Migration Outcome Gap: The Cost of Leaving Children Behind

Conghan Zheng*

May, 2023

Click here for the latest version

Abstract

Parents invest money and parental care in their children. Given the constraints on migration, migrant parents may trade one for the other. In China, in addition to financial constraints, there are also institutional mobility restrictions on rural-urban migration, which limit the ability of migrants to claim urban citizenship and thus pose the problem of multiple selectivity. Using a simultaneous equations model to deal with endogeneity, the results highlight the motive of households to migrate for better educational opportunities for children and predict a gain from rural-urban migration in children's educational outcomes. And a significant part of the income gains from migration is invested in the child's human capital accumulation. The results also suggest that *Hukou* restrictions in China don't prevent migrant families from going to urban areas, but only limit their access to local social services and lead to a loss of welfare.

Keywords: Rural-urban Mobility, Internal Migration, Immigrant Workers, Resettlement, Human Capital, Child Development

JEL codes: I15, I25, J61, R23

*Universitat Autonoma de Barcelona (UAB), Barcelona School of Economics (BSE), zhengconghan@outlook.com.

1 Introduction

1.1 The Restrictions on Rural-Urban Mobility

Over the past 40 years, China has experienced remarkable economic growth driven by the rapid development of its manufacturing sector. This growth has been made possible by the large influx of cheap rural labor that has migrated from villages to cities. While this process of urbanization has brought many benefits to people of rural origin, such as improved living standards and employment opportunities, it has also had negative consequences. One such consequence is the impact on migrant families, especially children.

For the children left behind by their migrant parents in the sending regions, while there is the possibility that remittances can alleviate families' liquidity constraints and thus improve children's educational outcomes, the absence of parental care and guidance can have long-lasting effects on children's emotional and cognitive performance that may outweigh the positive effects of remittances. According to recent estimates (Tang et al., 2021), there are approximately 61 million left-behind children in rural China, accounting for 37.7 percent of all rural children and 21.88 percent of all children nationwide.

Without a local household register, children brought to receiving areas by their migrant parents typically have limited access to urban public social services, including education and health care. Existing literature also documents a positive association between having migrant family members and children's mental health problems such as depression, anxiety, loneliness, and low self-esteem (Myerson, 2017; Ivlevs et al., 2019).

China's rural-urban household division has its roots in the introduction of the household registration system, commonly known as *Hukou*, more than half a century ago. Implemented in 1958, the system was designed to curb rural-urban migration and ensure adequate food production. The *Hukou* distinguishes between agricultural and non-agricultural households, commonly known as rural and urban households.

The *Hukou* system links access to certain local social services to the place of household registration, usually the place of birth[1]. Residents receive their *Hukou* booklets by birth. A new member born or married into a household is added to the *Hukou* booklet and has the same rural-urban classification as other members.

Large cities, especially megacities [2], set requirements for migrant applicants to meet

[1] Although the rural-urban classification has not been printed on newly issued *Hukou* booklets since 2014 due to a reform, all respondents in my data still know their *Hukou* type.

[2] According to the 2010 Chinese census, there are 7 cities in mainland China that have more than 10 million people living in each of their urban areas.

before they can obtain a local urban *Hukou*. Typically, requirements are set for social insurance participation, education level, investment and real estate purchase, and employment conditions, etc.

In the Chinese context, the majority of rural-urban migrants are low-skilled workers who often do not meet the eligibility criteria for urban Hukou status in large cities. According to my data constructed from *China Family Panel Studies* (CFPS, explained in section 2.1), waves 2010-2020, only 56% of the migrant parents (the parents of ever rural Hukou holder children) have completed middle school (the compulsory education in China). As a result, they are unable to access basic local social services such as health care and education for their families. This has created significant challenges for the resettlement of migrant families.

1.2 Household Migration on Child Education

It's well established in the literature that parental labor migration can have both positive and negative effects on children's educational outcomes.

On the one hand, remittances from migrant parents can improve the economic circumstances of their families and provide resources for their children's education and health. Studies have shown that these positive effects can occur through mechanisms such as alleviating household liquidity constraints (Edwards et al., 2003; Du et al., 2005) and promoting increased investment in children (McKenzie et al., 2011; Ambler et al., 2015).

On the other hand, parental migration, which inherently involves parental absence from home, may negatively affect children's educational attainment due to a lack of parental care (Lahaie et al., 2009) or an increase in the amount of time left-behind children spend working on farms or in households (Chang et al., 2011; Antman, 2011), although a corresponding decrease in child labor is expected as household budget constraints are eased.

As the institutional constraints on rural-urban upward mobility in China limit the ability of migrants to claim urban citizenship, theories of school aggregation and segmented assimilation, largely discussed in the context of immigration and minorities in international migration, predict adverse consequences for various aspects of child development. Substantial structural barriers keep a large proportion of migrant children segregated in low-quality informal schools at destination, Lu et al. (2013) documents a large difference in academic achievement and mental well-being between migrant children attending isolated migrant schools in urban areas of China and all children attending public schools that include both urban natives and migrant children.

All of these negative effects may be particularly pronounced for younger children, who require more direct care and supervision.

In terms of gender effects, because men are the ones who migrate or take the lead role in household migration in most contexts, the literature also examines the contributions of the father and links the resulting loss when the father migrates to the literature on father absence more broadly (Antman, 2012). Another possibility explored in the literature, related to within-household gender bargaining, is that the effects on children may differ by the age and gender of the child. McKenzie et al. (2011) find a negative effect of Mexico-US migration on the schooling of children left behind and attribute this behavior to increased housework for girls and migration for boys. Intra-household bargaining over the couple's migration is also associated with effects that vary by the age and gender of the child.

Because children in migrant-sending households experience both positive and negative factors, the overall impact of parental migration and the likely resulting family separation on children's development could be negative or positive, depending on the balance of these factors. The direction of the overall effects is therefore mainly an empirical question.

1.3 The Selectivity Issue in Household Migration

In the migration literature, it's generally accepted that migration is likely to be correlated with the same factors that affect the outcomes of children in the sending household. For example, in terms of children's education, if we agree that migration is costly, at least for rural people, then families that are better able to finance migration are likely to be better able to finance their children's education.

Almost all recent research on migration controls for selection into migration. Research has used propensity score matching (Kuhn et al., 2011), natural experiment (Gibson et al., 2011), randomized controlled trial (RCT, Bryan et al., 2014), and fixed effects estimators (Antman, 2012; Chang et al., 2019) to net out much of the observed and unobserved variation that is common within households. Longitudinal data, in which researchers can observe outcomes before and after the migration event, offer a potential solution to the problem of reverse causality, in which the observed child outcome actually causes the migration event, rather than the other way around.

However, these approaches are still vulnerable to the possibility that some unobserved time-varying factor is driving both migration and child outcomes. Given these problems, researchers have turned to instrumental variable (IV) estimators to correct for selection bias and reverse causality. The main goal of these studies is to find a valid instrument that affects the outcome of interest only through its effect on migration. Among the IVs used in the recent literature on the impact of household migration on left-behind children, there are generally two types: historical migration patterns (McKenzie et al., 2011; Böhme et al., 2015) and variables related to economic conditions in destination areas (Amuedo-Dorantes et al., 2010,Gao et al., 2022).

Gibson et al. (2013) examines the quadruple selectivity problem that complicates the comparison of households with and without migrants. First, households select into migration, which is considered in almost all recent literature. Second, households select into whether the whole household migrates. And therefore there could also be intra-household selection of migrants, which is explored by Murard, 2019, that after the household endogenously decides whether or not to send a migrant, there is a subsequent selection of which family members to send. But this intra-household bargaining is still between the potential migrants; the child, who is passive in the decision, is not involved.[3] Third, migrants choose to return, and their households could be considered less affected by migration. And fourth, migrants decide when to return, and the impact of migration varies with the duration of the migration period, so researchers also face selectivity in when households migrate.

The last three types of selection are also interrelated in the sense that returning migrants may decide to become repeat migrants (or seasonal migrants, which is the majority of internal migrant workers in developing countries) in the first place, and that's why they leave the child behind when they leave. Or it could be that the majority of these low-skilled migrant workers (according to the previous figure that only 56% of them have completed middle school) have decided to move before they get a job offer, and once they find a job in the destination city they begin to consider whether they can take their children with them, while the financial and institutional costs drive their decision to return and leave the child behind at the same time.

Even research on left-behind families that claims to address selection into migration typically does not address the latter three forms of selection, especially selection into whether the whole household migrates. For example, among all the literature on left-behind families in migration, Ivlevs et al. (2019) ignores all types of selection, Chang et al. (2019) deals only with selection into migration, Murard (2019) considers intra-household

[3]Shrestha (2017) discusses the spillover effect of international migrants on the educational attainment of non-migrants in the same country.

selection of migrants but does not include selection on the outcomes of children, who are passive in the household's migration decision process but may be the main driver of migration.

The primary objective of this study is to examine the impact of parental migration on children's school enrollment and to elucidate the underlying patterns and mechanisms through which these effects occur. To achieve this goal, I first employ a dynamic panel data methodology to address the empirical problem posed by age-dependent heterogeneity and the lack of within variation. To further account for the simultaneity and time-varying unobservables suggested by the dynamic panel results, I then develop a simultaneous equations structural model to address all of the aforementioned issues of selectivity. Several robustness checks are performed to ensure the reliability of the results, and counterfactual exercises are also conducted to assess how child enrollment would be affected by policies targeting internal migrants. Finally, I discuss the results in detail and draw conclusions regarding the implications of my research for understanding the complex relationship between parental migration and children's educational outcomes.

