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

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PRELIMINARY VERSION: PLEASE DO NOT CITE OR CIRCULATE WITHOUT THE AUTHOR'S PERMISSION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: September 2, 2020

Abstract

In this paper I investigate how parental out-migration affects the schooling outcomes of children left behind in rural China. I establish a theoretical framework to study three important and widely-studied mechanisms that migration could affect left-behind children's school performance: direct effect through parental accompaniment, and indirect effects through child's study time, and education spending. Motivated by the solution of the model, I apply the structural equation modeling to estimate the influence through different mechanisms. To handle the endogeneity caused by both confounder and sample selection, I propose an identification strategy based on instrumental variables, order condition, and Heckman correction. Applying the model on household-level survey data from 9 provinces, I find that the effects through all mechanisms are significantly negative, though the indirect effects are smaller. The subgroup analysis by child's gender and birth order shows that girls and older siblings suffer more from underinvestment in education. 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

*xmluo@ucdavis.edu

1 Introduction

In this paper I investigate how parental out-migration affects the educational performance of children who are left behind in rural China. “Left-behind children” refers to children between 0 and 15 years old who stay in the rural areas where their *hukou* (household registration) are located, with at least one parent moving from rural to urban areas.

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 labor demand in urban areas generated in the economic reform, the Chinese government gradually relaxed the restriction on *hukou* system to permit migration from rural to urban areas. However, the transfer of *hukou* status remains 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, 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” run by local entrepreneurs, though 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 are 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 importance of this problem, it remains under-studied to quantify the influence of migrant parents on left-behind children’s schooling outcomes through different mechanisms simultaneously, which is crucial for policymaking. Previous studies typically investigate the effect through the lens of a single mechanism. For instance, Antman (2013) studies the effect of reduced parental accompaniment, and finds that the absence of parents incurs psychological costs for left-behind children, thereby worsening their schooling performance. Chen (2013) and Chang et al. (2011) examine the effect of children's labor substitution caused by parental migration. Both studies 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 thus have less time for studying. Another widely-studied mechanism is income. Remittances sent home by migrating parents increase household income, alleviate household financial burdens, and improve children's living conditions, educational investment, and nutrition status. Extensive evidence has been found in Mexico (McKenzie and Rapoport, 2011), Indonesia, Thailand (Bryant et al., 2005), Philippines (Arguillas and Williams, 2010; Bryant et al., 2005) , and El Salvador (Edwards and Ureta, 2003). However, these studies only estimate the total effect of parental migration. While they have certain policy implications, they fail to provide insights on which target to intervene, and thus insufficient for policymaking.

In this paper, I take a step in addressing this problem to understand different mechanisms simultaneously through which the parental migration may affect the children's schooling performance. In particular, I consider three mechanisms – parental accompaniment, children's time allocation, and income. To tackle with this problem, I establish a theoretical framework to model the parents and child as two agents attempting to maximize utility under their own constraints. By solving the equilibrium, it clarifies how different mechanisms interact and contribute to the total effect. Motivated by the solution of the model, I apply the structural equation modeling to estimate the influence through different mechanisms. To handle the endogeneity caused by both confounder and sample selection, I propose an identification strategy based on instrumental variables, order condition and Heckman selection model. The identification strategy and the estimation techniques are not limited to this problem and can be easily extended to broader topics related to parent's labor market participation decision and child education (e.g. Agostinelli and Sorrenti, 2018; Blundell and Hoynes, 2004).

The rest of this paper is organized as follows. Section 2 starts with the most general form of the utility maximization for parent and child and solves the equilibrium, 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 analysis for subgroups. Section 6 explores more policy implications and potential future research questions, and Section 7 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 $\tilde{u}_t$ be the utility of the child in period t and s be the share of time that the child spends studying. Equivalently, $(1-s)$ denotes the share of time that the child spends on activities other than studying. Furthermore, I denote by e the human capital level of child in period 1 measured by schooling performance, by e_0 the ability gift, by $d \in [0,1]$ the parent's willingness to migrate and leave child behind, by W_p be parent income from work which depends on d, and by β_k the discount factor of the second-period utility. I assume the first-period utility depends on consumption c_1 and the study time s, while the second-period utility depends solely on the consumption c_2. Given all other variables, the child chooses an optimal study time s that maximizes the total utility, i.e.

$$\begin{aligned}\max_{s} \quad & \tilde{u}_1(s, c_1) + \beta_k \tilde{u}_2(c_2), \\ s.t. \quad & c_1 \leq W_p(d), \\ & c_2 \leq g(e), \\ & e \leq f(d, s, c_1, e_0).\end{aligned} \tag{1}$$

Similarly, for the parent, let u_t be the utility of parent in period t and β_p be the discounting factor. The parent maximizes total utility by choosing the optimal migration willingness d^*, i.e.

$$\begin{aligned} \max_{d} \quad & u_1(c_1) + \beta_p u_2(c_2), \\ s.t. \quad & c_1 \leq W_p(d), \\ & c_2 \leq g(e), \\ & e \leq f(d, s, c_1, e_0). \end{aligned} \tag{2}$$

To derive the equilibrium, I make the following assumptions.

- $\frac{\partial \tilde{u}_t}{\partial c_t} > 0$, $\frac{\partial u_t}{\partial c_t} > 0$ and $\frac{\partial^2 \tilde{u}_t}{\partial c_t^2} < 0$, $\frac{\partial^2 u_t}{\partial c_t^2} < 0$, implying that the utility in each period increases while the marginal utility decreases in consumption in that period.
- $\frac{\partial \tilde{u}_1}{\partial s} < 0$ and $\frac{\partial^2 \tilde{u}_1}{\partial s^2} < 0$, implying a fatigue effect of studying that is marginally decreasing.
- $\frac{\partial f}{\partial s} \geq 0$, $\frac{\partial f}{\partial c_1} \geq 0$, $\frac{\partial^2 f}{\partial s^2} \leq 0$, $\frac{\partial^2 f}{\partial c_1^2} \leq 0$, implying that the study time and consumption weakly increase the production of human capital, but with decreasing marginal returns.
- $\frac{\partial f}{\partial d} < 0$, implying that parental migration leads to less accompaniment with the child, which worsens child education performance.
- $\frac{\partial g}{\partial e} \geq 0$ and $\frac{\partial^2 g}{\partial e^2} \leq 0$, implying that higher human capital of the child leads to higher income in the future, though with a decreasing marginal returns.
- $\frac{\partial W_p}{\partial d} \geq 0$, implying the existence of monetary incentives to migrate.

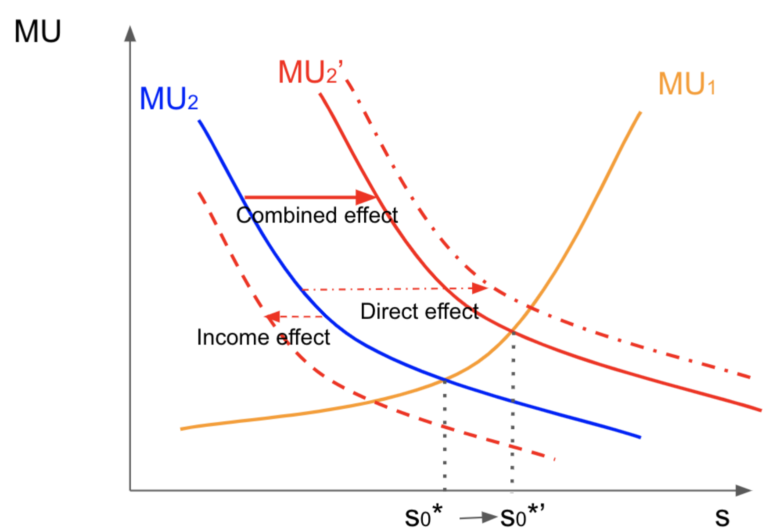
2.2 Child Optimal Decision

For child utility maximization, there is a trade-off between current and future utility. Holding parent migration status d fixed, if study time s increases, the first-period utility decreases due to the fatiguing effect of studying, while the second-period utility increases because the child's human capital will increase due to increased study time, resulting in a higher future consumption c_2.

Intuitively, the original equilibrium of child's optimal study time should be at the intersection of the marginal effect of study time on current utility ($MU_1 = -\frac{\partial u_1}{\partial s}$)

and its marginal effect on future utility ($MU_2 = \frac{\partial u_2}{\partial s}$). The marginal effect of study time on current utility only depends on the level of study time. Holding study time fixed, if the parent increases the willingness to migrate, the marginal effect of study time on first-period utility is not affected, but the increased migration willingness will have a negative direct effect on education performance and a positive indirect effect on education performance due to increased education investment. The final effect on child's education performance depends on the relative sizes of these direct and indirect effects. If the negative direct effect of migration dominates, then child's education performance worsens, leading to less income and consumption in the second period, so that the marginal utility from future consumption increases [1]. Graphically, the curve for marginal effect of study time on current utility remains unchanged, and the curve for marginal effect of study time on future utility shifts up, as shown in Figure 1. In general, the optimal child study time is increasing in migration willingness if the negative direct effect of migration dominates and is decreasing in migration willingness otherwise.

