ECMA 31320 Final Project

Childhood Family Residential Mobility and Long-run Education & Labor Outcomes

Sifei Liu, Bhavya Pandey, Yicheng Zhang

May 26, 2023

Abstract

This study analyses the impact of residential mobility in childhood on an individual's long-run education and labour outcomes. We use detailed data from the Panel Study of Income Dynamics (PSID, 1968-2019) to locate 7712 children who ever moved between ages 6 to 18 years, their household characteristics, and educational and income outcomes later in life. We deploy an instrumental variables strategy with 2SLS regression as a tool, and control for sources of selection into moving, to ascertain the impact of mobility and age at the time of move on outcomes. We find that the negative effect of childhood moving on educational outcomes is most prominent for younger children, with children between the ages 6 to 8 particularly vulnerable. Whereas, the coefficient on treatment dummy for income outcomes gives a negative yet non-significant result.

1 Introduction

1.1 Motivation

Over the past decades, as economic prosperity has grown around the world, so have differences in income and wealth. The impacts of rising inequality, and efforts to decompose them in conjunction with social phenomena and welfare policies, have been studied in varied contexts. On the other hand, particularly in western societies, policies have been geared to ensure that individual outcomes depend most on merit and effort, rather than family background, geographical location, and other supporting factors. In general, these policy efforts have translated to ensuring equitable access to resources and provisions of public goods, encouraging growth conducive to local contexts, and enabling those at the bottom of the pyramid.

Human capital formation is an important function that is being studied in the context of individual income, to understand the dynamism of inequalities pertaining to wealth and income. And within the ambit of human capital, childhood is a phase in an individual's life that is particularly relevant to closely understand. This is the time where many skills and traits are developed by an individual, which then capitalize during later stages in life, and inform key outcomes such as educational attainment and performance, social and interpersonal relationships, as well as labour outcomes throughout life. This is supported by the theoretical idea which underscores the importance of developing skills early on in childhood, and their impact on augmenting skills later in life, as well as increasing returns to later skill investments.

Further, evidence (Heidrich, 2015) shows that neighbourhood quality is an important determinant for ensuring positive outcomes for children's outcomes in health and education. This brings us to the question of mobility in childhood, and whether moving to different neighbourhoods, regions, states, or locations can have an impact on children's long-run outcomes.

In this paper, we aim to investigate how childhood experience of family moving influences education attainment and labor outcome in early career and mid-career age. From a theoretical perspective, both positive and negative associations between childhood re-locations and various outcomes may be expected (Black and Devereux, 2011; Chetty and Hendren, 2015; Chetty et al., 2015; Chetty et al., 2014). Households make mobility decisions to optimize the future outcomes, which may include better labor outcomes for the head of the household or his spouse, improving the living conditions for the whole family, and providing better educational resources to the child(ren). However, the moving experience is also found to be potentially related to many drawbacks, most prominently:

> Network Destruction (Keene et al., 2010): When families move, important relationships, information sources and networks that could be disrupted. Families tend to be less resourceful and are more vulnerable to external shocks. All members of the family need to build new connections and familiarize themselves with the environment. The process could be effortful and time-consuming, posing challenges to the financial stability of the family.
>
> Destruction of "intra-family" social capital (Démurger, 2015): Parent–child relationships may be distorted and re-established and could negatively impact the child's personal development, educational attainment, and labor outcomes.

1.2 Related Literature

Our paper complements the strand of literature estimating the long run costs of childhood moving incurred by individuals in the long run. The first of these estimates have been showcased in Heidrich, 2015 and Heidrich, 2016 by leveraging granular Swedish Register Data. Heidrich has defined these costs as the "effect of childhood moving on long run adult income" and has put forth that not only do such costs exist, but that they increase in the number of times a child moves while growing up, and also that the timing of the move matters.

With regards to educational attainment, instances in literature have shown that frequent moves can lead to poor academic performance, difficulty in making social connections, and lower levels of educational attainment (Schmidt et al., 2017). Children who move frequently may experience disruptions in their education, such as having to adjust to new schools and curricula, and may struggle to keep up with their peers academically (Friedman et al., 2017).

Further, Tonnessen et al., 2016 studied childhood residential mobility and long-term outcomes and found that children with more residential moves are more likely to drop out of high school, to have a lower adult income, and to experience early parenthood. Most of these associations are weaker in the sibling fixed effects models explored by the group, and also that the age at moving matters. There is also evidence of heterogeneity among age groups in education and health outcomes.

Evidence from Swedish Register Data corroborates this observation, by leveraging a study of the long-term effects of inter-municipal moving during childhood on income. Heidrich, 2016 finds that children's long-run incomes are significantly negatively affected by moving during childhood, and the effect is larger for those who move more often. Controlling for important sources of selection such as parent separation, parents' unemployment, education, long run income, and immigration background, it is also evidenced that the negative effect of childhood moving on adult income is increasing in age at move. This study concludes that: "children benefit economically from the quality of the region they move to only if they move before age 12 (sons) and age 16 (daughters)".

Considering the dynamic nature of human capital development, disruption during childhood or adolescence could potentially accumulate to have long-term effects, which leaves us with some economically relevant questions. Based on the review of literature, we can expect that disruptions such as those coming from mobility, can fall in the same line as shocks in skill formation process while children are growing up, to have an accumulated (potentially negative) effect on outcomes in adulthood. Existing literature shows a negative correlation between residential mobility and educational outcomes of children. We would like to take one step further in investigating the causal impact of such an effect on educational attainment, as well as, long-term labor outcomes.

With this motivation and background of literature in mind, we propose the following hypotheses as the object of our study.

1.3 Hypotheses

Based on existing literature on the possible effect of moving, we form the following hypotheses:

- The disruption due to moving induces negative effects on long-run education and labor outcomes.
- The effect of disruption would be stronger for teenagers than younger children.

The first hypothesis suggests that when children experience a significant disruption in their lives due to moving, it will have a detrimental impact on their long-term educational and labor outcomes. The disruption caused by moving may lead to various challenges, such as changing schools, adjusting to a new environment, and building new social connections. These challenges can potentially disrupt academic progress and hinder the development of skills and knowledge necessary for successful educational and occupational trajectories. Hence, we expect to observe negative effects on education and labor outcomes as a result of moving.

The second hypothesis proposes that the negative impact of moving-induced disruption on education and labor outcomes will be more pronounced among teenagers compared to younger children. This is based on Tonnessen et al., 2016's work which shows that "...those who moved prior to elementary school do not have severe long-term outcomes compared with children who did not move at that age, whereas children who moved during teens did have more adverse outcomes than those who did not move at that age". It is hypothesised that older children tend to typically have a more established social and educational networks, and their educational trajectories and career aspirations may already be forming. Moving during this critical developmental stage may disrupt their connections, educational continuity, and the stability required for pursuing long-term goals. Consequently, we expect the effect of disruption to be stronger for teenagers than for younger children.

To address our research question, we estimate the average effects of experiencing family mobility at a young age on the educational attainment of the child and the long-run labor outcomes, controlling for the improvement in family resources. To explore any potential causal link between residential mobility and long-run outcome of children, we use instrument variables to estimate the average effects. We take into consideration age and education of parents, the availability of living space in households, the age of parents at the time of moving, and the frequency of moving for the head of the household. Section 2 of the paper details our methods and empirical strategy, followed by a description of the data in Section 3. Results for education and labor outcomes are detailed in Section 4.

2 Empirical Strategy

2.1 Baseline model and challenges for identification

The baseline model can be represented by the following OLS estimation equation:

$$Y_i = \beta D_i + \alpha X_i + \gamma_i + \epsilon_i \tag{1}$$

Here Y_i represents the child's long run outcome we are interested in, namely highest education attainment and average annual income between 25 and 35 years old. D_i is the treatment Dummy indicating whether the child has moved during age 6-18 years old. We add individual and household-level controls in X_i, including parents' education, family income level during 0-5 years old, and basic demographic information such as the age when the child moves, gender and race. Finally, as we may anticipate some regional and birth cohort variation in long-run outcome, we add γ_i, which are the year fixed effect and region/state fixed effect when the child was five years old.

2.1.1 Multicollinearity

The multicollinearity between control and target variable would give inaccurate estimated coefficients and its statistical power. To address this potential issue, we tested the Variance Inflation Factors (VIF) between our target variable D_i and controls X_i. The VIF of D_i (*has moved*) within the sample of investigating the labor outcome is 2.50, indicating that there is moderate correlation between the target variable and other controls. To see the magnitude of the collinearity problem, we plot the correlation matrix across all the right-hand side variables (see Appendix Figure 3), and all the correlations between the treatment dummy and controls remain small. In interpreting the results, we would assume such collinearity does not severely bias the results.

2.1.2 Omitted variable bias (OVB)

There could be variables influencing the long-run outcome of children that we are unable to fully control their effects. Only when such omitted variables are uncorrelated with our target and outcome variable, we could safely ignore the bias caused. However, it is possible that moving is led by certain endogenous motivations or reasons that would decide future development of the children. The direction of OVB is challenging to predict as we are identifying the effect in a fairly long time frame and these unaccounted motives could lead to mixed effects on moving decisions and children's long-run outcome. For instance, people may move to better accommodate the household production or consumption need, which could also imply their effort of building better future for the next generation. On the other hand, external force such as displacement, divorce, or health issue could force household to move, and the future development of children could be disrupted mainly by these events accompanying the relocation.

2.1.3 Measurement error

We worry about the existence of (classic) measurement error would lead to attenuation bias to our estimated coefficients. A measurement error could be manifested in many ways in this study, and even arise due to external biases such as social desirability or recall. If the measurement error is random, it can attenuate the estimated coefficients towards zero, making the outcome relationships appear weaker leading to a problem of underestimation. The direction and magnitude of the bias introduced by non-random measurement error depend on the specific nature of the measurement error and its relationship to the true values. If there is a systematic underreporting or misclassification of moving, it may lead to an underestimation of the true effect of moving on education and labor outcomes. Conversely, if there is a systematic overreporting or misclassification, it may lead to an overestimation of the effect.