2 Data

2.1 Data and Terms

Data are collected from the China Family Panel Studies (CFPS) project[4]. Given the richness of the data project, I focus on rural children and compare those who are affected by parental migration with those who are not.

For the purposes of this study, I define several key variables that can be constructed from the available data. The term *children* refers to individuals between the ages of 0 and 17 (both inclusive) at the time of observation. Any individual who enters the survey at or before the age of 17 is classified as a child and contributes data points that aid in identification. *Rural children* are defined as children who have ever held a rural *Hukou* (data on *Hukou* status is available at birth or within the scope of the survey), which serves as a sufficient indicator of their rural origin.

My research focuses on the migration decision of the parents of those children of rural

[4]The CFPS project (Peking University, 2015) is a national household survey with 6 waves now available: 2010, 2012, 2014, 2016, 2018, 2020. The baseline target sample of the CFPS consists of 16,000 households in 25 out of a total of 31 provincial-level administrative divisions in mainland China, representing 95% of China's population. Follow-up surveys are conducted on all of these individuals and on the new members as they form new households.

origin. Children born in migrant families where the migrant parents can register the newborns as local urban residents (in the Chinese context, this requires that at least one of the parents holds a local urban *Hukou*) are not the target subjects of my analysis, because their core families must consist of an urban native parent, a rural-urban migrant parent, and the urban child, who is never observed as a rural-registered child. And the impact on these families is complicated by the presence of intermarriage between natives and migrants, which is outside the scope of my current research.

The term *migrant* at time t[5] refers to rural-urban migrants who live in urban areas but have a rural *Hukou* at the end of the period t. This variable captures temporary migrant status. The variable *migrant parent* indicates whether at least one parent is classified as a migrant. The term *left-behind child* at time t refers to children left behind by at least one migrant parent at time t. I test this definition and find that if a child is left behind, there is a 97% probability that it has been left behind by all living parents. Thus, to maximize the sample size and include more data points in my analysis, I use this definition.

To accurately classify the status of a rural child (or the event the child experiences, the group the child belongs to) in the context of migration, I divide the sample of rural children into three categories: children of non-migrant parents (NM), children who migrated with their parents (MWP), and children who were left behind (LFB).

When this categorical variable is used to describe the state of the child at a given point in time, the three categories are mutually exclusive. A child is not an active decision-maker in the household's migration decision process, but can be either the child of a non-migrant parent if both parents remain in the place of origin (NM), or the child of a migrant parent if at least one parent migrates (MWP or LFB). If the child is from a migrant family, then the parents make a further decision (although this decision may not be sequential in practice, I use this sequential interpretation only to distinguish between the two groups, which has no impact on my analysis later) about whether or not to take the child with the migrant parent to the destination, then the child will either stay with the migrant parents (MWP) or be left behind (LFB). And mutually exclusive statuses mean that the rural child can currently experience only one of the three statuses.

However, when the categorical variable describes the child's experience, the three categories are not necessarily mutually exclusive. Based on a child's history up to a given

[5]The baseline wave of the survey is conducted in 2010, and follow-up waves are conducted every two years from 2012 to 2020. The time variable in my analysis refers to the year of the interview, which may differ from the wave variable. For example, the 2018 wave could be conducted between late 2017 and early 2019.

Table 1: Sample sizes of children by group, pooled data

Children	Count	Children of Rural Origin (92% of total children)	Count
All	51897 (100.00%)	All	47682 (100.00%)
- from non-migrant families	- 35549 (68.50%)	- from non-migrant families	- 32001 (67.11%)
- from migrant families	- 16348 (31.50%)	- from migrant families	- 15681 (32.89%)
All children from migrant families	16348 (100.00%)	All children from migrant families	15681 (100.00%)
- not left-behind by parents	- 13701 (83.81%)	- not left-behind by parents	- 13051 (83.23%)
- left-behind by at least one parent	- 2647 (16.19%)	- left-behind by at least one parent	- 2630 (16.77%)

[a] A *child* is defined as a person between the ages of 0 and 17 (both inclusive).

point in time, a child may experience more than one event from the event pool (NM, MWP, or LFB). For example, the historical status LFB refers to whether the child was ever left behind by at least one migrant parent at any point in time. And the child may experience more than one event in his or her history; for example, a child may be left behind by his or her father at age 5, then migrate with his or her mother at age 10, and then be left behind by his or her mother at age 15.

The descriptive statistics discussed below are constructed based on the child's current state (temporary status, where the three categories are mutually exclusive) rather than his or her historical status (ever status, where the three categories are not necessarily mutually exclusive).

Table 1 shows the sample sizes for children by group (temporary status). The majority of children in the sample come from rural areas (more than 90%), with a significant proportion coming from migrant families (more than 30%) and a non-negligible proportion is separated from their parents (around 17%).

2.2 Descriptive Statistics

2.2.1 The Pattern of Household Migration

Given the availability of panel data, it is instructive to examine the dynamics of children's status over time, which will also reflect the migration pattern of their parents. Table 2 shows the law of motion calculated from empirical frequencies. The first two diagonal

Table 2: Transitions in migration modes

t \t+1	Non-migrant	Migrates w/o child	Migrates w/ child
Non-migrant	19990 (92%)	683 (3%)	1063 (5%)
Migrates w/o child	1141 (53%)	772 (36%)	247 (11%)
Migrates w/ child	2086 (22%)	195 (2%)	7387 (76%)

[a] $t \in \{2010, 2012, 2014, 2016, 2018\}$, $t+1 \in \{2012, 2014, 2016, 2018, 2020\}$
[b] Number of Observations (Child $\times$ $[t, t+1]$) = 33,564
[c] Number of Unique Children = 11,739
[d] Modes refers to the current/temporaray status.

states indicate that among all possibilities, children from non-migrant families and those who migrated with their parents are most likely to remain in their respective statuses in subsequent periods. In contrast, for left-behind children, the most likely event in the next period is reunification with their parents in the place of origin. This finding is consistent with evidence from China, which suggests that most migrants eventually repatriate and that the majority of migrants in urban destinations are repeat or seasonal migrants (Wang et al., 2014).

One potential concern that arises from Table 2 is the large overlap between the (historical) non-migrant group and the (historical) left-behind group of children, even though their current statuses are mutually exclusive, as can be seen from the lower left element of the table. The fact that children in the NM group have usually (the probability will be greater than 53%, which is the average of all $[t, t+1]$ pairs) experienced the LFB event in their very recent history (this is a more serious problem than the fact that most LFB children have experienced NM in their history) makes it even impossible to separate the effect of one event from the other. And there is no terminal action for the child's history.

Table 3 looks at the correlation between the two groups of migrant children: MWP and LBF. The table shows that 53% of the parents who left their children in the origin eventually return to the origin within two years, while only 22% of the parents who migrated with their children return to the origin within two years.

Table 3: Migration pattern by group

Child's Current/Temporary Staus	Parents Decision		Avg. rate of parents' return in two years (2010-2020)	Avg. rate of parents' migration in two years (2010-2020)
	Parents Migrate	Parents migrate with child		
Non-migrant	No	No		8%
Left behind by parents	Yes	No	53%	
Migrated with parents	Yes	Yes	22%	

Table 4: Descriptive statistics of rural children, pooled data 2010-2020

Children of Rural Origin	Age group	Count	Educational Outcomes: Share enrolled	Math test score [0,10]	Word test score [0,10]	Health Outcomes: Weight-for-age z-score	Height-for-age z-score	Depression scale*(-1) [-10,0]
of non-migrants	[6,11]	11533	94.50%	4.159	5.195	-0.83	-1.412	-0.197
			(0.20%)	(0.037)	(0.038)	(0.029)	(0.025)	(0.003)
migrates w/o child	[6,11]	1011	96.70%	4.204	5.674	-0.756	-0.933	-0.23 (0.01)
			(0.60%)	(0.142)	(0.155)	(0.098)	(0.072)	
migrates w/ child	[6,11]	4693	96.40%	4.331	5.743	-0.497	-0.615	-0.213
			(0.30%)	(0.052)	(0.056)	(0.041)	(0.033)	(0.004)
of non-migrants	[12,17]	10592	86.90%	5.979	6.366	-0.96	-0.737	-0.204
			(0.30%)	(0.023)	(0.025)	(0.023)	(0.015)	(0.002)
migrates w/o child	[12,17]	736	90.10%	6.452	6.594	-0.621	-0.509	-0.233
			(1.10%)	(0.088)	(0.098)	(0.066)	(0.058)	(0.006)
migrates w/ child	[12,17]	4286	91.20%	6.346	6.734	-0.642	-0.315	-0.218
			(0.40%)	(0.034)	(0.038)	(0.035)	(0.021)	(0.002)

[a] Standard errors in parentheses.
[b] Statuses refer to the child's current/temporaray status.
[c] The highest value of each indicator in each age group is in red, the lowest in green.

One possible explanation for this fact is that the two groups of parents differ in their migration intentions. They're more likely to choose seasonal migration and therefore more likely to leave their children behind. Another explanation is that the two groups of parents differ in their ability to settle in the destination city. Parents who are better able to stay longer before returning are also better able to bring their child to the destination. This confirms the earlier concern that any failure to address the issue of selectivity may lead to biased estimates.