Figure 1: Child decision



Due to the trade-off, under mild assumptions, we can find an interior solution $s^*(d)$. In Appendix C, I derive the closed-form solution of $s^*(d)$ assuming certain

[1]See Appendix A for a mathematical description.

functional forms of the utility function. In general, assuming the existence of the interior optimal solution, I derive that [2]

$$\frac{\partial s^*}{\partial d} \propto -\Bigg(\overbrace{\frac{\partial f}{\partial c_1}\frac{\partial W_p(d)}{\partial d}}^{\text{Income effect}} + \overbrace{\frac{\partial f}{\partial d}}^{\text{Direct effect}} \Bigg), \tag{3}$$

where $\propto$ denotes "proportional to", which hides a positive multiplicative factor. The decomposition of $\frac{\partial s^*}{\partial d}$ shows how left-behind child's optimal study time changes when parent migration status changes. $\frac{\partial f}{\partial c_1}\frac{\partial W_p(d)}{\partial d}$ represents the effect of migration on child education performance through income, or the income effect of migration. $\frac{\partial f}{\partial d}$ depicts the effect of parent migration on child education performance through accompaniment, or the direct effect of migration. Equation (3) shows that the signs of $\frac{\partial f}{\partial d}$ and $\frac{\partial W_p(d)}{\partial d}$ is unknown, the sign of $(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial W_p}{\partial d})$ is undetermined[3]. According to literature, it is reasonable to assume that $\frac{\partial f}{\partial d} \leq 0$, and $\frac{\partial W_p(d)}{\partial d} \geq 0$ so that $\frac{\partial f}{\partial c_1}\frac{\partial W_p(d)}{\partial d} \geq 0$. If the negative direct effect of being left-behind is greater than the positive indirect effect through income, then $\frac{\partial s^*}{\partial d} \geq 0$, suggesting that the child will increase study time to compensate for worse performance, and vice versa.

2.3 Parent Optimal Decision

For parent utility maximization, there is a trade-off between current and future consumption. Holding child study time s fixed, if parent's willingness to migrate d increases, then parent's first-period utility increases due to the increased consumption c_1, whereas the second-period utility decreases because the child's human capital will decrease due to the lack of parent accompaniment, resulting in a lower future consumption c_2.

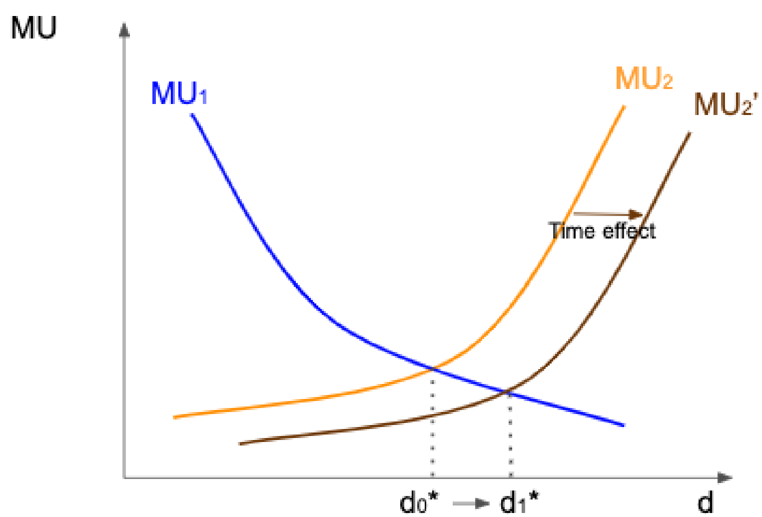
Intuitively, the original equilibrium of parent's willingness to migrate is at the intersection of the marginal effect of migration on current utility ($MU_1 = \frac{\partial u_1}{\partial d}$) and its marginal effect on future utility ($MU_2 = -\frac{\partial u_2}{\partial d}$). Parent utility in period 1 only depends on consumption levels. Holding parent migration status constant, if the child increases study time, it will not affect consumption or utility in

[2]See Appendix A for the derivation

[3]Derivation in Appendix A.

period 1, but will decrease the marginal utility in period 2. This is because the increased study time will lead to higher education performance, and thus higher income and higher consumption in period 2, so that the marginal utility from future consumption decreases [4]. Graphically, the curve for marginal effect of migration on current utility remains unchanged, and the curve for marginal effect of migration on future utility shifts down, as shown in Figure 2. That is, the optimal migration decision is increasing in child study time. Unlike the child optimal decision process, the MU_2 curve always shifts down as the willingness to migrate increases.

Figure 2: Parent decision



Due to the trade-off, under mild assumptions, we can find an interior solution $d^*(s)$. In Appendix C, I derive the closed-form solution of $d^*(s)$ assuming certain functional forms of the utility function. In general, assuming the existence of the interior optimal solution, I derive that [5]

$$\frac{\partial d^*}{\partial s} \propto -\left(\overbrace{\frac{\partial f}{\partial c_1} \frac{\partial W_p(d)}{\partial d}}^{\text{Income effect}} + \overbrace{\frac{\partial f}{\partial d}}^{\text{Direct effect}} \right), \tag{4}$$

[4]Please refer to Appendix B for further details.

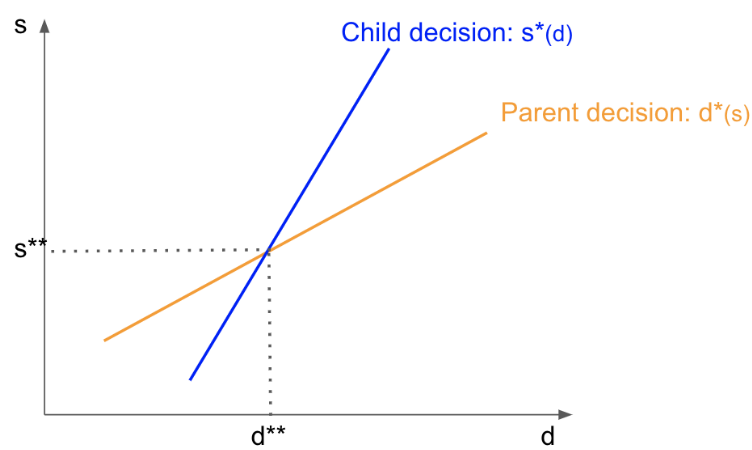
[5]See Appendix B for the derivation.

The decomposition of $\frac{\partial d^*}{\partial s}$ shows how parent's optimal migration decision changes as child study time changes. The meaning of each part of $\frac{\partial d^*}{\partial s}$ is the same as in Equation (3) [6]. The marginal effect of parental migration on current utility is $\frac{\partial u_1}{\partial c_1}\frac{\partial c_1}{\partial d}$, and its marginal effect on future utility is $-\beta_p\frac{\partial u_2}{\partial e}(\frac{\partial f}{\partial d}+\frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d})$. To guarantee an interior solution, we need the future marginal effect to be nonnegative, that is, $\frac{\partial f}{\partial d}+\frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d}\leq 0$. Therefore, $\frac{\partial d^*}{\partial s}>0$. This confirms our intuition that $\frac{\partial d^*}{\partial s}$ has a definite sign as opposed to $\frac{\partial s^*}{\partial d}$.

2.4 Equilibrium of Parent and Child Decision

In Section 2.2 and Section 2.3, I show that child's optimal decision on study time is a function of parental migration status d, and parent's optimal decision is a function of child study time s. Solving both equations will lead to the equilibrium. By giving specific functional forms to the utility function, production function, and wage function, I show that there is only one unique equilibrium solution [7]. Figure 3 is an illustration of the equilibrium solution.

Figure 3: Equilibrium



Given a migrant family, the theoretical model induces the following relationship among the observables – child's schooling performance e, education spending

[6]Derivation in Appendix B.
[7]Derivation in Appendix C.

