Selected Presentation at the 2020 Agricultural & Applied Economics Association Annual Meeting, Kansas City, Missouri, July 26-28

How Does Parental Out-migration Affect Left-behind Children's Schooling Outcomes? – Effect Sizes and Mechanisms

Xiaoman Luo *

University of California, Davis

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 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 child's gender and birth order 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

*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 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 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 investigated the effect through the lens of a single mechanism. Antman (2013) studied the effect of reduced parental accompaniment, and found that the absence of parents incurs psychological costs for left-behind children, thereby worsening their schooling performance. Chen (2013) and Chang et al. (2011) examined effect of children's labor substitution caused by parental migration. 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 that distract them from school studies. 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 rely on reduced-form empirical strategies which can only estimate the total effect of parental migration. While it has certain policy implications, it fails to provide insights on which target to intervene, and thus insufficient for policy making.

In this paper, I take a step in addressing this problem via a structural-form analysis to understand different mechanisms through which the parental migration affects 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 constraint. By solving the equilibrium, it clarifies different 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 confounders 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 (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, subgroups analysis, and sensitivity 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 $\tilde{u}_t$ be the utility of child in period t. Let s be the share of time that the child spends studying, so $(1-s)$ denotes the share of time that the child spends on activities other than studying. I assume the utility of child depends on consumption c_t, where $t \in \{1,2\}$, 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 $\frac{\partial^2 \tilde{u}_t}{\partial c_t^2} < 0$. This is because I assume that child's utility in each period increases with consumption in that period , and the marginal utility decreases with consumption. Furthermore, I assume that child get fatigued from studying so child utility in period 1 decreases in s, and marginal utility is decreasing with s, i.e., $\frac{\partial \tilde{u}_1}{\partial s} < 0$ and $\frac{\partial^2 \tilde{u}_1}{\partial s^2} < 0$.

Let e be the human capital level of child in period 1, and e_0 denote the ability gift. Let d be the proportion of days that the parent migrates out and leaves child behind, so $d \in [0,1]$. Let W_p be parent income from work, which depends on parent migration status. β_k is the discount factor of the child. The utility of the child is

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

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., $\frac{\partial^2 f}{\partial s^2} \leq 0$, $\frac{\partial^2 f}{\partial c_1^2} \leq 0$. In addition, I assume that higher human capital of the child in period 1 will lead to weakly higher income in period 2 but the returns to education is decreasing, that is, $\frac{\partial g}{\partial e} \geq 0$ and $\frac{\partial^2 g}{\partial e^2} \leq 0$. As for income in period 1, I assume that $\frac{\partial W_p}{\partial d} \geq 0$. The child maximizes utility by choosing the optimal study time s^*.

For the parent, let u_1 be the utility of parent in period 1 and u_2^0 be the utility in period 2, β_p be the discounting factor, and other notations are the same as for child. 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 parent's utility comes from consumption c_t in each period, where $\frac{\partial u_t}{\partial c_t} > 0$, $\frac{\partial^2 u_1}{\partial c_1^2} < 0$, and $\frac{\partial u_2^0}{\partial c_2} < 0$ [1]. Parent maximizes utility by choosing the optimal migration level d^*. The utility of parent is

$$
\begin{aligned}
\max_{d} \quad & u_1(c_1) + \beta_p u_2^0(c_2), \\
s.t. \quad & c_1 \leq W_p(d), \\
& c_2 \leq g(e), \\
& e \leq f(d, s, c_1, e_0).
\end{aligned} \tag{2}
$$

2.2 Child Optimal Decision

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

[1] The use of notation of u_2^0 is explained in Appendix B.

equilibrium,

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

$\frac{\partial s^*}{\partial d}$ shows how left-behind child's study time changes when parent migrates away. The meaning of each part of $\frac{\partial s^*}{\partial d}$ is marked in Equation (3)[2]. Since 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. 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 also a trade-off between current consumption and future consumption, which guarantees an interior solution. At an interior equilibrium,

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

$\frac{\partial d^*}{\partial s}$ shows how parent migration decision changes when child study time changes. The meaning of each part of $\frac{\partial d^*}{\partial s}$ is marked in Equation (4) [3]. The marginal benefit of parental migration is $\frac{\partial u_1}{\partial c_1}\frac{\partial c_1}{\partial d}$, and its marginal cost 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 marginal cost 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, $\frac{\partial d^*}{\partial s} > 0$.

2.4 Equilibrium of Parent and Child Decision

In Section 2.2 and Section 2.3, I assume the decision process to be a simultaneous process for the parent and child. Thus, child's optimal decision on study

[2]Derivation in Appendix A.

[3]Derivation in Appendix B.

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 simultaneously 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 [4]. The unique equilibrium suggests that it makes no difference whether assuming simultaneous or sequential decision process. Under these specified function forms, we can get a system of four equations: $\frac{\partial e}{\partial d}$ from the education production function, $\frac{\partial s}{\partial d}$ from child utility maximization, $\frac{\partial c_1}{\partial d}$ from how wage is determined, and a function of d from joint utility maximization. The structural model in Section 4.1 is based on these results.