2.2.2 The Child Outcomes

To provide an overview of the well-being of rural children affected by parental migration, I present descriptive statistics of child outcomes in Table 4, and two of them are shown in Figure 1 for a clear view of the comparison across groups. I include two different categories of child development indicators: health outcomes and educational outcomes. An explanation of the indicators can be found in Table A1.

Because the effects of experiencing each status may vary depending on a child's age and level of schooling, I divide the sample of children aged 0-17 into three age cohorts: primary school age (6-11 years), middle school age (12-14 years), and high school age (15-17 years).

In the primary school age group, children left behind perform best in terms of school enrolment, while children who migrated with their parents perform best in all other indicators except mental health (note that the lower the depression scale, the better the

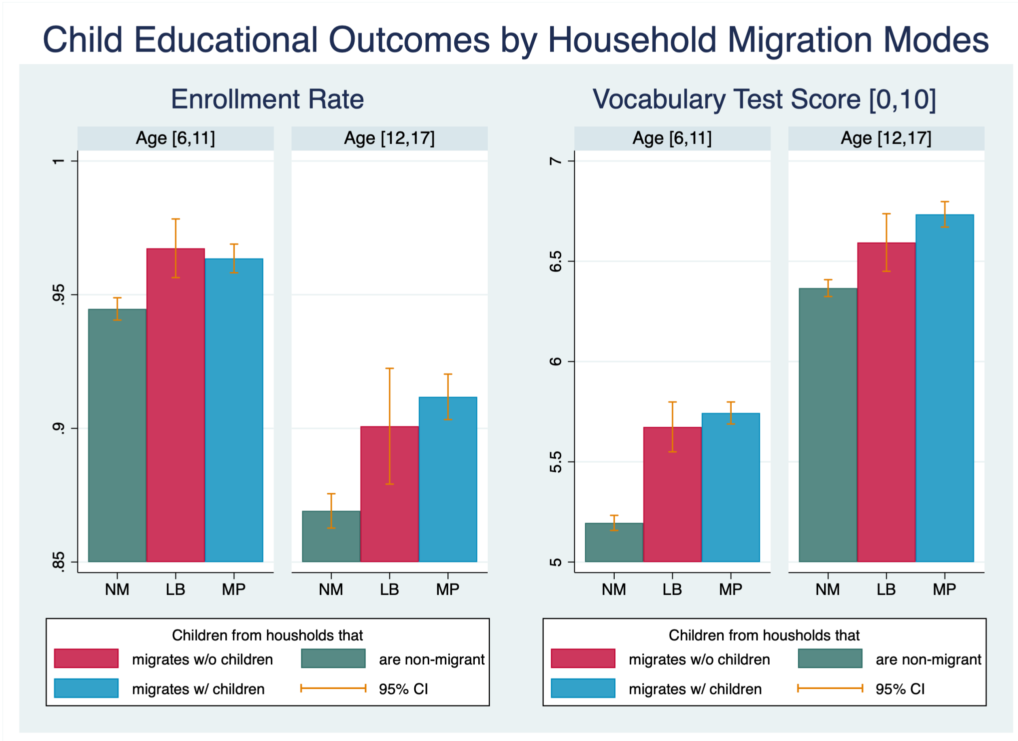

Figure 1: Comparison of Child Educational Outcomes by Group

mental health). The children of non-migrants have the best mental health and the worst in all other indicators. The other two age groups show similar patterns, with children who migrated with their parents doing best on some indicators and children who stayed behind doing best on others, except for mental health, and the children of non-migrants do the worst on all indicators except for mental health. This finding is consistent with existing literature documenting a positive association between having migrant family members and mental health problems such as depression, anxiety, loneliness, and low self-esteem (Myerson, 2017; Ivlevs et al., 2019). One concern, however, is that it may be the same reason children have poor mental health and parents leave their hometowns.

Because of the potential endogeneity that it could be the same factors that drive both parents' labor market outcomes and their migration and child care decisions, it is imperative to include parents' characteristics in my analysis. Table 5 provides a descriptive overview of the characteristics of parents in each group.

At first sight, migrant parents (parents of MWP and LFB children) always have higher incomes and are more likely to be employed in either blue-collar or white-collar occupations, where the base category is not employed, in all three age groups. Interestingly, fathers who leave their children behind tend to have the highest levels of education and

Table 5: Descriptive statistics of parents and counts of children by group, pooled data

			Father				Mother			
Children of Rural Origin	Age group	Count	Compulsory Edu. Completion Rate	Avg. log income, log k CNY	#Blue-collar	#White-collar	Compulsory Edu. Completion Rate	Avg. log income, log k CNY	#Blue-collar	#White-collar
of non-migrants	[0,5]	9876	64%	0.153	73.0%	6.1%	58%	0.085	48.2%	5.3%
				(0.003)	(7210)	(606)		(0.002)	(4765)	(522)
migrates w/o child	[0,5]	883	82%	0.261	68.3%	16.9%	75%	0.138	42.2%	10.3%
				(0.01)	(603)	(149)		(0.006)	(373)	(91)
migrates w/ child	[0,5]	4072	75%	0.22	69.7%	13.0%	77%	0.122	44.2%	11.0%
				(0.005)	(2837)	(530)		(0.003)	(1799)	(446)
of non-migrants	[6,11]	11533	56%	0.148	77.5%	5.9%	47%	0.084	64.5%	5.2%
				(0.003)	(8938)	(676)		(0.001)	(7442)	(605)
migrates w/o child	[6,11]	1011	76%	0.267	71.2%	15.4%	65%	0.165	56.0%	9.9%
				(0.01)	(720)	(156)		(0.02)	(566)	(100)
migrates w/ child	[6,11]	4693	72%	0.223	74.5%	12.8%	68%	0.13	57.6%	12.0%
				(0.005)	(3495)	(602)		(0.003)	(2703)	(565)
of non-migrants	[12,17]	10592	47%	0.126	79.0%	5.1%	33%	0.076	73.2%	3.6%
				(0.002)	(8366)	(537)		(0.001)	(7755)	(381)
migrates w/o child	[12,17]	736	65%	0.219	73.9%	14.1%	49%	0.12	65.1%	8.7%
				(0.015)	(544)	(104)		(0.006)	(479)	(64)
migrates w/ child	[12,17]	4286	65%	0.195	75.7%	11.5%	53%	0.12	64.9%	9.9%
				(0.004)	(3245)	(495)		(0.003)	(2783)	(426)

[a] Status refers to the child's current/temporaray status.
[b] For income data, standard errors in parentheses. For percentages, counts in parentheses.
[c] The base category for parents employment is not employed.
[d] The highest value of each indicator in each age group is in red, the lowest in green.

are most likely to be employed in white-collar occupations. Although more educated fathers with better jobs may earn higher wages in their destination country, they are less likely to bring their children with them than less educated blue-collar fathers. As the father's role as a breadwinner increases, his role as a caregiver for the child decreases. Conversely, mothers who migrate with their children are more likely to have higher levels of education and income. The mother's two roles as breadwinner and caregiver move in the same direction. This motivates further study of the father's and mother's contribution to child care with the variation brought about by the migration process. [6]

The data described provide a compelling motivation for further analysis aimed at disentangling the endogeneity between migration decisions and the child's educational outcome of interest. For the purposes of this analysis, I treat fertility as an exogenous decision. In addition, to eliminate possible interdependencies among children in the same household, I include only the youngest minor child from each household in my analysis, so that the effect of family size is included in the household fixed effects and not in the core of the current analysis.

[6] Some studies have also attempted to distinguish effects based on the gender of the migrant parent, but because the extent of female migration is limited in many countries, few studies have produced meaningful results. One exception is Cortes (2015), which finds that maternal absence is more detrimental than paternal absence.

3 The Migration Outcome Gap

Using panel data from all waves (2010-2020), where each observation refers to each household in each wave, it is natural to start by estimating a panel data model with fixed effects, such as the individual (household) fixed effects or the mother fixed effects common in the child development literature (Blau, 1999).

A potential concern with the fixed effects regression is the identification of the effects of status. The fixed effects specifications are more empirically correct in the sense that they are more robust, while the identification of the linear parameter of child status (NM, MWP, or LFB) on child outcome relies on having enough observations of children changing status across periods, otherwise the individual or the mother fixed effects will absorb almost all of the variation. With the short panel data I'm using, there are not enough observations to ensure the precision of the estimation.

One solution is to estimate the outcome curves using dynamic panel strategy, similar to the strategy used in Guner et al. (2018). The following model is estimated:

$$y_{it} = \phi y_{it-1} + \alpha(a_{it}) + \sum_{j} \beta_j(a_{it}) \cdot D_{jit} + X_{it}\gamma + \delta_{region} + \delta_{mother} + \varepsilon_{it} \tag{3.1}$$

where y_{it} is the outcome of the child in household i in period t, and j indicates the child's group, a is the child's age, and α and β_j are linear functions of age (estimated using age dummies and the interaction of age dummies and D), D_j is on if the child belongs to group j, vector X is a vector of household demographics, including the child's sex and the parents' education, occupation, and income. It also includes the region (province) fixed effects and the mother fixed effects (which may differ from the household effect because the same mother may form different households over time). And $\beta(a)_j$ are the age curves of interest.

