The Effect of Parental Migration on the Human Capital Investment of Children: Evidence from Indonesia

Alberto Ramírez*

June 1, 2015

Abstract

I document evidence of a relationship between parental migration and childhood educational attainment in Indonesia. Using all the currently available waves of the Indonesian Family Life Survey (IFLS) I fit a discrete-time probit duration model to exploit the time-variation of the parents' migration, as well as a linear model of schooling attainment in the repeated cross-section. I find that parental migration positively influences children's educational attainment in the cross-section. The duration results indicate a decrease in the hazard of dropping out of school with all types of parental migration. The greatest impact is observed when a parent's migration coincides with their child enrolled in secondary schooling, conditional on having lasted long enough to enter this level. Since the selection problem of parental migration is not yet dealt with, the results of the analysis are sensitive but indicate a positive relationship. I take these results as motivating evidence to commence the modeling of a dynamic discrete choice model that structurally estimates the effect of parental migration decisions on their child's human capital investment.

[0] I would like to thank my supervisor, Joan Llull, for his input and guidance; Lidia Farré for her guidance, input, and help with the IFLS data set; and Caterina Calsamiglia and Nezih Guner for their comments and input.

1 Introduction

Returns to education in the developing world are arguably higher than in the developed world (Montenegro and Patrinos, 2013), yet household educational investment in their children can be affected by many factors. One of the threats to schooling attainment (and arguably why returns to schooling are higher in the developing world) is that child labor may substitute adult labor, unless a sufficiently high wage is attained by parents to meet the sustenance level of household consumption so that child labor is not engaged (see Basu and Van, 1998; Baland and Robinson, 2000). As these theoretical treatments predict two possible scenarios, empirical analysis of the prevailing equilibrium, and the context and determinants of its existence is necessary. Notwithstanding, these theoretical results are not always deterministic as other motives may influence parental decisions. Kruger et al. (2012) find that although higher parental wages are associated with higher schooling in Brazil, greater local economic activity may lead to increases in child labor even when conditioning on socioeconomic status.

A natural extension is to analyze whether parental migration is sufficient to relax household budget constraints so that enough resources are available to free the child for schooling, as parents are not necessarily confined to their local labor markets and migration offers a mechanism to spatially reallocate their labor supply. But with migration comes great uncertainty. The intuition in the theoretical framework of Basu and Van and Baland and Robinson lends itself to this question, as it is *a priori* ambiguous what the effect of a parents migration into a different labor market would have on a child's educational attainment given the uncertainty associated with migration. The current research question concerns itself with the extensive internal migration in Indonesia, and its effect on the human capital development of the children of these migrants. The question I wish to address then is whether the positive association between parental migration and childhood educational attainment observed in the Indonesian data supplied by the Rand Corp. (the Indonesian Family Life Survey - IFLS) is due to the effect of expected income from migration, accounting for the uncertainty and cost of such events, where parents are motivated to invest in schooling because they care about the future well-being of their child.

The topic of children left behind is important because the added disruption to the household may place the child at greater risk vis-à-vis a simple budget constraint issue, as these households may face more and varying constraints on consumption allocation. Here, the literature has focused primarily on external migration with little said on internal migration. One the one hand, the literature finds positive outcomes for those children whose parent (or sometimes even a relative) migrates[1]. Others have reported, however, that results can be more nuanced. Ferrone and Giannelli (2015) find that children's schooling in Uganda is negatively impacted by being left behind, and benefit more from partaking in the parental migration than they do from solely the remittances received by their migrant parents.[2] Others have also found that the effect of parental migration can spill over to non-migrant children. Hunt (2012) finds that native-born black US high school students complete more years of schooling in the presence of immigrant classmates, indicating a spillover that induces these native-born to complete more schooling to forego labor market competition with immigrant dropouts; and more recently Clifton-Sprigg (2015) finds that lower secondary students in Poland perform higher when in the presence of classmates whose parents are currently in a migration event.