2.2 Identification strategy and empirical model

Given the limitations of OLS in identifying causal relationship, we seek the use of an Instrumental Variables (IV) strategy to eliminate the influence of measurement error and omitted variable bias. The empirical model we choose to estimate is represented by the following two-stage system of equations:

$$D_i = \delta Z_i + \theta X_i + \gamma_i + u_i \quad (2)$$

$$Y_i = \beta D_i + \alpha X_i + \gamma_i + v_i \quad (3)$$

With constraints of the available data, we tried to identify some potential Instrumental Variables (IVs):

- Presence of new-born within two years of moving
- Room number/family member ratio
- Parent's age when given birth to the child
- Number of states or regions the head/reference person has lived in

Those factors are possible predictors of people making mobility decisions (when their children were at age 6-18), but are not directly related to the educational outcomes or labor market outcomes of the children when they grow up.

From the survey, the most popular reason for moving is the consumption need for housing concerns. For instance, the presence of a recent newborn and the room-per-family-member ratio tends to be correlated with the average space within the housing unit that each family member is enjoying. So the presence of a newborn or lower room-per-family-member ratio may be correlated with a higher possibility of moving. For parent's age, we expect that younger parents may be more likely to relocate, but parents' age may not have direct impact on children's future outcome. Younger parents may be less established in career, but may have more energy devoting to children's development. On the other hand, relatively older parents may make more mature parenting decisions and have better financial conditions.

For the last instrument, we conjecture that the experience of household heads moving across states may lead to the household being more flexible in residential relocation. Although high frequency of moving may be related to the profession of the household head, there seems to be no prominent concern of these variation of profession casting decisive influence on their children's future outcome.

Admittedly, each of these variables has flaws. In the following analysis we use mother's age when the child was five years old as the main instrument.

2.3 Assumptions for identification strategies

Here we recapitulate the assumptions for identifying the target variable regression estimates in baseline OLS and 2SLS models.

- **Ordinary Least Squares (OLS)**

  A multivariate linear regression model is given by:

  $$Y_i = \beta_0 + \beta_1 X_1 i + \beta_2 X_2 i + ... + \beta_k X_k i + u_i$$

  where $i = 1, 2, ..., n$.

  The Ordinary Least Squares (OLS) assumptions for linear regression are as follows:

  - A1: Regressors $(X_1 i, X_2 i, ..., X_k i)$ are drawn such that the assumption of independent and identical distribution $(i.i.d)$ holds.
  - A2: u_i represents the error term with conditional mean zero, given the regressors: $E(u_i | X_1 i, X_2 i, ..., X_k i) = 0$.
  - A3: There is unlikeliness in the occurrence of larger outliers ensuring that Y_i and $X_1 i, ..., X_k i$ have finite fourth moments.
  - A4: There is no incidence of perfect multicollinearity.

  As mentioned above, an emerging challenge is that of the Omitted Variable Bias (OVB) that arises when the regressor, X, is correlated with an omitted variable. This issue violates the second OLS assumption. For omitted variable bias to occur, two conditions are generally fulfilled:

  - X is correlated with the omitted variable.
  - The omitted variable is a determinant of the dependent variable Y.

  Thus we seek to use instrumental variable to correct the OVB problem assuming the following assumptions hold.

- **Instrumental Variables (IV)**

  Regression models may suffer from problems like omitted variables, measurement errors and simultaneous causality. We focus on the Two-Stage Least Squares (2SLS) tool to deploy an Instrumental Variables (IV) technique in this paper. To discuss the assumptions of this technique, we consider a simple linear regression model as follows:

  $$Y_i = \beta_0 + \beta_1 X_1 + u_i, i = 1, ..., n$$

where the error term is correlated with the regressor X_1 (X is endogenous) such that OLS is inconsistent for the true β_1. In its simplest form, an IV regression makes use of a single instrumental variable Z to obtain the value of a consistent estimator for β_1. In this case, Z must satisfy the following two conditions to be a valid instrument

- A5: Instrument relevance condition: X and its instrument Z must be correlated.
- A6: instrument exogeneity condition: Instrument Z must not be correlated with the error term u_i.

Further, an extension of the OLS assumptions also holds for the IV set up, for the method to delineate causality. Additionally, regarding A1, specific to our context, we assume the incidence of randomized treatment across age groups in our sample, but we allow for slackness of such randomization across regions. We apply clustered standard error to address the within-group correlation of error at regional level.

3 Data

The Panel Study of Income Dynamics (PSID) by the Survey Research Center, Institute for Social Research, University of Michigan, 2023 has a data structure that enables us to test our hypotheses. The study surveyed 4802 households in 1968 and ran follow-up surveys to these households and their derived households from 1969 to 2019. As of 2019, there are 41 rounds of surveys and a total of 82573 individuals recorded.

In this project, we directly requested all rounds of available surveys on both the individual level and household levels from the PSID website. We then identify the children who were between six to eighteen years old when their family relocate. In total, we locate 7712 children who ever moved in 6-18 from a sample of 11820 children. For all the children in our sample, we identify their parents with the help of household identifiers. The following list of variables was extracted from the survey.

- Outcome Variable
  - Educational Attainment: The children's educational attainment taken at the age of 27 (measured by years of schooling).
  - Income: All available income of the child from young adulthood (25 yrs) to mid-career age (35 yrs). The extracted list of incomes are then converted to $Income25-30p$, $Income25-35p$, $Income30-35p$ measuring the partial average of the child's income during the corresponding ages.

- Treatment Variable
  - *hasmove* = 1 if the household moved when the child was between the age 6 and 18.
  - *nmove*: Total number of family relocation when the child was between the age 6 and 18.
- Covariates
  - Child's race and gender, recorder with the dummy *white* and *male*.
  - Child's birth year and the *age* when moving
  - Both parents' available incomes when the child was between age 0 and 5. We measure family income for each year as the sum of parental incomes. Then we take the five-year partial average of all sensible family incomes (between 200 and 100k) as the control for pre-treatment family wealth.
  - Parents' years of schooling
  - Living state and region of the household head for all covered years. These two variables allow us to identify cross-state- and cross-region-relocation.
  - Reasons to move (0 to 4, dummies)
- Potential Instrumental Variables (IVs)
  - Parents' age when their child was five years old
  - Number of states the head/reference person has lived in
  - Rooms to number of family members ratio at child age of 5.

3.1 Limitations of Data

When wrangling the survey data, we find the following major limitations that could potentially hazard our regression estimation and causal inference. Firstly, compared to the data availability of educational data, there are much lesser years of data available for the measurement of income. As we include the income measurement in the regression estimation, the sample size typically shrinks to less than 1/5. For the income data that is provided, we remain doubtful about the validity of the measurement, especially when some individuals appear to have large fluctuations in some subsequent years. To try to remedy the issue, we filter out the income record with annual income below 200 dollars or above 100k. Secondly, we are oversampling younger children in our regression since we are recording the child the first time he/she experienced family mobility. This could be a potential issue when interpreting regression coefficients related to a child's age. Thirdly, we couldn't find a more solid measurement of pre-treatment controls of the family. Interesting options for such controls are, for instance, the living neighborhood of the household and the neighborhood's mean income, the number of schools in the living region, or micro-level measurements such as the child's test results before the treatment period. Some of these variables are potentially accessible with IRB approval.

4 Results

4.1 Education outcome

Table 1: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
has moved	6,720	0.649	0.477	0	1
edu	6,720	13.544	2.255	1	17
education (father)	6,720	13.020	2.610	1	17
education (mother)	6,720	13.161	2.286	3	17
male	6,720	0.519	0.500	0	1
white	6,720	0.640	0.480	0	1
age	6,720	8.339	3.214	6	18
family income	6,720	19,169.510	15,969.160	2,007.000	99,000.000
year at 5yrs old	6,720	1,981.779	8.252	1,968	2,009
state at 5yrs old	6,720	24.367	13.596	1	51

Table 2 presents our estimated average treatment effects of a child's moving experience on his / her educational attainment. We specify the outcome variable as the educational attainment of the child when he/she is 27 years old or above. *age* captures the age of the child when the family moved. We measure family income using the partial average of the sum of parents' income during the child's age 0 to 5.

As reported in column (1), when coarsely regressing on the treatment dummy *has moved* with the child's birth cohort and the living state at the age of five as fixed effects, the estimated coefficient is negative and significant. As we add demographic controls and controls on households' characteristics, the magnitude of the coefficient continues to shrink and lose a bit of its power but remains negative and significant. The regression results reported in column (3) indicate that a child's experience of family mobility during the years 6 to 18 is associated with a 0.167 decrease in the child's future educational attainment. It is also found that the child being male is associated with 0.57 fewer years of schooling. Moreover, family income and both parents' years of education are positively and significantly associated with the child's educational attainment.

To investigate the potential heterogeneous effects of moving experience, we start by interacting the treatment dummy with *male* and *age* respectively in columns (4) and (5). The results show that there are no significant heterogeneous effects on the sex of a child, but strong heterogeneous effects on children's moving age. Column (5) indicates that children of younger age are more vulnerable to the mobility experience as far as educational attainment is concerned.

Table 2: Effects on Education Attainment

	Dependent variable: Education Attainment				
	(1)	(2)	(3)	(4)	(5)
has moved	-0.487^{***}	-0.373^{***}	-0.167^{*}	-0.182^{*}	-1.118^{***}
	(0.056)	(0.088)	(0.084)	(0.103)	(0.185)
male		-0.551^{***}	-0.571^{***}	-0.589^{***}	-0.569^{***}
		(0.048)	(0.050)	(0.079)	(0.050)
white		0.525^{***}	-0.044	-0.044	-0.034
		(0.084)	(0.083)	(0.083)	(0.082)
log(family income)			0.691^{***}	0.692^{***}	0.659^{***}
			(0.050)	(0.050)	(0.050)
education (father)			0.144^{***}	0.144^{***}	0.144^{***}
			(0.014)	(0.014)	(0.015)
education (mother)			0.180^{***}	0.180^{***}	0.182^{***}
			(0.017)	(0.017)	(0.017)
has moved:male				0.028	
				(0.095)	
age					-0.147^{***}
					(0.019)
has moved:age					0.125^{***}
					(0.022)
Age dummies	No	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Observations	6,720	6,720	6,720	6,720	6,720
R^2	0.103	0.137	0.271	0.271	0.272
Adjusted R^2	0.092	0.125	0.260	0.260	0.262
Note:			*p<0.1; **p<0.05; ***p<0.01		

4.1.1 Heterogeneity across age groups

To see detailed breakdowns of the heterogeneous treatment effects, we add a series of age group dummies after dividing the age groups as 6-year-olds, 7 to 8-year-olds, 9 to 12 years-olds, 13 to 15-year-olds, and 16-18 year-olds. Table 3 reports the regression results with and without interacting age group dummies with the treatment dummy. The coefficient of the interactions for age group 6-year-olds and 7 to 8-year-olds are negative and statistically significant, suggesting that children of age 6-8 are particularly vulnerable.