The D variable here refers to the current/temporary status of the child. The choice of this variable is based on two reasons. First, in the dynamic panel setting, this variable makes more sense than the historical/ever status to explain the accumulation of child outcomes over time, while the historical status is more appropriate for the cross-sectional setting. And a more important empirical reason is that since this approach to the migration outcome gap is motivated by the fact that the time variation in children's status is not large enough to identify the effect of status, the historical status has even less time variation than the current status, making it even more difficult to identify the effect of status.

For each status, I estimate an outcome curve $\beta(a)$ across all ages. Then, the pooled observations of children who do not change status across time periods (e.g., who remain a left-behind child across all observed time periods) contribute to the identification of the individual curves for each status, while the pooled observations of children who do change status contribute to the identification of the gap between the curves. There are three statuses, while the non-migrant children are taken as the baseline, which generates two gaps.

Because I use the dynamic panel method to account for the cumulative nature of children's outcomes, and because the support of the historical status variable for most children does not change much between any two consecutive periods[7], I present only the estimation results using the child's current status, which makes more sense in the dynamic panel setting, in Figure 2 and Figure 3, both using the specification (3.1) but with different child outcome indicators y. The regression results of the full equation are reported in Table A3. The vocabulary test scores are reported together with the enrollment rate to account for differences in the quality of education between rural and urban areas.

Looking at the left panel of figure 2, we see that the event of *migrating with parents* has no significant effect before about age 15, which is roughly the age at which compulsory schooling ends. Compared to non-migrant children, migrant children living with their parents in urban areas are more likely to attend high school. From figure 3 we can see that the *currently migrating with parents* event has a positive immediate effect on the vocabulary test score. If we compare these two left panels and focus on children under 15, the event *currently migrating with parents* doesn't change the frequency of children going to school, but still increases their performance on the vocabulary tests, suggesting that the quality of education is better in urban areas than in rural areas. Given that parental characteristics are controlled for in the regressions, this difference in test scores is less likely to be the result of differences in the amount of help their parents provided for their academic performance.

Comparing the right panels of figure 2 and figure 3, left-behind children are generally less likely to attend school before high school age, while the effect of being left-behind on vocabulary test scores is not different from zero, meaning that although they attend school less than the baseline non-migrant children, their scores are not significantly affected. Even though the left-behind children live in rural areas, attend similar schools as the

[7]Obviously, the time variation of this support is less than the time variation of the current status.

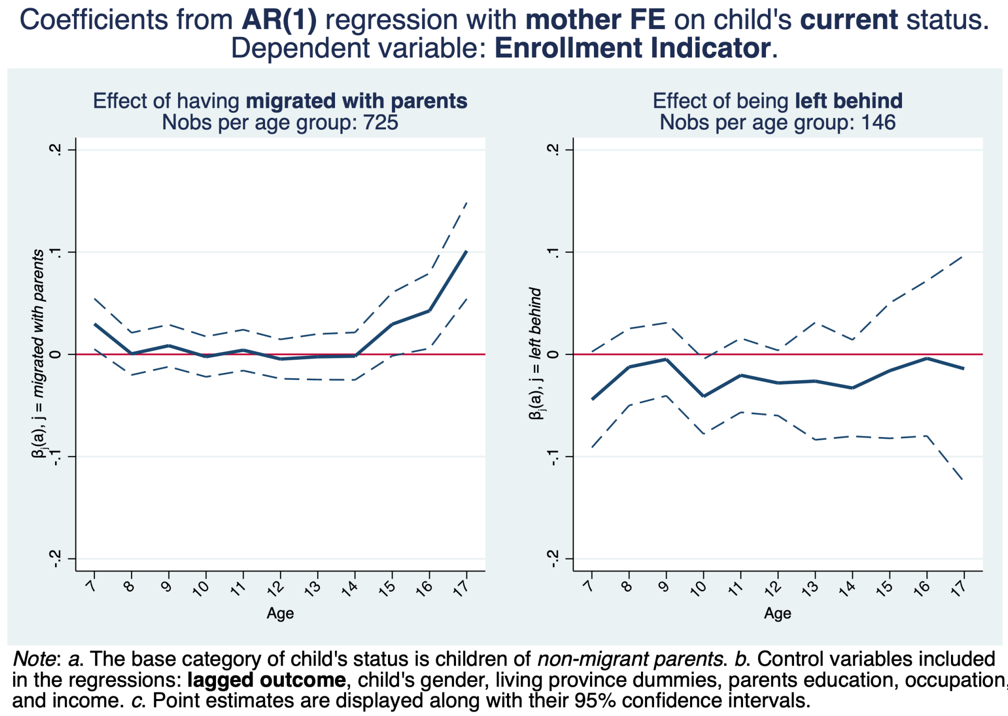

Figure 2: Migration Outcome Gap on Enrollment Rate: Estimates for $\beta_j(a_{it})$ from the dynamic panel

non-migrant children, and don't have their parents living nearby, they are not necessarily worse off than the non-migrant children in terms of academic performance, suggesting that with the remittances their parents send home, they (or more likely their grandparents) still have the ability to translate monetary resources into educational resources in rural areas.

4 A Simultaneous Equations Model on Household Decisions

To better deal with the problems of multiple selectivity, potential time-varying unobservables, and reverse causality, my next step is to include household decisions about migration and child education in a structural model.

As illustrated in section 2.2.1, the overlap in experiences between the two groups of children: children of non-migrant parents and children who left-behind their parents, is large. And this contributes to the problem that it's difficult to identify the effect of being

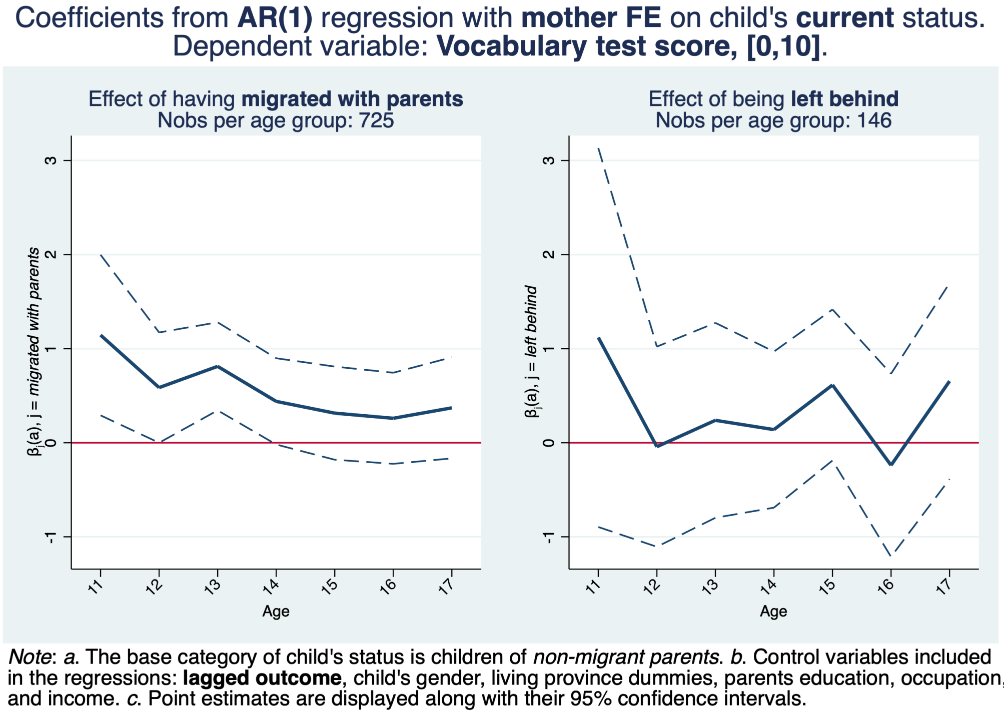

Figure 3: Migration Outcome Gap on Vocabulary Test Score: Estimates for $\beta_j(a_{it})$ from the dynamic panel

left behind from being the child of non-migrant parents. My response to this problem in the following model is to put non-migrant children and left-behind children in the same pool, and to define the household's migration decision as whether the whole household migrates. An argument in favor of this solution is that the effects of permanent migration and temporary could be very different in magnitude, the return migrants and their families are less affected by the migration.

Another possible solution to the above problem is to calculate the total number of periods that each child is under each event in its entire history. But the empirical difficulty is that for two children of the same age who have experienced the two events (although there are actually three events, the degree of freedom is 2) both for the same number of periods, these events may still be spaced differently in time, which may result in different effects on them, but this strategy can't distinguish between them.

One solution that can somehow solve the heterogeneous experience problem is a discrete choice model of the migration decision and the decision to bring children, with the child outcome indicator as a state variable. The problem with this solution, however, is that the cumulative nature of child outcomes is omitted because child outcomes (especially test scores or z-scores for weight and height) are not necessarily linear in time.

In the current analysis, I simply omit the temporal spacing of events and start with a static model that considers the binary enrollment indicator as the child outcome variable, the cumulative nature of which can be forgotten for a second. This model will be extended to a dynamic discrete choice model in the future.

In this section, I present a simultaneous equations model of household decisions: the gains and costs of migration and enrollment, and the correlation between household migration and child enrollment. There is no dynamics in the model.

