Estimating the Effects of Family Relocation on Children's Education and Youth Risky Behavior

Hui Mai*

Job Market Paper

Oct. 16th, 2014

Abstract: Using individual-level data from the NLSY79 and the NLSY79 Children and Young Adults, we empirically investigate the role of family relocation on children's schooling and youth behavior problems. By exploiting the variation in sibling's age at the time of family relocation, we find no detectable negative effects of family relocation on various children's outcomes. In addition, while the OLS estimates vary by gender and ethnicity, this variety disappears in the sibling fixed effects estimates. Our empirical results indicate that the unobserved family characteristics that drive the decision of family relocation are responsible for children's schooling and behavior outcomes in the long run.

Keywords: family relocation, education, youth risky behavior, sibling fixed effects

JEL classification: D10, J13

* Contact Information: Hui Mai, Department of Economics, 305 Savery Hall, University of Washington, Seattle
Email: huimai@uw.edu

I. Introduction

Mobility is a common occurrence for American families. Approximately 20% of the U.S. households change the location of their residences in any given year. Despite the fact that mobility rates differ by age group, the figures for young-school age children are about the same as for the population generally. American children have the highest rate of residential and school mobility (Long, 1992). Residential mobility, especially its consequence on children has captured the attention of social scientists for more than a half century. However, opinions have cycled about whether residential mobility has positive or negative implications for children's life course.

This paper investigates the impact of family relocation on children's education outcome and their youth risky behavior. It also addresses the question that whether the timing of family relocation matters for children's outcomes. In most cases, families are selected into relocation for positive reasons as upward social mobility and increased economic opportunity for the family (Kopf, 1977) or negative reasons as job loss, divorce or other forms of family disruption. Longitudinal data from the Panel Study of Income Dynamics (PSID) shows that parental divorce sharply increases the annual probability that children will move out of their neighborhoods (South, Crowder and Trent, 1998). Conditional upon moving, they move to significantly poorer neighborhoods than do children in stable two parent families (South, Crowder and Trent, 1998). We consider the endogeneity and self-selection nature of family relocation to pose both a challenge and potentially a clue in explaining the positive or negative correlation between family relocation and various children's outcomes later on.

Family relocation is closely tied to children's growing environment and critically contributes to their mental and physical development. Existing studies link residential mobility to a range of child and adolescent outcomes and develop much diversified conclusions (Astone and McLanahan, 1994; Simpson and Fowler, 1994; Kerbow, 1996; Alexander, Entwisle and Dauber, 1996 Tucker, Marx and Long, 1998; Scanlon and Devine, 2001). Most of these researches document a clear pattern that residential and school moves are associated with poor academic performance, poor health condition and delinquent youth behavior, whereas other researches point that the association between moving and children's outcome may be spurious. The negative correlation may be a function of other characteristics of people who move often. Using longitudinal

data, they are able to identify most of the negative effect of moving is due to preexisting differences between the movers and non-movers (Pribesh and Downey, 1999).

Most of the previous researches take the naïve approach to evaluate how family relocation would affect children's outcome. That is to estimate the parameters of a regression equation in which the dependent variable is children's outcome (measured at a specific age), and the explanatory variables include an indicator for whether the family relocates, demographic variables, and at times, variables such as family income and labor market participation of the mother. The coefficient of the relocation indicator is meant to capture the effect of family relocation on children's outcome. The prerequisite for the coefficient from this naïve approach to make economic sense is random sampling. However, as families mostly self-select into relocation, the results from this approach are subject to selection bias. The negative effects of family relocation on children's outcome could be amplified when choosing families with low social economic status and vice versa. While families selecting into relocation creates measurement error for the direct effects of family relocation on children's outcomes, it also indicates that all the children that are present in the household are jointly affected by this event which leaves us with natural variation in terms of sibling's age at family relocation.

This paper uses the NLSY79 and the NLSY79 Children and Young Adults to study the impact of residential mobility on children's schooling and youth risky behavior. By implementing sibling fixed effects regression model, the model permits us to hold constant effects which are common to all siblings, we can then control for all sources of observed and unobserved heterogeneity at the family level. The main virtue of this strategy is that it allows us to circumvent problems of selection in a clean and straightforward way. Since the within-family strategy relies on differences in ages of children when family relocates, we control for birth order and birth cohort effect in all specifications We find that the seemingly significant negative impact of family relocation for children at school age (13-18) on children's schooling and risky behavior during young adulthood based on the naïve OLS estimation disappear after controlling for sibling fixed effects. This finding points to the potential bias existing in previous literature due to the selection of family relocation and unobserved family characteristics.

II. Literature

Despite over decades of research on the topic of family relocation on children's outcomes, current research in this area has been limited in several ways:

First, most research in this area fails to consider that it is not moving per se but rather the underlying reasons why the mobility occurs in the first place leads to negative or positive educational or behavioral outcomes. Since families do not choose to relocate randomly, there may be important differences between mobile and non-mobile families which account for the observed relationship between mobility and academic and behavioral outcomes and those underlying differences might explain the deleterious effects of mobility commonly found in previous literature.

Second, lots of research has focused on fragile families with low social economic status which makes the negative relation between mobility and children's academic performance and youth risky behavior more pronounced. Youth in those families who move may already be performing worse academically, at a higher risk of dropping out and being more involved in a variety of delinquent and problem behaviors. In these cases, any observed relationship between mobility and delinquency may be spurious rather than the causal effect of moving. Therefore, negative relation captures not only the effects of residential mobility but also the unobserved family characteristics that initiate the move at the first place.

A large body of literature focuses on residential instability and children's school performance. These studies find that, on average, students who experience residential moves perform less well than students who do not. Specifically, moving is related to reduced academic performance (Ingersoll, Camman and Eckerlin, 1989; Haveman, Wolfe and Spaulding, 1991; Pribesh and Downey, 1999; Wood, Halfon, Scarlata, Newacheck and Nessim, 1993). Coleman (1988) suggests that geographic mobility is a strong predictor of high school dropout. Residential mobility can lead to school mobility, especially in condensed urban area. Astone and McLanahan (1994) use data from High School and Beyond to show that residential mobility is associated with a greater probability of school dropout, after controlling for a number of family and demographic factors. Rumberger and Larson (1998) and Swanson and Schneider (1999) use data from the National Education Longitudinal Study to examine the relationship between school mobility and high school dropout and find that students changing schools frequently are also at greater risk of dropping out.

While the correlation between residential mobility and school performance is well established in previous research, parallel developments in criminology have also been documenting links between residential mobility and delinquency at the individual and community levels (Crutchfield, Geerken and Gove 1982). Moving may bring on feelings of loss, caused by separation from loved ones, friends, or community supports. As children deal with feelings of loss, they require extra support from parents in order to adequately transit to new environment. A child's emotional needs may be overlooked as parents and caretakers are faced with their own emotional, physical, and social demands of moving (Simpson and Fowler, 1994). Therefore, moving could affect important aspects of youth development, especially problem behaviors like high school dropout, running away from home, smoking, drinking and drug use. Inquiry into the relationship between psychological and behavior problems and moving has also been undertaken (Mundy, Robertson, Greenblatt and Robertson, 1989; Simpson and Fowler, 1994; Stacks, 1994; Tooley, 1970). Almost all of the researches done in this area show negative associations between residential mobility and youth outcomes (Adam and Chase-Lansdale, 2002; Astone and McLanahan, 1994; DeWit, 1998; Haynie and South, 2005; Hoffmann and Johnson, 1998; Wood , Halfon, Scarlata, Newacheck, and Nessim, 1993). Mobile children are more likely to be psychiatrically hospitalized, more likely to initiate drug and alcohol use (Catalano, Hawkins, White and Pandina, 1985), and more likely engage in premarital sexual behavior (Stacks, 1994).

It is of great importance to understand whether residential mobility directly contributes to the reduced academic performance or youth risky behavior such as running away from home, smoking, drinking, or using drugs at an early age. On the one hand, parents relocate in an effort to improve family conditions, especially if job opportunity arises. If the move is associated with increasing family income and upward mobility in social economic status, ex ante, we would not expect too much harmful impact on children's academic performance and behavior problem. Many parents are even hesitant to relocate their children precisely because they think it will be traumatic or harmful to educational performance (DeLuca and Rosenblatt, 2010). Likewise, recent policy initiatives to relocate poor families from urban ghettos are premised on the idea that moving to safer neighborhoods relieves family stress, reduces the exposure of adolescents to violence, and even improves youth development through access to higher quality schools and positive adult role models (Ainsworth, 2006; Clampet-Lundquist, Edin, Kling and Duncan, 2006; Sanbonmatsu, Kling, Duncan and Brooks-Gunn, 2006). Such moves to improved contexts are expected to benefit

youth in the long run. On the other hand, residential instability may influence educational achievement and behavior problems through its relationship with increased school mobility (Kerbow, 1996). This school mobility, or frequently changing schools, is associated with worse academic outcomes (Crowder and South, 2003; Swanson and Schneider, 1999; South, Haynie and Bose, 2007) and emotional and behavioral problems (Pittman and Bowen, 1994), as well as reduced social competence and self-esteem which lead to youth risky behavior. Children's emotional competence to deal with mobility may not be well expected when parents make the decision about moving. The lack of participation of children in the decision process may compromise the potential benefits of moving, especially for those less advantaged groups such as fragile and low income families.