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,

[4]Derivation in Appendix C.

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. 2891 children in 2242 households are left in the data. The parents in the dataset for analysis come from 81 cities in 9 provinces, and their migration destinations cover 176 cities in 31 Provinces.

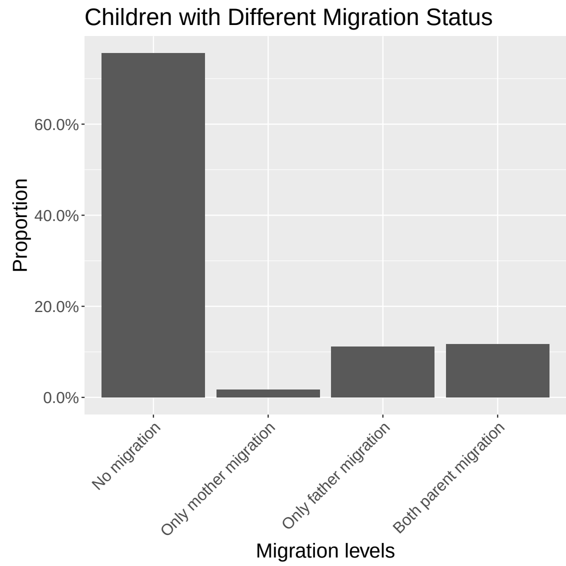
3.2 Descriptive Statistics

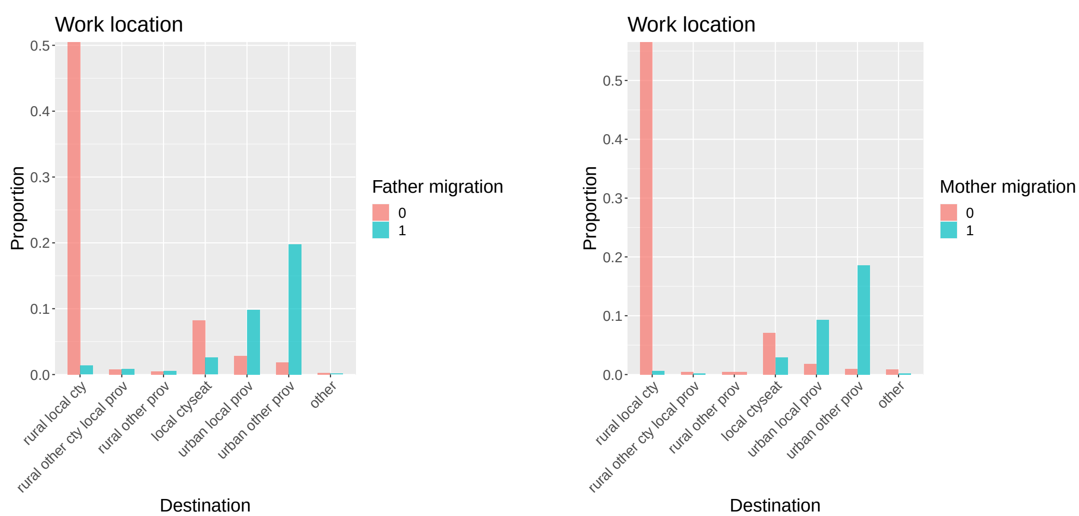
In this section, I will use data visualization to briefly show what the data looks like. Figure 1 shows children with different parental migration status. Left-behind children account for roughly 30% of children in rural areas. As introduced in the following section for defining the treatment variable, 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 2a, 2b show the migration destinations for parents. Since father and mother could have different destinations, 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 provinces. 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 3 depicts the reasons why parents do not bring children when migrating

Figure 1





(a) Father migration

(b) Mother migration

Figure 2: Destination of Work

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 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 children.