4.1 Household Decisions

The subjects are rural households with one child. Each household makes two binary decisions: whether or not to migrate, denoted by m, and whether or not to enroll the child in school, denoted by e.

$$m = \mathbb{1}\{\text{the whole household migrates}\},$$
$$e = \mathbb{1}\{\text{the child is enrolled in school}\}$$

The main model consists of two Probit regressions:

$$\begin{aligned} \mathbb{P}(m_t = 1) = F\Big\{ & \phi^m \cdot m_{t-1} + \eta^m \cdot (\bar{w}_t - w_t^\circ) + \mu_1^m \cdot \big(\bar{c}_{h,t} - c_{h,t}^\circ\big) + \mu_2^m \cdot \bar{I}_t \\ & + \beta^m(a_t) \cdot e_t + z_t \cdot \gamma^m + \delta_\circ^m \Big\} \end{aligned} \tag{4.2}$$

$$\begin{aligned} \mathbb{P}(e_t = 1) = F\Big\{ & \phi^e \cdot e_{t-1} + \eta^e(a_t) \cdot [(1 - m_t) \cdot w_t^\circ + m_t \cdot \bar{w}_t] \\ & + \mu^e \cdot \big[(1 - m_t) \cdot c_{edu,t}^\circ + m_t \cdot \bar{c}_{edu,t}\big] + z_t \cdot \gamma^e + x_t \cdot \alpha^e + \delta_\circ^e \Big\} \end{aligned} \tag{4.3}$$

where the superscripts are used to distinguish the two equations.

In equation (4.2), $\bar{w} - w^\circ$ is the migration gain, which is the predicted age- and education-specific urban income premium; $\bar{c}_h - c_h^\circ$ is the housing cost of migration to urban areas; $\bar{I}^\circ$ is the institutional cost faced by the household. The construction of these terms is explained in later sections. The variable a indicates whether the child is of primary school age (6-11), middle school age (12-14), or high school age (15-17). The inclusion of the term $\beta^m(a_t) \cdot e_t$ is based on the assumption that the household's migration decision is affected by whether the child is enrolled or not, and its parameter reflects how much the household will migrate to have a higher chance of enrolling the child, and this effect is assumed to be heterogeneous across different age groups of children.

In equation (4.3), $(1-m) \cdot w^{\circ} + m \cdot \bar{w}$ is the expected wage, taking into account migration decision, whose age cohort-specific effect reflects how much of the monetary resources earned by the household, taking into account the urban wage premium from migration, will be invested in the child's educational resources, which may differ depending on the child's level of schooling. And the expression $(1-m) \cdot c^{\circ}_{edu} + m \cdot \bar{c}_{edu}$ is the expected age- and location-specific educational expenditure of the child, taking into account the additional cost of education in urban areas.

In both equations, the parent characteristic vector z is included, and the child characteristic vector x is included in the enrollment equation. This accounts for the fact that the active decision makers in the migration decision are actually the parents.

The decision variable e appears directly in the migration equation, and m appears somewhat indirectly in the enrollment equation. Although a term m (or its interaction with age group) in the enrollment equation can account for the effect of migration on enrollment, the linear inclusion of the two decisions on the right-hand side of the two equations will just makes the system of equations fully recursive. In terms of identification, it's equivalent to either linearly including an e in the migration equation or linearly including an m in the enrollment equation. The former is chosen in the current analysis to indicate how much parents migrate for higher enrollment of the child.

4.2 Institutional Costs

The term $\bar{I}^{\circ}$ is a proxy for the institutional costs of migration. This measure $I^{\mathtt{d}}$ is derived from Zhang et al. (2019) and measures the ease with which the potential migrant can move to the destination $\mathtt{d}$; the more difficult it is to obtain the *Hukou* of $\mathtt{d}$, the higher the index $I^{\mathtt{d}}$.

The index $I^{\mathtt{d}}$ has two values for each place $\mathtt{d}$ (province[8]) - one before 2014 and one

[8]There are a total of 31 provincial administrations (for simplicity, I will refer to all of them as "provinces" in my text, although not all of them are called "provinces" in Chinese, e.g. some of them are autonomous regions or directly administered municipalities) in mainland China. Each province consists of many cities (municipalities), and most city-level administrations have both rural and urban areas. In fact, immigration regulations are issued at the city level to restrict mobility to the urban areas of that city. And only large cities, there are usually one or at most two "large" cities from each province, set requirements for immigrants. That's one reason why I aggregate the data at the provincial level. Another reason is the level at which the social insurance fund is financed, which in my model is the key factor influencing the decision to migrate. Usually it's at the provincial level or the city level, and the national trend is moving from the city fund to the provincial fund, this fact supports that I aggregate the data at the provincial level.

from and after 2014 - and remains fixed for the duration of each time interval.

The index measures the ease of obtaining a local *Hukou* based on the migrant's employment (job and length, contribution to pension system, etc.), educational background (high-tech migrants are more welcomed), local investment, and real estate purchase. A higher index indicates a more restrictive policy. The index is highest for the capital of China, followed by the other major cities.

Given the index I° observed at the destination level, I construct a *shift-share instrument* $\bar{I}^{\circ}$ for the institutional cost of migration from origin $\mathtt{o}$, which is the weighted sum of the *Hukou* index across all potential destinations:

$$\bar{I}_t = \sum_{\mathtt{d}} \frac{\mathtt{weight(o,d)}}{\sum_{\mathtt{d}} \mathtt{weight(o,d)}} \cdot I_t^{\mathtt{d}}$$

where the weight is either *geographic distance* or *historical settlement patterns.*

Using geographic distance as weights generally makes sense for capturing relocation costs, but it has the disadvantage that in mountainous areas or other complex terrain, travel costs are not necessarily linear in distance. Therefore, I also include the stock of migrants at the baseline period (year 2010), which represents the historical settlement patterns, to more accurately capture the physical costs and also to capture the idiosyncratic variation in migrant networks and preferences for particular destinations (Imbert et al., 2022).

In a nutshell, $\bar{I}^{\circ}$ is the institutional cost that households at origin $\mathtt{o}$ face when migrating internally to any destination.

Recent literature discusses identification and inference in shift-share designs (Adão et al., 2019), and suggests that consistency can be achieved if either the shares (Goldsmith-Pinkham et al., 2020) or the shifts (Borusyak et al., 2022) are exogenous. In my setting, the shares-historical settlement patterns-reflect migrant workers' expectations about the evolution of labor demand across destinations and are likely endogenous to labor outcomes in cities. To address this concern, I also attempted to use distance as a proxy for shares.

The validity of my shift-share design using historical settlement patterns rests primarily on the assumption that shifts-the institutional costs-are exogenous to child outcomes in cities. This assumption is plausible because the national policy issued in 2014 on the *Hukou* index [9] is less likely to be correlated with local available school seats and edu-

[9] From 2014, migration to megacities becomes more restrictive, and migration to other cities becomes less restrictive.

cational quality. And also the exclusion restriction is required, i.e. that the shifts only affect the child outcome of migrant households through migration, and holds because all child-related outcomes are not included in the construction of the index. The exclusion restrictions on the exogenous regressor $\bar{I}^{\circ}$ contribute to parameter identification.

4.3 Financial Costs

The financial constraints on migration in my setting include higher housing and education costs in urban areas.

Although in the $m = 0$ group, parents from households that leave their children behind also suffer from higher living costs in urban areas, there could still be higher living costs for households that bring their children to urban areas, because bringing children may require larger housing space. And this space requirement could vary depending on the age of the child and the destination.

Therefore, the housing costs (c_h) as well as the education costs (c_{edu}) in urban areas (l=1) and in rural areas (l=0) are predicted using the same specification[10]:

$$c_{ilt} = \theta^c \cdot a_{it} + \delta^c_{prov} \times \delta^c_l \times \delta^c_t + \epsilon_{ilt}$$

where a_{it} denotes the age of the child in household i at time t. A location-specific time effect $\delta^c_{prov} \times \delta^c_l \times \delta^c_t$ is also included, where l is the type of location, either rural (l=0) or urban (l=1).

Then the expected expenses at destination and at origin are

$$\begin{aligned} \bar{c} &\equiv \sum_{prov \in \{\mathtt{d}\}} \frac{\mathtt{weight}(\mathtt{o}, prov) \cdot \hat{c}_{i1t,prov}}{\sum_{prov \in \{\mathtt{d}\}} \mathtt{weight}(\mathtt{o}, prov) \cdot \hat{c}_{i1t,prov}} \\ c^{\circ} &\equiv \hat{c}_{i0t,\mathtt{o}} \end{aligned}$$

respectively. The same method of "shifting" shares used to construct institutional costs is used here, since the unit costs of different destinations do not necessarily affect potential migrants in the same way.

In a binary model, including the cost difference (or income difference for migration

[10]The expenditure variables are in logs. To avoid zeros when taking the log, I first add one Chinese yuan (CNY) to the expenses and then take the log. One CNY is about 0.15 USD in 2010, so the effect of adding it is almost negligible. After predicting the two costs, I convert the units from logs to 100 CNY. This is also true for the income variable w.

gains) is equivalent to including costs as an alternative-specific variable. And although I observe the reported costs, the alternative case counterfactual is still needed. I run the above regressions to predict the age- and location-specific costs for each household, to eliminate endogeneity of the information contained in the household's decision, and to construct the alternative-specific (l-specific) costs for each household.