Our reading of these seminal and influential work is that they well document the correlation between mobility and a number of children's outcomes, however, we believe prior research has not done an adequate job of examining whether mobile youth are selected into both mobility and problem behaviors. We are not the first, however, to raise the possibility that important selection effects may be driving the association between mobility and youth outcomes. Pribesh and Downey (1999) find that preexisting differences accounted for 90% of the difference in test scores between movers and non-movers. Francisca M. Antman(2012) claims that migrants and nonimmigrants are likely to differ in unobservable ways that also affect children's educational outcomes. Her paper uses Mexico data and addresses selection problem by looking within the family to exploit variation in siblings' ages at the time of parental migration to the US and finds positive effects of father migration to the U.S. on daughters' education outcome. Using randomized housing-mobility experiment, Jens Ludwig and others (Ludwig, Duncan, Gennetian, Katz, Kessler, Kling and Sanbonmatsu, 2013) closely examine the impact of residential mobility on low-income families using data from the Moving to Opportunity (MTO), which offers some public-housing families but not others the chance to move to less-disadvantaged neighborhoods. Their results show after 10-15 years MTO has no detectable effect on economic outcomes, youth schooling and youth physical health and they discover mixed results by gender on other youth outcomes, with girls doing better on some measures. Their empirical results resonate with our findings from the NLSY79 that family relocation itself should not have influential impact for children's outcome in the long run whereas the underlying family characteristics that trigger the decision of moving should be responsible for the differential outcomes between mobile and non mobile children.

Using individual level data from the NLSY79, we examine whether residential mobility has detrimental effects on children's academic performance and whether it leads to risky youth behavior as smoking, drinking and substance use at an early age and find no detectable negative effects of family relocation on children's outcomes. We believe that our paper represents a more rigorous attempt to adjudicate between causal and selection hypotheses about the effects of mobility than what we have seen in previous work.

Our scholarly contribution is twofold. First, our paper addresses the endogeneity problem of family relocation that are overlooked by prior research and proposes an empirical strategy that could potentially control for selection bias of family relocation. It is a relatively novel idea to use the natural age variation generated by sibling order at the event of moving to evaluate the effects of family relocation on children's outcomes. There hasn't been enough discussion in terms of how to choose the right timing of family relocation. We believe our econometrics strategy constitutes an important contribution to the improved understanding of whether and when family relocation would lead to negative or positive children's outcomes.

Second, we use dummy variables indicating different age period when estimating how age at family relocation would affect children's academic performance and risky youth behavior. Previous research has been vague about how they handle multiple relocations happened at different age when evaluating the impact of family relocation on children's outcomes and many of them choose to treat this as one time shock to children. Our paper tackles this problem by including dummy variables on different age interval at family relocation. We will return to the specific details subsequently in interpreting the results of our econometric analysis.

III. Empirical Specification and Identification Strategy

In this research, we first evaluate the long run impact of relocation on children's academic achievements of children. The academic outcome variables include whether a child finishes high school, whether a child ever repeated grade, and the highest grade completed by a certain age. This certain age is set to be 20 in our specification so that the outcome variables measure long run effects and the sample is sufficiently large. We then examine the second set of outcome variables which are youth risky behavior problems. The outcome variables are whether a child starts running away from home, smoking, drinking, or using drug (primarily marijuana use) by age 16.

We begin by introducing the baseline model for this paper illustrated by equation (1):

$$y_{ij}=\beta' X_{ij}+\delta_1 \text{BirthCohort}_{ij}+\delta_2 \text{SiblingOrder}_{ij}+\gamma 1_{\text{Move\_ij}}+\alpha_j+\varepsilon_{ij}$$

$$\text{for } i=1,2 \text{ and } j=1,2,\ldots,n, \qquad (1)$$

where y_{ij} is the outcome variable of interest for sibling i in family j, X_{ij} is a vector of covariates which can vary between siblings in family j, $BirthCohort_{ij}$ is a birth cohort dummy variable, 1_{Move_ij} is an indicator variable which equals to 1 if sibling i in family j ever relocates throughout the sample period from 1986 to 2010 and equals to zero otherwise, α_j denotes unobserved family characteristics, and ε_{ij} is idiosyncratic error term. In this model, γ represents the impact of family relocation on outcome variables and it is the parameter we want to identify for this particular exercise. The vector of covariates, X_{ij}, may include variables that representing shared family characteristics such as ethnicity, mother's education, mother's marital history information and total number of children in the household and variables that representing non shared individual characteristics such as gender, whether mother was less than 19 at the birth of children, and mother's employment status when children's outcome is evaluated.

We include sibling order in model (1) as younger sibling may pick up some behavior pattern from elder siblings. This is because first born children will experience family relocation at later age than their younger siblings. Failure to account for birth order could entail a bias in the estimate of the effect of age at family relocation[1]. Moreover, if there is any spillover effect among siblings, for instance, younger children pick up the negative behavior problem from older siblings and therefore perform worse in terms of outcome variables. For instance, if the first child starts smoking at an early age, the second sibling might start smoking at an early age as well. We can capture this learning experience by controlling sibling order in our model. This model allows 1_{Move} interact with X_{ij} variables such as gender and ethnicity. To account for any general time trend in children's education outcomes and behavior pattern, we also control for the birth cohort in our analysis.

It is important to note that the conventional OLS estimation using the above equation may suffer from the endogeneity problem because families choose to relocate based on family income, job opportunities, marital status and other family characteristics. As these reasons are likely to be family specific rather than individual specific, we make additional assumptions that individual

[1] Recent studies show that among families with more than one child firstborn children on average outperform their younger siblings in terms of educational outcomes (Kristensen and Bjerkedal, 2007, Booth and Kee, 2009).

specific components which are not included by the family specific effects are idiosyncratic. To eliminate the fixed unobserved family heterogeneity, this paper proposes taking the difference between sibling 1 and sibling 2 equations. Suppose that only one of the siblings is the child who has ever relocated from 1986 to 2010. Then, we have

$$E\left[y_{1j}\text{-}y_{2j}\middle|X_{1j}=\widetilde{X_{1j}},\ X_{2j}=\widetilde{X_{2j}},\ 1_{1,Move}=1,1_{2,Move}=0\right]$$
$$=\beta'\widetilde{\Delta X_j}+\delta_1\Delta\widetilde{BirthCohort_j}+\delta_1\Delta\widetilde{SiblingOrder_j}+\gamma \quad (2)$$

To identify γ, we further need to take care of the difference in the birth cohort and sibling order dummies. This can be done by including families that do not change residential address, but have identical $\widetilde{\Delta X_j}$, $\Delta\widetilde{BirthCohort_j}$, and $\Delta SiblingOrder_j$ to family j. Including non-mobile families in estimation, we are essentially comparing the difference in outcomes between siblings with the same age difference in mobile families and non-mobile families and relating this to the children's different age at post relocation experiences. To incorporate this idea, we apply the following approach:

$$E\left[y_{1j}\text{-}y_{2j}\middle|X_{1j}=\widetilde{X_{1j}},\ X_{2j}=\widetilde{X_{2j}},\ 1_{1,Move}=1,1_{2,Move}=0\right]\text{-}\ E\left[y_{1j'}\text{-}y_{2j'}\middle|X_{1j}=\widetilde{X_{1j'}},\ X_{2j}=\widetilde{X_{2j'}},\ 1_{1,Move}=0,1_{2,Move}=0\right]$$
$$=(\beta'\widetilde{\Delta X_j}+\delta_1\Delta\widetilde{BirthChort_j}+\delta_1\Delta\widetilde{SiblingOrder_j}+\gamma)\text{-}$$
$$(\beta'\widetilde{\Delta X_{j'}}+\delta_1\Delta\widetilde{BirthCohort_{j'}}+\delta_1\Delta\widetilde{SiblingOrder_{j'}})$$
$$=\gamma \quad (3)^2$$

Now γ can be identified from equation (3). This illustration takes families with 2 children as an example. However, this identification approach can be easily generalized to situations with more than two siblings. To sum up, the main virtue of this strategy is that it allows us to account for family specific characteristics that might be correlated with educational outcomes of children and parental relocation patterns. Our main specification for the model would be using sibling fixed effects to identify γ . Meanwhile, we will also present results based on the pooled population without sibling fixed effects.

One important virtue of the estimation strategy used here is that it can be easily extended to allow the impact of family relocation to vary depending on the age of the child at the time of the event. Distinguishing effects based on the child's age at the time of the family relocation also brings this paper into relation with the literature on child development and family dynamics which

[2] Taryn Ann Galloway (2012) applies similar identification strategy in the study of timing of divorce and its impact on children's crime-related and educational outcomes.

investigates the effects of residential mobility on children at different age groups in other context such as divorce and separation. Despite these two topics share lots of similarities, an important feature of family relocation is that it happens more frequently over the life course of individual therefore creating multiple shocks at different age for individuals. To control for the impact of multiple relocation happened at different age for the same individual, we allow age at relocation indicator variables to be non exclusive to each other.

Equation (4) represents our extended model:

$$y_{ij}=\beta' X_{ij}+\delta_1 \text{BirthCohort}_{ij}+\delta_2 \text{Sibling Order}_{ij}+\sum_{s=1}^{6}\gamma_s I_{\text{MoveAge}[3s-2,3s]ij}+\alpha_j+\varepsilon_{ij}$$

$$i=1,2,...,m;\quad j=1,2,...,n;\quad s=1,2,...,6 \text{ or } s=1,2,...,5 \qquad (4)$$

where $I_{MoveAge[3s-2,3s]}$ represents a dummy variable indicating whether a child i in family j relocate at age between $3s-2$ and $3s$ for $s=1,2,3,4,5$. For example, if a child relocate at age between 1 and 3, we have $I_{MoveAge[1,3]}=1$ for s=1. Similar explanations apply to other index variables representing child age group when family relocates. We include 6 dummy variables (or 5 age interval dummies depending on the outcome variables) indicating children's age at relocation in a much smaller interval intending to capture impacts of the multiple relocations happened during different age period. The model (4) can capture the potentially discontinuous nature of age at family relocation on children's outcome variables. The baseline group is chosen as relocation age greater than 18 (or 16) or children who never relocate throughout the sample period[3].