W_p, child's study hour s, parent's willingness to migrate d, as well as other covariates X that account for the heterogeneity of families:

$$\begin{aligned} e &= f(d, s, W_p; X) \\ W_p &= W_p(d; X) \\ s &= s^*(d; X) \\ d &= d^*(s; X) \end{aligned} \tag{5}$$

The main goal of this work is to study the effects of migration on child's schooling performance, i.e. $\frac{\partial e}{\partial d}$. By definition,

$$\frac{\partial e}{\partial d} = \underbrace{\frac{\partial f}{\partial d}}_{\text{Direct effect}} + \underbrace{\frac{\partial f}{\partial s}\frac{\partial s}{\partial d} + \frac{\partial f}{\partial W_p}\frac{\partial W_p}{\partial d}}_{\text{Indirect effects}}.$$

This yields the decomposition of the total effect into the direct effect and indirect effects. All parameters in the decomposition are necessary for answering my research question, and thus the first three equations in (5) must be estimated. By contrast, the effect of s on d, i.e. $\frac{\partial d}{\partial s}$, does not have to be known. As a result, we can simplify the last equation by solving d^{**} as shown in this subsection, thereby facilitating the system. To summarize, I focus on the following system:

$$\begin{aligned} e &= f(d, s, W_p; X) \\ W_p &= W_p(d; X) \\ s &= s^*(d; X) \\ d &= d^{**}(X) \end{aligned} \tag{6}$$

It is worth emphasizing that the way to simplify (5) hinges on the research question and whether the simplification works depends on whether an effective identification strategy exists. For (6), I found a promising identification strategy as detailed in the next section. In principle, we can also study how child's schooling performance affects parent's migration decision, which can be answered by (5) in theory. However, it is arguably more challenging to find a convincing identification strategy.

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 outperforms 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., and 2666 children in 2112 households are left in the data. The parents in the dataset for analysis come from 81 cities in 9 provinces, and their migration destinations spread over 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 4a shows children with different parental migration status. Adding up the proportions of left-behind children with migrating mother only, migrating father only, and both parent migrating, left-behind children account for roughly 30% of children in rural China. 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 4b shows the work locations for parents. On the horizontal axis, the first three categories are working in rural areas, which are rural areas in the local county, rural areas in other counties in the same province, and rural areas in other provinces. The last two categories are working in urban areas, which are cities of the local province and cities of other provinces. The middle category, local county seat, is in between rural and urban areas, which is less developed than cities but more developed than the rural areas. From Figure 4b, most non-migrants work in the rural areas in the local city or the local county seat, and most migrants move from rural to urban areas in other provinces or in the local provinces. Thus, the majority of migrants move from rural to urban areas, which is the focus of my analysis.

Figure 4c 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 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 their children if bringing the children 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 their children.

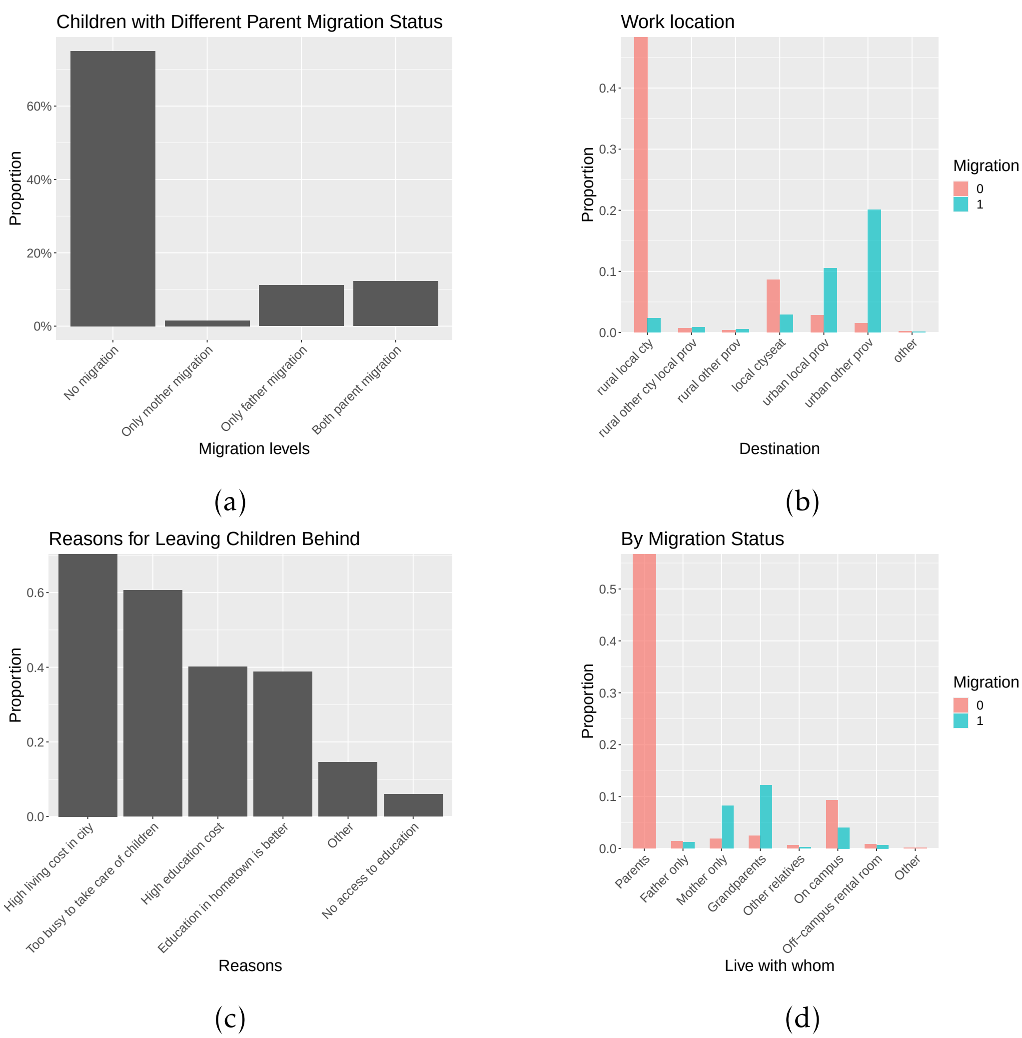


Figure 4: Distribution of normalized exam scores

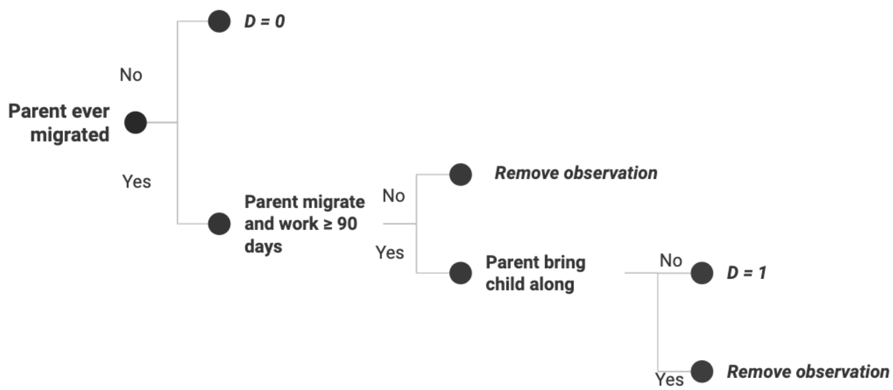
Figure 4d depicts whom the children in rural areas live with. Usually, we assume it's best for the children to live with their parents, so the first three categories on the horizontal-axis are the best-case scenarios, where the child lives with both parents, or with either father or mother. In the next three categories, children are taken care of by other people, such as their grandparents, other relatives, or by teachers at boarding schools. In the last case, children live by themselves in off-campus rental rooms. We could see that when parents migrate away, children are most likely taken care of by grandparents, who are generally not quite well-educated or have much modern parenting knowledge or skills as the children's parents do.

In the next subsections, I will introduce in more details about how the treatment variable, 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. The dummy variable D stands for the migration status of child's parents, which corresponds to d in my model setup. $D = 1$ if at least one parent of the child migrates out and leaves the child behind, and $D = 0$ only if neither father or mother migrates out. Figure 5 shows the detailed definition of the dummy variable for child's left-behind status. Since migration destination is only recorded if migrants work in that place for more than 90 days during the last year, and this information is needed for my identification strategy, 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, rather than children who migrate away together with their migrant parents, I exclude the migrant children from this 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 Mediation Variables

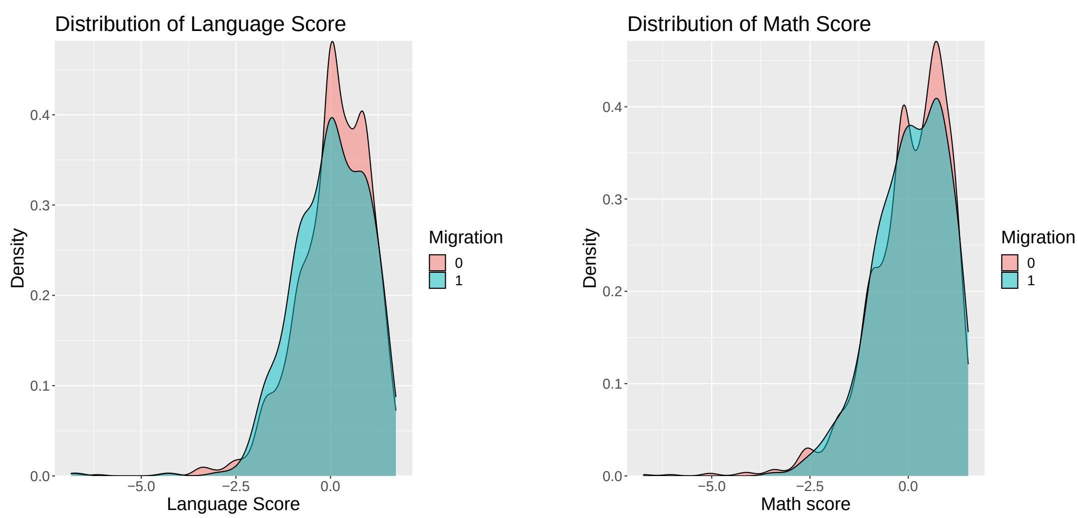
For the measure of child study time T, I use the variable recording child's weekly study hour reported by their guardians, which corresponds to s in my model set up. For the measure of spending on child education W, it corresponds to c_1 in my model set up, and includes spending on child's tuition at school, supplemental classes inside and outside of school, food, and accommodation cost in the year 2008 reported by guardians.