Table 3: Effects on Education Attainment: Heterogeneity on Age Group

	Dependent variable: Education Attainment	
	(1)	(2)
has moved	−0.200**	0.606*
	(0.079)	(0.302)
age 6	0.508***	1.261***
	(0.128)	(0.261)
age 7-8	0.297**	0.875***
	(0.142)	(0.300)
age 9-12	0.331***	−0.276
	(0.120)	(0.350)
age 13-15	−0.114	−0.089
	(0.170)	(0.332)
has moved:age 6		−1.097***
		(0.307)
has moved:age 7-8		−0.745**
		(0.327)
has moved:age 9-12		0.498
		(0.380)
has moved:age 13-15		−0.087
		(0.380)
Baseline Controls	Yes	Yes
Birth Year FE	Yes	Yes
State FE	Yes	Yes
Observations	7,078	7,078
R^2	0.271	0.278
Adjusted R^2	0.262	0.268
Note:		*p<0.1; **p<0.05; ***p<0.01

4.1.2 IV identification

In Figure 1, we present a graphical representation of our IV approach using mother's age when the child was five years old. In the background of each graph is a histogram for the density of the mother's age. The mean of the mother age variable is 30.50 with a standard deviation of 5.32.

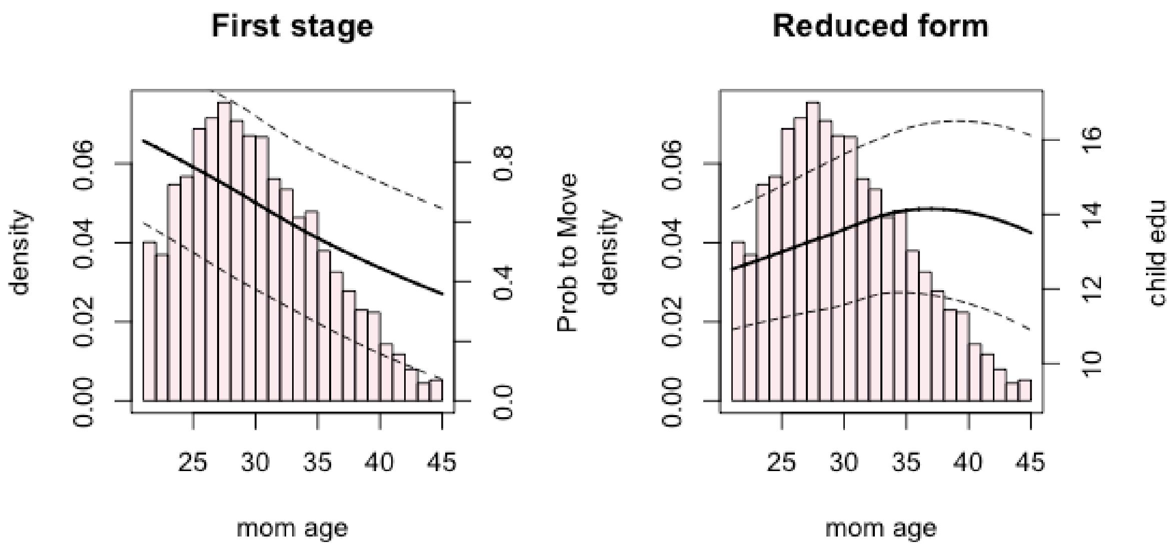

Figure 1: Effect of Mother's age on Probability to move (First Stage) and Children's Education Attainment (Reduced Form)

Plot on *First stage* shows the effect of younger mothers on the probability of relocation. The graph is a flexible analog to the first-stage equation, where we plot a local linear regression of actual moving status against the mother's age, with full control of demographic and household financial characteristics, and with birth year, living states fixed effects. The probability to move is monotonically declining in our mother's age measure and is close to linear. A 10 years decrease in the mother's age when carrying the child is associated with an approximately 19 percentage point increase in the probability the family moves.

Plot on *Reduced form* displays the reduced-form effect of a five-year-old child's mother's age against the child's future educational attainment, again using local linear regression with the full set of controls and fixed effects. The child's educational attainment is monotonically increasing in the mother's age for the first 80th percentile of the mother's age while slightly skew downwards. Dashed lines represent 90% confidence intervals.

The first column in Table 4 is a recall of our baseline model estimated on a slightly shrunk sample due to data availability of the variable $mother's\, age$. The second column reports first-stage estimates which regress a dummy variable for whether the family moved when the child was 6-18 with a full set of controls and fixed effects included. The coefficient implies that when a child's mother was one year younger, the probability of the child experiencing family mobility during 6-18 years old will increase by 1.4 percentage points. The third column reports the IV estimate, indicating that family mobility during a child's age 6-18 lowers the child's educational attainment by 1.44 years on average. The effect is both statistically and economically significant.

Table 4: Effects on Education Attainment: Mother's age as IV

	Dependent variable: Education Attainment		
	OLS	1st stage	2nd stage
has moved	-0.185^{**}		-1.441^{***}
	(0.084)		(0.398)
mother's age		-0.014^{***}	
		(0.001)	
male	-0.566^{***}	-0.010	-0.576^{***}
	(0.051)	(0.010)	(0.047)
white	-0.036	0.026	-0.001
	(0.088)	(0.022)	(0.081)
log(family income)	0.683^{***}	-0.082^{***}	0.532^{***}
	(0.052)	(0.011)	(0.071)
education (father)	0.151^{***}	0.003	0.157^{***}
	(0.014)	(0.004)	(0.016)
education (mother)	0.177^{***}	0.002	0.179^{***}
	(0.016)	(0.003)	(0.018)
Age dummies	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Observations	6,541	6,541	6,541
R^2	0.271	0.361	0.225
Adjusted R^2	0.260	0.351	0.213
F Statistic (df = 99; 6441)	24.167***	36.767***	22.783***
Note:		*p<0.1; **p<0.05; ***p<0.01	

4.2 Labor outcome

Extending on the education outcome, we conducted additional analysis in understanding how would moving experience may cast an influence on the labor outcome. The baseline empirical model of OLS and 2SLS remains the same as Equation (1) - (3), with the only difference of the outcome variable being the average taxable income of the children between 25 and 35 years old.

Table 5 describes the sample we used to investigate the effect of childhood residential mobility on long-run income. Note that due to the data deficiency in years between 1994 and 2003, we only have 1267 observations. The sampling is thus biased towards individuals in earlier birth cohort and we have more observations of income before 30 years than after 30 years old, which may further bias the average income level.

Table 5: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
income 25-35	1,267	18,427.14	13,813.72	130.00	115,000.00
has moved	1,267	0.62	0.48	0	1
education (child)	1,267	13.94	2.01	8	17
family income	1,267	16,218.79	15,596.73	411.00	122,437.50
education (father)	1,267	12.50	2.76	3	17
education (mother)	1,267	12.80	2.27	3	17
mother's age	1,267	31.71	5.83	21	48
male	1,267	0.54	0.50	0	1
white	1,267	0.61	0.49	0	1
age	1,267	7.93	2.93	6	18
year at 5yrs old	1,267	1,977.72	9.64	1,968	1,997
region at 5yrs old	1,267	2.52	0.95	1	4

Table 6 presents our estimated average treatment effects of a child's moving experience on long-run labor outcome, measured by his/her average annual income during 25-35 years old. All the controls are consistent with our analysis on education attainment, with the only difference being that we control for region fixed effect instead of state fixed effect due to the limited sample size. As reported in column (1), as a motivational setup, simply regressing on the treatment dummy *has moved* with the child's birth cohort fixed effect, the estimated co-efficient is negative and significant. Similar as the case with education outcome, as we add more controls and fixed effects, the magnitude and statistical power of the coefficient decreases but remains negative and significant.

The regression results reported in column (3) indicate that a child's experience of family mobility during the years 6 to 18 is associated with a 11.7% decrease in the child's future income during ages 25-35. It is also found that the child being male is correlated with 30% more average annual income.[1] Family income and both parents' of education attainment are positively and significantly correlated with the child's educational attainment. Following the analysis on education outcome, we conduct similar heterogeneous analysis of moving experience, by interacting the *has moved* with male and age respectively in columns (4) and (5), yet neither of them show significant results.

Table 6: Effects on Long-run Income

	Dependent variable: log(income 25-35)				
	(1)	(2)	(3)	(4)	(5)
has moved	−0.177***	−0.131**	−0.117*	−0.080	0.220
	(0.053)	(0.065)	(0.065)	(0.086)	(0.209)
age					0.043
					(0.029)
log(family income)		0.167***	0.160***	0.159***	0.161***
		(0.042)	(0.042)	(0.042)	(0.042)
education (mother)		0.048***	0.049***	0.049***	0.048***
		(0.014)	(0.014)	(0.014)	(0.014)
education (father)		0.040***	0.038***	0.038***	0.037***
		(0.011)	(0.011)	(0.011)	(0.012)
male		0.298***	0.298***	0.338***	
		(0.049)	(0.049)	(0.079)	
white		0.137**	0.117**	0.119**	0.125**
		(0.053)	(0.058)	(0.058)	(0.058)
has moved:male				−0.066	
				(0.101)	
has moved:age					−0.046
					(0.030)
Age dummies	No	Yes	Yes	Yes	No
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	Yes	Yes	Yes
Observations	1,267	1,267	1,267	1,267	1,267
R^2	0.036	0.155	0.160	0.160	0.132
Adjusted R^2	0.014	0.123	0.127	0.127	0.106
Note:			$^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Apart from the average annual income between 25 and 35 years old, we have conducted similar baseline analysis for the average annual income between 25 to 30 and 31 to 35 years old, respectively.