Figure 3

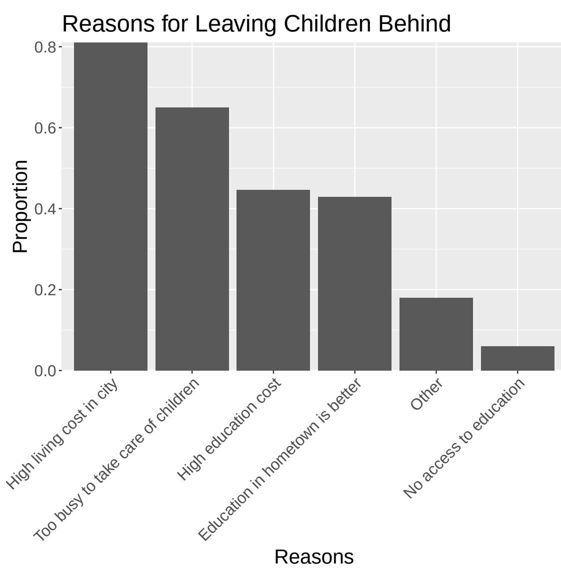
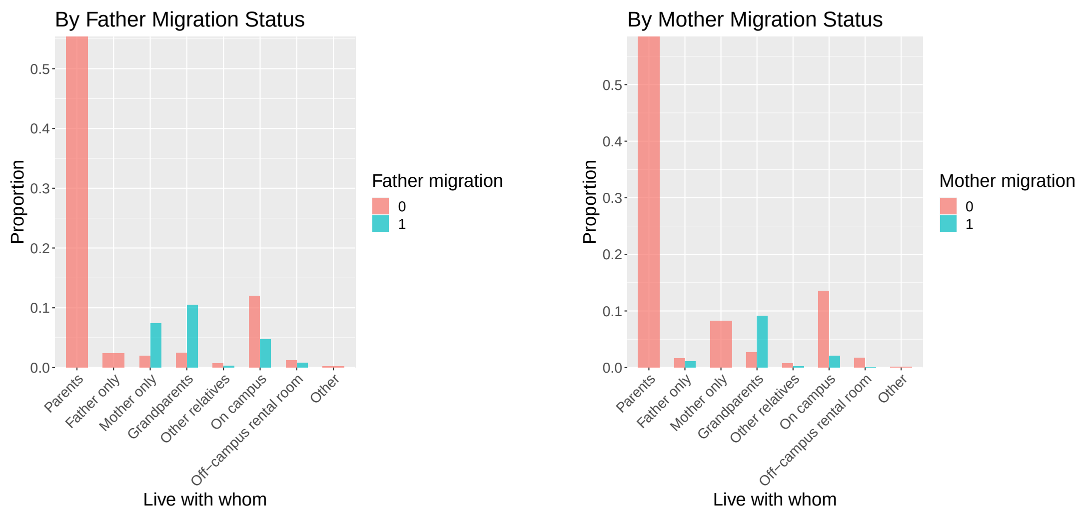


Figure 4a, 4b depict whom the children in rural areas lives with. Usually, we assume it's best for children to live with parents, so the first three categories on the x-axis are the best case scenario, where children live with both parents, or with either father or mother. In the next three categories, 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 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 children's parents do.

In the next subsections, I will introduce in more details about how the treat-



(a) Father migration (b) Mother migration

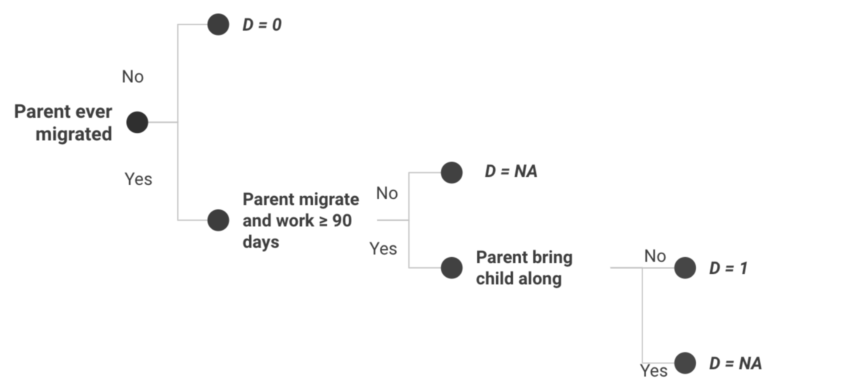
Figure 4: Guardians

ment 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 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. However, the cultural background of different regions might also influence exam 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.

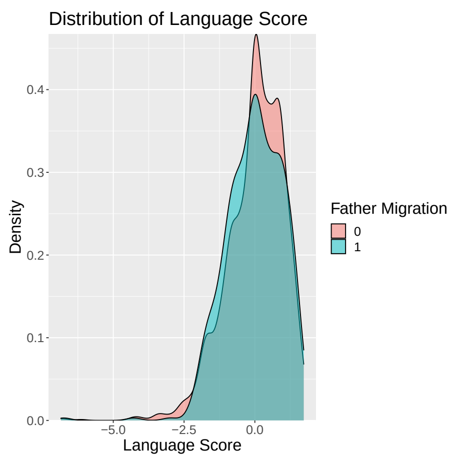
Figure 6 shows the distribution of exam scores. We could see that for left-behind children, the distribution is more right skewed, suggesting that these children perform worse in exams in general. And the difference is more obvious in language scores.

3.6 Covariate Variables

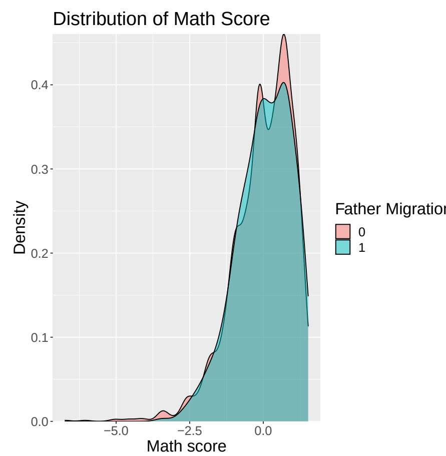
As for other covariates, I first include the personal characteristics of child, such as age, gender, height and weight, birth weight, health status, and whether the child goes to boarding school. I also include the parent characteristics such as the age and years of education. In addition, I include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