3.5 Dependent Variables

In the model setup, I define child human capital as e. In the RUMiC data, I choose the standardized child exam scores P as a measure of 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.

The exam scores are reported by parents or other 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 converting them to the 100-point scale, and then deduct its means and divide by its standard deviation. After normalization, the outcome variables are more comparable and make more sense in interpreting effect sizes.

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



(a) Language score

(b) Math score

Figure 6: Score Distribution

3.6 Covariate Variables

As for other covariates, I first include the personal characteristics of child, such as the age, gender, height and weight, birth weight, health status, and whether the child goes to boarding school. I also include parent-level characteristics such as the age and years of education of their parents.

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 shows the summary statistics of dependent variables and some important independent variables. From the table, left-behind children perform much worse than children with non-migrant parents in language exams and slightly better in math exams, though neither of the differences is statistically significant at the 1% level. Left-behind children are also significantly younger, lighter, and shorter than their counterparts. The difference in weight and height is probably due to 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 Structural Equation Model

Under specific function forms [8], I show that the system of equations (6) have the following linear forms:

$$P_i = \gamma_0 + \gamma_T \cdot T_i + \gamma_W \cdot W_i + \gamma_D \cdot D_i + \xi \cdot X_i + \phi_i, \tag{7}$$

$$T_i = a_T + b_T \cdot D_i + \xi \cdot X_i + u_i, \tag{8}$$

$$W_i = a_W + b_W \cdot D_i + \xi \cdot X_i + v_i, \tag{9}$$

$$D_i = \mathbb{1}(a_D + \xi \cdot X_i + \zeta_i \geq 0), \tag{10}$$

where P_i denotes the schooling performance of child i, measured by normalized final exam scores in language and mathematics as described in Section 3.5, T_i

[8]See assumptions in Appendix C.

Table 1: Summary Statistics

Variable	Migrant Parents	Non-migrant Parents	Difference (P-value)
Dependent Variables			
Language score	0.01	0.08	0.14
Math score	0.08	0.06	0.73
Covariates: Child			
Male	0.53	0.55	0.52
Age	11.27	11.52	0.07^{*}
Height	135.92	142.44	$< 0.001^{***}$
Weight	38.96	41.46	$< 0.001^{***}$
Birthweight	32.50	32.51	0.94
Covariates: Parents			
Mother age	35.65	38.13	$< 0.001^{***}$
Father age	37.29	39.88	$< 0.001^{***}$
Mother edu year	7.50	7.36	0.21
Father edu year	8.23	8.19	0.70
Note:	*p<0.1; **p<0.05; ***p<0.01		

denotes the study time, and W_i denotes the education spending. I choose slightly different notation from the theoretical model to highlight that these variables are specific measurements of the abstract variables in the model. D_i is a binary variable of the parental migration decision. In the theoretical model, the parent chooses the willingness to migrate while in practice we can only observe a binary decision. This makes the last equation different from the other three. To account for individual heterogeneity, other covariates and error terms are included. X_i 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). The error terms ϕ_i, u_i, v_i, and ζ_i are random errors, and we assume that they are correlated due to unobserved confounder in order to account for the endogeneity. When ζ_i is normal, the last equation is equivalent to a Probit model.

With an identified 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:

$$\delta = \underbrace{\gamma_D(\text{parental accompaniment})}_{\text{Direct effect}} + \underbrace{\gamma_T b_T(\text{time allocation}) + \gamma_W b_W(\text{income})}_{\text{Indirect effects}}, \tag{11}$$

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. δ is the effect estimated with reduced-form models. With the structural model, we are able to decompose it into the three parts of interest.

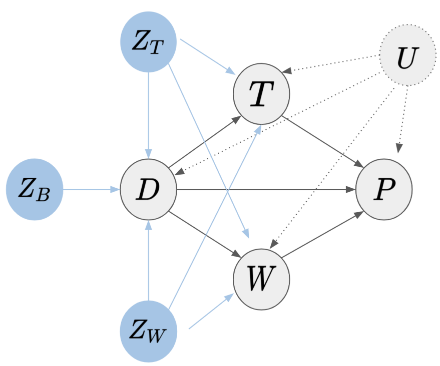
4.2 Identification of Coefficients

Since there are many unobserved factors that correlate with both parental migration decisions and children's school performance, migration decision is an endogenous variable and 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. In addition, child's study time and education spending are also endogenous. To identify the coefficients in the structural equation model, I need at least three exogenous variables.

A necessary condition for identification for structural equation modeling is the order condition, that is, for each equation in the system, the number of excluded exogenous variables should be larger or equal to the number of included endogenous variables minus one. Suppose we are able to find three instrumental variables: Z_B that can can only affect P through D, and Z_T, and Z_W that can affect P through D or T or W, then we can estimate both direct and indirect effects of migration because the order condition will be satisfied. I will illustrate this using Figure 7 and Table 2.

Figure 7a 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, and total education spending W, and it is equivalent to Equation 1 in Figure 7b. By the same reasoning, the four equations in Figure 7b is equivalent to the diagram.

Variables in white circles and the connecting arrows in Figure 7a form a simplified path diagram corresponding to the structural equations in Equation (7) to Equation (10). In Figure 7a, U in the dashed circle represents the omitted variables that could correlate with D, P, T, and W, as described above. P, T, and W are endogenous variables, and instrumental variables Z_B, Z_T, and Z_W in blue circles are exogenous variables that satisfy the requirements above. The order condition of the system in Figure 7b is listed in Table 2.



(a) Path diagram

$$Eq1 : P \sim D + T + W$$
$$Eq2 : T \sim D + Z_T + Z_W$$
$$Eq3 : W \sim D + Z_T + Z_W$$
$$Eq4 : D \sim Z_B + Z_T + Z_W$$

(b) Simplified structural equation model

Figure 7

Table 2: Order Condition of Structural Equation Model

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

In Table 2, taking Equation 2 in Figure 7b as an example, D, T are the included endogenous variables in this equation, so the number of included endogenous variables minus 1 is 1. Z_B is the only excluded exogenous variable in this equation, so the number of excluded exogenous variables is 1. This is how we check the order condition for Equation 2 in Table 2. The same method applies for the other three equations. The order conditions are satisfied for all equations, and thus all the coefficients in the structural equation model in Equation (7) to Equation (10) are identifiable.

4.3 Choice of Instrumental Variables

In Section 4.2, I show that as long as exogenous instrumental variables Z_B, Z_T, and Z_W are found, the coefficients in the structural equation model will be identified. So in this section, I will describe the instrumental variables I use that satisfy the requirements for Z_B, Z_T, and Z_W.

Z_B should be an instrumental variable that only affects child performance through migration status. Some popular candidates for Z_B 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.

This paper follows the method of Bartik (1991) and uses a Bartik-like instrument as Z_B. The Bartik-like instrument 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.

As for Z_T, and Z_W, I use the size of farmable land in the household as Z_W, and adult male share in the household as Z_T:

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

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. This satisfies the requirement for Z_T, and Z_W that they can affect performance through migration status, study time, or education spending.

4.4 Nonrandom Missing Patterns

The above sections address one common source of endogeneity in variables of interest. In this section, I will focus on another source of endogeneity, one that originates from nonrandom missing patterns in the study time and education spending, especially the former. It is unlikely that these two variables are missing at random conditionally on observed covariates because less caring parents may not care about the child's education and thus fail to report the information.

Previous studies simply remove observations with missing values in empirical analysis without accounting for nonrandom missing patterns. However, simply removing the observations with missing values in these variables may yield underestimation or overestimation of the negative effect of migration. Instead, I assume that the parent reports the study time or education spending only when a certain utility is above zero. When the utility is a linear function of the covariates with normal errors, this is precisely a Heckman model. In principle, the Heckman model can be added into the structural equation model directly and estimated using methods maximum likelihood. However, the non-standard form will significantly complicate the structure, making the estimation overly challenging. Therefore I apply a two-step procedure in which I first estimate the Heckman model for the study time and education spending separately to impute the missing values, and then estimate the structural equation model using the imputed data.