The propensity to migrate internally, however, provides a country the mechanism to reallocate labor efficiently. And access to more local labor markets might have differential impacts to those children whose parents migrate internally vs. an external migration.[3] What makes Indonesia an interesting study is its extensive internal migration relative to external migration, where in 2000 about 10% of the population lived in a province different from birth whereas only 1.5% lived abroad (Ducanes and Abella, 2009). I find that internal migration rates in Indonesia average about 2.09%, growing at about 0.04%/year.[4] In Indonesia, where barriers to migration are non

[1]Antman (2012) and Hanson and Woodruff (2003) report that father's migration to the US has positive effects on the young children left behind versus an internal migration for Mexican migrants, a fact that may be due to a more compressed wage distribution in the US (Borjas, 1987); while Hildebrandt and McKenzie (2005) find positive health outcomes for these children. Similar results are found in El Salvador (Cox and Ureta, 2003); for Polish children (Clifton-Sprigg, 2014); and Klemp et al. (2013) find similar results in pre-industrial England, with parents differentially allocating apprenticeship opportunities in favor of their eldest children. Interestingly McKenzie and Rapoport (2006) report negative effects on the educational attainment of Mexican children left behind. This study differs from Antman, and Hanson and Woodruff in the data used, but employs similar railroad networks instruments - implying that this IV may suffer from weakness due to an autocorrelation to past shocks as other types of network instruments have similarly displayed (Borjas, 1999).

[2]Hu (2013) finds a likewise negative effect for Chinese children, though remittances here have the opposite effect in that they offset the parental absence. Differences between countries are many, and relative skill sets and external networks that groups can tap may play important roles.

[3]Individual motivations may be driven by more personal factors: rural and urban inequality and how it relates to the education of migrants (Lucas, 1997); and more interestingly the quality of information available about the rest of the country may dampen the propensity to migrate (Farré and Fasani., 2012).

[4]While this topic is important in many countries, it is most intriguing in China, which every year experiences the largest of all known migration events in the world. Through the *Hukou* policy of migration barriers China attempts to control this flow of humanity. A policy reform in 1998 allowed Pan (2012) the ability to determine the effect on human capital generation under migration controls. Pan finds that the relaxation of the policy that allowed rural parents to re-categorize their child as urban (conditional on at least one parent having the urban classification) lead to a negative investment in secondary education of rural children, as these children were now able to access labor

existent and the government actively promoted transmigration programs, I find that migrant parents have higher education than their non-migrant counterparts, with roughly equal shares living in rural and urban areas.

In the current analysis conducted on the IFLS data set I find positive associations between parental migration and childhood educational attainment. I conduct the analysis on a repeated cross-section of the data set and a duration analysis on the panel dimension, where my data is drawn from those children who have already completed their pre-tertiary education (and thus, fully observed). Children in Indonesia have high rates of primary school completion, where roughly 1/3 of children only have up to primary schooling. The repeated cross-sections yield an average of 2 months more of educational attainment for each migration event, with similar results for both closer types of migrations (intradistrict) and farther types (inter-district). I also find that at least 30% of mothers who migrate engage in staggered migration: either she or the father leave in a previous period and the other follows afterwards (the correlation between father's and mother's migration is estimated at 0.64 and highly significant, where it is presumed the children follow). The duration results reveal that parental migration also decreases the hazard that children drop out of school, where the effect of migration may be strongest at the secondary schooling level (which is non-compulsory). The analysis displays some sensitivity as a result of not yet accounting for the selection mechanism that sorts parents into migration and the inherent endogeneity of the human capital investment in children, as well as uncontrolled unobservables. Despite this, the results point to a significant positive association.

My analysis is related to the current literature on migration and human capital development. The Basu and Van framework implies that unless a minimum threshold wage is attained parents make a trade-off, which can be explored by shutting down the parental wage channel that may impact a child's educational production function. The principal motivation for the investment in a child's human capital development is then that parents are altruistic toward the well-being of their children.[5] Gayle et al.

migration into urban areas.