IV. Data

The data sets we use for the study is the National Longitudinal Study of Youth 1979 (the NLSY79) and the NLSY79 Children and Young Adults. The National Longitudinal Survey of Youth 1979 cohort (the NLSY79) is a multi-purpose panel survey that originally included a nationally representative sample of 12,686 men and women who were all 14 to 21 years of age on December 31, 1978. Annual interviews have been conducted with the NLSY79 main Youth respondents since 1979, with a shift to a biennial interview mode after 1994. To acquire child specific information, we also employ the NLSY79 Children and Young Adults data which is a

[3] For the education outcomes, we leave out children who relocate above age 18 and children who never relocate during our sample period as control group for our sibling effects estimation. For the risky behavior outcomes, we leave out children who relocate above age 16 and children who never relocate during our sample period as control group as the cut off age for risky behavior is 16.

separate survey on all children born to the NLSY79 female respondents. The child survey includes assessments of each child as well as additional demographic and development information collected from either the mother or child.

The advantage of using the NLSY79 over the PSID (Panel Study of Income Dynamics) or other national survey data is that mother's marital and relocation history can be linked with her children through a unique ID, therefore, we can identify siblings from the same household and track down a complete family relocation history based on mother's response to questions on their residence each survey year. Moreover, the panel structure of the NLSY data can be useful in terms of tracing children's academic performance and the development of youth risky behavior for the entire youth period (up to age 20). Instead of looking for the intermediate impact of family relocation, we can study the impact of family relocation on children's outcome when children reach a certain age later on.

Our main sample is constructed from the NLSY79 Children and Young Adults. Since this ongoing project surveys all children born by mothers from the NLSY79 on a biennial basis from 1986, by the year of 2010, we have 11,498 children from 4,390 households. The birth cohort of our samples ranges from 1970 to 2010. For our analysis, we construct two samples with different age eligibility. The first sample is for the study of family relocation and children's education outcomes. The age eligibility for being included in this sample is age 20 by year 2010. 7,638 individuals meet this age criterion. We drop observations whose mothers are in active military force then we obtain 4,926 individuals with non missing values on education and other independent variables. The second sample is for the study of youth risky behavior such as smoking, drinking and using drugs such as marijuana and running away from home. The age eligibility for second sample is 16 by year 2010. 10,223 Individuals achieve age 16 in our data. We retain 8,247 individuals for the second analysis.

In terms of sample attrition, the NLSY79 children and young adult is linked with the NLSY79 mother, the effective attrition rate for the NLSY79 is 18 percent[4]. Since there is an increased attrition rate in more recent wave, so we only include individual observation up to 2010. If attrition is non random, then that could bias our estimation results in a systematic way. By comparing the demographic characteristics of the sub samples and the master sample, the gender composition,

[4] "Excluding these subsets of respondents means that effective attrition for those who would otherwise be eligible for interview is about 18 percent" - NLSY79 Children and Young Adult User Guide, 2002.

household size, mother's education level are quite similar, though the sub sample for education has a relative higher proportion of Hispanic and Black samples[5].

1) Descriptive Statistics

(Table 1: Approximately Here)

Table 1 provides the summary statistics of the master sample and two subsamples for the key variables used for this study. In total, we have 11,498 individual observations from 4,390 households from 1984 to 2010. Based on the regression specification later, we restrict our sample to a subset of all the individual observations available. The figures in the table describe children's demographic and family characteristics. From table 1, in terms of gender distribution, we have a relatively balanced sample with about 50% of the observation being female. The ethnic composition represents the over sampling of Hispanic and Black Americans in the NLSY[6]. In the subsample, Hispanic and Black take larger component than the master sample. The number of children in the household reveals the information about the average family size for a typical American family would be 3. Our table also shows family characteristics that are shared or not shared among siblings. For instance, mother's education is shared feature among siblings in the same household while whether mother was less than 19 at the birth of the child may be different for individual child. One important feature on the NLSY79 data is it directs children's information with their biological mother which makes it easy to make inference under an intergenerational context. One the other hand, this data doesn't have all the similar information regarding to children's biological father as their mothers.

On the issue of family relocation, about 77 % of children experience family relocation before age 18 and the average number of family relocation (at county level) is about twice per child by 18. Since we don't have the direct information from children on their relocation history, we infer the relocation history from their mothers' and we assume that children live with their mothers before age 18. We believe this assumption is reasonable in the sense that the vast majority of

[5] Including more disadvantaged individuals could bias downward the impact of family relocation on children's outcome. However, even with is potential downward bias, we do not detect the negative impact of family relocation on children's outcome using sibling fixed effects.

[6] To account for the over sampling of Hispanic and Black Americans, we apply sample weight to all the OLS regression and sibling fixed effects regression and record the results in the appendix.

children reside with their parents before going to college and especially with their mother following divorce.

2) Numbers of Family Relocation, Age at Family Relocation and Children's Outcome Variables

In order to show how frequent children move before age 20, we construct the following bar plot indicating the percentage of children who relocated with family (at county level) by children's age. As is presented by this graph, at any given age, 20 % of the children relocate at the county level with family.

(Graph 1: Approximately Here)

(Table 2: Approximately Here)

The descriptive statistics on the dependent variables from the above tables begin to shed light on the questions motivating our study. The key outcome variables for this analysis include measures for children's academic achievements and their youth risky behavior such as smoking, drinking, using drugs (primarily marijuana) and running away from home. More specifically, we use individual's highest education achieved and whether the individual finishes high school by age 20 and whether ever repeated grade by age 20 to represent academic achievement. As is mentioned above, individuals who are less than age 20 by year 2010 are excluded from this part of analysis. And we use whether children start smoking, drinking, using drug and running away from home before age 16 as indicators for children's youth risky behavior. Table 2 gives us a sense of the number of families on which our main identification strategy rests. As we can see, comparing with children who never experience family relocation before age 18, children relocate before 18 on average received about 0.13 year less education, in general have lower possibility (1 % less) to finish high school by the age of 20[7] and display higher probability repeating grade by age 20 (3% more), engaging youth risky behavior such as smoking, drinking using drugs and running away from home before age16. In general, children who relocate before age 16 have 18% higher

[7] In this paper, high school graduation rate is defined by finishing high school by age 20. GED recipients are not counted as high school graduates. See Heckman and LaFontaine (2010) "The American High School Graduation Rate: Trends and Levels" for the detailed discussion of using different sources of data (the NLSY 79, Census and the PSID) to get comparable high school graduate rate.

probability smoking, 26% higher probability of drinking and 18% higher likelihood to use drugs and 6% higher probability of running away from home comparing to the baseline group.

(Table 3: Approximately Here)

(Table 4: Approximately Here)

Table 3 summarizes our outcome variables of interest by number of family relocation experienced before age 18. As the number of family relocation increases, children on average receive less years of education by age 20, have lower probability of finishing high school by age 20 and they have higher probability of repeating grade by age 20 , engaging in risky behavior such as smoking, drinking, using drugs and running away from home before age 16. This table demonstrates the importance of accounting for multiple relocation going through by children as the impact of family relocation could be accumulating over time.

Table 4 compares children's academic performance and risky youth behavior across 7 (or 6) groups based on their age at family relocation. We use children who experience family relocation above age 18 (or age 16) and children who never relocate over our sample period as a benchmark for comparison. Noticing that children who relocate during age 13-15 are mostly affected by family relocation as their school performance is poorer and they show a higher probability of engaging in risky youth behavior such as smoking, drinking, using drugs and running away from home before age 16 comparing to other age groups.

The NLSY79 also gathers series information on children's behavior development. The behavior problems index is asked parents to children from age 4-14. There are 26 questions asked for all children and 2 questions asked only for children have been to school. For each question, parents reply that the statement is "often true", "sometimes true", or "not true". To convert into a total score, the NLSY sets "not true" equal to zero and "often true" or "sometimes true" equal to one, then sums the answers to the questions (so the maximum score is either 26 or 28). The NLSY then standardizes the total score by child's age. We use the standardized behavior index. Since the questions are formed in a way to detect potential behavior problems for children, therefore a higher numerical score representing worse behavior problems.

(Graph 2: Approximately Here)

(Graph 3: Approximately Here)

From graph 2, there is some persistent pattern existing for children who go through different numbers of family relocation. The behavior problem indexes for children who experience family relocation twice and above rises above the other groups indicating frequent relocation would be positively related with increasing behavior problem.

Similarly noted from graph 3, the behavior problem index for children who ever relocate before age 18 is much higher comparing with those who never relocated before age 18 for all different age and the discrepancy increases with children's age.

V. Estimation Results

Before evaluating the results of the estimation of equation 7 with sibling fixed effects, a useful benchmark for comparison is the standard OLS regression with no family fixed effects. We employ a linear probability model (LPM) for discrete outcome variables. The following tables report these results for the overall sample using OLS.

(Table 5-6: Approximately Here)

In the above tables, model 1 displays the estimation results from our baseline model when we only include one dummy variable showing child relocation by age 18 for education outcome and by age 16 for youth risky behavior. Model 2 shows results when we include multiple dummy variables indicating relocation at different age intervals.

In table 5, regressions for the first 4 columns of education outcomes exclude samples that are below age 20 since we primarily use years of education by age 20 and whether finishing high school by age 20. There are other variables such as grade score in school or school drop rate available used in relevant studies. They are not our primary choice for the following reasons: 1) Grade scores are usually not standardized and not comparably across different schools, regions. 2) School drop rate doesn't preserve enough information on children's education level and therefore not informative enough to use in estimation. 3) We are more interested in the long run effect of family relocation on children's education attainment rather than some transient and short term effects. If in the long run, the assumed negative family relocation effect would go away, then we would have proper reason to believe family relocation could work out for children in the long run despite some temporally undesirable effects. Therefore, using highest education achieved by age 20 and a dummy variable indicating finishing high school by age 20 would be superior as outcome

variables for education attainment than others. Besides, we also employ another outcome variable for education which is whether individual ever repeats grade in school. For risky behavior, we exclude samples that are below age 16 by 2010 as we primarily focus on youth behavior problem such as smoking, drinking, using drugs and running away from home by age 16. Studies show early involvement in these activities during young adulthood would lead to substantial high probability of youth delinquent behavior and even criminal behavior. As a starting point, it would make sense for us to start from identify the "gateway" effects of those youth behavior problem.