5 Empirical Results

5.1 Main Results on All Samples

The structural form is estimated with maximum likelihood based on Equation (7) to Equation (10). Table 3 shows the direct and indirect effects of migration using all samples after imputing for missing values in study time and education spending. The two columns represent the effect on normalized language scores and math scores. First-stage results show whether the instrumental variables are highly correlated to the endogenous variables discussed in the empirical framework. These correlations are significant and the signs are as expected. I expect the coefficient on inter-city Bartik instrument to be positive, because better employment in other cities will make people more likely migrate away. The coefficient on within-city Bartik instrument should be negative, since better employment in the place of residence will make people less likely migrate away to work. The coefficients on the share of adult males in the household is expected to be negative. This is because I take the negative log of the adult male share in the regression, so a higher share of adult males is equivalent to a smaller value of negative log share. Households with more adult males will have less workload per capita, making people more likely to migrate out, so this coefficient should be negative. The coefficient on arable farm size in the household should be positive because I expect that people from households with more resources will be more likely to migrate away. The p-values of coefficients are reported in parentheses. For indirect effects, we report the p-value from the joint significance test. Although this p-value cannot be inverted to a confidence interval as opposed to Sobel test, it is valid for testing the null effect. In addition, it is found to be powerful compared to alternative methods (e.g. Fritz and MacKinnon, 2007; Hayes and Scharkow, 2013). Note that the first-stage results are the same for both subjects because I use exactly the same variables and same samples in the first stage.

Recall that the direct effect of migration is γ_D, the effect of parent accompaniment. The indirect effect of migration through child study time is $\gamma_T b_T$, and the indirect effect through education spending is $\gamma_W b_W$. Usually, we consider 0.3 standard deviations away from the mean to be a large difference in size. In Table 3, for the direct effect on language scores, the performance of left-behind children is almost 0.5 standard deviations lower than children whose parents don't migrate out, and this difference is significant at the 1% level. The direct effect on math scores is smaller in size. The performance of left-behind children is roughly 0.3 standard deviations lower than children with non-migrating parents, and this difference is also significant at the 1% level.

As for the indirect effects shown in Table 3, parental migration has significant negative indirect effect on left-behind children's language scores through reduced study time and reduced education spending, but the effect on math scores is only

Table 3: Effect of Parental Migration on Child Schooling Outcomes (All Sample)

	Language Score	Math Score
Direct Effect		
Parental Accompany	-0.487***	-0.327***
	(0.000)	(0.001)
Indirect Effect		
Study time	-0.015**	-0.008
	(0.021)	(0.160)
Education spending	-0.162***	-0.131***
	(0.004)	(0.004)
First Stage		
Inter Bartik	1.986***	1.986***
	(0.003)	(0.003)
Within Bartik	-1.499***	-1.499***
	(0.000)	(0.000)
Adult male share	-0.166*	-0.166*
	(0.067)	(0.067)
Farm size	0.148***	0.148***
	(0.000)	(0.000)
Obs.	1952	1952

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

significant through education spending. The effect of migration on exam scores through the study time mechanism is much smaller in size compared to that through education spending.

5.2 Exploring Heterogeneous Treatment Effects

In the previous section, I explore the effect of migration on all children's exam scores, but the treatment effects might vary for different subgroups of children. In this section, I do subgroup analysis to investigate the heterogeneous treatment effects. In particular, I am interested in subgroups partitioned by gender, age, and child order. The different attitude of guardians toward boys and girls, the age of the children, as well as the role that the eldest and younger children play in multiple-child families will probably lead to heterogeneous treatment effects when parents migrate away. For each subgroup, I repeat the process of estimating the structural equation models using the data with imputed study time and education spending, and the results are presented in Table 4, Table 5, and Table 6.

Table 4 shows the effect of parental migration for left-behind boys and girls separately. Considering the direct effect, left-behind children are more negatively affected in language scores, which is consistent with our finding in Table 3. Due

Table 4: Effect of Parental Migration on Child Schooling Outcomes (Subgroup by Gender)

Panel A: Boys		
	Language Score	Math Score
Direct Effect		
Parental Accompany	-0.411***	-0.381***
	(0.003)	(0.004)
Indirect Effect		
Study time	-0.086**	-0.104***
	(0.021)	(0.004)
Education spending	-0.041**	-0.035
	(0.048)	(0.100)
First Stage		
Inter Bartik	3.181***	3.181***
	(0.001)	(0.001)
Within Bartik	-1.139***	-1.139***
	(0.003)	(0.003)
Adult male share	-0.137	-0.137
	(0.445)	(0.445)
Farm size	0.214***	0.214***
	(0.000)	(0.000)
Obs.	1077	1077
Panel B: Girls		
Direct Effect		
Parental Accompany	-0.653***	-0.386***
	(0.000)	(0.007)
Indirect Effect		
Study time	-0.008	-0.005
	(0.376)	(0.376)
Education spending	-0.354**	-0.216*
	(0.039)	(0.058)
First Stage		
Inter Bartik	0.310	0.310
	(0.696)	(0.696)
Within Bartik	-1.165***	-1.165***
	(0.009)	(0.009)
Adult male share	-0.245**	-0.245**
	(0.027)	(0.027)
Farm size	0.087	0.087
	(0.109)	(0.109)
Obs.	875	875

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

to the lack of accompaniment by parents, left-behind boys perform 0.4 standard deviations lower than children with non-migrating parents on language exams, whereas left-behind girls perform more than 0.65 standard deviations lower, and both effects are significant at the 1% level. The direct effect on left-behind girl's language score is 62.5% worse than that of the boys.

As for the indirect effects, left-behind boys are mainly affected through the significant reduction in study time rather than the reduced education spending, although both effects are small in size. This might be partially explained by the difference in time management skills and study habits between boys and girls. For left-behind girls, although not significantly affected through reduced study time, their school performances significantly worsens through reduced education spending. Because of the underinvestment in education when their parents migrate away, their language scores are 0.35 standard deviations lower and their math scores are more than 0.2 standard deviations lower than children with non-migrating parents. 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.

Table 5 shows the effect of parent migration on children in different age groups. I group all children with age ≤ 12 into one category, and the rest into the other category. The cut-off age is set at 12 because this is the age to graduate from elementary schools and enter middle schools in China. In general, I expect that when parents migrate away, younger children are more sensitive to the lack of parental accompaniment and are less adapt at managing their lives, so they tend to suffer more both directly and indirectly when left-behind. This is corroborated with results in Table 5. The direct effect of parental migration is slightly larger on younger children's language scores. It is more worth noticing that younger children suffer significantly from the underinvestment in their education spending. The negative effect of migration through reduced education spending is both larger in size and more statistically significant (at least at the 5% level) for the younger children than that for the older children.

Table 5 shows the effect of parent migration on younger children and older children when pooling together all children in each age group. But I am also interested in the effect on older and younger children within the same household.

Table 5: Effect of Parental Migration on Child Schooling Outcomes (Subgroup by Child Age)

Panel A: Age ≤ 12		
	Language Score	Math Score
Direct Effect		
Parental Accompany	-0.475***	-0.269***
	(0.000)	(0.005)
Indirect Effect		
Study time	-0.011	-0.004
	(0.107)	(0.437)
Education spending	-0.183**	-0.142***
	(0.011)	(0.005)
First Stage		
Inter Bartik	2.680***	2.680***
	(0.002)	(0.002)
Within Bartik	-1.616***	-1.616***
	(0.000)	(0.000)
Adult male share	-0.112	-0.112
	(0.334)	(0.334)
Farm size	0.183***	0.183***
	(0.001)	(0.001)
Obs.	1114	1114
Panel B: Age > 12		
Direct Effect		
Parental Accompany	-0.433**	-0.368*
	(0.030)	(0.074)
Indirect Effect		
Study time	-0.015	-0.010
	(0.223)	(0.391)
Education spending	-0.117	-0.109
	(0.157)	(0.192)
First Stage		
Inter Bartik	1.003	1.003
	(0.375)	(0.375)
Within Bartik	-1.354**	-1.354**
	(0.010)	(0.010)
Adult male share	-0.211	-0.211
	(0.157)	(0.157)
Farm size	0.118*	0.118*
	(0.062)	(0.062)
Obs.	838	838

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

For households with multiple children, usually the elder children could take care of the younger children in terms of providing emotional support, life management, and help with studying. Therefore, when parents migrate away, I expect that the role of parent will partially shift to the eldest child and the eldest child will take care their younger siblings, so the subsequent children will actually suffer less than the eldest child. This is confirmed with results in Table 6. In Table 6, I group all children with birth order greater or equal to 2 into the "subsequent child" category, because otherwise each subgroup will be too small to report valid standard errors of coefficient estimates. Due to the direct effect of parental migration, on average the eldest children achieve almost 0.55 standard deviations lower in language and 0.36 standard deviations lower in math than their non-migrant counterparts, and these effects are significant at at least the 5% level. The direct effect of migration on the subsequent children is 0.37 standard deviations lower in language and almost non-existent in math than children with non-migrating parents, which is much smaller than that of the eldest children in the household. As for indirect effects, subsequent children also suffer less from reduced study time and reduced education spending compared with the eldest children, both in effect sizes and in significance levels.