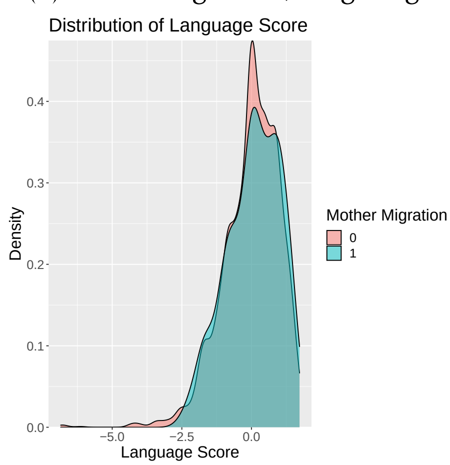
Table 1 shows the summary statistics of some important dependent and independent variables. From the table, left-behind children perform significantly worse than children with non-migrant parents in language exam, but not significantly different in math exam. Left-behind children are also significantly younger,



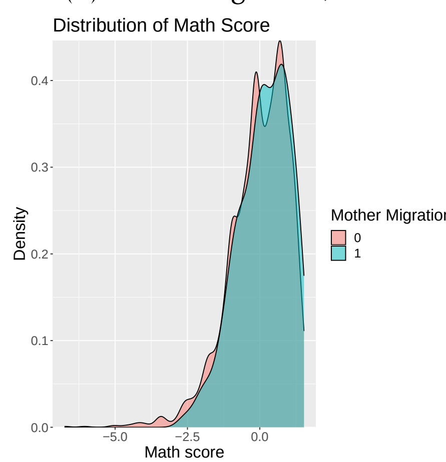
(a) Father migration, languange



(b) Father migration, math



(c) Mother migration, language



(d) Mother migration, math

Figure 6: Distribution of normalized exam scores

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.

Table 1: Summary Statistics

Variable	Migrant Parents	Non-migrant Parents	Difference (P-value)
Dependent Variables			
Language score	-0.03	0.01	0.36
Math score	0.03	-0.01	0.37
Covariates: Child			
Male	0.53	0.55	0.49
Age	11.77	12.10	0.02**
Height	138.60	144.97	< 0.001***
Weight	40.31	42.83	< 0.001***
Birthweight	3255.01	3241.87	0.54
Covariates: Parents			
Mother age	36.15	38.60	< 0.001***
Father age	37.83	40.35	< 0.001***
Mother edu year	7.48	7.31	0.12
Father edu year	8.24	8.18	0.56
Note:	*p<0.1; **p<0.05; ***p<0.01		

4 Empirical Framework

4.1 Structural Form Regression Model

Recall from Section 2.4 that under specific function forms [5], we can get a system of four equations: $\frac{\partial e}{\partial d}$ from the education production function, $\frac{\partial s}{\partial d}$ from child utility maximization, $\frac{\partial c_1}{\partial d}$ from how wage is determined, and a function of d from joint utility maximization. These results are consistent to the structural

[5] See assumptions in Appendix C.

form model I use to study the direct and indirect effect of parental migration on the educational outcomes of left-behind children:

$$P_{ij} = \gamma_0 + \gamma_T \cdot T_{S_{ij}} + \gamma_W \cdot W_{T_{ij}} + \gamma_D \cdot D_{ij} + \xi \cdot X_{ij} + \omega_j + \phi_{ij}, \tag{5}$$

$$T_{S_{ij}} = a_T + b_T \cdot D_{ij} + \xi \cdot X_{ij} + \omega_j + u_{ij}, \tag{6}$$

$$W_{T_{ij}} = a_W + b_W \cdot D_{ij} + \xi \cdot X_{ij} + \omega_j + v_{ij}, \tag{7}$$

$$D_{ij} = \mathbb{1}(a_D + \xi \cdot X_{ij} + \omega_j + \zeta_{ij} \geq 0), \tag{8}$$

where P_{ij} is the schooling performance of child i in province j, measured by normalized final exam scores in language and mathematics as described in Section 3.5. 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). ω_j is province fixed-effect. The error terms ϕ_{ij}, u_{ij}, v_{ij}, and ζ_{ij} are random errors.

In this model, the direct effect of parental migration on performance is denoted by γ_D, the coefficient on D_{ij} in Equation (5). The indirect effect of migration on performance through child's study time is denoted by $b_T \cdot \gamma_T$, which is the multiplication of the coefficient on D_{ij} in Equation (6) and the coefficient on $T_{S_{ij}}$ in Equation (5). The indirect effect of migration on performance through education spending is denoted by $b_W \cdot \gamma_W$, which is the multiplication of the coefficient on D_{ij} in Equation (7) and the coefficient on $W_{T_{ij}}$ in Equation (5). These paths depict the mechanisms of interest.

4.2 Identification of Coefficients

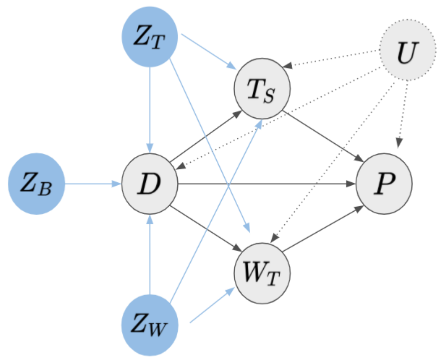
Since there are many unobserved factors that correlate with both parental migration decisions and children's school performance, migration decision is an endogenous variable and ordinary least squares (OLS) estimates tend to have omitted variable bias. For instance, parents who highly value children's 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 spendings are also endogenous. To identify the coefficients in the structural equation model, I need at least three exogenous variables.