4.4 Urban Income Premium

I also predict the alternative-specific income for each household and construct the age-, education-, and province-specific urban income premium using the specification:

$$w_{ilt,prov} = \theta_1^w \cdot a_{it}^h + \theta_{2l}^w \cdot edu_i^h + \delta_{prov}^w \times \delta_l^w \times \delta_t^w + \varepsilon_{ilt}$$

where w is the log income, a^h is the age of the household head[11] and the superscript h is used to distinguish between the age of the child and the ages of the parents. The effect of the education level of the household head edu^h is allowed to differ between rural and urban areas, so the parameter θ_{2l} has a subscript l. A location-specific time effect $\delta_{prov}^w \times \delta_l^w \times \delta_t^w$ is included.

The household-specific urban wage premium is the difference between the urban (l=1) and rural (l=0) income of a household with the same characteristics:

$$\bar{w} \equiv \sum_{prov \in \{\mathtt{d}\}} \frac{\mathtt{weight}(\mathtt{o}, prov) \cdot \hat{w}_{i1t,prov}}{\sum_{prov \in \{\mathtt{d}\}} \mathtt{weight}(\mathtt{o}, prov) \cdot \hat{w}_{i1t,prov}},$$
$$\text{and} \quad w^{\circ} \equiv \hat{w}_{i0t,\mathtt{o}}$$

4.5 Estimation Results

The estimation results are presented in Table 6, while the full regression table of different specifications is presented in Table A4. Results using alternative weights are presented in Table 6 and Table A4.

The system of simultaneous equations consists of two Probits, the parameter estimates of which are used as initial values for the joint system in the Full Information Maximum Likelihood (FIML) estimation, where the consistent estimation relies on the errors from the two equations being jointly normally distributed.

[11] The household head is defined on the *Hukou* booklet of each household, and is usually the father.

Table 6: Parameter Estimates, Gains and Costs Weighted by the Baseline Migration Stock

Migration Equation			Enrollment Equation		
Coefficient	Estimate	Std.Err.	Coefficient	Estimate	Std.Err.
η^m	0.7394	0.7504	$\eta^e(a=[6,11])$	10.2857	0.7218
μ_1^m	28.4402	11.3079	$\eta^e(a=[12,14])$	18.5135	0.9291
μ_2^m	-1.1701	0.7904	$\eta^e(a=[15,17])$	13.2745	0.5781
ψ^m(Agr.)	0.0009	0.0034	μ^e	42.6730	1.7576
ψ^m(Non-Agr.)	0.0056	0.0011	$\alpha^e \cdot \mathbb{1}(a=[12,14])$	-7.7172	1.0234
ϕ^m	-0.0446	0.0051	$\alpha^e \cdot \mathbb{1}(a=[15,17])$	-3.6493	0.8339
$\beta^m(a=[12,14],e=0)$	-0.2236	0.1287			
$\beta^m(a=[15,17],e=0)$	0.5584	0.0835			
$\beta^m(a=[6,11],e=1)$	1.6500	0.0620			
$\beta^m(a=[12,14],e=1)$	1.6736	0.0690			
$\beta^m(a=[15,17],e=1)$	1.9896	0.0735			

[a] Standard errors are clustered by household.

From Table 6, the estimates for $\beta_m(a_t)$ suggest that parents migrate for better educational opportunities for their children. The effect increases as the child ages which suggests that households are more likely to migrate if they have middle and high school-aged children, and that the goal of financing a middle school child's education contributes the most to their motivation to migrate[12].

The higher institutional costs of mobility restriction policies do not really work to reduce the flow of unskilled migrants (μ_2^m). A possible reason for the insignificant parameter μ_2^m not distinguished from zero for institutional costs is that institutional barriers are usually associated with better amenities at the destination, which are not included as part of the gains from migration in the model. And the value of amenities is not reflected in the income premium either, because the income of an attractive (in terms of amenities) destination may be compromised. Although restrictions are imposed on receiving regions to limit the inflow of unskilled migrants, their effect on migration is simply not significant.

After compulsory education, the enrollment rate drops significantly (α^e), and the increase in income from migration somehow mitigates this trend (η^e). More importantly, a positive η^e suggests that migrant households invest the urban income premium earned from migration in their children's education, with the effect becoming significant after the children enter middle school.

The estimates for μ_1^m and μ^e tell us that parents are sensitive to both education and

[12] Although middle and high school seats are much more limited, especially in urban areas, according to Gao et al. (2022).

Table 7: Effect of Removing the *Hukou* Restrictions ($\bar{I} \equiv 0$) , Gains and Costs Weighted by the Baseline Migration Stock

Probability	Estimate (actual)	Estimate ($\bar{I} \equiv 0$)	Relative Change ($\bar{I} \equiv 0$)
Prob.(m=1,e=1)	0.39 (0.0025)	0.43 (0.0023)	10.3%
Prob.(m=1,e=0)	0.03 (0.0003)	0.04 (0.0003)	33.3%
Prob.(m=0,e=1)	0.75 (0.0011)	0.71 (0.0012)	-5.3%
Prob.(m=0,e=0)	0.09 (0.0007)	0.09 (0.0007)	0.0%

[a] Mean coefficients; SE in parentheses.

housing costs. Most manufacturing workers are paid on a daily or a weekly basis, and it may be easy for them to quit a job and return to their hometown, it's not easy for them to quit the apartment, which is usually paid on a monthly basis, or it might be even harder for them to quit their children's school, where fees are paid every semester.

4.6 Counterfactual Exercises

To better capture the effect of migration on child enrollment and to assess how child enrollment responds to different policies targeting migrant households, two counterfactual exercises are conducted.

In the first exercise, the results of which are shown in Table **??**, the *Hukou* restriction is removed, which is equivalent to setting the institutional cost to zero. When there are no restrictions on applying for citizenship, the probability of the entire household migrating decreases for households with primary and middle school children and increases for households with high school children. Households with high school children have the highest willingness to migrate and are the typical migrants restricted by the policy.

In this counterfactual, the enrollment rate increases for primary and high school-aged children (although the latter increase is almost negligible). For primary school children, this increase is likely due to the elimination of the delay in enrolling the child in school. For high school students, this increase is made possible by preventing them from entering the labor market too early. For middle school children, most of whom are already enrolled in th compulsory schooling, the enrollment rate of which is approaching the "natural" ceiling, for whom are still not enrolled, the effect of preventing delay or dropout would also not work.

In the second exercise, the results of which are shown in Table 8, the education expenses of migrant households are subsidized by 1000 CNY per year[13]. The likelihood of

[13]The average babysitter salary across all cities in my 2010 data is about 1200 CNY per month.

Table 8: Effect of Tuition Reduction at Destination for Migrant Households ($\Delta\bar{c} = -1000$ CNY/Year)

	Age Group		
	6-11	12-14	15-17
Prob.(Migration)			
Actual	0.295	0.308	0.286
	(0.004)	(0.007)	(0.007)
Predicted	0.311	0.327	0.338
	(0.004)	(0.006)	(0.007)
Prob.(Enrollment)			
Actual	0.954	0.973	0.800
	(0.002)	(0.002)	(0.006)
Predicted	0.971	0.968	0.854
	(0.001)	(0.001)	(0.003)
Observations	10871	4573	4139

SE in parentheses.
Gains and costs weighted by baseline migration stock.
Currency is CNY in 2010, 1000 CNY is around 150 USD.

the entire household migrating increases, more so for households with older children. The school enrollment rate increases slightly for primary and high school-aged children, with almost the same effect size of removing the *Hukou* restriction. The effect is significantly larger for households with high school-age children, because post-compulsory schooling is costly and more sensitive to household budgets.

Results from counterfactual exercises using alternative weights can be found in Table A6 and Table A7.

5 Conclusion

In the context of rural-urban migration in China, using a dynamic panel data model and a simultaneous equations system, the analysis highlights the motive to migrate for higher educational opportunities of the child in the household migration decision and predicts a gain in educational outcomes from rural-urban migration.

For the sending regions, from the results of the dynamic panel model, even if the children stay behind, there is a gain in their educational outcomes from parental migration. However, the loss in their mental health observed in the descriptive statistics is not included in the model. The results of the structural model confirm this finding and also indicate that the income premium from household migration is invested in the child's education and contributes to the child's human capital accumulation.

For receiving regions, institutional costs do not work to curb the inflow of unskilled migrants, and housing costs are the main concern for migrant households, although the benefits of amenities are not included in the model due to a lack of policies. Mobility restrictions prevent migrant households from legally accessing quality public services associated with a local urban *Hukou*, but don't prevent them from going to destination areas. This only leads to a loss of welfare.

As discussed in the previous sections, if the data allow, a measure of amenities is needed to test the validity of the current analysis, and other measures (test scores and health outcomes) can also be considered in the main model. And more importantly, as discussed in section 4, if concerns regarding the nonlinearity of the cumulative outcomes over time can be addressed, a dynamic discrete choice model may be a better fit for incorporating the full history of the child into the analysis.

References

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales (Nov. 1, 2019). "Shift-Share Designs: Theory and Inference*". In: *The Quarterly Journal of Economics* 134.4, pp. 1949–2010.

Ambler, Kate, Diego Aycinena, and Dean Yang (Apr. 2015). "Channeling Remittances to Education: A Field Experiment among Migrants from El Salvador". In: *American Economic Journal: Applied Economics* 7.2, pp. 207–232.