In both tables, the OLS estimates show persistent statistically negative effects of family relocation before age 18 (or 16) on children's education attainment, youth risky behavior such as smoking, drinking, using drugs and running away from home. The coefficients for our γ (ever relocated by age 18 or 16) are all statistically significant pointing to the potential correlation between family relocation and children's outcome. When we include multiple time intervals on family relocation in model 2, the coefficients are negative and statistically significant for children who experience family relocation ranging from age 13-15 and age 16-18 in terms of education attainment and positive for children who experience family relocation at age 13-15. The negative coefficients suggest potential negative impacts of family relocation on children's education outcome, while the positive coefficients for children's age at relocation on those behavior outcome variables signify that comparing with the benchmark group, they have a higher probability to engage in youth risky behavior at an early age. This result pertains most of the previous research done this area that is to present the negative relation between residential mobility and children's development. However, whether this correlation stands by causal correlation between these factors needs to be further examined.

To visualize the point estimates for the coefficients of interest and their confidence interval, we hereby present the following graph with the estimates of the coefficients and their 95% confidence interval. While the negative impact of family relocation during different age periods for education outcomes are mostly significant for age group 13-15, we can see a clear pattern for youth risky behavior outcomes with an increasing negative impact of family relocation as age at relocate increases. For education outcomes, this rising trend of increased impact of family relocation by relocation age is hardly detectable, however, for risky behavior, the OLS results suggest that relocation at later age causes larger negative impact on children's risky behavior. This trend seems to suggest that parents could choose the proper timing for family relocation in order to minimize

the potential negative impacts of family relocation. We will see whether this phenomenon would be preserved in our sibling fixed effects.

(Graph 4: Approximately Here)

(Graph 5: Approximately Here)

In addition, all the other explanatory variables have the expected sign and are mostly statistically significant. From the regression results, ethnicity difference exists in the sense that comparing to non-Hispanic and no-Black group, Hispanic and Black American groups receive less education and have higher probability of engaging in risky behavior in general. Female children in the household have an average higher education years and higher probability to finish high school by the age of 20. The female children are more prone to report running away from home comparing with male children. The coefficients of sibling order for education outcome are negative and statistically significant indicating younger children have lower education achievement meanwhile the coefficients for risky behavior are significant positive implying younger children have higher probability of engaging risky behavior and there are some spillover effects of youth risky behavior from older sibling to younger sibling in the family. Family size also matters as it reduces education achievement for children. An interesting result for the coefficient of mother's working status (when outcome variable is evaluated) is that mother currently working would increase children's education achievement but it may also increase the risky behavior of children at the same time. Mother's involvement in work would reduce the time spent with children this would potentially increase the chance for them to pick up youth risky behavior. However, we should be careful when making causal inference based on this piece of information since these negative features from mother may be correlated with other unobserved variables which cause the increase of youth risky behavior. In other words, for instance, the dummy variable representing whether mother's age is less than 19 at birth of children would be correlated with the error term and therefore the coefficient from this regression could be biased. Therefore, the coefficient from this naïve regression method is subject to further investigation.

Here we present our main results from sibling fixed effects after controlling series of non shared characteristics among siblings such as birth cohort, birth order; mother's working status and whether mother is less than 19 when give birth of the child.

(Table 7: Approximately Here)

(Table 8: Approximately Here)[8]

Table 7 demonstrate sibling fixed effects estimates on children's education outcomes and behavior problem such as smoking, drinking, using drugs and running away from home. Surprisingly, the significant positive effects of family relocation on education outcomes and negative effects on youth behavior disappear and get substantially reduced after we use the within family variation as an estimation strategy. Comparing with groups who relocate above age 16 and those who never relocate, children who relocate at an early age display non detectable differences from the former group in terms of the relative outcome variables. And this result persists in model 2 when we include several relocation age intervals to address the timing issue of family relocation. We couldn't detect any consistent pattern or trend in terms of the timing of family relocation. The coefficients from the relocation age interval estimates are closely centered at zero.

From the estimation results of the sibling fixed effects, we find some interesting results. As for the coefficient for female, it is still statistically significant positive implying female children have higher education (about 0.4 years more education) and show higher probability (12% more) to finish high school and lower probability of repeating grade by age 20. Other family characteristics that are not shared among siblings become statistically insignificant in our sibling fixed effects suggesting that our concern on the coefficient from OLS is addressed. The previous statistically significant coefficients for family characteristics do not imply causal correlation but inferring a correlation between repressors and residual terms. Therefore, it is within our expectation that the coefficients for those variables after controlling the variation within family using sibling fixed effects become insignificant as they are differenced out by our estimation strategy.

(Graph 6: Approximately Here)

(Graph 7: Approximately Here)

Despite the fact that most of the regression coefficients are not statistically significant, the coefficients for sibling order are mostly statistically significant indicating there are indeed spillover effects from older siblings to younger ones. To make the estimated coefficients comparable with the ones we get from previous OLS regression, we have the above graphs showing the point estimates and their relative confidence interval. As is shown from the pictures,

[8] For sibling fixed effects model, we only retain families with at least two children and eliminate all the twin samples.

all the estimated coefficients after controlling for sibling fixed effects are not statistically significant at a 5% level. The point estimates are all quite near to the zero vertical line.

We believe the following reasons may contribute to the regression results. Firstly, the negative effects of family relocation on children's education and youth behavior are overly exaggerated especially if we put the question under a bigger picture where we are not focusing on those disadvantaged families. When we use sibling fixed effects, we purge out family characteristics that are similar among siblings and only use the variation for age at relocation, therefore, it provides us with cleaner estimation strategy but not necessarily more significant results. By taking the difference among siblings we increase the noise signal ratio which leads to larger standard error. Secondly, children before school age (0-6) should not be affected by family relocation happened during this period since they have not yet attended school, therefore the school mobility story doesn't hold and their behavior patter has not been formed. On the contrary, we would argue that previous researches who find significant negative effects on young children actually capture other family characteristics rather than the effects of family relocation on children's outcome.

VI. Conclusion

By applying a sibling fixed effects regression model to get around the endogeneity problem of family relocation and selection bias, this paper has re-examined the link between family relocation and educational attainment and youth risky behavior. Unlike some of the previous research which emphasizes the negative impacts of family relocation for young children on their education attainment and youth risky behavior, our results imply that family relocation which happens at different stage of children's life could have non detectable impact on children's long run development. To sum up, for school age children ranging from age 13-18, we discover no evidence of the detrimental effects of family relocation on children's education achievement, risky behavior such as smoking, drinking, using drugs and running away from home in our sibling fixed effects model. Our research leads to an increased understanding of why, despite active intervention such as the Moving to Opportunity (MTO) project, the expected outcomes of improving the long term outcome for children who are prone to family shocks have often not been materialized. The specific family characteristics that derived the decision of family relocation could be contributing to children's outcomes in the long run. Our findings could help policy makers to form a good understanding on why family relocation is often associated with poor academic performance and

increased youth risky behavior, therefore they could design solutions that address the genuine driving force of family relocation and offset the potential negative family shocks on children in the long run.

References

1 Francisca M. Antman (2010). "International migration, spousal control, and gender discrimination in the allocation of household resources". University of Colorado at Boulder Department of Economics Working Paper No. 10-15.

2 Karl L. Alexander, Doris R. Entwisle and Susan L. Dauber (1996). "Children in motion: School transfers and elementary school performance". *Journal of Educational Research,* 90 (1), 3-12.

3 Nan M. Astone and Sara S. McLanahan (1994). "Family structure, residential mobility, and school dropout: A research note". *Demography,* 31 (4), 575-584.

4 Robert Audette, Robert Algozzine and Michele Warden(1993). "Mobility and school achievement". *Psychological Reports,* 72(1), 701-702.

5 Margeret Z. Booth (1995). "Children of Migrant Fathers: The Effects of father absence on Swazi children's preparedness for school*". Comparative Education Review*, 39(2), 195-210.

6 Kyle Crowder, Scott J. South (2003). "Neighborhood distress and school dropout: The variable significance of community context". *Social Science Research*, 32(1), 659-698.

7 Gordon B.Dahl, Enrico Moretti (2004). "The Demand for sons: Evidence from divorce, fertility and shotgun Marriage". *National Bureau of Economic Research Working Paper* No. 10281.

8 Fiona Steele, Wendy Sigle-Rushton and Fystein Kravdal(2009). "Consequences of family disruption on children's educational outcomes in Norway". *Demography*, 46(3), 553 -574.

9 Donna K.Ginther and Robert A. Pollak (2004). "Family structure and children's educational outcomes: blended families, stylized facts, and descriptive regressions". *Demography* 41(4), 671-696.

10 Elizabeth Rowe, Molly Laird and Jacqueline Brathwaite (1995). "Mobility as a mediator of the effects of child maltreatment on academic performance". *Child Development,* 66(4), 1130–1142.

11 Kevin Lang and Jay L. Zagorsky (2001). "Does growing up with a parent absent really hurt? ". *Journal of Human Resources,* 36(2), 253-273.

12 Larry Long (1992). “International perspectives on the residential mobility of America’s children”. *Journal of Marriage and the Family*, 54(4), 861–869.

13 John Ermisch, Macro Francesconi, and David J. Pevalin.(2004) “Parental partnership and joblessness in childhood and their influence on young people's outcomes”. *Journal of the Royal Statistical Society*: Series A (Statistics in Society), 167(1), 69-101.

14 John Ermisch and Macro Francesconi (2001). “Family structure and children’s achievements”. *Journal of Population Economics*, 14(2), 249-270.