The necessary and almost sufficient condition for identification in 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_S, and total education spending W_T, 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 (5) to Equation (8). Although the covariates X are left out of the equations and the path diagram in Figure 7 for simplicity, it will not change the result of order conditions. In Figure 7a, U in the dashed circle represents the omitted variables that could correlate with D, P,T_S, and W_T, as described above. P, T_S, and W_T 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.

In Table 2, take Equation 2 in Figure 7b as an example. D, T_S 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



(a) Path diagram

$Eq1 : P \sim D + T_S + W_T$

$Eq2 : T_S \sim D + Z_T + Z_W$

$Eq3 : W_T \sim D + Z_T + Z_W$

$Eq4 : D \sim Z_B + Z_T + Z_W$

(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

three equations. The order conditions are satisfied for all equations, and thus all the coefficients in the structural equation model in Equation (5) to Equation (8) are identifiable. 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:

$$\delta = \gamma_D(\text{parental accompaniment}) + \gamma_T b_T(\text{time allocation}) + \gamma_W b_W(\text{income}), \quad (9)$$

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.

4.3 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 variables of interest. 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 underestimation of the negative 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 presented in Section 5.

5 Empirical Results

5.1 Main Results on All Samples

The structural form is estimated with maximum likelihood based on Equation (5) to Equation (8). Table 3 shows the direct and indirect effects of migration using all samples after imputing for missing values in study time and education spendings. The first two columns represent the effect on normalized language scores by father migration and mother migration separately, and the last two columns are for math scores. First-stage results show whether the instrumental variables are highly correlated to the endogenous variable–migration decision. 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 settlement will make people less likely migrate away to work. The coefficients on adult male share and farm size in the household should both be positive, because I expect that people from households with more resources and less workload concern will be more likely migrating away. The p-values of coefficients are reported in parentheses.

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 sizes. For the direct effect on language scores, the performance of children whose father out-migrates is roughly 0.4 standard deviations lower than children whose father do not migrate, and this difference is significant at the 1% level. Performance of children whose mother out-migrates is more than 0.6 standard deviations lower than children whose mother do not migrate, and this difference is also significant at 1% level. In Table 3, for direct effect on math score, children whose father out-migrates achieve 0.2 standard deviations lower than children whose fathers do not migrate, and this difference is significant at the 5% level. Children whose mother out-migrates achieve almost 0.5 standard deviations lower than children whose mothers do not migrate, and this effect is 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 score through these mechanisms are relatively small in size. Mother migration generally affects left-behind children through children's reduced study time, but father migration mainly affects through reduced educational spendings.

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

	Language Score		Math Score	
	(1) Father	(2) Mother	(3) Father	(4) Mother
Direct Effect				
Parental Accompany	-0.442***	-0.642***	-0.220**	-0.468***
	(0.000)	(0.000)	(0.017)	(0.001)
Indirect Effect				
Study time	-0.019**	-0.255***	-0.006	-0.232***
	(0.021)	(0.009)	(0.475)	(0.009)
Education spending	-0.184***	-0.024	-0.124***	-0.012
	(0.003)	(0.144)	(0.005)	(0.149)
First Stage				
Inter Bartik	1.437**	1.687***	1.542**	2.030***
	(0.021)	(0.007)	(0.033)	(0.004)
Within Bartik	-1.753***	-0.914***	-1.856***	-0.749**
	(0.000)	(0.002)	(0.000)	(0.013)
Adult male share	0.420**	0.381	0.491**	0.367
	(0.029)	(0.117)	(0.025)	(0.176)
Farm size	0.017***	0.020***	0.009	0.021***
	(0.009)	(0.003)	(0.232)	(0.003)
Obs.	1991	2190	1991	2190

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

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 this table. Comparing with Table 3, the number of observations immediately shrink by almost 900, and the direct and indirect effects in structural form analysis generally shrink in size and become less significant. Again, this confirms 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.

Table 4: Effect of Parental Migration on Child Schooling Outcomes (All Samples, Non-imputed)

	Language Score		Math Score	
	(1) Father	(2) Mother	(3) Father	(4) Mother
Direct Effect				
Parental Accompany	-0.294***	-0.295***	-0.121	-0.140*
	(0.003)	(0.001)	(0.174)	(0.097)
Indirect Effect				
Study time	-0.009	-0.010	0.005	0.000
	(0.426)	(0.205)	(0.618)	(0.976)
Education spending	-0.164**	-0.232***	-0.124**	-0.161***
	(0.011)	(0.003)	(0.015)	(0.003)
First Stage				
Inter Bartik	1.785*	2.450***	1.926*	2.124**
	(0.061)	(0.007)	(0.059)	(0.015)
Within Bartik	-1.918***	-1.776***	-1.995***	-1.928***
	(0.000)	(0.000)	(0.000)	(0.000)
Adult male share	0.659**	0.645**	0.645**	0.680**
	(0.025)	(0.029)	(0.038)	(0.034)
Farm size	0.019**	0.027**	0.013	0.023**
	(0.046)	(0.016)	(0.234)	(0.046)
Obs.	1119	1238	1119	1238