[5]This motivation is conceptually compatible even when parents make the trade-off to engage the child in the labor market as opposed to schooling, if the household income is not sufficient to substitute for other childhood activites. In this case the current well-being of the child is deemed more important. Basu and Van show this is perfectly rational. An insurance motivation is also perfectly in line with the concept of investing in a child's education. Here the parents may have a future expectation on some portion of their child's income insofar as the child commits to engaging in an old-age transfer to their parent in return for the resources transfered to them in the form of human capital in their youth (see Raut and Tran, 2005, and footnote 7). However, it would seem parents' insurance motives would be contingent on first meeting short-term requirements of the household and here one would not expect an insurance motive to hold.

(2015)[6] highlights that altruistic parental models can be estimated by considering that parents are altruistic towards their dynasty, foregoing the complication of how parents allocate welfare among their children. The authors consider parental inputs as time invested in children and determine the dynastic discount factors associated with human capital transmission. A similar framework can be employed here to help model the mechanism by which migration transmits human capital.

The rest of this paper is organized as follows: section 2 introduces the educational system of Indonesia; section 3 briefly describes the Indonesian Family LIfe Survey data; section 4 introduces the empirical evidence from the data that motivates the research question; section 5 analyses the data to condition the means of interest on observables and the hazard associated with migration; section 6 concludes.

2 Education in Indonesia

Education became a central focus in 1973 when then president Suharto issued a decree to combat the low enrollment in primary schooling and the then 20% youth illiteracy rate. This decree, the Sekolah Dasar INPRES Program, set aside oil revenues to start a process of building new primary schools across the country, with the amount in each region to be determined by the rate of children not enrolled in the educational system. The variation induced by the program is the central focus of Duflo (2001).[7]

Education in Indonesia is characterized by a three tier system comparable to the education levels in most countries: primary, lower secondary, and upper secondary, which under Indonesian Law 20 of the National Education System (Part I, Chapter 4, Article 6) is declared compulsory up to age 15. This corresponds to completion of primary and lower secondary schooling (corresponding to grade 9). Additionally, parents may choose to send their children to kindergarten and to community play groups (comparable to pre-school). Children may conduct their education in their local language up to grade 3 in primary schooling, where instruction switches to Bhasa Indonesian thereafter. Further, this three-tier system is offered in secular form (governed by the Ministry of Education and Culture) and religious form (governed by the muslim-dominated Ministry of Religious affairs) (Kuiper, 2011). The secular route also offers the option of completing lower and upper secondary in vocational schools, which precludes the ability to then enter tertiary education. In my research I

[6]In a companion paper the authors detail how to estimate these models through a new estimator they propose for dynastic models.

[7]Raut and Tran (2005) further find that investment in children's education in Indonesia may be motivated by the reciprocally, self-reinforced insurance motive for old-age transfers from children. In analyzing my data set of Indonesia I find evidence that may be consistent with this via an uptick in migration for those aged 65+, who are overwhelmingly female in composition.

do not distinguish between the secular, religious, or vocational routes of educational attainment and treat them all equally, as my concern is the acquisition of the full cycle of pre-tertiary education. Tertiary education consists of choosing between 2-4 years of diploma studies (analogous to associates level college) or a 4 year university, whereupon entrance into graduate studies is then allowed.

At the end of each level of pre-tertiary education students sit a national exam which must be completed to enter the next level. Decentralized examining of students occurred from 1965 through 1980 via the *Ujian Negara* (State Exam), when a switch to a more centralized exam structure, the *Evaluasi Belajar Tahap Akhir National* (National Final Learning Evaluation - commonly abbreviated to EBTANAS), was made. In practice local governments retained much control over the structure of these tests. Due to this heterogeneity the government switched to a fully centralized testing system in 2003, the *Ujian Akhir Nasional* (National Final Examination - UAN) (Rahmi, 2011). In my IFLS data set I have these scores for those who were able to recall them or produce the certificate of their results. Given that up to 2003 there is regional heterogeneity in its implementation the use of these scores requires such a correction.