[1] Together with what we observed in the analysis on education outcome, where male tends to be associated with lower years of education, the data seem to reveal the persistent gender wage gap across generation.

The baseline results for other specifications of measuring income is displayed in Appendix Table A1.[2]

4.2.1 Heterogeneity across age groups

Similar as what we observed in the effects of mobility on education attainment, we find a strong divergence of effect for children moved before and after 12 years old. Interestingly, we observe an opposite effect across age groups from the results on education attainment. To see detailed breakdowns of the heterogeneous treatment effects, we decompose the 6 to 12 years-old group by add a series of age group dummies for 6-year-olds, 7 to 8-year-olds, and 9 to 12 years-olds, respectively. Table 7 reports the regression results with and without interacting age group dummies with the treatment dummy. The coefficient of the interactions for age group 6-year-olds and 7 to 8-year-olds are positive and statistically significant, suggesting that in terms of the future income, children of age 6-8 are less negatively impacted by moving.

4.2.2 IV identification

Figure 2 is again, the graphical representation of our IV approach. In the background of each graph is a histogram for the density of the mother's age at the time that the child was five years old. The mean of the mother age variable is 31.71 with a standard deviation of 5.83.

Plot on $First\,stage$ shows the effect of younger mothers on the probability of relocation. With full set of baseline controls, the local linear regression of actual moving status against mother's age can be viewed a representing the first-stage estimation equation (2). Dashed lines still represent 90% confidence intervals. The first stage within this sample is still monotonically declining in our mother's age measure yet with wider confidence interval.

Plot on $Reduced\,form$ displays the reduced-form effect of a five-year-old child's mother's age against the child's future income during 25 and 35 years old. Using same plotting strategy, the child's income level remains rather flat and we do not observe a significant increasing trend as in the sample for analysing education attainment.

The first column in Table 8 repeats our baseline OLS model and with a slight decrease in sample size constrained by data availability of mother's age. Column (2) reports the first stage of mother's

[2]The estimated coefficients for the treatment dummy remains negative across both age groups, yet compared with the 25 to 35 age group, the statistical power becomes weaker possibly due to the less observations in each group and higher measurement error.

Table 7: Effects on Long-run income: Heterogeneity on Age Group

	Dependent variable: log(income 25-35)			
	(1)	(2)	(3)	(4)
has moved	-0.077	-0.628^{**}	-0.107^{*}	-0.629^{**}
	(0.053)	(0.297)	(0.064)	(0.298)
age 6-12	0.011	-0.513^{*}		
	(0.083)	(0.290)		
has moved:age 6-12		0.571^{*}		
		(0.303)		
age 6			-0.020	-0.513^{*}
			(0.092)	(0.291)
age 7-8			-0.010	-0.529
			(0.093)	(0.323)
age 9-10			0.120	-0.525
			(0.112)	(0.675)
age 11-12			0.022	-0.392
			(0.118)	(0.680)
has moved:age 6				0.547^{*}
				(0.308)
has moved:age 7-8				0.560^{*}
				(0.336)
has moved:age 9-10				0.686
				(0.684)
has moved:age 11-12				0.449
				(0.689)
Controls	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	1,267	1,267	1,267	1,267
R^2	0.156	0.158	0.158	0.160
Adjusted R^2	0.131	0.132	0.130	0.130

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

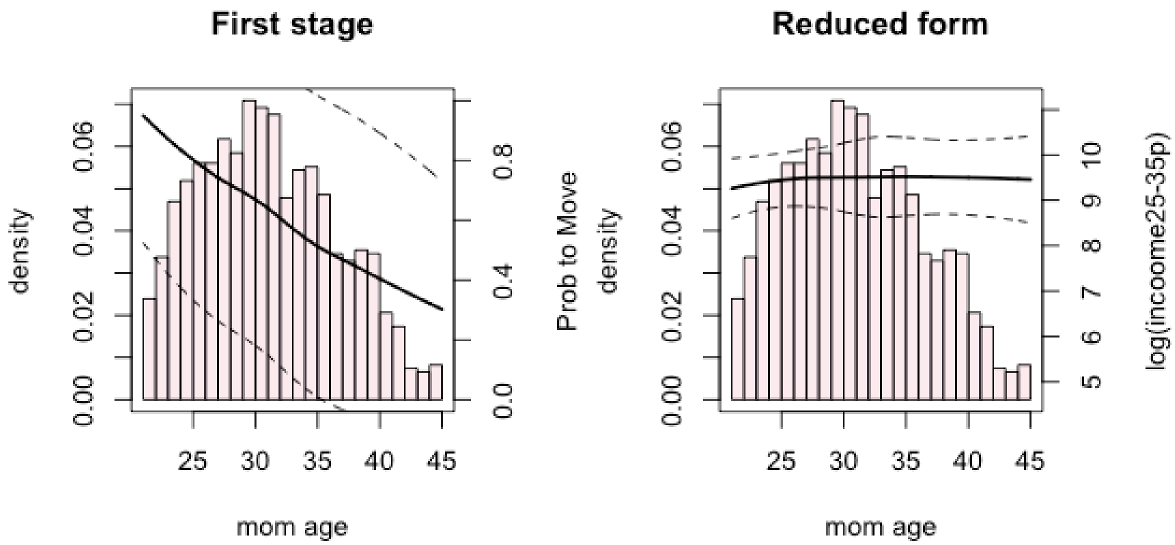

Figure 2: Effect of Mother's age on Probability to move (First Stage) and Children's Long-run income (Reduced Form)

age as the instrument. It can be shown that younger mother is associated with higher probability of moving. The coefficient implies that when a child's mother was one year younger, the probability of the child experiencing family mobility during 6-18 years old will increase by 1.5 percentage points. Column (3) in Table 8 report the 2SLS estimates using mother's age as the instrument. The estimated coefficient gives a positive sign with high standard error. To account for potential non-linearity of the treatment of the instrument, we split mother's age into small subgroups and found that the treatment effect is only prominent among age 23-30 and age ≥ 35. Among them, younger age group positively correlated with probability of moving and older age group has a negative correlation with moving, which aligns with our expectation. We then ran another 2SLS estimation using mother's age between 23-30 as a dummy. As shown in column (5), the estimated coefficient on treatment dummy gives a negative yet non-significant result. Just by observing the magnitude of the estimates, we also find 2SLS estimates to be larger than the OLS baseline result. This is consistent with our analysis for education outcome, but the inflating effect of bias correction is less strong.

We also conducted similar 2SLS analysis using number of states or regions the household head has lived in when the child was five years old. Both gives not so strong first stage and for detailed results, see Appendix Section B.

Table 8: Effects of Mobility on Long-run Income: Mother's age as IV

	Dependent variable: log(income 25-35)				
	OLS	1st stage	2nd stage	1st stage	2nd stage
	(1)	(2)	(3)	(4)	(5)
has moved	-0.117^{*}		0.108		-0.138
	(0.065)		(0.305)		(0.340)
mother's age		-0.015^{***}			
		(0.002)			
mother's age within 23-30				0.150^{***}	
				(0.022)	
log(family income)	0.160^{***}	-0.084^{***}	0.188^{***}	-0.095^{***}	0.158^{***}
	(0.042)	(0.018)	(0.056)	(0.018)	(0.059)
education (mother)	0.049^{***}	0.002	0.048^{***}	0.001	0.049^{***}
	(0.014)	(0.006)	(0.014)	(0.006)	(0.014)
education (father)	0.038^{***}	-0.007	0.038^{***}	-0.004	0.038^{***}
	(0.011)	(0.005)	(0.012)	(0.005)	(0.011)
male	0.298^{***}	0.013	0.293^{***}	0.012	0.298^{***}
	(0.049)	(0.021)	(0.049)	(0.021)	(0.049)
white	0.117^{**}	0.021	0.113^{*}	0.015	0.118^{**}
	(0.058)	(0.025)	(0.058)	(0.025)	(0.058)
Age dummies	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	1,267	1,267	1,267	1,267	1,267
R^2	0.160	0.453	0.152	0.447	0.160
Adjusted R^2	0.127	0.431	0.118	0.426	0.127
F Statistic (df = 48; 1218)	4.836^{***}	20.982^{***}	4.725^{***}	20.537^{***}	4.772^{***}
Note:					$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

4.3 Alternative specifications and Auxiliary Regressions

4.3.1 Number of times moved

In Table 9 we identify the number of moves the child experienced during the ages 6 to 18 and denote it as $move\,times$. Column (2) and (3) reports the regression results on $move\,times$ with and without a square term. The estimation of $move\,times$ is negative and significant, indicating that one additional moving experience is associated with a 0.087 decrease in years of schooling. The estimation with the square of the number-of-move is positive but not significant. In column (4), we add four dummies to explore the treatment effects of different numbers of moves. The variable $move1$ indicates the child has one moving experience in total, and the variable $move4+$ indicates the child has more than three moving experiences. The regression results suggest that children living in families with frequent moving have an average of fewer years of schooling. For example, children that experience three times moving attain 0.366 lesser years of schooling than those that never move. Moreover, moving only once seems to be associated with slightly more years of schooling but the estimation is not significant.

We then conduct parallel analysis for labour outcome. In Table 10 we report the regression results using $move\,times$ as the target variable. Baseline OLS results are in Column (1) to (3). The estimation of $move\,times$ is negative and significant, indicating that one additional moving experience is associated with a 0.027 decrease in income. Column (4) aims to investigate second order relationship between moving times and long-run income. The estimation with the square of the number-of-move is positive but with very small magnitude and not significant. In column (5), we decompose the number of moving times and explore the treatment effects of different numbers of moves. The regression results suggest that children living in families with frequent moving have an average of lower future income. For example, children that experience moving twice is 0.189 lesser income than those who have never moved. However, such effect is more noisy for children have moved three times.

Table 11 repeats the 2SLS analysis on labour outcome as in using the dummy treatment variable $hasmoved$. The negative relationship persist but the magnitude of the 2SLS estimates become slightly below the OLS estimate. We also note that, the F-statistics of the first stage decreases significantly, making the results less reliable.