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

5.3 Exploring Heterogeneous Treatment Effects

I also investigate the heterogeneous effects for different subgroups. In particular, I am interested in subgroups partitioned by gender and by child order, because I think the different attitude of guardians toward boys and girls as well as the role that the eldest and younger children play in the household will lead to heterogeneous treatment effects when parents migrate away. For each subgroup, I repeat the process of estimating the structural equation models using data with imputed study time and education spending, and the results are presented in Table 5 and Table 6.

Table 5 shows the effect of parent migration for boys and girls separately. Parental migration has a much larger negative direct effect on left-behind girls when their fathers migrate away. Children are more negatively affected in language scores, which is consistent with our finding in Table 3. As for indirect effects, parental migration has a larger 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 larger negative indirect effect on their exam scores through reduction in educational spendings, and this is significant at at least the 5% level no matter it is father or mother who migrates. 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 6 shows the effect of parent migration for the eldest child and subsequent children separately. I group all children with birth order ≥ 2 into the "subseuent child" category, because otherwise each subgroup will be too small to report valid standard errors of coefficient estimates. I expect that when parents migrate away, 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 then the eldest child. This is confirmed with results in Table 6. The direct effect of parental migration is much larger for the eldest child, more than 0.3 standard deviation lower than their non-migrant counterparts, and it's significant in both math and language scores. The direct effect of migration on the subsequent children is marginally large in language scores, and almost non-existent in math scores. As for indirect effect, subsequent children also suffer less from reduced study time and reduced education spending.

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

	Language Score		Math Score	
Panel A: Boys				
	(1) Father	(2) Mother	(3) Father	(4) Mother
Direct Effect				
Parental Accompany	-0.373***	-0.508***	-0.187	-0.443***
	(0.006)	(0.003)	(0.110)	(0.005)
Indirect Effect				
Study time	-0.040**	-0.187**	-0.013	-0.219**
	(0.037)	(0.021)	(0.415)	(0.026)
Education spending	-0.151**	-0.035	-0.127***	-0.022
	(0.013)	(0.108)	(0.007)	(0.188)
First Stage				
Inter Bartik	2.183**	2.848***	2.075**	2.889***
	(0.015)	(0.002)	(0.034)	(0.002)
Within Bartik	-2.091***	-1.140***	-2.095***	-0.781*
	(0.000)	(0.009)	(0.000)	(0.052)
Adult male share	0.066	-0.003	0.062	-0.116
	(0.806)	(0.994)	(0.828)	(0.754)
Farm size	0.016*	0.022**	0.018**	0.027***
	(0.055)	(0.014)	(0.029)	(0.004)
Obs.	1098	1211	1098	1211
Panel B: Girls				
Direct Effect				
Parental Accompany	-0.588***	-0.503***	-0.341***	-0.273***
	(0.000)	(0.000)	(0.004)	(0.008)
Indirect Effect				
Study time	-0.004	-0.019*	-0.003	-0.011
	(0.432)	(0.076)	(0.376)	(0.266)
Education spending	-0.489**	-0.379**	-0.279**	-0.266***
	(0.029)	(0.015)	(0.027)	(0.004)
First Stage				
Inter Bartik	0.112	0.543	0.544	1.134
	(0.862)	(0.480)	(0.520)	(0.230)
Within Bartik	-1.093**	-1.259***	-1.336***	-1.485***
	(0.013)	(0.004)	(0.005)	(0.002)
Adult male share	0.694**	1.088***	0.904**	1.306***
	(0.026)	(0.005)	(0.010)	(0.002)
Farm size	0.019*	0.034***	-0.006	0.008
	(0.093)	(0.007)	(0.646)	(0.589)
Obs.	893	979	893	979