15 Jeffrey R.Kling, Jeffrey B.Liebman and Lawrence F. Katz(2007). “Experimental analysis of neighborhood effects”. *Econometrica* , 75(1), 83-119.

16 Shana Pribesh and Douglas B. Downey (1999). “Why are residential and school moves associated with poor school performance?” *Demography*, 36(4), 521-534.

17 Sara McLanahan (2004). “Diverging destinies: How children are faring under the second demographic transition”. *Demography*, 41(4), 607-627.

18 Gloria A. Simpson, and Mary Glenn Fowler (1994). “Geographic mobility and children’s emotional/behavioral adjustment and school functioning”. *Pediatrics*, 93(2), 303–309.

19 John W. Santrock (1972). “Relation of type and onset of father absence to cognitive development.” *Child Development*, 43(2),455-469.

20 Duncan Thomas (1994). “Like father, like son; like Mother, like daughter: parental resources and child height”. *Journal of Human Resources*, 29(4), 950-988.

Tables and Graphs

Table 1: Summary Statistics for Different Samples

	Master Sample		Subsample for Education		Subsample for Risky Behavior	
Variables	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Female	0.49	0.005	0.49	0.006	0.49	0.005
Sibling Order	2.08	0.012	1.92	0.013	1.99	0.011
Number of Children	2.93	0.013	3.01	0.018	2.93	0.014
Hispanic	0.20	0.004	0.23	0.005	0.20	0.004
Black	0.26	0.004	0.32	0.006	0.26	0.005
Non-Hispanic/Non-Black	0.54	0.005	0.45	0.006	0.54	0.005
Number of Relocation before 18	2.08	0.007	2.48	0.010	2.13	0.008
Mother's Age at Birth	26.5	0.055	23.9	0.055	25.1	0.046
Mother less than 19 at Birth	0.08	0.004	0.12	0.005	0.10	0.004
Mother Never Married	0.09	0.003	0.09	0.004	0.09	0.003
Mother Always Remain Married	0.48	0.004	0.41	0.006	0.47	0.004
Mother Ever Divorced	0.43	0.004	0.50	0.006	0.44	0.005
Mother – Finished High School	0.50	0.005	0.50	0.006	0.52	0.005
Mother – Some College	0.28	0.004	0.32	0.006	0.28	0.005
Mother – College and Above	0.22	0.004	0.18	0.005	0.20	0.004
Ever Relocated before age 18	0.77	0.004	0.87	0.006	0.78	0.004
Number of Observations	11,498		4,926		8,247	

Table 2: Summary Statistics for Dependent Variables

	Mig18=0		Mig18=1	
	Mean	Standard Deviation	Mean	Standard Deviation
Total Years of Education	11.98	0.066	11.85	0.024
Whether Completed High School by Age 20	0.72	0.018	0.71	0.007
Whether ever Repeated Grade by Age 20	0.20	0.011	0.24	0.005
Whether Started Smoking by Age16	0.19	0.009	0.37	0.006
Whether Started Drinking by Age16	0.27	0.010	0.53	0.006
Whether Started Using Drug by Age16	0.15	0.008	0.33	0.006
Whether Started Running away from Home by Age 16	0.04	0.005	0.10	0.004

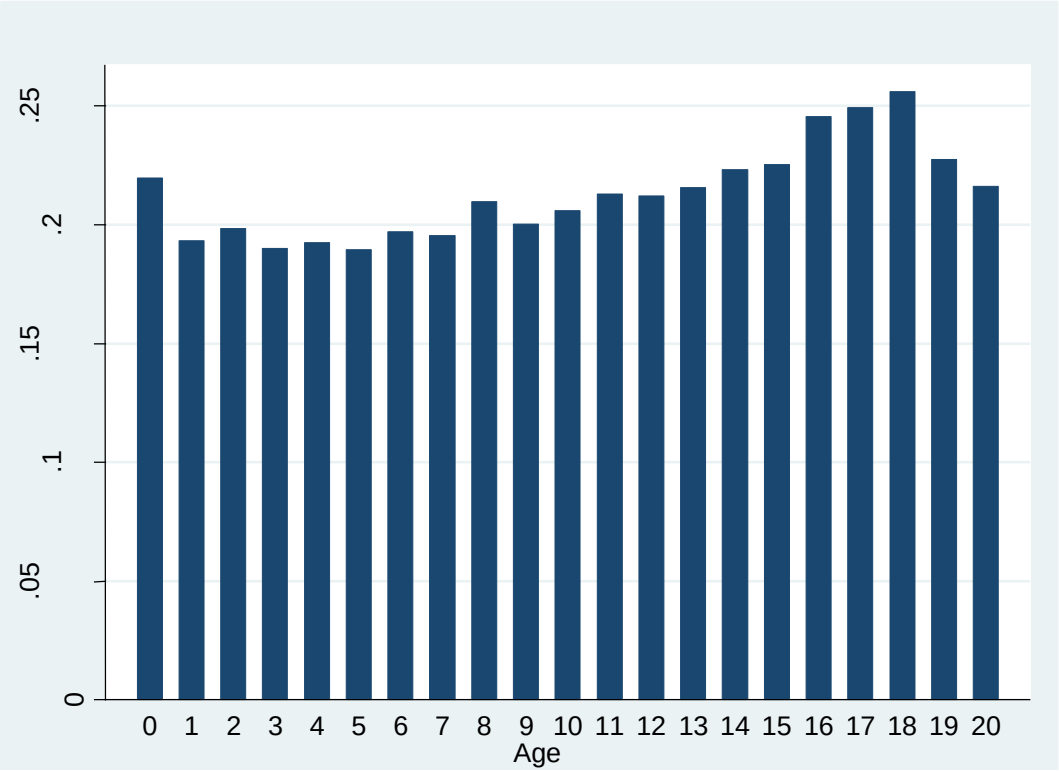
Table 3: Summary Statistics on Dependent Variables by

Numbers of Family Relocation before Age 18

	Sample 1-Education			Sample 2- Risky Behavior			
	Highest Education by 20	Finish High School by 20	Ever Repeated Grade by 20	Smoking Before 16	Drinking before 16	Using Drug before 16	Running away before 16
Never Relocate	11.98	0.72	0.08	0.19	0.27	0.15	0.04
Relocate Once	11.92	0.72	0.12	0.31	0.46	0.29	0.09
Relocate 2-3 times	11.89	0.71	0.10	0.37	0.53	0.33	0.10
Relocate 4 times and Above	11.74	0.68	0.13	0.43	0.59	0.40	0.13
Number of Observation	4,926			8,247			

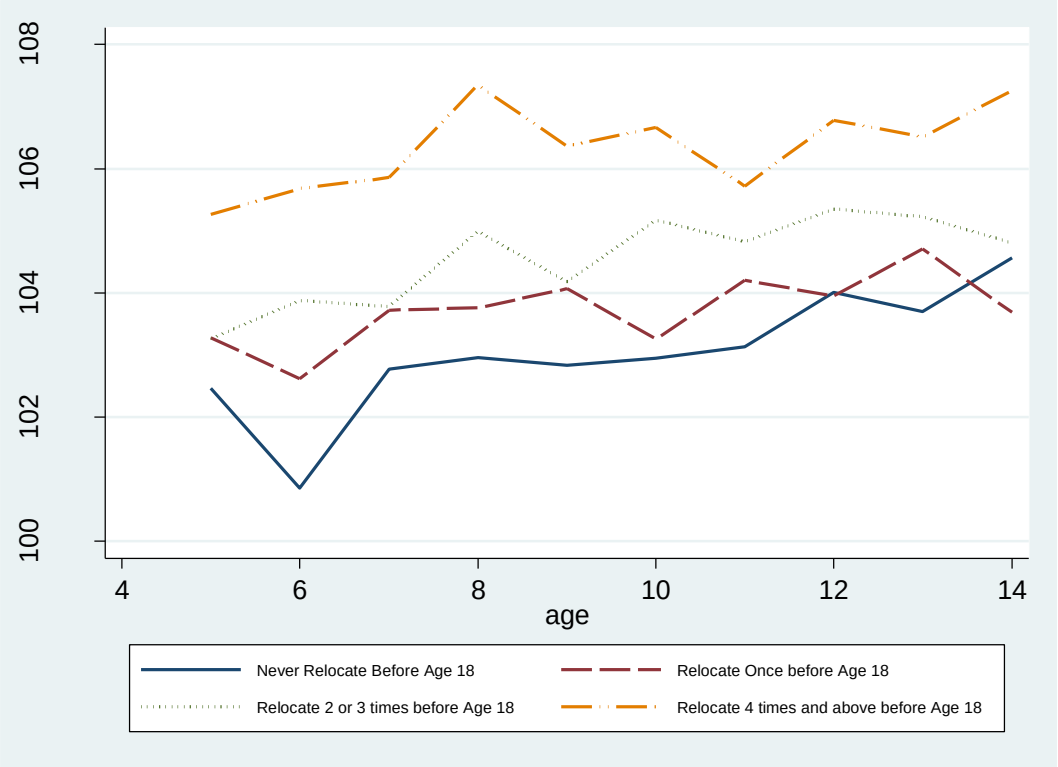
Table 4: Summary Statistics on Dependent Variables by Age at Relocation

	Sample 1-Education			Sample 2- Risky Behavior			
	Highest Education	Finish High School before 20	Ever Repeated Grade	Smoking before Age 16	Drinking before Age 16	Using Drug before Age 16	Running away before Age 16
Relocate at Age 0-3	11.80	0.69	0.21	0.36	0.49	0.33	0.10
Relocate at Age 4-6	11.80	0.69	0.22	0.34	0.46	0.30	0.09
Relocate at Age 7-9	11.95	0.73	0.20	0.34	0.52	0.34	0.10
Relocate at Age 10-12	11.90	0.72	0.20	0.38	0.57	0.34	0.10
Relocate at Age 13-15	11.84	0.69	0.22	0.43	0.61	0.39	0.12
Relocate at Age 16-18	11.68	0.67	0.25	0.48	0.65	0.44	0.13
All the others	11.76	0.68	0.22	0.37	0.50	0.33	0.09
Number of Observation	4,926			8,247			