3 Data

The Indonesian Family Life Survey is an ongoing, longitudinal survey of Indonesian households conducted by the RAND corporation in conjunction with SurveyMETER.[8] Four waves were fielded between 1993, 1997, 2000, and 2007, with the next wave to be released in 2016. It is a very rich data set, unique in that it contains a battery of surveys including: full retrospectives of pregnancies, contraceptive use, marriages, migrations, labor force participation, education; surveys on household consumption, expenditure, production, decision making, and remittances; subjective wellbeing and socioeconomic perceptions; community participation; health batteries and biomarkers conducted by a trained nurse; and some cognitive tests of adults and children. It further contains batteries of surveys conducted on the communities in which IFLS households are located, of which there are 321 in 1993. These community surveys yield detailed information on the regional heterogeneity that exists within Indonesia. Location information is hidden at the village level for privacy concerns. Further details on this data set can be found in appendix A.

Histories are captured through retrospectives, where new survey participants are

[8] A second, recently available data set, the IFLS-East, was conducted in 2012 by SurveyMETER. I do not use this data set in the current paper but will be incorporating this data into my future analysis. I refer the reader to Appendix A for more information on this survey.

asked to give detailed accounts and thereafter are asked to update information from the previous waves. I use this rule to combine all 4 waves of household data and to identify redundancies. This procedure generates a database of 15,086 households comprising 66,778 individuals. Since the analysis requires those who have finished schooling, the majority of the children in my sample consists of adults, of which I identify 47,159. It follows then that children who have finished schooling are also predominantly adults, and from this set I identify 16,948 adult children. A further 241 individuals from the child data sets who finished their schooling are identified, for a total of 17,189 "children". These children correspond to 6545 fathers and 7331 mothers. Although I restrict the data set to the biological children of the head of the household, because the coding of the head of household can change across waves some multiple generations creep in. I keep these children in the set even though the research question does not currently pertain to intergenerational mobility.[9]

I generate an unbalanced panel of these children from the first year they enter school to the year of their exit or graduation, linking to parents' information (including the year of their migration events and the type, whether an inter-district or intradistrict type of migration). Given that 20% of the children repeat at least one grade I expand the panel to accommodate these events when they occur. For the analysis on the repeated cross-sections I collapse the panel dimension.

4 Empirical Evidence

4.1 Internal Migration

Internal migration is prevalent in Indonesia and has been fostered by government transmigration programs (initiated in 1950, through transmigration plans) to relieve pressures on densely populated islands. Between 1950 and 1968 some 90,000 households were relocated, and by 1997 6.5 million individuals (Farré and Fasani., 2012). The IFLS data evidences this substantial migration. As reported in figure A.4, roughly half of the identified adults have migrated, and figure A.6 reports that nearly 1/3 of the sample has had an inter-district migration.[10] Because variation in migration types is necessary for the future estimation of a migration cost function, I also elaborate on the subsample of adults who have ever migrated across the islands in Indonesia - in

[9]Hertz and Jayasundera (2007) extended Duflo's analysis to estimate the intergenerational mobility in Indonesia using the IFLS on account of the INPRES program, using the data supplied by Duflo. A future research question is being explored to utilize the dynastic structure of the survey to help estimate the impact of migration on intergenerational mobility.

[10]As appendix A.1.1 elaborates, there are 4 types of geographical divisions: provinces, districts (*kabupatan*), sub-district (*kecamatan*), and at the lowest level villages (*kelurahan*, or the more frequent usage *desa*) and municipalities (*kotamadya*).

the IFLS data this is about 10% of migrants.[11]

Figure A.7 in appendix A.2.2 graphs the average migration events per cohort. As is typical in the migration literature I also find in Indonesia that younger cohorts have higher migration events than older cohorts. Interestingly there is a reversal in this trend with the oldest cohort, which may be due to migrations into their children's household. There is also a fair amount of dynamics. The shares reported in figure A.11 would imply that about 20% of the whole sample has migrated more than once, while about 10% of the sample has migrated more than once between districts. A next step is to understand how much migration concerns returning to the home location. Finally, figure A.12 graphs the migration rates over time according to inter-district and intradistrict migration events. The average yearly migration (summing the two series) has been about 2,09% with a growth rate of 0.04%/year.