6 Discussion

6.1 Policy implications

Results from Section 5 show the relative effectiveness of policies targeting at left-behind children. For example, in order to increase left-behind girls' school performances, policies targeted at compensating their education expenditure should in general be much more effective than policies targeting at increasing their study time. On the contrary, left-behind boys tend to benefit more from policies that increase their study time. For the eldest children in the household, it is necessary to design policies that guarantee both their study time and education spending.

Since education-related spending can be decomposed into different categories of expenditures, taking one step further, I am able to perform more refined

Table 6: Effect of Parental Migration on Child Schooling Outcomes (Subgroup by Child Order)

Panel A: Eldest child		
	Language Score	Math Score
Direct Effect		
Parental Accompany	-0.547***	-0.363**
	(0.002)	(0.013)
Indirect Effect		
Study time	-0.025	-0.015
	(0.187)	(0.187)
Education spending	-0.131	-0.113*
	(0.124)	(0.082)
First Stage		
Inter Bartik	2.156**	2.156**
	(0.041)	(0.041)
Within Bartik	-1.458***	-1.458***
	(0.001)	(0.001)
Adult male share	0.024	0.024
	(0.867)	(0.867)
Farm size	0.161***	0.161***
	(0.009)	(0.009)
Obs.	838	838
Panel B: Subsequent child		
Direct Effect		
Parental Accompany	-0.373***	-0.148
	(0.004)	(0.200)
Indirect Effect		
Study time	-0.015	0.005
	(0.233)	(0.661)
Education spending	-0.095**	-0.052
	(0.044)	(0.164)
First Stage		
Inter Bartik	0.267	0.267
	(0.803)	(0.803)
Within Bartik	-2.141***	-2.141***
	(0.000)	(0.000)
Adult male share	-0.419***	-0.419***
	(0.008)	(0.008)
Farm size	0.119*	0.119*
	(0.061)	(0.061)
Obs.	851	851

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

analysis based on this decomposition and explore more complicated policies. Table 7 also shows the direct and indirect effects of migration on left-behind children's school outcomes. In this analysis, education spending is divided into two parts: nutrition spending and course-related spending. Course-related spending include expenditure on tuition and remedial classes at school and outside of school, and I simply call it "tuition spending". Table 7 shows that the direct effect and indirect effect of migration through study time is quite consistent with the results from Table 3, and the indirect effect of migration through reduced education expenditure mainly kicks in through reduced nutrition spending. Therefore, policies targeting at increasing left-behind children's education spending should lean more towards improving children's nutrition. For instance, conditional cash transfer programs that improve these children's food intakes and nutrition status should be more effective in increasing their school performances than tuition waivers. This analysis can be carried out on different child subgroups based on different policy interest.

Table 7: Effect of Parental Migration on Child Schooling Outcomes (Nutrition and Tuition)

	Language Score	Math Score
Direct Effect		
Parental Accompany	-0.468***	-0.306***
	(0.000)	(0.001)
Indirect Effect		
Study time	-0.010**	-0.005
	(0.042)	(0.301)
Tuition	-0.044*	-0.041**
	(0.068)	(0.032)
Nutrition	-0.095***	-0.034**
	(0.009)	(0.013)
First Stage		
Inter Bartik	1.770***	1.770***
	(0.007)	(0.007)
Within Bartik	-1.657***	-1.657***
	(0.000)	(0.000)
Adult male share	-0.176*	-0.176*
	(0.054)	(0.054)
Farm size	0.151***	0.151***
	(0.000)	(0.000)
Obs.		

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

6.2 Future research questions

In the empirical analysis of this paper, the child is regard as left behind if either father or mother migrates out and leaves the child behind. However, it is worth studying the multi-level treatment effect – the effect of being left-behind when one parent migrates out and when both parents migrate out might be different. The marginal effect and gender patterns is worth exploring.

The empirical framework and analysis in this paper focuses on how parent's migration decision affects left-behind children's school performances. The other direction of influence – the effect of child performance on parent migration decision, is also interesting and of no less policy importance. Once this direction of question is studied, we will have a more general view of the interaction between child performance and parent migration decision, and be able to design more comprehensive policies targeting at migration or left-behind children's education. For this research question, child's education performance is the main endogenous variable, but it is much more complicated than the endogenous migration decision in my current research question. To address the endogeneity of child performance, more information on children, such as child-level networking information and class-level or school-level information is needed. However, such information is not available in my current dataset, because the RUMiC survey is focused on adult migrants rather than their children. I leave this question to new child-centered data and future research.

7 Conclusion and Remarks

In this paper I establish 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. Analysis based on the structural model reveals significant negative direct effect of parent migration on left-behind children's language scores and math scores, and language scores tend to be more negatively affected. The significant indirect effect through reduced education investment, and reduced nutrition inputs in particular, has significant migration policy implications.

Structural form results by subgroups reveals how parental migration affect left-behind children differently through different pathways. Results from subgroup analysis by gender draw attention to the lack of study time for left-behind boys and severe underinvestment in education for left-behind girls in rural China. Subgroup analysis by age groups shows that younger children are more vulnerable when left behind by parents, both through reduced parent accompaniment and through reduced education spending. Subgroup analysis by birth order reveals that within the same household, younger siblings are less affected because of the buffering role played by the eldest child. Understanding these pathways helps 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 function, education production function, 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 equation modeling. 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 parents' labor market participation on child education.

Appendix

Appendix A. Child Utility Maximization

The utility of child is

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

Plugging constraints to utility function

$$\tilde{L} = \tilde{u}_1(s, W_p) + \beta_k \tilde{u}_2(g(f(d, s, c_1, e_0)))$$

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

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

The marginal effect of studying time on current utility is $MU_1 = -\frac{\partial \tilde{u}_1}{\partial s}$, and its marginal effect on future utility is $MU_2 = \beta_k \frac{\partial \tilde{u}_2}{\partial c_2} \frac{\partial g}{\partial e} \frac{\partial f}{\partial s}$. The goal is to study the effect of d on s^*, so further take the derivative of $\frac{\partial \tilde{L}}{\partial s}$ with respect to d,

$$\begin{aligned}
\frac{\partial^2 \tilde{L}}{\partial s \partial d} =& \frac{\partial^2 \tilde{u}_1}{\partial s^2} \frac{\partial s}{\partial d} + \frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1} \frac{\partial c_1}{\partial d} + \beta_k A \left(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial s} \frac{\partial s}{\partial d} + \frac{\partial f}{\partial c_1} \frac{\partial W_p(d)}{\partial d} \right) + \\
& \beta_k \frac{\partial \tilde{u}_2}{\partial c_2} \frac{\partial g}{\partial e} \frac{\partial^2 f}{\partial s^2} \frac{\partial s}{\partial d} + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2} \frac{\partial g}{\partial e} \left(\frac{\partial^2 f}{\partial s \partial d} + \frac{\partial^2 f}{\partial s \partial c_1} \frac{\partial W_p(d)}{\partial d} \right) = 0,
\end{aligned}$$

where

$$A = \frac{\partial^2 \tilde{u}_2}{\partial c_2^2} \left(\frac{\partial g}{\partial e} \right)^2 \frac{\partial f}{\partial s} + \frac{\partial \tilde{u}_2}{\partial c_2} \frac{\partial f}{\partial s} \frac{\partial^2 g}{\partial e^2} < 0.$$

Therefore,

$$\frac{\partial s^*}{\partial d} = -\frac{\beta_k A \left(\overbrace{\frac{\partial f}{\partial c_1} \frac{\partial W_p(d)}{\partial d}}^{\text{Income effect}} + \overbrace{\frac{\partial f}{\partial d}}^{\text{Direct effect}} \right) + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2} \frac{\partial g}{\partial e} \left(\frac{\partial^2 f}{\partial s \partial d} + \frac{\partial^2 f}{\partial s \partial c_1} \frac{\partial W_p(d)}{\partial d} \right) + \frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1} \frac{\partial W_p(d)}{\partial d}}{\frac{\partial^2 \tilde{u}_1}{\partial s^2} + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2} \frac{\partial g}{\partial e} \frac{\partial^2 f}{\partial s^2} + \beta_k A \frac{\partial f}{\partial s}}.$$