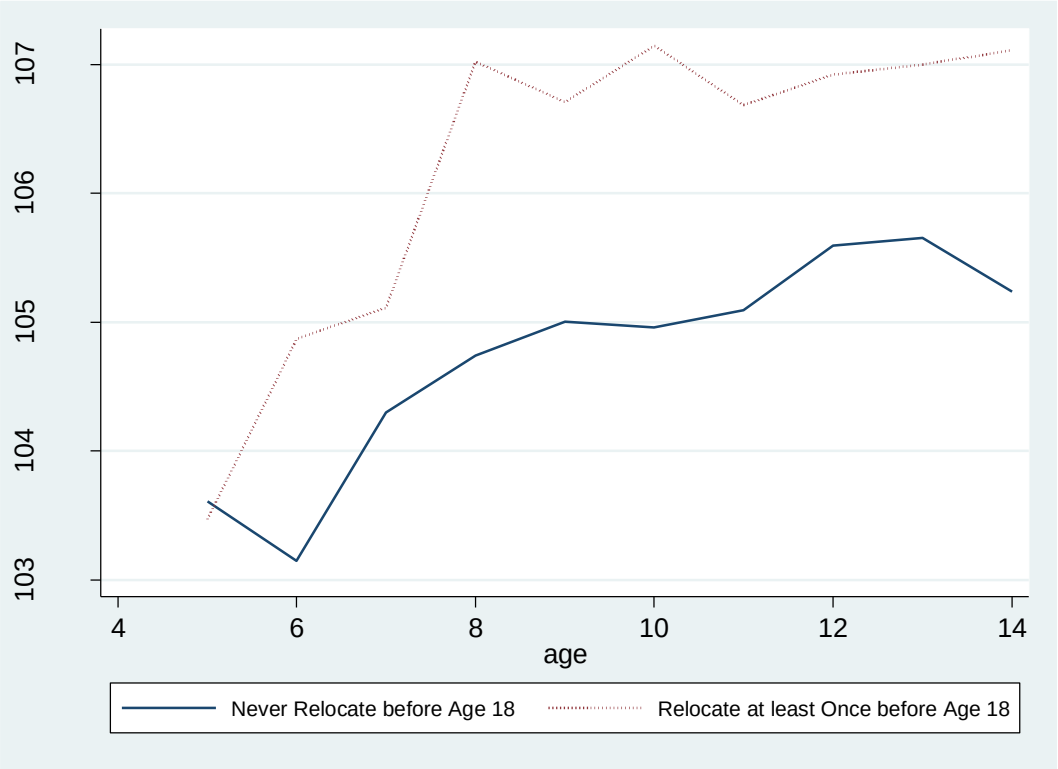
Graph 1: Percentage of Children who change residence by Age

(Source: The NLSY79 Children and Young Adults)



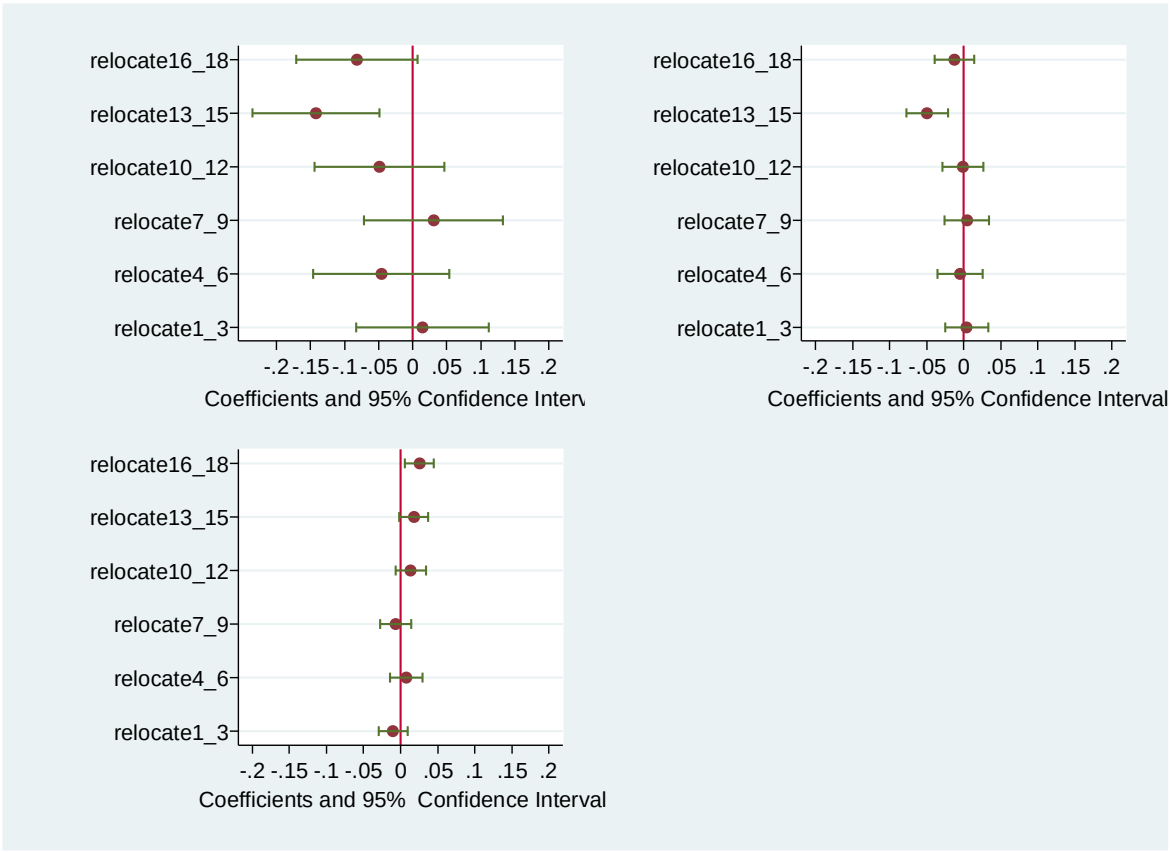
Graph 2: Behavior Index by Numbers of Family Relocation before Age 18

(Source: The NLSY79 Children and Young Adults)