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

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

	Language Score		Math Score	
	(1) Father	(2) Mother	(3) Father	(4) Mother
Panel A: Eldest child				
Direct Effect				
Parental Accompany	-0.509***	-0.474***	-0.339**	-0.339**
	(0.005)	(0.001)	(0.037)	(0.013)
Indirect Effect				
Study time	-0.046*	-0.045**	-0.033	-0.033*
	(0.075)	(0.021)	(0.177)	(0.095)
Education spending	-0.126*	-0.215**	-0.115**	-0.189***
	(0.077)	(0.015)	(0.025)	(0.005)
First Stage				
Inter Bartik	2.091**	2.259**	2.265**	2.500**
	(0.027)	(0.011)	(0.029)	(0.010)
Within Bartik	-1.564***	-1.648***	-1.627***	-1.751***
	(0.000)	(0.000)	(0.000)	(0.000)
Adult male share	-0.079	0.063	-0.059	0.003
	(0.797)	(0.832)	(0.862)	(0.994)
Farm size	0.014*	0.026**	0.009	0.019*
	(0.098)	(0.010)	(0.375)	(0.084)
Obs.	1012	1102	1012	1102
Panel B: Subsequent child				
Direct Effect				
Parental Accompany	-0.347**	-0.270***	0.032	-0.050
	(0.023)	(0.009)	(0.801)	(0.590)
Indirect Effect				
Study time	-0.004	-0.024***	0.032*	-0.002***
	(0.316)	(0.002)	(0.056)	(0.001)
Education spending	-0.130**	-0.111***	-0.006**	-0.083***
	(0.021)	(0.001)	(0.025)	(0.001)
First Stage				
Inter Bartik	0.372	0.479	-1.512	0.381
	(0.709)	(0.649)	(0.273)	(0.745)
Within Bartik	-1.810***	-2.172***	-2.298***	-2.373***
	(0.001)	(0.000)	(0.000)	(0.000)
Adult male share	0.572*	1.403***	0.508	1.509***
	(0.069)	(0.001)	(0.218)	(0.001)
Farm size	0.020	0.021	-0.012	-0.002
	(0.113)	(0.139)	(0.338)	(0.907)
Obs.	919	1097	919	1097

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

5.4 Sensitivity Analysis

Sensitivity analysis is conducted in this section. In particular, I check whether the results are mainly driven by some specific destinations. Jiangsu Province is the most popular destination of migration in the RUMiC rural survey dataset. Table 7 shows the results for all samples after removing Jiangsu Province. We could see that the sizes of direct and indirect effects remain almost the same compared with results in Table 3, and the significance levels also remain to be the same. First-stage results are also quite stable in sizes and significance levels. This confirms that the results are not driven by the most popular destination.

Table 7: Effect of Parental Migration on Child Schooling Outcomes (Leave One Dest Out, Imputed)

	Language Score		Math Score	
	(1) Father	(2) Mother	(3) Father	(4) Mother
Direct Effect				
Parental Accompany	-0.410***	-0.621***	-0.180**	-0.442***
	(0.000)	(0.000)	(0.030)	(0.001)
Indirect Effect				
Study time	-0.023**	-0.254***	-0.008	-0.235***
	(0.022)	(0.008)	(0.380)	(0.008)
Education spending	-0.132***	-0.024	-0.087***	-0.012
	(0.004)	(0.141)	(0.004)	(0.155)
First Stage				
Inter Bartik	1.237*	1.711***	1.099	1.982***
	(0.055)	(0.007)	(0.138)	(0.006)
Within Bartik	-2.210***	-1.028***	-2.383***	-0.850**
	(0.000)	(0.002)	(0.000)	(0.010)
Adult male share	0.489**	0.433*	0.585**	0.459
	(0.024)	(0.092)	(0.018)	(0.113)
Farm size	0.017***	0.021***	0.007	0.022***
	(0.008)	(0.003)	(0.407)	(0.004)
Obs.	1894	2074	1894	2074

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

6 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 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 reveals how parental migration affect left-behind children differently through 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. Subgroup analysis by birth order reveals that 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 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 parents' labor market participation on child education.

Appendix

Appendix A. Child Utility Maximization

The utility of child is

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

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 benefit of studying time is $\beta_k \frac{\partial \tilde{u}_2}{\partial c_2} \frac{\partial g}{\partial e} \frac{\partial f}{\partial s}$, and its marginal cost is $-\frac{\partial \tilde{u}_1}{\partial s}$. 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,

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

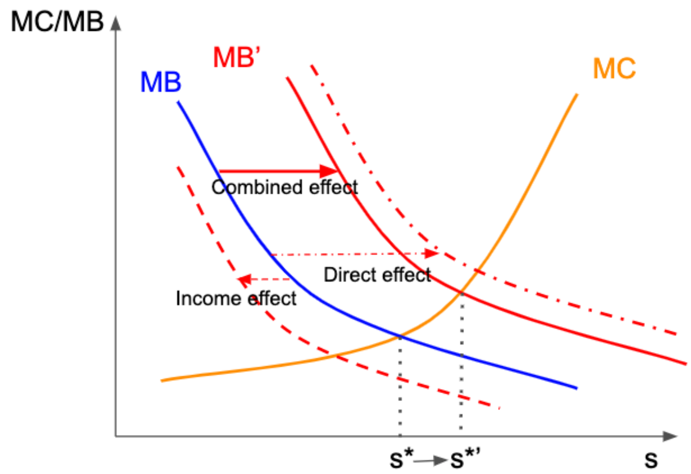
If we further assume the separability of the child utility function and human capital production function, we will get rid of terms of $\frac{\partial^2 \tilde{u}_1}{\partial s \partial c_1}$,$\frac{\partial^2 f}{\partial s \partial d}$, and $\frac{\partial^2 f}{\partial s \partial c_1}$, then $\frac{\partial s}{\partial d}$ 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 benefit and marginal cost of studying, which is s^* in Figure 8. For the child, the change in parent migration status d will only affect the marginal benefit of study. If d increases, then due to the direct effect of migration, f will decrease, thus c_2 decreases, leading to the increase in the marginal benefit of study time. Therefore, the marginal benefit curve will shift to the right due to the direct effect of migration. On the other hand, when d increases, c_1 will increase. Due to the indirect effect of migration through income, f will increase, thus c_2 increases, leading to the decrease in the marginal benefit of study time. Thus, the marginal benefit curve will shift to the left due to the indirect effect of migration through income. The final direction of shift will depend on the relative sizes of these two effects. If the negative direct effect outweighs the positive indirect effect through income, the marginal benefit curve will finally shift to the right and the new equilibrium study time will increase to $s^{*'}$, suggesting that if d increases, s^* is expected to increase. Figure 8 corresponds to this case. This is consistent to our finding above.