Amuedo-Dorantes, Catalina and Susan Pozo (Dec. 1, 2010). "Accounting for Remittance and Migration Effects on Children's Schooling". In: *World Development* 38.12, pp. 1747–1759.

Antman, Francisca M. (May 2011). "International Migration and Gender Discrimination among Children Left Behind". In: *American Economic Review* 101.3, pp. 645–649.

— (Oct. 1, 2012). "Gender, Educational Attainment, and the Impact of Parental Migration on Children Left Behind". In: *Journal of Population Economics* 25.4, pp. 1187–1214.

Blau, David M. (1999). "The Effect of Income on Child Development". In: *The Review of Economics and Statistics* 81.2, pp. 261–276. JSTOR: 2646864.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel (Jan. 1, 2022). "Quasi-Experimental Shift-Share Research Designs". In: *The Review of Economic Studies* 89.1, pp. 181–213.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak (2014). "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh". In: *Econometrica* 82.5, pp. 1671–1748. JSTOR: 24029292.

Böhme, Marcus H., Ruth Persian, and Tobias Stöhr (Jan. 1, 2015). "Alone but Better off? Adult Child Migration and Health of Elderly Parents in Moldova". In: *Journal of Health Economics* 39, pp. 211–227.

Chang, Fang et al. (Apr. 2019). "Parental Migration, Educational Achievement, and Mental Health of Junior High School Students in Rural China". In: *China Economic Review* 54, pp. 337–349.

Chang, Hongqin, Xiao-yuan Dong, and Fiona MacPhail (Dec. 2011). "Labor Migration and Time Use Patterns of the Left-behind Children and Elderly in Rural China". In: *World Development* 39.12, pp. 2199–2210.

Cortes, Patricia (Jan. 1, 2015). “The Feminization of International Migration and Its Effects on the Children Left Behind: Evidence from the Philippines”. In: *World Development*. Migration and Development 65, pp. 62–78.

Du, Yang, Albert Park, and Sangui Wang (Dec. 1, 2005). “Migration and Rural Poverty in China”. In: *Journal of Comparative Economics*. Symposium: Poverty and Labor Markets in China 33.4, pp. 688–709.

Edwards, Alejandra Cox and Manuelita Ureta (Dec. 1, 2003). “International Migration, Remittances, and Schooling: Evidence from El Salvador”. In: *Journal of Development Economics*. 14th Inter-American Seminar on Economics 72.2, pp. 429–461.

Gao, Xuwen et al. (2022). “Restrictions on Migration Create Gender Inequality: The Story of China’s Left-Behind Children”. In.

Gibson, John, David McKenzie, and Steven Stillman (2011). “THE IMPACTS OF INTERNATIONAL MIGRATION ON REMAINING HOUSEHOLD MEMBERS: OMNIBUS RESULTS FROM A MIGRATION LOTTERY PROGRAM”. In: *The Review of Economics and Statistics*.

— (Jan. 2013). “Accounting for Selectivity and Duration-Dependent Heterogeneity When Estimating the Impact of Emigration on Incomes and Poverty in Sending Areas”. In: *Economic Development and Cultural Change* 61.2, pp. 247–280.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (Aug. 2020). “Bartik Instruments: What, When, Why, and How”. In: *American Economic Review* 110.8, pp. 2586–2624.

Guner, Nezih, Yuliya Kulikova, and Joan Llull (May 2018). “Marriage and Health: Selection, Protection, and Assortative Mating”. In: *European Economic Review* 104, pp. 138–166.

Imbert, Clement et al. (June 1, 2022). “Migrants and Firms: Evidence from China”. In: *American Economic Review* 112.6, pp. 1885–1914.

Ivlevs, Artjoms, Milena Nikolova, and Carol Graham (2019). “Emigration, Remittances, and the Subjective Well-Being of Those Staying Behind”. In.

Kuhn, Randall, Bethany Everett, and Rachel Silvey (2011). “The Effects of Children’s Migration on Elderly Kin’s Health: A Counterfactual Approach”. In: *Demography* 48.1, pp. 183–209. JSTOR: 41237717.

Lahaie, Claudia et al. (Aug. 1, 2009). “Work and Family Divided across Borders: The Impact of Parental Migration on Mexican Children in Transnational Families”. In: *Community, Work & Family* 12.3, pp. 299–312.

Lu, Yao and Hao Zhou (Feb. 2013). "Academic Achievement and Loneliness of Migrant Children in China: School Segregation and Segmented Assimilation". In: *Comparative Education Review* 57.1, pp. 85–116.

McKenzie, David and Hillel Rapoport (Oct. 1, 2011). "Can Migration Reduce Educational Attainment? Evidence from Mexico". In: *Journal of Population Economics* 24.4, pp. 1331–1358.

Murard, Elie (2019). "The Impact of Migration on Family Left Behind: Estimation in Presence of Intra-Household Selection of Migrants". In: *SSRN Electronic Journal*.

Myerson, Rebecca (Apr. 5, 2017). *Policy Analysis with Endogenous Migration Decisions: The Case of Left-Behind Children in China*. URL: https://papers.ssrn.com/abstract=3058134 (visited on 03/20/2023). Pre-published.

Peking University, Institute of Social Science Survey (2015). *China Family Panel Studies (CFPS)*. Version V42.

Shrestha, Slesh A. (Mar. 2017). "No Man Left Behind: Effects of Emigration Prospects on Educational and Labour Outcomes of Non-migrants". In: *The Economic Journal* 127.600, pp. 495–521.

Tang, Zequn and Ning Wang (Oct. 1, 2021). "School Disruption of Children in China: The Influence of Parents' Rural–Urban Migration". In: *Children and Youth Services Review* 129, p. 106167.

Wang, Xianwen et al. (2014). "Tracing the Largest Seasonal Migration on Earth". In.

Zhang, Jipeng, Ru Wang, and Chong Lu (Mar. 2019). "A Quantitative Analysis of *Hukou* Reform in Chinese Cities: 2000-2016". In: *Growth and Change* 50.1, pp. 201–221.

Appendix: Additional Tables and Figures

Table A1: List of Child Development Indicators

Data	Description	Wave coverage	Age coverage
Enrollment	Binary: whether the child is currently enrolled in school?	All five waves (2010-2020)	[0,18)
Years of schooling	Integer: years of full-time schooling	All five waves (2010-2020)	[0,18)
Cognitive test, set A	(1) Mathematical problems (2) Vocabulary problems	2010, 2014, 2018	[10,18)
Cognitive test, set B	(1) Number series test (2) Word recall test	2012, 2016, 2020	[10,18)
Growth reference	(1) Weight-for-age z-score (2) Height-for-age z-score	All five waves (2010-2020)	[0,18)
Depression scale	CES-D, measures how often the respondent experienced symptoms associated with depression over the past week.	All five waves (2010-2020)	[10,18)

Table A2: Parameter Estimates, Gains and Costs Weighted by Distance

Migration Equation			Enrollment Equation		
Coefficient	Estimate	Std.Err.	Coefficient	Estimate	Std.Err.
η^m	0.6020	0.7397	$\eta^e(a=[6,11])$	10.3919	0.7220
μ_1^m	41.2758	11.4773	$\eta^e(a=[12,14])$	18.5636	0.9248
μ_2^m	-1.2650	0.7979	$\eta^e(a=[15,17])$	13.3221	0.5735
ψ^m (Agr.)	0.0009	0.0034	μ^e	40.6723	1.7433
ψ^m (Non-Agr.)	0.0055	0.0011	$\alpha^e \cdot \mathbb{1}(a=[12,14])$	-7.6597	1.0164
ϕ^m	-0.0448	0.0051	$\alpha^e \cdot \mathbb{1}(a=[15,17])$	-3.5856	0.8294
$\beta^m(a=[12,14], e=0)$	-0.2285	0.1279			
$\beta^m(a=[15,17], e=0)$	0.5591	0.0831			
$\beta^m(a=[6,11], e=1)$	1.6477	0.0615			
$\beta^m(a=[12,14], e=1)$	1.6722	0.0687			
$\beta^m(a=[15,17], e=1)$	1.9894	0.0732			

[a] Standard errors are clustered by household.

Table A3: Mother Fixed Effects Regressions on Lagged Outcomes and Child's *Current* Status × Age Dummies. Sample: 10 to 17 Years Olds.