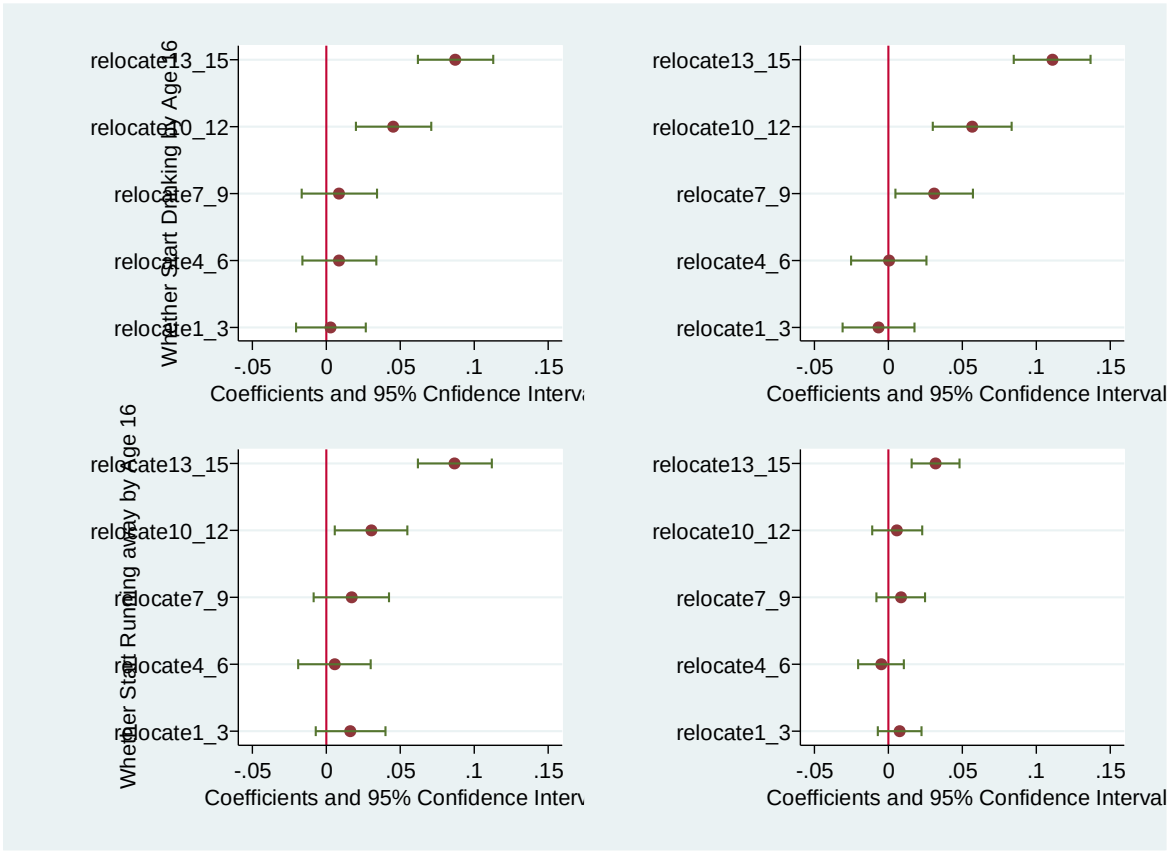


Graph 3: Behavior Index by Whether Relocate before Age 18

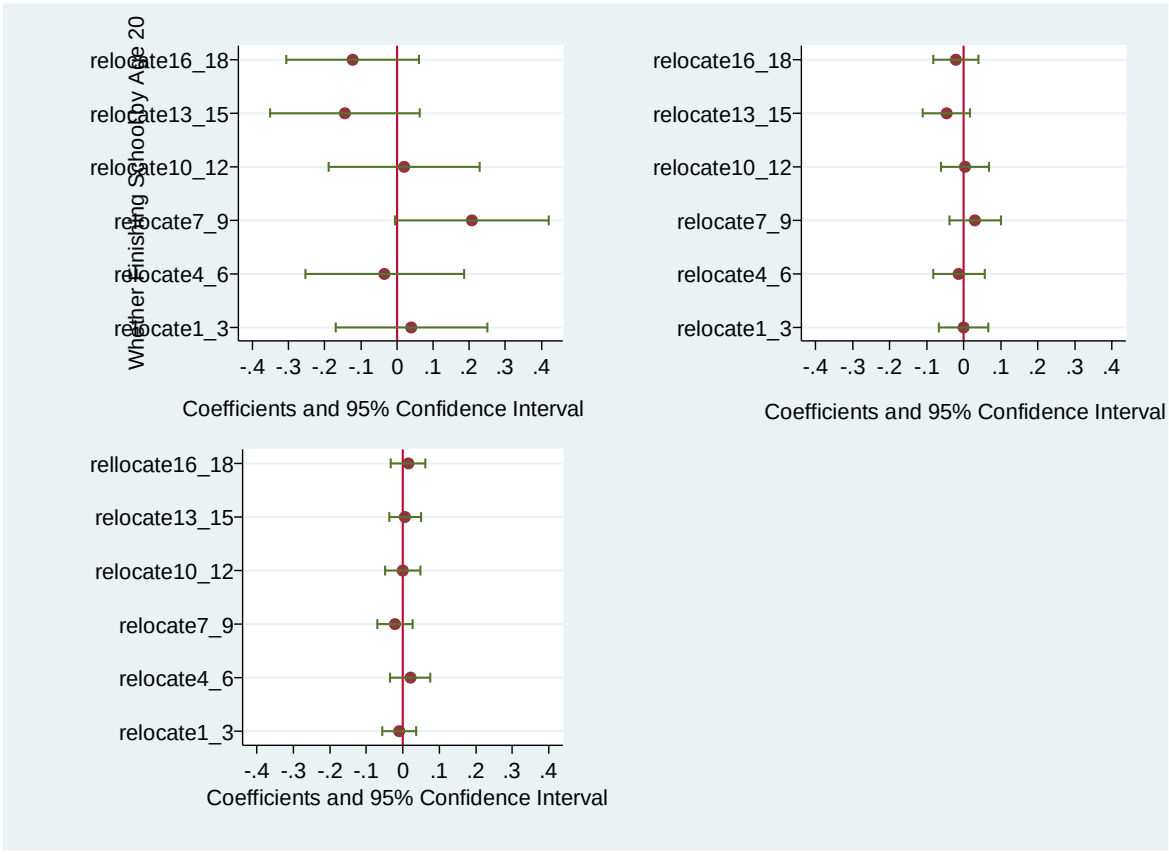
(Source: The NLSY79 Children and Young Adults)



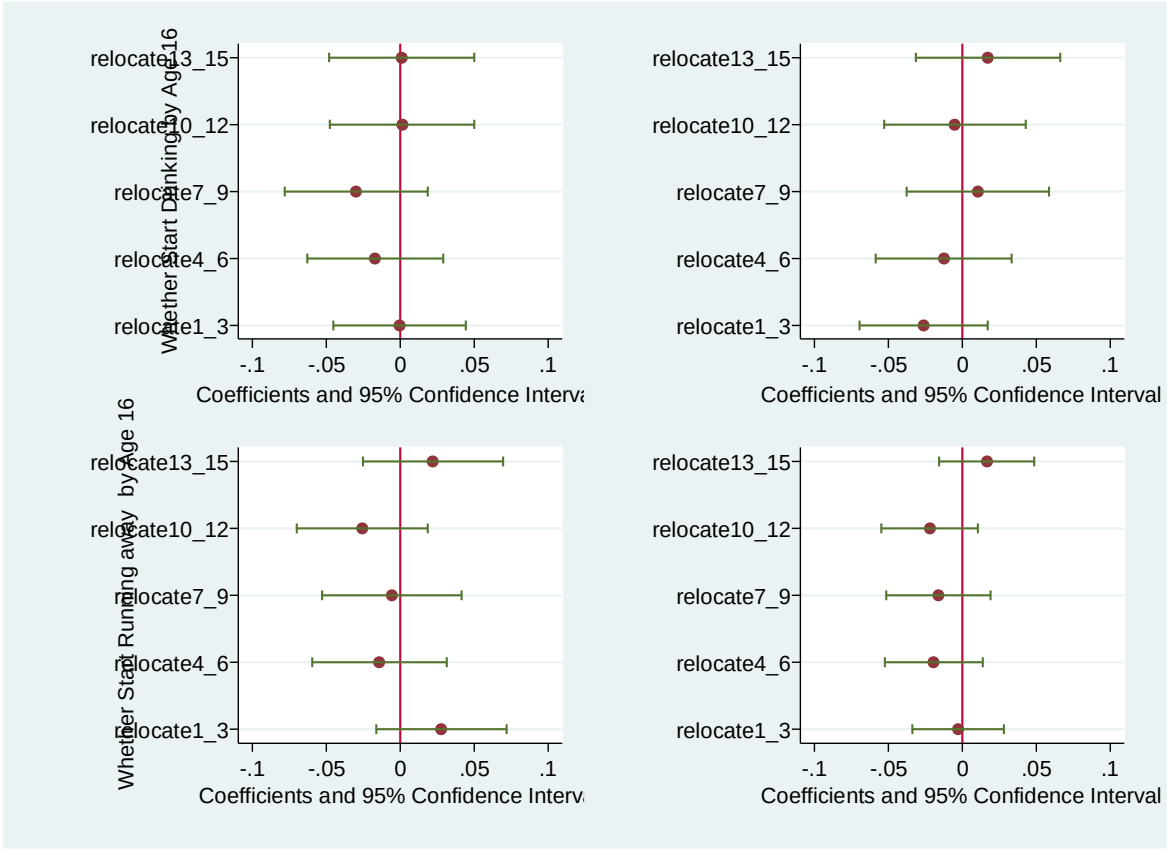
Graph 4: Point Estimates of the Coefficients and Confidence Intervals from OLS Regression
Education Outcomes



Graph 5: Point Estimates of the Coefficients and Confidence Intervals from OLS Regression
-Youth Risky Behavior Outcomes



Graph 6: Point Estimates of the Coefficients and Confidence Interval from Sibling Fixed Regression -Education Outcomes



Graph 7: Point Estimates of the Coefficients and Confidence Interval from Sibling Fixed Regression -Youth Risky Behavior Outcomes

Table 5: OLS Estimation for Children's Outcome Variables –Education Outcomes

	Highest Education		Finish High School by Age 20		Ever Repeated Grade by Age 20	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Female	0.371***	0.369***	0.105***	0.105***	-0.0512***	-0.0510***
	(0.0414)	(0.0415)	(0.0123)	(0.0122)	(0.00880)	(0.00877)
Hispanic	-0.132**	-0.132**	-0.0322*	-0.0324*	0.0199	0.0197
	(0.0641)	(0.0640)	(0.0180)	(0.0180)	(0.0123)	(0.0123)
Black	-0.0354	-0.0295	-0.00925	-0.00730	0.0617***	0.0607***
	(0.0567)	(0.0567)	(0.0166)	(0.0166)	(0.0121)	(0.0121)
Sibling Order	-0.129***	-0.128***	-0.0298***	-0.0297***	0.0221***	0.0216***
	(0.0274)	(0.0272)	(0.00846)	(0.00842)	(0.00663)	(0.00660)
Mother less than 19 at Birth	-0.237***	-0.235***	-0.0696***	-0.0699***	0.0254	0.0242
	(0.0863)	(0.0865)	(0.0253)	(0.0254)	(0.0177)	(0.0177)
Mother Working	0.401***	0.402***	0.112***	0.112***	-0.0162	-0.0165
	(0.0586)	(0.0588)	(0.0163)	(0.0163)	(0.0112)	(0.0111)
No. of Children	-0.0962***	-0.0948***	-0.0240***	-0.0235***	0.00997*	0.00981*
	(0.0223)	(0.0223)	(0.00657)	(0.00656)	(0.00566)	(0.00563)
Always Marry	0.491***	0.477***	0.129***	0.126***	-0.0535**	-0.0496**
	(0.0875)	(0.0876)	(0.0260)	(0.0260)	(0.0235)	(0.0233)
Ever Divorce	0.146*	0.144*	0.0532**	0.0530**	-0.0164	-0.0147
	(0.0842)	(0.0846)	(0.0253)	(0.0254)	(0.0235)	(0.0234)
Ever Relocate by 18	-0.148**		-0.0290		0.0296**	
	(0.0711)		(0.0189)		(0.0117)	
Age Relocate 1-3		0.0143		0.00384		-0.0102
		(0.0495)		(0.0148)		(0.0100)
Age Relocate 4-6		-0.0459		-0.00465		0.00737
		(0.0509)		(0.0156)		(0.0112)

Age Relocate 7-9		0.0307		0.00440		-0.00661
		(0.0521)		(0.0154)		(0.0108)
Age Relocate 10-12		-0.0486		-0.000734		0.0135
		(0.0487)		(0.0141)		(0.0104)
Age Relocate 13-15		-0.142***		-0.0491***		0.0177*
		(0.0475)		(0.0144)		(0.01000)
Age Relocate 16-18		-0.0817*		-0.0127		0.0256**
		(0.0455)		(0.0137)		(0.00997)
Mo_SomeCollege	0.441***	0.439***	0.100***	0.0997***	-0.0631***	-0.0626***
	(0.0526)	(0.0525)	(0.0157)	(0.0157)	(0.0105)	(0.0104)
Mo_CollegeAbove	0.794***	0.794***	0.161***	0.161***	-0.0799***	-0.0798***
	(0.0628)	(0.0630)	(0.0174)	(0.0174)	(0.0109)	(0.0109)
_cons	11.18***	11.14***	0.488***	0.482***	0.0771**	0.0840***
	(0.156)	(0.148)	(0.0455)	(0.0440)	(0.0323)	(0.0323)
N	4926	4,926	4926	4,926	4926	4,926