The selection associated with migration is quite evident. The two subsamples in figure A.4 are systematically different across many dimensions. For the entire sample I find that adults have about as much education as reported by Duflo, of around 7.9 years; and the age of the sample is consistent with what Bryan and Morten (2015) finds using a different data set. Migrant adults skew slightly more male and have higher educational attainment in both levels and years. Interestingly, those who don't have any migration events in their histories are no more older, but are more than twice as likely to be unmarried. Probing this group shows that they tend to be more male, skew towards lower educational attainment, but are overly represented in the youngest cohort of 15-24. Figure A.5 presents a marriage contingency table, which shows that given the assortative matching present in Indonesia women may be less likely to marry these lower educated males. Across the cohorts, it is generally the case as well that those with more migration events tend to be more educated, as reported in figure A.10.

4.2 Adult Children

As reported in figure A.17, children whose parents migrate have generally higher education than those whose parents don't migrate and no more likely to repeat a grade. They also complete more levels of schooling (on average, at least half of high school - indicating a pattern that parental migration might allow children to go beyond compulsory schooling) and more likely to go to college versus those whose parents don't migrate. These children are also more likely to have attended Kindergarten and born in urban areas. It will be important to control for these factors as they may

[11] Figure A.13 reports the distribution of these movements and shows that, in general, migrants move to nearby islands.

also be correlated with educational attainment. And as would be expected, children of migrant parents have higher frequencies of having moved (where I observe their age 12 location to be different than their birth location). A smaller observed proportion of children whose parents don't migrate while they were in school have also moved before age 12. However this must have occurred prior to primary enrollment given that I do not observe parental movement while they are in school.

Literacy in Bahsa Indonesian among all the children is very high, and very little systematic difference is observed here. Interestingly the proficiency of speaking Indonesian, at around 47%, seems quite low relative to the proficiencies of reading and writing. I find that this is a problem, in fact, with the survey question. It does not ask respondents their proficiency in speaking Indonesian as the other questions do, rather it asks which language is most frequently spoken at home. For this reason I omit this variable as a control in subsequent analysis. Figure A.18 reports the descriptive statistics of their parents.

As previously discussed, the educational attainment of children is systematically different between those whose parents migrate and those who don't, which indicates a selection bias and a mechanism that needs to account for this. As further evidence, I decompose the educational attainment of children across several relevant dimensions to guide analysis and future modeling of the problem. Along the dimension of inter- & intradistrict migration events, children have an associated larger attainment for those whose parents migrate versus those that don't migrate (figure A.19). However, between the two migrant groups there isn't a systematic difference. This is fortunate as it indicates that in building a model I can restrict myself to the inter-district migration type of events. Figures A.20 and A.21 report the cross-tab of educational attainment across the urbanization of birth. Rural children obtain roughly 2 years less of schooling, and rural girls slightly less than rural boys. I refrain from parsing the data along the rural-urban-sex dimensions as the observations become small.

5 Empirical Analysis

As the evidence of parental migration and children's educational attainment presented in the previous section was based on simple univariate and bivariate exploration, I now analyze the relationship of parent's migration and children's schooling with the data set I currently have available to condition the mean on current observables. To do so I construct mutually exclusive migration dummies for the father's migration, the mother's migration, and the parent's migration for both inter-district and intradistrict migrations. Since I am not yet accounting for the selection induced by parental migration this analysis is to understand the patterns in the data to motivate the modeling and for the structural estimation.

I therefore conduct two types of analysis: a cross-sectional analysis on the collapsed, unbalanced panel to obtain estimates of the average years of schooling attained conditional on the parental migration types; and a survival analysis with a discrete-time probit exploiting the duration nature of the panel dimension. I report results in appendix B.