Table 9: Effects of Number of times moved on Education Attainment

	Dependent variable: Education Attainment			
	(1)	(2)	(3)	(4)
has moved	-0.167^{*}			
	(0.084)			
move times		-0.087^{***}	-0.094^{**}	
		(0.018)	(0.041)	
move times squared			0.001	
			(0.005)	
move1				0.160
				(0.099)
move2				-0.253^{**}
				(0.107)
move3				-0.366^{***}
				(0.107)
move4+				-0.357^{***}
				(0.107)
male	-0.571^{***}	-0.571^{***}	-0.571^{***}	-0.569^{***}
	(0.050)	(0.048)	(0.048)	(0.049)
white	-0.044	-0.030	-0.030	-0.035
	(0.083)	(0.082)	(0.082)	(0.082)
log(family income)	0.691^{***}	0.645^{***}	0.644^{***}	0.642^{***}
	(0.050)	(0.049)	(0.050)	(0.051)
education (father)	0.144^{***}	0.144^{***}	0.144^{***}	0.145^{***}
	(0.014)	(0.014)	(0.014)	(0.014)
education (mother)	0.180^{***}	0.181^{***}	0.181^{***}	0.181^{***}
	(0.017)	(0.017)	(0.017)	(0.017)
Age dummies	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	6,720	6,720	6,720	6,720
R^2	0.271	0.275	0.275	0.277
Adjusted R^2	0.260	0.264	0.264	0.266
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 10: Effects of Number of times moved on Long-run Income

	Dependent variable: log(income 25-35)				
	(1)	(2)	(3)	(4)	(5)
move times	-0.047^{***}	-0.028^{**}	-0.027^{*}	-0.050	
	(0.013)	(0.014)	(0.014)	(0.036)	
move times squared				0.004	
				(0.005)	
move1					-0.032
					(0.085)
move2					-0.189^{**}
					(0.088)
move3					-0.099
					(0.094)
move4+					-0.142^{*}
					(0.083)
log(family income)		0.160^{***}	0.152^{***}	0.150^{***}	0.151^{***}
		(0.042)	(0.042)	(0.043)	(0.043)
education (mother)		0.049^{***}	0.049^{***}	0.049^{***}	0.049^{***}
		(0.014)	(0.014)	(0.014)	(0.014)
education (father)		0.040^{***}	0.038^{***}	0.038^{***}	0.038^{***}
		(0.011)	(0.011)	(0.011)	(0.012)
male		0.298^{***}	0.299^{***}	0.299^{***}	0.298^{***}
		(0.049)	(0.049)	(0.049)	(0.049)
white		0.140^{***}	0.120^{**}	0.119^{**}	0.117^{**}
		(0.053)	(0.058)	(0.058)	(0.058)
Age dummies	No	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	Yes	Yes	Yes
Observations	1,267	1,267	1,267	1,267	1,267
R^2	0.037	0.155	0.160	0.161	0.163
Adjusted R^2	0.015	0.124	0.127	0.127	0.128

Note: *p<0.1; **p<0.05; ***p<0.01

Table 11: Effects of Mobility on Long-run Income: Mother's age as IV

	Dependent variable: log(income 25-35)				
	OLS	1st stage	2nd stage	1st stage	2nd stage
	(1)	(2)	(3)	(4)	(5)
move times	-0.027^{*}		0.020		-0.025
	(0.014)		(0.058)		(0.061)
mother's age		-0.077^{***}			
		(0.009)			
mother's age_23-30				0.839^{***}	
				(0.104)	
log(family income)	0.152^{***}	-0.634^{***}	0.192^{***}	-0.681^{***}	0.154^{**}
	(0.042)	(0.087)	(0.064)	(0.086)	(0.065)
education (mother)	0.049^{***}	0.036	0.048^{***}	0.030	0.049^{***}
	(0.014)	(0.028)	(0.014)	(0.028)	(0.014)
education (father)	0.038^{***}	-0.021	0.037^{***}	-0.007	0.038^{***}
	(0.011)	(0.024)	(0.012)	(0.024)	(0.012)
male	0.299^{***}	0.080	0.293^{***}	0.074	0.298^{***}
	(0.049)	(0.100)	(0.050)	(0.100)	(0.049)
white	0.120^{**}	0.202^{*}	0.111^{*}	0.167	0.120^{**}
	(0.058)	(0.119)	(0.059)	(0.119)	(0.059)
Age dummies	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	1,267	1,267	1,267	1,267	1,267
R^2	0.160	0.234	0.152	0.229	0.160
Adjusted R^2	0.127	0.203	0.119	0.199	0.127
F Statistic (df = 48; 1218)	4.85***	7.734***	4.727***	7.544***	4.774***
Note:					*p<0.1; **p<0.05; ***p<0.01

4.3.2 Move across states or regions

Table 12 and 13 reports the results when regressing education and income respectively on *has moved* from five different samples with the same control group (households that never move) but different treatment groups as in the full sample, the sample of households that move across states, the sample of households that move within states, the sample of households that move across regions and a sample of households that move within regions.

Table 12: Effects on Education outcome: Subgroup of whether moved across states/regions

	Dependent variable: Education Attainment				
	(1)	(2)	(3)	(4)	(5)
mov	−0.203**	0.141	−0.225**	0.211**	−0.216**
	(0.079)	(0.089)	(0.089)	(0.096)	(0.088)
sex	−0.549***	−0.531***	−0.571***	−0.559***	−0.556***
	(0.052)	(0.066)	(0.064)	(0.064)	(0.062)
race	0.0005	0.160*	−0.048	0.163	−0.034
	(0.076)	(0.093)	(0.080)	(0.107)	(0.080)
log(fam_income)	0.474***	0.432***	0.478***	0.450***	0.464***
	(0.032)	(0.056)	(0.037)	(0.052)	(0.039)
dad_edu	0.155***	0.164***	0.140***	0.158***	0.145***
	(0.016)	(0.023)	(0.017)	(0.023)	(0.018)
mom_edu	0.185***	0.187***	0.190***	0.180***	0.193***
	(0.017)	(0.022)	(0.018)	(0.024)	(0.018)
Type of move	Full sample	Across state	Within state	Across region	Within region
Age dummies	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Observations	7,078	3,487	6,053	3,261	6,279
R^2	0.273	0.300	0.275	0.311	0.272
Adjusted R^2	0.263	0.280	0.263	0.289	0.260
Note:				*p<0.1; **p<0.05; ***p<0.01	

Table 12 shows a distinctive difference in the treatment effects between the across-states vs within-states, and the across-regions vs within-regions, demonstrating that long-distance relocation is linked to higher educational attainment. For instance, Columns (4) and (5) indicate that children whose families relocated across regions likely have 0.42 more years of education than their peers whose families relocated within areas. This finding is particularly intriguing as it may suggest different associated mechanisms of family mobility on children's educational attainment. Moreover, these results seem to be countering the theoretically arguments pointing the negative effects of network destruction caused by relocation. Further research into the relationship between relocating distance and family wealth would be worth exploring.

However, as seen in Table 13, this distinction is less pronounced when considering labor outcome. As we noted in the discussion, the estimation's precision is compromised by the omitted variable bias as well as our noisy measurement of labor outcomes.

Table 13: Effects on Labor outcome: Subgroup of whether moved across states/regions

	Dependent variable: log(income 25-35)				
	(1)	(2)	(3)	(4)	(5)
has moved	-0.117^{*}	-0.086	-0.118^{*}	-0.118	-0.109
	(0.065)	(0.117)	(0.070)	(0.136)	(0.069)
log(family income)	0.160^{***}	0.188^{***}	0.165^{***}	0.199^{***}	0.157^{***}
	(0.042)	(0.059)	(0.048)	(0.065)	(0.046)
education (mother)	0.049^{***}	0.057^{***}	0.048^{***}	0.058^{***}	0.048^{***}
	(0.014)	(0.020)	(0.015)	(0.021)	(0.014)
education (father)	0.038^{***}	0.021	0.035^{***}	0.014	0.039^{***}
	(0.011)	(0.016)	(0.013)	(0.017)	(0.012)
male	0.298^{***}	0.214^{***}	0.350^{***}	0.242^{***}	0.336^{***}
	(0.049)	(0.067)	(0.053)	(0.071)	(0.052)
white	0.117^{**}	0.042	0.148^{**}	0.088	0.132^{**}
	(0.058)	(0.086)	(0.064)	(0.091)	(0.062)
Type of move	Full sample	Across state	Within state	Across region	Within region
Age dummies	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	1,267	624	1,120	589	1,155
R^2	0.160	0.193	0.168	0.193	0.168
Adjusted R^2	0.127	0.127	0.131	0.123	0.132
Note:					*p<0.1; **p<0.05; ***p<0.01

4.3.3 Reasons to move

Table 14 and 15 shows the baseline OLS regression results using different subgroups move for different reasons. Column (1) in both tables show a positive but non-significant association between moving and children outcome. As for moving for consumption reason, for education outcome, we find that there is a significant and negative relationship between moving and education attainment, which seems counter-intuitive. Similar negative results can be seen in labour outcome as well, only that is less pronounced in statistical power. Interestingly, moving due to outside force seems to be especially detrimental for long-run income but not necessarily for education outcome. Although we lack enough statistical power to make firm conclusion on such relationships, it is a starting point and motivates further study on different mechanism of moving and their long-term effects.