If we further assume the separability of the child utility function and human capital production function, we will get rid of terms of $\frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1}$,$\frac{\partial^2 f}{\partial s \partial d}$, and $\frac{\partial^2 f}{\partial s \partial c_1}$, then $\frac{\partial s}{\partial d}$ is simplified to

$$\frac{\partial s^*}{\partial d} = -\frac{\beta_k A\left(\overbrace{\frac{\partial f}{\partial c_1}\frac{\partial W_p(d)}{\partial d}}^{\text{Income effect}} + \overbrace{\frac{\partial f}{\partial d}}^{\text{Direct effect}}\right)}{\frac{\partial^2 \tilde{u}_1}{\partial s^2} + \beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}\frac{\partial^2 f}{\partial s^2} + \beta_k A \frac{\partial f}{\partial s}}.$$

The denominator of $\frac{\partial s^*}{\partial d}$ is negative, so the sign of $\frac{\partial s^*}{\partial d}$ depends on its numerator, and specifically depends on the relative size of $\frac{\partial f}{\partial d}$ and $\frac{\partial f}{\partial c_1}\frac{\partial W_p(d)}{\partial d}$. Assuming that $\frac{\partial f}{\partial d} \leq 0$ and that $\frac{\partial f}{\partial c_1}\frac{\partial W_p(d)}{\partial d} \geq 0$, if the negative direct effect of being left-behind is larger than the positive indirect effect through income, then $\frac{\partial s^*}{\partial d} \geq 0$, suggesting that the child will increase study time to compensate for worse performance due to the absence of parent, and vice versa.

Graphically, the original equilibrium of child study time should be at the intersection of the marginal utility of studying in period 1 and period 2, which is s^* in Figure 1. For the child, holding study time fixed, if there is a change in parent migration status d, then the child's education performance will change due to the direct effect of migration and its indirect effect through income/consumption, thereby affecting the income and consumption in the future. However, it will not affect consumption in the first period. This is equivalent to stating that holding s fixed, when d changes, c_1 will remain unchanged but e and c_2 will be affected. Recall that the marginal effect of study on current utility only depends on study time s, so even if d changes, the current marginal effect won't change because s is held fixed. The marginal utility of migration in period 2 depends on the discounted marginal effect of study time on future utility through consumption in that period ($\beta_k \frac{\partial \tilde{u}_2}{\partial c_2}\frac{\partial g}{\partial e}\frac{\partial f}{\partial s}$). When d changes, the discount rate β_k won't change, and education production through the indirect effect of study time won' change either because s is held constant. However, education production will be directly affected by migration status and indirectly affected by migration through income, and thus the returns to education will change. In addition, the change in education performance leads to change in the consumption level, which will affect the marginal utility of consumption in the second period. Considering the changes in returns to education and marginal utility of consumption, the marginal utility of study time in period 2 will be affected.

If d increases, on the one hand, the direct effect of migration will lead to worse educational performance. Since the return to education decreases as education performance increases, we would expect an increase in the return to education. As educational performance becomes worse, future income and future consumption will drop, which leads to an increase in the marginal utility of consumption since it's decreasing in consumption levels. Therefore, considering the direct effect of migration, we expect the return to education and the marginal utility of consumption to increase as d increases, and thus the marginal utility of study time in period 2 increases as d increases. Graphically, the curve for marginal utility of study time in period 2 shifts up since the new marginal effect of studying on future utility becomes higher for every level of s. If d increases, on the other hand, the indirect effect of migration through income will lead to better educational performance. Since the return to education decreases as education performance increases, we would expect to see a drop in the return to education. As educational performance becomes better, future income and future consumption will increase, leading to a decrease in the marginal utility of consumption since it's decreasing in consumption levels. Due to the indirect effect of migration through income, as d increases, we expect the return to education and the marginal utility of consumption to decrease, and thus the marginal effect of study time on future utility decreases. Graphically, the future marginal effect curve shifts down since the new marginal effect of study on future utility becomes lower for every level of s.

In summary, when migration status increases, although the current marginal effect curve of study time remains unchanged, the shift of the future marginal effect curve will depend on the relative sizes of the two opposite forces from the direct and indirect effect of migration. If the direct effect outweighs the indirect effect through income, then the curve will finally shift up and the new equilibrium study time will increase to $s^{*'}$, suggesting that if d increases, s^* is expected to increase, as shown in Figure 1. This suggests that the child will have to study longer to compensate for the large detrimental effect of migration on their school performances. This is consistent to our findings in Appendix A.

Appendix B. Parental Utility Maximization

The utility of parent is

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

Plugging constraints to the utility function

$$L = u_1(W_p) + \beta_p \ u_2(g(e))$$

With a slight abuse of notation, we write $u_2(e)$ for $u_2(g(e))$. Then we know that

$$\begin{aligned} \frac{\partial u_2}{\partial e} &= \frac{\partial u_2}{\partial c_2}\frac{\partial g}{\partial e} \geq 0, \\ \frac{\partial^2 u_2}{\partial e^2} &= \frac{\partial^2 \ u_2}{\partial c_2^2}(\frac{\partial g}{\partial e})^2 + \frac{\partial^2 g}{\partial e^2}\frac{\partial u_2}{\partial c_2} \leq 0. \end{aligned}$$

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

$$\frac{\partial L}{\partial d} = \frac{\partial u_1}{\partial c_1}\frac{\partial c_1}{\partial d} + \beta_p \frac{\partial u_2}{\partial e}(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d}) = 0.$$

For the parent, when they migrate out, they are potentially benefiting from higher consumption due to higher income in the first period, but at the cost of their child's educational performance and thus their future income and consumption. From the first-order condition, we know the marginal effect of parental migration on current utility is $MU_1 = \frac{\partial u_1}{\partial c_1}\frac{\partial c_1}{\partial d}$, and its marginal effect on future utility is $MU_2 = -\beta_p \frac{\partial u_2}{\partial e}(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d})$. To guarantee an interior solution, we need the marginal effect on future utility to be nonnegative, that is, $\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d} \leq 0$.

From the first-order condition, we can derive parent's optimal migration decision d^* as a function of child's study time s. Our goal is to study the effect of s on

d^*, so further take the derivative of $\frac{\partial L}{\partial d}$ with respect to s,

$$\begin{aligned}\frac{\partial^2 L}{\partial s \partial d} =& \frac{\partial u_1}{\partial c_1}\frac{\partial^2 c_1}{\partial d^2}\frac{\partial d}{\partial s} + \frac{\partial^2 u_1}{\partial c_1^2}(\frac{\partial c_1}{\partial d})^2\frac{\partial d}{\partial s} + \\ & \beta_p(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d})\frac{\partial^2 u_2}{\partial e^2}(\frac{\partial f}{\partial d}\frac{\partial d}{\partial s} + \frac{\partial f}{\partial s} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d}\frac{\partial d}{\partial s}) + \\ & \beta_p\frac{\partial u_2}{\partial e}[\frac{\partial c_1}{\partial d}(\frac{\partial^2 f}{\partial c_1 \partial d}\frac{\partial d}{\partial s} + \frac{\partial^2 f}{\partial c_1 \partial s} + \frac{\partial^2 f}{\partial c_1^2}\frac{\partial c_1}{\partial d}\frac{\partial d}{\partial s}) + \frac{\partial f}{\partial c_1}\frac{\partial^2 c_1}{\partial d^2}\frac{\partial d}{\partial s}] = 0.\end{aligned}$$

Since we assume the separability of human capital production function, i.e., $\frac{\partial^2 f}{\partial s \partial d} = \frac{\partial^2 f}{\partial s \partial c_1} = \frac{\partial^2 f}{\partial c_1 \partial d} = 0$, the second-order condition can be simplified, and thus

$$\frac{\partial d^*}{\partial s} = \frac{-\beta_p \frac{\partial^2 u_2}{\partial e^2}\frac{\partial f}{\partial s}(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d})}{\frac{\partial u_1}{\partial c_1}\frac{\partial^2 c_1}{\partial d^2} + \frac{\partial^2 u_1}{\partial c_1^2}(\frac{\partial c_1}{\partial d})^2 + \beta_p\frac{\partial^2 u_2}{\partial e^2}(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d})^2 + \beta_p\frac{\partial u_2}{\partial e}[\frac{\partial^2 f}{\partial d^2} + \frac{\partial^2 f}{\partial c_1^2}(\frac{\partial c_1}{\partial d})^2 + \frac{\partial f}{\partial c_1}\frac{\partial^2 c_1}{\partial d^2}]}.$$

Since $\frac{\partial u_1}{\partial c_1} \geq 0, \frac{\partial^2 u_1}{\partial c_1^2} \leq 0; \frac{\partial u_2}{\partial e} \geq 0, \frac{\partial^2 u_2}{\partial e^2} \leq 0; \frac{\partial c1}{\partial d} \geq 0, \frac{\partial^2 c_1}{\partial d^2} \leq 0; \frac{\partial f}{\partial c_1} \geq 0, \frac{\partial^2 f}{\partial c_1^2} \leq 0; \frac{\partial f}{\partial d} \leq 0, \frac{\partial^2 f}{\partial d^2} \leq 0$, and $\beta_p > 0$, the denominator of $\frac{\partial d}{\partial s}$ is negative. The numerator is also negative since $\frac{\partial f}{\partial s} \geq 0$ and $\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d} \leq 0$. Thus, $\frac{\partial d^*}{\partial s} \geq 0$ as long as there is an interior solution. This suggests that if the child is willing to study for longer times, parent will be more "assured" and more likely to migrate out. In addition, $\frac{\partial s^*}{\partial d} \geq 0$ due to $\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d} \leq 0$.