Figure 8: Child decision



Appendix B. Parental Utility Maximization

The utility of parent is

$$\begin{aligned} \max_{d} \quad & u_1(c_1) + \beta_p u_2^0(c_2), \\ s.t. \quad & c_1 \le W_p(d), \\ & c_2 \le g(e), \\ & e \le f(d, s, c_1, e_0). \end{aligned}$$

Plugging constraints to the utility function

$$L = u_1(W_p) + \beta_p \; u_2^0(g(e))$$

To simplify the derivation, now define a function $u_2(\cdot)$ such that

$$u_2(e) = \; u_2^0(g(e)),$$

then we know that

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

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

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

From the first-order condition, we know the marginal benefit of parental migration is $\frac{\partial u_1}{\partial c_1}\frac{\partial c_1}{\partial d}$, and its marginal cost 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 marginal cost to be nonnegative, that is, $\frac{\partial f}{\partial d} + \frac{\partial f}{\partial c_1}\frac{\partial c_1}{\partial d} \leq 0$.

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

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

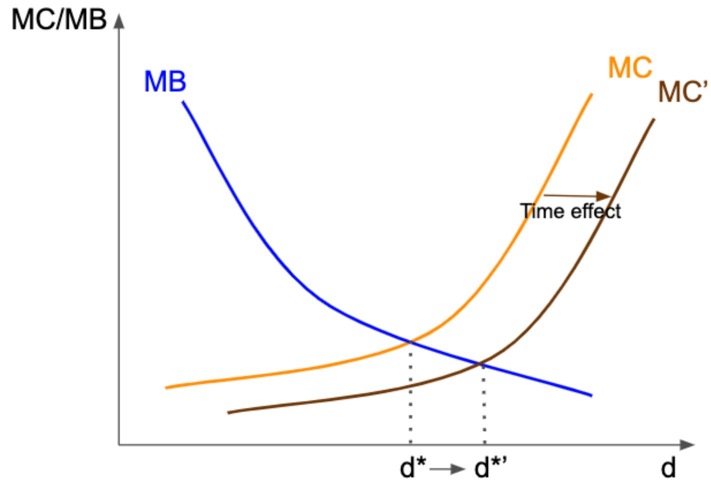
Since we assume the separability of human capital production function, i.e., $\frac{\partial^2 f}{\partial s \partial d} = \frac{\partial^2 f}{\partial s \partial c_1} = \frac{\partial^2 f}{\partial c_1 \partial d} = 0$, the second-order condition can be simplified, and thus

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

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 benefit and marginal cost of migration, which is d^* in Figure 9. For the parent, the change in child study time s will only affect the marginal cost of migration. If s increases, then due to the indirect effect of migration through reduced study time, f will increase, thus c_2 increase, leading to the decrease in the marginal cost of migration. Therefore, the marginal cost curve will shift to the right due to the indirect effect of migration through study time and the new equilibrium migration status should increase to $d^{*'}$. That means if s increases, d^* is expected to increase. This is consistent to our finding above.

Figure 9: Parent decision



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

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

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

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

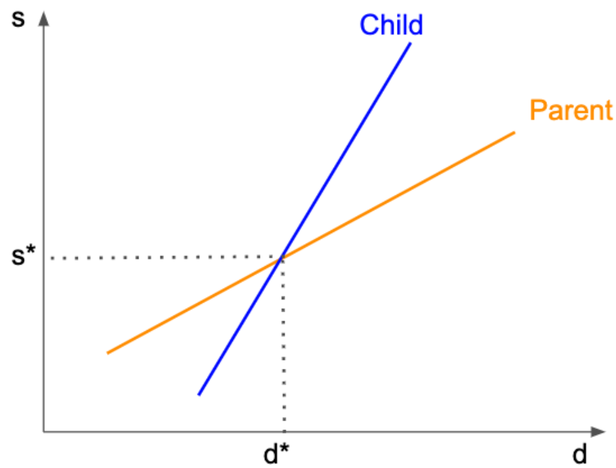
Since $w_1\gamma_T T_0 \geq 0$, and $w_1\gamma_D(\beta_p - 1) > 0$ because $\gamma_D < 0$ and $0 \leq \beta_p < 1$, we know D^* is non-decreasing as s increases, which is consistent to our finding in Appendix B.

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 10. 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.

Figure 10: Equilibrium



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.

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.