Table 14: Effects on Education outcome: Subgroup based on moving reason

	Dependent variable: educational attainment			
	(1)	(2)	(3)	(4)
has moved	0.162	-0.160^{*}	0.054	-0.091
	(0.109)	(0.082)	(0.192)	(0.106)
log(family income)	0.396^{***}	0.484^{***}	0.398^{***}	0.378^{***}
	(0.053)	(0.048)	(0.062)	(0.059)
education (father)	0.163^{***}	0.142^{***}	0.147^{***}	0.145^{***}
	(0.021)	(0.020)	(0.027)	(0.025)
education (mother)	0.197^{***}	0.178^{***}	0.196^{***}	0.205^{***}
	(0.025)	(0.021)	(0.022)	(0.024)
male	-0.561^{***}	-0.580^{***}	-0.582^{***}	-0.586^{***}
	(0.077)	(0.064)	(0.078)	(0.078)
white	0.053	-0.026	0.075	0.127
	(0.119)	(0.097)	(0.128)	(0.099)
Reason for moving	Production	Consumption	Outside force	Other
Age dummies	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	3,163	4,918	2,766	3,008
R^2	0.300	0.274	0.307	0.316
Adjusted R^2	0.278	0.260	0.281	0.292
Residual Std. Error	1.916	1.968	1.941	1.914
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 15: Effects on Labor outcome: Subgroup based on moving reasons

	Dependent variable: log(income 25-35)			
	(1)	(2)	(3)	(4)
has moved	0.107	−0.101	−0.380***	−0.090
	(0.108)	(0.078)	(0.142)	(0.145)
log(family income)	0.206***	0.185***	0.174**	0.150**
	(0.066)	(0.049)	(0.075)	(0.073)
education (mother)	0.058***	0.042***	0.063***	0.063***
	(0.020)	(0.016)	(0.022)	(0.021)
education (father)	0.024	0.028**	0.020	0.011
	(0.016)	(0.013)	(0.018)	(0.018)
male	0.327***	0.283***	0.325***	0.288***
	(0.068)	(0.056)	(0.078)	(0.073)
white	0.132	0.097	0.094	0.106
	(0.088)	(0.068)	(0.099)	(0.096)
Reason for moving	Production	Consumption	Outside force	Other
Age dummies	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	622	960	557	554
R^2	0.219	0.160	0.246	0.201
Adjusted R^2	0.155	0.116	0.176	0.127
Note:			*p<0.1; **p<0.05; ***p<0.01	

5 Discussion

5.1 Comparison between IV and OLS results

For both set of analysis on the education and labour outcome, the OLS regression coefficients is smaller than the 2SLS estimates. Such difference is especially pronounced for education attainment, with OLS estimates accounting for around 1/8 of the 2SLS estimates. The implication of the difference can be explained as follows:

- Omitted variable bias (OVB): Using instrument help correct the OVB of the empirical model. A more significantly negative result imply that there is a large positive bias caused by omitted variables. This is likely to happen if the motivation of moving is mainly related to improving living and production quality which further contributes to children's future outcome.
- Measurement error: As IV corrects the attenuation bias of the measurement error of the treatment variable, it is also likely to have a 2SLS estimate with larger magnitude.

5.1.1 Interpretation of linear IV estimand

In the regression results of education outcome, as we use a linear IV, it is difficult to directly interpret the results as the local average treatment effect (LATE) for compliers. Following Dahl et al., 2014's discussion of comparing the LATE and OLS interpretation, the linear IV estimand can be decomposed to evenly stepped subgroups and collectively enter a over-identified model estimated using 2SLS regression. Theoretically, the 2SLS regression should produce the same results as the just identified model. Table 16 and Table 18 display the first stage of our decomposed IV's. For education outcome analysis, we then use the three subgroups, each as a separate dummy instrument, to re-estimate the model. As shown in Table 17, the coefficients remain negative but lose some statistical power. This reveals that our IV results may still be valid for investigating the effects for education outcome. For labor outcome, it can be seen that although used as a linear IV, the effect of mother's age on moving is more like a inversed U-shape for the subsample investigating the effect on, possibly leading to the insignificant result in Table 8 column (3). Hence, we conducted further analysis in Table 8 column (4)-(5) as mentioned in previous sections to estimated the empirical model and we gain consistent results.

Table 16: Sample for education attainment - First stage: Discretized Mother's age

	Dependent variable: has moved		
	(1)	(2)	(3)
mother's age group	0.127***	0.034**	−0.039***
	(0.014)	(0.014)	(0.013)
log(family income)	−0.043***	−0.056***	−0.055***
	(0.007)	(0.007)	(0.008)
education (father)	0.002	0.003	0.003
	(0.004)	(0.004)	(0.004)
education (mother)	−0.001	−0.001	−0.001
	(0.003)	(0.003)	(0.003)
male	−0.005	−0.006	−0.005
	(0.009)	(0.010)	(0.009)
white	0.013	0.009	0.010
	(0.021)	(0.022)	(0.022)
Mom's age group	20-25	26-30	31-35
Age dummies	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Observations	6,879	6,879	6,879
R^2	0.340	0.332	0.332
Adjusted R^2	0.331	0.322	0.322
Residual Std. Error	0.389	0.391	0.391
F Statistic (df = 99; 6779)	35.351***	34.045***	34.061***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 17: Effects on education attainment: Mother's age (discrete) as IV

	Dependent variable: educational attainment					
	1st stage	*2nd stage*	*1st stage*	*2nd stage*	*1st stage*	*2nd stage*
has moved		-1.356^{*}		-2.693		-2.344
		(0.723)		(1.706)		(1.864)
mother's age_20-25	0.127^{***}					
	(0.014)					
mother's age_26-30			0.034^{**}			
			(0.014)			
mother's age_31-35					-0.039^{***}	
					(0.013)	
log(family income)	-0.043^{***}	0.400^{***}	-0.056^{***}	0.323^{***}	-0.055^{***}	0.343^{***}
	(0.007)	(0.060)	(0.007)	(0.090)	(0.008)	(0.122)
education (father)	0.002	0.165^{***}	0.003	0.168^{***}	0.003	0.168^{***}
	(0.004)	(0.018)	(0.004)	(0.020)	(0.004)	(0.022)
education (mother)	-0.001	0.181^{***}	-0.001	0.179^{***}	-0.001	0.180^{***}
	(0.003)	(0.017)	(0.003)	(0.020)	(0.003)	(0.020)
male	-0.005	-0.550^{***}	-0.006	-0.557^{***}	-0.005	-0.555^{***}
	(0.009)	(0.050)	(0.010)	(0.050)	(0.009)	(0.050)
white	0.013	0.024	0.009	0.038	0.010	0.034
	(0.021)	(0.079)	(0.022)	(0.071)	(0.022)	(0.082)
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,879	6,879	6,879	6,879	6,879	6,879
R^2	0.340	0.234	0.332	0.091	0.332	0.138
Adjusted R^2	0.331	0.222	0.322	0.078	0.322	0.126
Note:				*p<0.1; **p<0.05; ***p<0.01		

Table 18: Sample for Average Income - First stage: Discretized Mother's age

	Dependent variable: has moved				
	(1)	(2)	(3)	(4)	(5)
mother's age group	−0.009	0.170***	0.053**	-0.009	− 0.162***
	(0.070)	(0.028)	(0.025)	(0.026)	(0.024)
log(family income)	−0.124***	−0.101***	−0.120***	−0.123***	−0.097***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
education (father)	−0.0001	−0.003	−0.001	−0.0001	−0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
education (mother)	0.001	0.002	0.001	0.001	0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
male	0.020	0.015	0.019	0.021	0.011
	(0.022)	(0.021)	(0.022)	(0.022)	(0.021)
white	0.019	0.013	0.020	0.019	0.019
	(0.026)	(0.025)	(0.025)	(0.026)	(0.025)
Mom's age group	20-22	23-26	27-30	31-34	35+
Age dummies	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	1,267	1,267	1,267	1,267	1,267
R^2	0.426	0.443	0.428	0.426	0.448
Adjusted R^2	0.404	0.421	0.406	0.404	0.426
F Statistic (df = 48; 1218)	18.861***	20.21***	19.025***	18.865***	20.586***
Note:					*p<0.1; **p<0.05; ***p<0.01

5.2 Clustered standard error

The change of clustering would change the estimated standard error (SE) of regression coefficients. Across the analysis for education attainment as the outcome variable, we cluster the SE at the state level, which gives 50 clusters, fulfilling the rule of thumb of at least 30 - 50 clusters. On the other hand, for the sample investigating the effects on labor outcome, due the smaller sample size, we only control for the region fixed effects along with birth year fixed effects and moving age dummies. Due to the fact that there are only four clusters, we deem the results being too conservative. We also try cluster at birth year (cohort) level, which gives 28 clusters. However, such clustering level increase the standard errors significantly. Table 19 compares the change of estimated standard errors across different clustering levels. Note that the increase in clusters does not bring a linear change in the maginitude of standard errors. Having difficulties in understanding the reason behind this, in above sections, we only reported the regression results with standard errors without clustered adjustment. We suspect that the small sample size leads to the fairly noisy estimates. Although the negativity of the effect seems consistent across OLS and 2SLS analysis and under different specifications, we would have reservations in claiming these results robust. It would be of our future steps to look into the reasons and if possible, seek other panel data to repeat the analysis.

6 Conclusion

Our crude regression results for education outcomes find that child's experience with family mobility between the ages of 6 and 18 is connected with a 1.67 year decrease in the child's future educational achievement. We also find that children between the ages 6 to 8 are particularly vulnerable to poor educational outcomes as a result of ill-effects of moving. This effect is corroborated by our IV analysis, where we conclude that when a child's mother was one year younger, the probability of the child experiencing family mobility during 6-18 years old will increase by 1.4 percentage points, furthering that family mobility during a child's age 6-18 lowers the child's educational attainment by 1.44 years on average.

Similarly, the regression results for labor outcomes indicate that a child's experience of family mobility during the years 6 to 18 is associated with a 11.7% decrease in the child's income during ages 25 to 35 years. However, interestingly we find an opposite effect when considering heterogeneity of age groups with regards to labour outcomes. The coefficient of the interactions for age group 6-year-olds and 7 to 8-year-olds are positive and statistically significant, suggesting that in terms of the future income, children of age 6-8 are less negatively impacted by moving. Upon undertaking an IV analysis, the coefficient implies that when a child's mother was one year younger, the probability of the child experiencing family mobility during 6-18 years old will increase by 1.5 percentage points.