Graphically, the original equilibrium of parent migration decision should be at the intersection of the marginal utility of migration in the first period and the marginal utility in the second period, which is d_0^* in Figure 2. Holding parent migration status constant, if there is a change in child's study time, then the child's education performance will be affected, thereby affecting the income and consumption in the second period. However, it will not affect the consumption in the first period. This is equivalent to stating that when holding d fixed and changing s, c_1 will remain unchanged but e and c_2 will be affected. Recall that the marginal utility from migration in the first period is the marginal effect of migration status on current utility through consumption in that period ($\frac{\partial u_1}{\partial c_1}\frac{\partial c_1}{\partial d}$), so even if s changes, the marginal utility in the first period won't change because d and c_1 remain the same. The marginal utility in the second period is the discounted marginal effect of migration status on future utility through consumption in that period ($-\beta_p \frac{\partial u_2}{\partial e}(\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d})$). future consumption depends only on child education performance. When child study time s changes, education production will be affected [9], the marginal utility in the second period from consumption/educational performance will be affected, so the marginal utility of migration in the second period will be affected. If s increases, e will increase as the marginal effect of study time on educational performance is positive. Therefore, consumption in the second period increases as educational performance increases. Since the marginal utility is decreasing in consumption, we expect to see a decrease in marginal utility from future consumption, so the marginal utility of migration in Period 2 will decrease. Therefore, when the child increases study time, although parent's marginal utility in Period 1 will not change, the curve for marginal utility in Period 2 will shift down since it becomes lower for every level of d. This results in the new equilibrium of migration status increasing from d_0^* to d_1^*, as shown in Figure 2. That is, d^* is increasing in s. This is consistent to our findings in Appendix B.

Appendix C. Example with Specific Functional Forms

There might be some concern in the above decision making process since I am assuming simultaneous decisions. In this section, I will use specific functional forms to show that the joint decision process of parent and child will lead to one unique equilibrium. In that case, it makes no difference if we are assuming a simultaneous decision process or a sequential one. In addition, the specific functional forms I choose is also consistent with my empirical model.

For the child decision process, the utility maximization could be depicted by:

$$\begin{aligned}
\max_{s} \quad & \log[(1-s)T_0] + \log(c_1) + \beta_k \log(c_2), \\
s.t. \quad & c_1 \leq a + w_1 \cdot D, \\
& c_2 \leq w_2 \cdot e, \\
& e \leq \gamma_T \cdot s \cdot T_0 + \gamma_w(a + w_1 D) + \gamma_D \cdot D.
\end{aligned}$$

Plugging the constraints into the objective function, we have

$$\tilde{L} = \log[(1-s)T_0] + \log(a + w_1 \cdot D) + \beta_k \log[w_2(\gamma_T \cdot s \cdot T_0 + \gamma_w(a + w_1 D) + \gamma_D D)].$$

[9] Education production through the direct effect of migration or the indirect effect through current consumption ($\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d}$) won't change because c_1 and d remain the same.

Taking its first-order derivative with respect to s, we have

$$\frac{\partial \tilde{L}}{\partial s} = -\frac{1}{1-s} + \frac{\gamma_T \cdot T_0 \cdot \beta_k}{\gamma_T \cdot s \cdot T_0 + \gamma_w(a + w_1 D) + \gamma_D D}$$

Setting the first-order condition to 0, we could solve for s^*, the optimal time decision of children:

$$s^* = \frac{\gamma_T \cdot T_0 \cdot \beta_k + a \cdot \gamma_w + (\gamma_D + w_1 \cdot \gamma_w) \cdot D}{\gamma_T T_0(\beta_k - 1)}$$

Since $\gamma_D + \gamma_w w_1 = \frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d} \leq 0$, and $\gamma_T T_0(\beta_k - 1) < 0$ due to the fact that discount factor $0 \leq \beta_k < 1$, we know that s^* is non-decreasing as D increases, which is consistent to our findings in Appendix A.

For the parent decision process, the utility maximization process is depicted by:

$$\begin{aligned}
\max_{D} \quad & \log(c_1) + \beta_p \log(c_2), \\
s.t. \quad & c_1 \leq a + w_1 \cdot D, \\
& c_2 \leq w_2 \cdot e, \\
& e \leq \gamma_T \cdot s \cdot T_0 + \gamma_w(a + w_1 D) + \gamma_D \cdot D.
\end{aligned}$$

Plugging in the constraints to the objective function,

$$L = \log(a + w_1 \cdot D) + \beta_p \log[w_2(\gamma_T \cdot s \cdot T_0 + \gamma_w(a + w_1 D) + \gamma_D \cdot D)].$$

Taking the first-order derivative with respect to D, we have

$$\frac{\partial L}{\partial D} = \frac{w_1}{a + w_1 D} - \frac{\gamma_w w_1 + \gamma_D \beta_p}{\gamma_T \cdot s \cdot T_0 + \gamma_w(a + w_1 D) + \gamma_D \cdot D}.$$

Setting the first-order condition to 0, we have

$$D^* = \frac{-a\gamma_D \beta_p + w_1 \gamma_T T_0 s}{w_1 \gamma_D(\beta_p - 1)}.$$

Since $w_1 \gamma_T T_0 \geq 0$, and $w_1 \gamma_D(\beta_p - 1) > 0$ because $\gamma_D < 0$ and $0 \leq \beta_p < 1$, we know D^* is non-decreasing as s increases, which is consistent to our finding in Appendix B.

Next I will show there is a unique equilibrium. If we draw the reaction function of the parent and the child on one graph with d on the horizontal axis and s on the vertical axis, this is equivalent to show that the reactions curves have different slopes. The slope of child's reaction function is $\frac{\gamma_D + w_1\gamma_w}{\gamma_T T_0(\beta_k - 1)}$, and the slope of parent's reaction function is $\frac{w_1\gamma_D(\beta_p - 1)}{w_1\gamma_T T_0}$:

$$
\begin{aligned}
&\frac{\gamma_D + w_1\gamma_w}{\gamma_T T_0(\beta_k - 1)} - \frac{w_1\gamma_D(\beta_p - 1)}{w_1\gamma_T T_0} \\
=&\frac{w_1(\gamma_D + w_1\gamma_w) - w_1\gamma_D(\beta_p - 1)(\beta_k - 1)}{w_1\gamma_T T_0(\beta_k - 1)} \\
=&\frac{w_1^2\gamma_w - w_1\gamma_D(\beta_p\beta_k - \beta_p - \beta_k)}{w_1\gamma_T T_0(\beta_k - 1)} \\
=&\frac{w_1[w_1\gamma_w - \gamma_D(\beta_p\beta_k - \beta_p - \beta_k)]}{w_1\gamma_T T_0(\beta_k - 1)}
\end{aligned}
$$

We already know the denominator of the difference is negative since $\beta_k - 1 < 0$, so the value of the difference only depends on the numerator. As long as we have $w_1\gamma_w \neq \gamma_D(\beta_p\beta_k - \beta_p - \beta_k)$, the slopes will be different. Since $-1 < \beta_p\beta_k - \beta_p - \beta_k \leq 0$, if the negative direct effect γ_D is very large, then the numerator of the difference would be negative so the difference would be positive, suggesting that the slope of child's reaction function would be steeper. This also makes intuitive sense because if γ_D is very large, then based on the MC-MB graph, to compensate for the negative direct effect, the child tends to increase study time by a lot, and the reaction is stronger than parent's. The graph for this case is depicted in Figure 3. The equilibrium study time s^* and equilibrium migration decision d^* is unique. This suggests that it doesn't make any difference whether we are assuming sequential or simultaneous decision process. The equilibrium outcomes are what we observe in our data.

References

Agostinelli, F. and Sorrenti, G. (2018). Money vs. time: family income, maternal labor supply, and child development. *University of Zurich, Department of Economics, Working Paper*, (273).

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?

Blundell, R. and Hoynes, H. W. (2004). Has' in-work'benefit reform helped the labor market? In *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000*, pages 411–460. University of Chicago Press.

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.

Fritz, M. S. and MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological science*, 18(3):233–239.

Hayes, A. F. and Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological science*, 24(10):1918–1927.

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.

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.