(Note: Birth Cohorts are included in the regression. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1)

Table 6: OLS Estimation for Children's Outcome Variables –Youth Risky Behavior

	Start Smoking by 16		Start Drinking by 16		Start Using Drug by 16		Start Running away by 16	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Female	-0.0417***	-0.0405***	-0.0244**	-0.0228**	-0.0854***	-0.0842***	0.0175***	0.0178***
	(0.0100)	(0.00999)	(0.0104)	(0.0103)	(0.00967)	(0.00964)	(0.00633)	(0.00632)
Hispanic	0.0428***	0.0392**	0.113***	0.109***	0.119***	0.116***	0.0280***	0.0273***
	(0.0158)	(0.0157)	(0.0161)	(0.0160)	(0.0158)	(0.0158)	(0.00978)	(0.00980)
Black	-0.0496***	-0.0569***	-0.0158	-0.0250	0.0148	0.00886	-0.00918	-0.0109
	(0.0150)	(0.0150)	(0.0156)	(0.0156)	(0.0148)	(0.0148)	(0.00874)	(0.00878)
Sibling Order	0.0191***	0.0193***	0.00883	0.00868	0.0228***	0.0231***	-0.00378	-0.00378
	(0.00726)	(0.00723)	(0.00726)	(0.00724)	(0.00687)	(0.00685)	(0.00433)	(0.00433)
Mother less than 19	-0.00205	-0.00228	0.0152	0.0149	0.0148	0.0149	0.0230	0.0234
	(0.0215)	(0.0213)	(0.0214)	(0.0212)	(0.0213)	(0.0211)	(0.0145)	(0.0145)
Mother Working	0.194***	0.185***	0.293***	0.280***	0.159***	0.151***	0.0542***	0.0518***
	(0.0116)	(0.0117)	(0.0125)	(0.0124)	(0.0113)	(0.0113)	(0.00681)	(0.00696)
No. of Children	0.00227	0.00119	0.00375	0.00263	-0.00364	-0.00471	0.0132***	0.0130***
	(0.00585)	(0.00580)	(0.00637)	(0.00625)	(0.00602)	(0.00596)	(0.00360)	(0.00359)
Always Marry	-0.100***	-0.0952***	-0.0619**	-0.0549**	-0.0854***	-0.0818***	-0.0484***	-0.0470***
	(0.0229)	(0.0227)	(0.0241)	(0.0239)	(0.0221)	(0.0219)	(0.0136)	(0.0135)
Ever Divorce	0.0346	0.0330	0.0523**	0.0518**	0.0332	0.0304	-0.00649	-0.00688
	(0.0233)	(0.0230)	(0.0240)	(0.0239)	(0.0225)	(0.0223)	(0.0141)	(0.0141)
Ever Relocate by 16	0.0669***		0.0923***		0.0654***		0.0238***	
	(0.0126)		(0.0132)		(0.0123)		(0.00687)	
Age Relocate 1-3		0.00298		-0.00666		0.0164		0.00764
		(0.0121)		(0.0123)		(0.0119)		(0.00758)
Age Relocate 4-6		0.00881		0.000255		0.00559		-0.00484
		(0.0127)		(0.0130)		(0.0125)		(0.00789)

Age Relocate 7-9		0.00877		0.0311**		0.0170		0.00845
		(0.0130)		(0.0133)		(0.0130)		(0.00836)
Age Relocate 10-12		0.0454***		0.0566***		0.0303**		0.00592
		(0.0131)		(0.0135)		(0.0125)		(0.00855)
Age Relocate 13-15		0.0872***		0.111***		0.0867***		0.0319***
		(0.0130)		(0.0133)		(0.0128)		(0.00825)
Mo_SomeCollege	0.0167	0.0162	0.0426***	0.0423***	0.0108	0.0102	0.0126	0.0126
	(0.0140)	(0.0139)	(0.0142)	(0.0141)	(0.0137)	(0.0137)	(0.00861)	(0.00860)
Mo_CollegeAbove	-0.0482***	-0.0485***	-0.0163	-0.0167	-0.0325**	-0.0330**	0.00320	0.00300
	(0.0156)	(0.0155)	(0.0169)	(0.0168)	(0.0150)	(0.0150)	(0.00898)	(0.00900)
_cons	0.343***	0.370***	0.299***	0.341***	0.279***	0.299***	0.0477**	0.0548**
	(0.0342)	(0.0340)	(0.0356)	(0.0351)	(0.0336)	(0.0337)	(0.0213)	(0.0214)
N	8247	8,247	8247	8,247	8247	8,247	8247	8,247

(Note: Birth Cohorts are included in the regression. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1)

Table 7: Sibling Fixed Effects Estimation for Children's Outcome Variables
Education Outcomes

	Highest Education		Finish High School by 20		Ever Repeated Grade by 20	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Female	0.339***	0.348***	0.119***	0.121***	-0.0389**	-0.0389**
	(0.0704)	(0.0706)	(0.0225)	(0.0225)	(0.0168)	(0.0168)
Sibling Order	-0.162**	-0.164**	-0.0470*	-0.0476*	0.0307	0.0300
	(0.0775)	(0.0765)	(0.0263)	(0.0261)	(0.0213)	(0.0213)
Mother Single	-0.000314	-0.00586	0.0166	0.0140	-0.0160	-0.0161
	(0.158)	(0.157)	(0.0476)	(0.0476)	(0.0361)	(0.0361)
Mother Working	0.214	0.204	0.0744*	0.0712*	-0.0211	-0.0212
	(0.154)	(0.154)	(0.0427)	(0.0428)	(0.0338)	(0.0340)
Ever Relocate by 18	-0.347		-0.0768		0.0119	
	(0.236)		(0.0698)		(0.0449)	
Age Relocate 1-3		0.0403		-0.000316		-0.0103
		(0.107)		(0.0335)		(0.0235)
Age Relocate 4-6		-0.0338		-0.0132		0.0203
		(0.112)		(0.0353)		(0.0283)
Age Relocate 7-9		0.207*		0.0310		-0.0223
		(0.109)		(0.0357)		(0.0248)
Age Relocate 10-12		0.0204		0.00302		-0.000311
		(0.107)		(0.0331)		(0.0246)
Age Relocate 13-15		-0.144		-0.0466		0.00584
		(0.106)		(0.0323)		(0.0223)
Age Relocate 16-18		-0.123		-0.0217		0.0146
		(0.0940)		(0.0311)		(0.0241)
_cons	11.60***	11.35***	0.583***	0.541***	0.0890	0.0960*
	(0.279)	(0.236)	(0.0862)	(0.0737)	(0.0564)	(0.0514)
N	3606	3,606	3606	3,606	3606	3,606

(Note: Birth Cohorts are included in the regression. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1)

Table 8: Sibling Fixed Effects Estimation for Children's Outcome Variables
Youth Risky Behavior

	Start Smoking by 16		Start Drinking by 16		Start Using Drug by 16		Start Running away by 16	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Female	-0.0506***	-0.0496***	-0.0399**	-0.0411**	-0.0879***	-0.0889***	0.0364***	0.0354***
	(0.0162)	(0.0162)	(0.0163)	(0.0163)	(0.0155)	(0.0154)	(0.0114)	(0.0114)
Sibling Order	0.0380**	0.0361**	0.0519***	0.0525***	0.0379**	0.0382**	0.0146	0.0125
	(0.0170)	(0.0169)	(0.0159)	(0.0158)	(0.0156)	(0.0157)	(0.0104)	(0.0103)
Mother Single	-0.0448	-0.0459	0.0272	0.0284	-0.0291	-0.0293	0.0185	0.0163
	(0.0342)	(0.0343)	(0.0335)	(0.0334)	(0.0325)	(0.0326)	(0.0268)	(0.0268)
Mother Working	-0.0104	-0.00713	0.00866	0.00445	-0.0396	-0.0390	0.0402**	0.0412**
	(0.0286)	(0.0287)	(0.0273)	(0.0274)	(0.0262)	(0.0263)	(0.0202)	(0.0200)
Ever Relocate by 16	-0.0229		-0.0173		-0.0156		-0.00933	
	(0.0304)		(0.0287)		(0.0288)		(0.0196)	
Age Relocate 1-3		-0.00152		-0.0260		0.0289		-0.00420
		(0.0226)		(0.0219)		(0.0225)		(0.0157)
Age Relocate 4-6		-0.0168		-0.0112		-0.0136		-0.0195
		(0.0233)		(0.0234)		(0.0232)		(0.0169)
Age Relocate 7-9		-0.0329		0.00993		-0.00718		-0.0201
		(0.0247)		(0.0244)		(0.0240)		(0.0178)
Age Relocate 10-12		-0.00307		-0.00221		-0.0290		-0.0229
		(0.0247)		(0.0243)		(0.0225)		(0.0166)
Age Relocate 13-15		0.00196		0.0184		0.0202		0.0175
		(0.0251)		(0.0249)		(0.0241)		(0.0163)
_cons	0.510***	0.505***	0.524***	0.524***	0.428***	0.408***	0.0986***	0.106***
	(0.0482)	(0.0460)	(0.0442)	(0.0426)	(0.0439)	(0.0443)	(0.0333)	(0.0326)
N	6677	6,677	6677	6,677	6677	6,677	6677	6,677

(Note: Birth Cohorts are included in the regression. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1)