Table 19: Effects on Long-run Income: Difference in Clustered SE

	Dependent variable: log(income 25-35)			
	(1)	(2)	(3)	(4)
has moved	-0.117^{*}	-0.117^{*}	-0.117^{***}	-0.117
	(0.065)	(0.048)	(0.033)	(0.080)
log(family income)	0.160^{***}	0.160^{**}	0.160^{***}	0.160^{***}
	(0.042)	(0.031)	(0.034)	(0.036)
education (mother)	0.049^{***}	0.049	0.049^{***}	0.049^{***}
	(0.014)	(0.022)	(0.011)	(0.013)
education (father)	0.038^{***}	0.038^{*}	0.038^{**}	0.038^{***}
	(0.011)	(0.012)	(0.015)	(0.008)
male	0.298^{***}	0.298^{***}	0.298^{***}	0.298^{***}
	(0.049)	(0.032)	(0.056)	(0.061)
white	0.117^{**}	0.117	0.117^{***}	0.117^{**}
	(0.058)	(0.096)	(0.030)	(0.054)
Clustering level	None	Region	Age	Birth year
Clusters	None	4	13	28
Age dummies	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	1,267	1,267	1,267	1,267
R^2	0.160	0.160	0.160	0.160
Adjusted R^2	0.127	0.127	0.127	0.127

Note: *p<0.1; **p<0.05; ***p<0.01

The coefficient on treatment dummy for income outcomes gives a negative yet non-significant result.

With respect to education outcomes, our results echo our first hypothesis that disruption due to moving induces negative effects on long-run outcomes and refutes the second one, that the effect of disruption would be stronger for teenagers than younger children. Instead, we observe that moving at younger age is more disruptive than moving in older years. With regards to long-run labor outcomes, younger ages are less negatively impacted by moving and the negative effect impact of moving is muted for younger children.

As next steps to further this study, we plan to control for parents' divorce and account for marital status of the head of the household in our analysis, to address the effect of separation as a reason for moving to further understand the mechanism behind mobility-induced long-run effects. We also plan to further granularize the level of investigation by splitting up the subgroup for moving age to isolate the impacts.

References

Black, S. E., & Devereux, P. J. (2011). Recent developments in intergenerational mobility. In *Handbook of labor economics* (pp. 1487–1541). Elsevier.

Chetty, R., & Hendren, N. (2015). *The impacts of neighborhoods on intergenerational mobility: Childhood exposure effects and county-level estimates* (tech. rep.). Working paper.

Chetty, R., Hendren, N., & Katz, L. F. (2015). *The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment* (tech. rep. No. 21156). NBER Working Paper.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, *129*(4), i1.

Dahl, G., Kostøl, A. R., & Magne, M. (2014). Family welfare cultures. *The Quarterly Journal of Economics*, *129*(4), 1711–1752. https://doi.org/10.1093/qje/qju019

Démurger, S. (2015). Migration and families left behind. *144*. https://doi.org/https://doi.org/10.15185/izawol.144

Friedman, E. M., Chase-Lansdale, P. L., & Sulaiman, A. (2017). Childhood and adolescent relocation following residential instability: The role of family transitions in educational success. *Child Development Perspectives*, *11*(2), 87–93.

Heidrich, S. (2015). Intergenerational mobility in sweden: A regional perspective. *916*.

Heidrich, S. (2016). *The Effect of Moving during Childhood on Long Run Income: Evidence from Swedish Register Data* (Umeå Economic Studies No. 929). Umeå University, Department of Economics. https://ideas.repec.org/p/hhs/umnees/0929.html

Keene, D. E., Padilla, M. B., & Geronimus, A. T. (2010). Leaving chicago for iowa's "fields of opportunity": Community dispossession, rootlessness, and the quest for somewhere to "be ok". *Human organization*, *69*(3). https://doi.org/https://doi.org/10.17730/humo.69.3.gr851617m015064m

Schmidt, L. A., Padilla, C., & Mora, M. (2017). Adolescent academic achievement following residential mobility: The role of individual, household, and neighborhood characteristics. *Journal of Youth and Adolescence*, *46*(11), 2258–2270.

Survey Research Center, Institute for Social Research, University of Michigan. (2023). Panel study of income dynamics, public use dataset [restricted use data, if appropriate] [Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI].

Tonnessen, M., Telle, K., & Syse, A. (2016). Childhood residential mobility and long-term outcomes. *Acta Sociologica*, *59*(2), 113–129. https://doi.org/10.1177/0001699316628614

Appendix

A Other measures of labor outcome

Figure 3: Correlation matrix of RHS variables in the sample for labor outcome

Table A1: Effects on Long-run Income (other specifications)

	Dependent variable:					
	avg income			log(avg income)		
	(1)	(2)	(3)	(4)	(5)	(6)
has moved	−1,830.578*	−8,712.105*	−2,501.706***	−0.083	−0.157	−0.110*
	(947.839)	(4,687.377)	(953.402)	(0.067)	(0.239)	(0.065)
family income	0.172***	0.286	0.166***			
	(0.035)	(0.183)	(0.035)			
log(family income)				0.169***	0.341**	0.166***
				(0.042)	(0.171)	(0.041)
education (mother)	733.873***	2,845.600***	764.698***	0.047***	0.119**	0.048***
	(199.859)	(996.305)	(201.191)	(0.014)	(0.051)	(0.013)
education (father)	507.316***	−1,417.880	504.355***	0.038***	−0.067	0.039***
	(165.999)	(1,068.023)	(167.907)	(0.012)	(0.055)	(0.011)
male	4,114.143***	1,524.141	3,974.785***	0.313***	−0.090	0.300***
	(716.872)	(3,883.147)	(722.168)	(0.050)	(0.199)	(0.049)
white	1,774.917**	6,105.721	2,295.735***	0.109*	0.176	0.122**
	(848.814)	(4,835.243)	(854.520)	(0.059)	(0.250)	(0.058)
Age range	20-30	31-35	25-35	20-30	31-35	25-35
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,222	147	1,280	1,222	147	1,280
R^2	0.199	0.352	0.196	0.170	0.283	0.166
Adjusted R^2	0.169	0.123	0.165	0.139	0.030	0.133

Note: *p<0.1; **p<0.05; ***p<0.01

B Other potential IVs

B.1 Number of Regions as IV

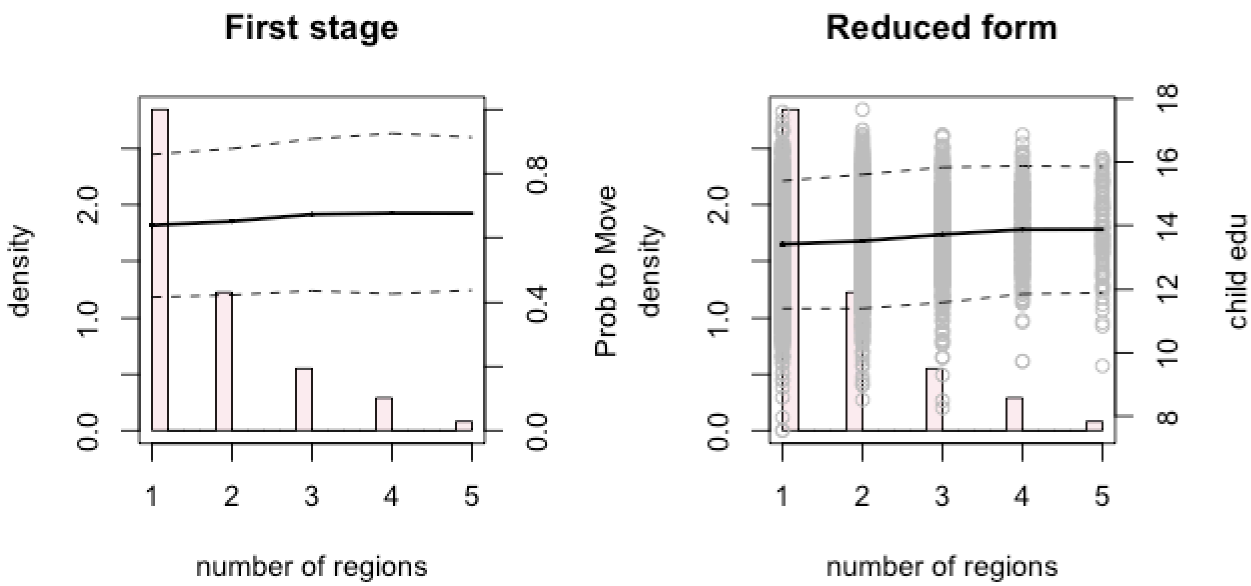

Figure 4: Effect of Number of regions on Probability to move (First Stage) and Children's Education Attainment (Reduced Form). Intuitive we are presuming a positive relationship between the number of regions the household head had been in and the probability of family mobility for teenage children. However, visually, the regression lines in both graphs are flat with wide confidence bandwidth.

Table A2: Effects on Education Attainment: Number of Regions as IV

	Dependent variable: Education Attainment		
	OLS	*1ststage*	*2ndstage*
has moved	-0.206^{**}		0.034
	(0.095)		(1.913)
nreg		0.021^{***}	
		(0.007)	
male	-0.549^{***}	-0.008	-0.547^{***}
	(0.052)	(0.011)	(0.053)
white	-0.094	0.039^{*}	-0.103
	(0.098)	(0.023)	(0.124)
log(family income)	0.727^{***}	-0.130^{***}	0.758^{***}
	(0.051)	(0.010)	(0.262)
education (father)	0.150^{***}	0.001	0.150^{***}
	(0.017)	(0.004)	(0.019)
education (mother)	0.165^{***}	0.003	0.164^{***}
	(0.021)	(0.004)	(0.020)
F Statistic (df = 90; 5365)	20.762***	33.201***	20.618***
Age dummies	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Observations	5,456	5,456	5,456
R^2	0.258	0.358	0.257
Adjusted R^2	0.246	0.347	0.244
Note:			*p<0.1; **p<0.05; ***p<0.01

Table A3: Effects on Labor outcome: Number of Regions as IV.