	Dependent Variable:					
	Enrollment	Math Test	Word Test	Weight z	Height z	Depression
Lagged Enrollment Indicator	0.025**					
	(0.008)					
Lagged Math Test Score, [0,10]		-0.300***				
		(0.017)				
Lagged Vocabulary test score, [0,10]			-0.422***			
			(0.021)			
Lagged Weight-for-age z-score				-0.152***		
				(0.026)		
Lagged Height-for-age z-score					-0.030***	
					(0.009)	
Lagged Depression scale, [0,10]						-0.176***
						(0.023)
Child's gender (male = 1)	-0.016**	-0.053	-0.317**	0.232***	0.074	0.123
	(0.005)	(0.101)	(0.118)	(0.058)	(0.046)	(0.080)
Father completed 9-year compulsory edu	0.020*	0.541***	-0.293	-0.016	0.299***	0.433**
	(0.009)	(0.143)	(0.182)	(0.109)	(0.084)	(0.144)
Mother completed 9-year compulsory edu	0.024**	0.249	0.256	0.015	0.274**	0.183
	(0.009)	(0.164)	(0.174)	(0.095)	(0.084)	(0.132)
Father's employment, base: not employed						
- Blue-collar	0.006	-0.005	0.026	0.077	0.093	0.065
	(0.007)	(0.136)	(0.152)	(0.066)	(0.065)	(0.128)
- White-collar	-0.011	-0.118	0.295	0.074	0.019	-0.130
	(0.010)	(0.209)	(0.254)	(0.087)	(0.090)	(0.206)
Mother's employment, base: not employed						
- Blue-collar	-0.002	0.132	0.141	0.055	0.022	0.179
	(0.006)	(0.114)	(0.133)	(0.050)	(0.051)	(0.109)
- White-collar	-0.000	-0.030	-0.074	0.090	-0.130	-0.106
	(0.009)	(0.229)	(0.262)	(0.077)	(0.083)	(0.185)
Father's log-income	0.011	-0.073	0.378	0.293***	0.476***	0.767***
	(0.007)	(0.185)	(0.285)	(0.075)	(0.066)	(0.225)
Mother's log-income	-0.023	-0.192	1.459***	-0.135	-0.186*	-0.502
	(0.017)	(0.393)	(0.425)	(0.097)	(0.090)	(0.408)
Constant	0.616***	6.816***	8.227***	1.245	-2.967**	3.306*
	(0.101)	(0.696)	(0.896)	(1.881)	(1.138)	(1.409)
Province Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Age Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Age Dummies × Current Status	Yes	Yes	Yes	Yes	Yes	Yes
N	22984	6380	7038	27161	26531	5915
R-squared	0.500	0.683	0.586	0.371	0.503	0.664

a. Standard errors in parentheses. Robust standard errors.
b. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A4: Simultaneous Equations Model Results. Income and Expenses Weighted by Baseline Migration Stock

	(1)		(2)	
Migration Equation				
Migration indicator, t-1	2.184***	(0.042)		
Income diff.	0.512*	(0.234)	-0.210***	(0.063)
Housing expenses diff.	9.753**	(3.423)	14.712***	(1.966)
Hukou Index	-0.086	(1.334)	-4.311***	(0.941)
Child				
- 6-11 × Enrolled=1	1.846***	(0.187)	1.694***	(0.044)
- 12-14 × Enrolled=0	-0.065	(0.320)	-0.177	(0.098)
- 12-14 × Enrolled=1	1.949***	(0.189)	1.700***	(0.050)
- 15-17 × Enrolled=0	0.743***	(0.201)	0.519***	(0.066)
- 15-17 × Enrolled=1	2.093***	(0.197)	1.990***	(0.056)
Household Head				
Completes 9-year compulsory edu.	0.156***	(0.032)	0.209***	(0.032)
Age	-0.001	(0.002)	-0.000	(0.003)
Constant	-3.067***	(0.282)	-1.684***	(0.168)
Enrollment Equation				
Enrollment indicator, t-1	0.842***	(0.056)		
Child's Age × Expected Income				
- 6-11 × Expected income	14.704***	(4.387)	-0.485***	(0.108)
- 12-14 × Expected income	19.231**	(7.240)	3.778**	(1.269)
- 15-17 × Expected income	8.255***	(2.128)	2.592***	(0.402)
Expected edu. expenses	0.514**	(0.195)	1.324***	(0.196)
Child				
Gender	0.030	(0.038)	0.011	(0.030)
Age Group (base = 6-11)				
- 12-14	-0.382***	(0.072)	0.004	(0.057)
- 15-17	-1.222***	(0.062)	-1.009***	(0.039)
#Siblings	0.013	(0.035)	-0.058*	(0.024)
Household Head				
Completes 9-year compulsory edu.	0.201***	(0.042)	0.161***	(0.037)
Constant	0.970***	(0.065)	1.285***	(0.051)
N	14745		18618	
Log-likelihood	-6444.2		-13263.2	

a. Standard errors in parentheses, clustered by household. *b.* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
c. The units of income and expenses are 10k CNY (about 1500 USD in 2010).

Table A5: Simultaneous Equations Model Results.

	Income and Expenses Weighted by							
	Geographical Distance		Geographical Distance		Historical Migration		Historical Migration	
Migration Equation								
d_m_1	2.184***	(0.041)	2.184***	(0.042)				
income_diff_dist	0.616**	(0.237)			-0.173**	(0.061)		
income_diff_migsto			0.512*	(0.234)			-0.210***	(0.063)
mrent_diff_dist	3.366	(2.900)			9.538***	(1.704)		
mrent_diff_migsto			9.753**	(3.423)			14.712***	(1.966)
index_dist	-0.705	(0.695)			-3.933***	(0.658)		
index_migsto			-0.086	(1.334)			-4.311***	(0.941)
Child								
1.agegroup × 1.d_e	1.822***	(0.185)	1.846***	(0.187)	1.683***	(0.045)	1.694***	(0.044)
2.agegroup × 0.d_e	-0.095	(0.318)	-0.065	(0.320)	-0.188	(0.100)	-0.177	(0.098)
2.agegroup × 1.d_e	1.925***	(0.187)	1.949***	(0.189)	1.687***	(0.051)	1.700***	(0.050)
3.agegroup × 0.d_e	0.723***	(0.198)	0.743***	(0.201)	0.497***	(0.067)	0.519***	(0.066)
3.agegroup × 1.d_e	2.072***	(0.195)	2.093***	(0.197)	1.977***	(0.057)	1.990***	(0.056)
Household Head								
edu9_h	0.150***	(0.032)	0.156***	(0.032)	0.188***	(0.032)	0.209***	(0.032)
age_h	-0.001	(0.002)	-0.001	(0.002)	0.000	(0.003)	-0.000	(0.003)
Constant	-2.954***	(0.222)	-3.067***	(0.282)	-1.762***	(0.135)	-1.684***	(0.168)
Enrollment Equation								
d_e_1	0.844***	(0.056)	0.842***	(0.056)				
1.agegroup × c.income_wsum_dist	13.492**	(4.201)			-0.527***	(0.100)		
2.agegroup × c.income_wsum_dist	20.166**	(7.282)			3.988**	(1.333)		
3.agegroup × c.income_wsum_dist	7.718***	(1.988)			2.510***	(0.401)		
Child's Age × Expected Income								
1.agegroup × c.income_wsum_migsto			14.704***	(4.387)			-0.485***	(0.108)
2.agegroup × c.income_wsum_migsto			19.231**	(7.240)			3.778**	(1.269)
3.agegroup × c.income_wsum_migsto			8.255***	(2.128)			2.592***	(0.402)
eduexp_wsum_dist	0.577**	(0.188)			1.395***	(0.178)		
eduexp_wsum_migsto			0.514**	(0.195)			1.324***	(0.196)
Child								
gender	0.029	(0.038)	0.030	(0.038)	0.010	(0.030)	0.011	(0.030)
Age Group (base = 6-11)								
2.agegroup	-0.395***	(0.071)	-0.382***	(0.072)	0.002	(0.056)	0.004	(0.057)
3.agegroup	-1.232***	(0.061)	-1.222***	(0.062)	-1.005***	(0.039)	-1.009***	(0.039)
nsibling	0.012	(0.035)	0.013	(0.035)	-0.049*	(0.024)	-0.058*	(0.024)
Household Head								
edu9_h	0.204***	(0.042)	0.201***	(0.042)	0.171***	(0.036)	0.161***	(0.037)
Constant	0.966***	(0.064)	0.970***	(0.065)	1.274***	(0.049)	1.285***	(0.051)
N	14745		14745		18618		18618	
Log-likelihood	-6451.5		-6444.2		-13250.1		-13263.2	

a. Standard errors in parentheses, clustered by household. *b.* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
c. The units of income and expenses are 10k CNY (about 1500 USD in 2010).

Table A6: Effect of Removing the *Hukou* Restrictions ($\bar{I} \equiv 0$) , Gains and Costs Weighted by Distance

Probability	Estimate (actual)	Estimate ($\bar{I} \equiv 0$)	Relative Change ($\bar{I} \equiv 0$)
Prob.(m=1,e=1)	0.39 (0.0025)	0.43 (0.0024)	10.3%
Prob.(m=1,e=0)	0.03 (0.0003)	0.03 (0.0003)	0.0%
Prob.(m=0,e=1)	0.75 (0.0011)	0.71 (0.0012)	-5.3%
Prob.(m=0,e=0)	0.09 (0.0007)	0.09 (0.0007)	0.0%

[a] Mean coefficients; SE in parentheses.

Table A7: Effect of Tuition Reduction at Destination for Migrant Households ($\Delta\bar{c} = -1000$ CNY/Year)

	Age Group		
	6-11	12-14	15-17
$Prob.(Migration)$			
Actual	0.295 (0.004)	0.308 (0.007)	0.286 (0.007)
Predicted	0.311 (0.004)	0.327 (0.006)	0.338 (0.007)
$Prob.(Enrollment)$			
Actual	0.954 (0.002)	0.973 (0.002)	0.800 (0.006)
Predicted	0.972 (0.001)	0.968 (0.001)	0.853 (0.003)
Observations	10871	4573	4139

SE in parentheses.
Gains and costs weighted by geographical distance.
Currency is CNY in 2010, 1000 CNY is around 150 USD.