher risk of poverty in rural areas: Is rural residence endogenous to poverty? *Journal of Agricultural and Resource Economics*, pages 185–199.

McKenzie, D. and Rapoport, H. (2011). Can migration reduce educational attainment? evidence from mexico. *Journal of Population Economics*, 24(4):1331–1358.

Meng, X. and Yamauchi, C. (2015). Children of migrants: The impact of parental migration on their children's education and health outcomes.

Wang, P. (2010). *2010 Report on China's Migration Population Development*. China Population Publishing House.

Xiang, A., Jiang, D., and Zhong, Z. (2016). The impact of rural–urban migration on the health of the left-behind parents. *China Economic Review*, 37:126–139.