	Dependent variable: log(income 25-35)				
	OLS (1)	1st stage (2)	2nd stage (3)	1st stage (4)	2nd stage (5)
has moved	-0.116		0.037		-0.590
	(0.078)		(1.018)		(1.358)
nregion		0.034^{**}			
		(0.014)			
dummy: nregion≥2				0.045^{*}	
				(0.025)	
log(family income)	0.184^{***}	-0.133^{***}	0.205	-0.133^{***}	0.122
	(0.049)	(0.020)	(0.143)	(0.020)	(0.186)
education (mother)	0.043^{***}	0.001	0.042^{***}	0.002	0.044^{***}
	(0.016)	(0.007)	(0.016)	(0.007)	(0.016)
education (father)	0.039^{***}	-0.004	0.040^{***}	-0.004	0.038^{***}
	(0.013)	(0.006)	(0.014)	(0.006)	(0.014)
male	0.305^{***}	0.001	0.305^{***}	0.001	0.306^{***}
	(0.058)	(0.024)	(0.058)	(0.024)	(0.059)
white	0.153^{**}	0.046	0.146^{*}	0.045	0.174^{*}
	(0.068)	(0.028)	(0.082)	(0.028)	(0.092)
F Statistic (df = 42; 944)	4.524***	18.221***	4.453***	18.121***	4.306***
Mom's age group	20-22	23-26	27-30	31-34	35+
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	987	987	987	987	987
R^2	0.168	0.448	0.164	0.446	0.135
Adjusted R^2	0.131	0.423	0.127	0.422	0.096
Note:					*p<0.1; **p<0.05; ***p<0.01

Table A4: Effects on Labor outcome: Number of Regions as IV

	Dependent variable: log(income 25-35)				
	OLS	1st stage	2nd stage	1st stage	2nd stage
	(1)	(2)	(3)	(4)	(5)
move times	-0.032^{**}		0.004		-0.080
	(0.016)		(0.121)		(0.183)
nregion		0.284^{***}			
		(0.069)			
dummy: nregion$\geq$2				0.329^{***}	
				(0.121)	
log(family income)	0.175^{***}	-0.807^{***}	0.203^{*}	-0.808^{***}	0.136
	(0.050)	(0.097)	(0.108)	(0.098)	(0.153)
education (mother)	0.043^{***}	0.025	0.042^{***}	0.027	0.045^{***}
	(0.016)	(0.032)	(0.016)	(0.032)	(0.017)
education (father)	0.040^{***}	-0.010	0.040^{***}	-0.007	0.040^{***}
	(0.013)	(0.027)	(0.013)	(0.027)	(0.013)
male	0.307^{***}	0.067	0.304^{***}	0.069	0.311^{***}
	(0.057)	(0.116)	(0.058)	(0.117)	(0.060)
white	0.155^{**}	0.243^{*}	0.147^{**}	0.240^{*}	0.166^{**}
	(0.068)	(0.137)	(0.074)	(0.138)	(0.080)
F Statistic (df = 42; 944)	4.571***	6.094***	4.455***	5.813***	4.44***
Age dummies	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	987	987	987	987	987
R^2	0.169	0.213	0.165	0.205	0.161
Adjusted R^2	0.132	0.178	0.127	0.170	0.123
Note:					*p<0.1; **p<0.05; ***p<0.01

Table A5: First stage: Discretized number of regions

	Dependent variable:			
	has moved		move times	
	(1)	(2)	(3)	(4)
nregion	0.034**		0.284***	
	(0.014)		(0.069)	
dummy: nregion=1		−0.045*		−0.329***
		(0.025)		(0.121)
log(family income)	−0.133***	−0.133***	−0.807***	−0.808***
	(0.020)	(0.020)	(0.097)	(0.098)
education (father)	−0.004	−0.004	−0.010	−0.007
	(0.006)	(0.006)	(0.027)	(0.027)
education (mother)	0.001	0.002	0.025	0.027
	(0.007)	(0.007)	(0.032)	(0.032)
male	0.001	0.001	0.067	0.069
	(0.024)	(0.024)	(0.116)	(0.117)
white	0.046	0.045	0.243*	0.240*
	(0.028)	(0.028)	(0.137)	(0.138)
Age dummies	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	987	987	987	987
R^2	0.448	0.446	0.213	0.205
Adjusted R^2	0.423	0.422	0.178	0.170
F Statistic (df = 42; 944)	18.221***	18.121***	6.094***	5.813***

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

B.2 Number of States as IV

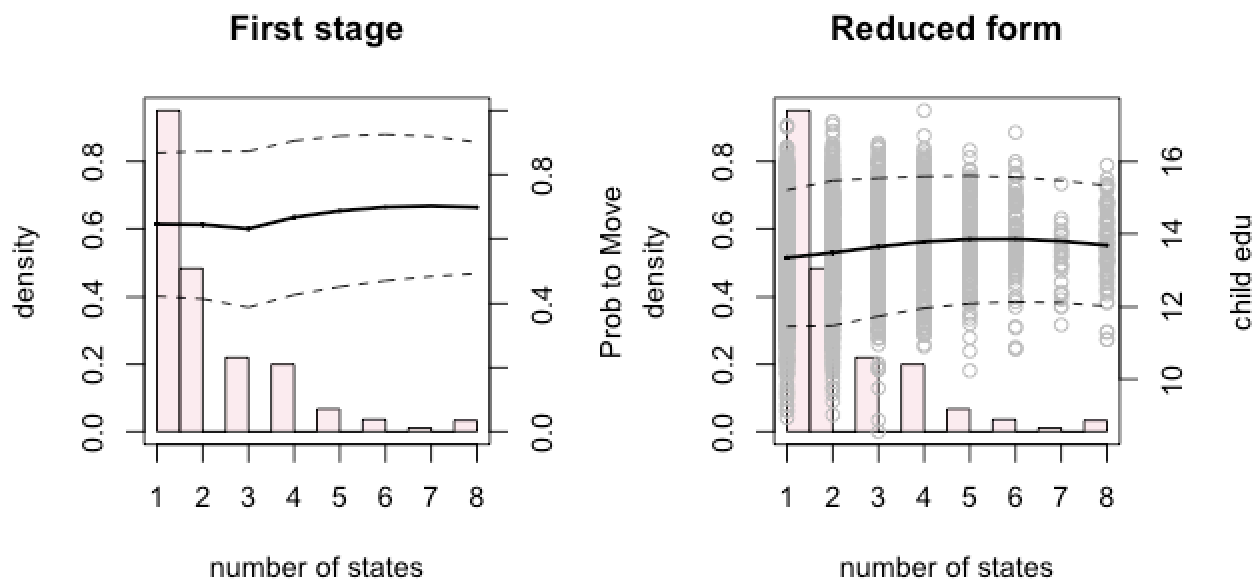

Figure 5: Effect of Number of states on Probability to move (First Stage) and Children's Education Attainment (Reduced Form)

Table A6: Effects on Education Attainment: Number of States as IV

	Dependent variable: Education Attainment		
	OLS	1st stage	2nd stage
has moved	-0.239^{**}		0.497
	(0.094)		(2.387)
dummy: nstate = 1 or 2		-0.011	
		(0.022)	
dummy: nstate = 3 or 4		-0.038	
		(0.026)	
male	-0.570^{***}	-0.006	-0.565^{***}
	(0.053)	(0.011)	(0.056)
white	-0.105	0.039^{*}	-0.134
	(0.086)	(0.023)	(0.131)
log(family income)	0.700^{***}	-0.130^{***}	0.796^{**}
	(0.053)	(0.011)	(0.324)
education (father)	0.160^{***}	0.001	0.159^{***}
	(0.018)	(0.004)	(0.020)
education (mother)	0.172^{***}	0.002	0.170^{***}
	(0.019)	(0.003)	(0.019)
Age dummies	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Observations	5,086	5,086	5,086
R^2	0.275	0.348	0.259
Adjusted R^2	0.262	0.336	0.246
F Statistic (df = 90; 4995)	21.282***	29.616***	20.704***
Note:		*p<0.1; **p<0.05; ***p<0.01	

Table A7: Effects on Labor outcome: Number of States as IV

	Dependent variable: log(income 25-35)				
	OLS	1st stage	2nd stage	1st stage	2nd stage
	(1)	(2)	(3)	(4)	(5)
has moved	−0.100		0.463		−0.363
	(0.076)		(0.951)		(1.351)
nstate		0.024**			
		(0.010)			
dummy: nstate≥4				0.060*	
				(0.034)	
log(family income)	0.173***	−0.131***	0.244*	−0.129***	0.139
	(0.048)	(0.020)	(0.130)	(0.020)	(0.178)
education (mother)	0.042***	0.002	0.041**	0.001	0.042***
	(0.016)	(0.007)	(0.016)	(0.007)	(0.016)
education (father)	0.040***	−0.006	0.042***	−0.005	0.039***
	(0.013)	(0.006)	(0.014)	(0.006)	(0.014)
male	0.315***	0.002	0.313***	0.003	0.316***
	(0.057)	(0.024)	(0.059)	(0.024)	(0.057)
white	0.142**	0.050*	0.114	0.049*	0.155
	(0.068)	(0.028)	(0.084)	(0.029)	(0.096)
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	998	998	998	998	998
R^2	0.166	0.447	0.119	0.445	0.156
Adjusted R^2	0.131	0.423	0.081	0.421	0.120
F Statistic (df = 41; 956)	4.652***	18.814***	4.368***	18.663***	4.556***
Note:					*p<0.1; **p<0.05; ***p<0.01

Table A8: First stage: Discretized Number of states

	Dependent variable: Has moved				
	(1)	(2)	(3)	(4)	(5)
nstate	0.024**	−0.052**	−0.007	0.066*	0.060*
	(0.010)	(0.025)	(0.027)	(0.039)	(0.034)
log(family income)	−0.131***	−0.130***	−0.127***	−0.128***	−0.129***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
education (father)	−0.006	−0.005	−0.004	−0.004	−0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
education (mother)	0.002	0.002	0.001	0.002	0.001
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
male	0.002	0.002	0.003	0.002	0.003
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
white	0.050*	0.051*	0.050*	0.050*	0.049*
	(0.028)	(0.028)	(0.029)	(0.029)	(0.029)
Num of states	Full sample	1	2	3	>=4
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	998	998	998	998	998
R^2	0.447	0.445	0.443	0.444	0.445
Adjusted R^2	0.423	0.422	0.419	0.421	0.421
F Statistic (df = 41; 956)	18.814***	18.721***	18.53***	18.653***	18.663***
Note:					*p<0.1; **p<0.05; ***p<0.01