5.1 Cross-Sectional Analysis

For this analysis I forgo the panel dimension since the only time varying covariates available are parental migration and the school grade at the given year (which makes the panel dimension more suited for a survival analysis rather than a fixed effect estimation). The dependent variable I consider, total school years, is a cross-sectional variable and so panel techniques can not be used.

I estimate regressions of the form

$$SchYrs_i = \alpha + \sum_k \beta_k(ParentMig_i^k) + \delta(ParentMig_i^{1,2}) + \sum_j \gamma_j X_{ij} + u_i \tag{1}$$

where $SchYrs_i$ is the total amount of schooling of child i; $ParentMig_i^k$ is an indicator variable that takes the value 1 if the parent has ever had a migration event of type $k = 1$ (*inter-district*) or $k = 2$ (*intradistrict*) while the child was in school, and ($ParentMig_i^{1,2}$) is an interaction of the parental migration types; X_i are controls that include birth year fixed effects, parent's education, an interaction of the birth year fixed effect, and other observables as described in the figures of appendix A[12]; and u_i is the unobserved random component (which I assume exhibits correlational clustering at the family level).

I estimate separate specifications of equation (1) for each parent (and another for both parents) to keep consistent interpretations considering the collapsed panel dimension; and I condition the regression on the other parent not migrating (omitting this for the case of both parents migrating). I interact the two types of migration events for although k is mutually exclusive in the panel dimension it may not be in the collapsed cross-section. A check of the data reveals that concurrence is in fact quite rare, with less than 1% of children whose parent's migrate displaying this.[13] I incorporate birth year dummies and the interactions of birth year with parental education to account for possible time trends that may exist with schooling for different

[12] I have not yet constructed regional fixed effects for the birth urbanization since the cleaning of the regional codes has not yet been conducted.

[13] Because I am only concerned with "ever migrated while in school" I reconstruct these dummies such that they are mutually exclusive in the collapsed data set.

cohorts and the possible association of parental education on the different cohorts.

5.2 Duration Analysis

The duration analysis of the data is conducted with a discrete-time probit estimation on my single spell data.[14] This method for conducting the duration analysis is quite flexible and allows me to incorporate repeated grades into the analysis. It also allows me to estimate the baseline hazard[15] nonparametrically, generating a dummy for each grade level and omitting the final year (grade 12) to anchor the estimation.

Because this is a proportional hazard model, it makes an assumption on the baseline hazard, This assumption states that the baseline hazard is constant in each time period, although it may vary from period to period. A problem with implementing these models occurs when a time varying covariate is introduced, as it may bias the identification of the coefficient of interest (Heckman and Singer, 1986). This arises if the covariate does not satisfy the constant hazard assumption of the proportional hazard model - that is, the covariate is varying not only across each time period (which is not a problem) but also *within* each time period - and so fails to be constant. From the outset this seems to apply to migration. In reducing migration events to discreet, yearly observations I collapse the within year variation that may exist (for example, seasonal migration). A test of the assumption is discussed in Keele (2010), which requires fitting a Cox model and testing whether the slope of Schoenfeld residuals is systematically different from 0. I indeed find that the parent's migration does not satisfy the constant hazard assumption of these proportional hazard models.[16] A general workaround to this problem is to interact the coefficient of interest with an assumed functional form of the baseline hazard. Instead I interact the parent's migration dummy with the time dummies, as I make no functional norm assumptions.

Another issue that must be dealt with is the initial conditions problem - how individuals enter into the sample. For example, if there's a selection problem that

[14]This is essentially a proportional hazard model. Proportional hazard models are a class of survival models where the covariates enter multiplicatively into the hazard function, and thus shift the baseline hazard proportionally. I consider the data single spell because I assume that once an individual exits the school system they do not reenter, a reasonable assumption.

[15]The baseline hazard is the fundamental risk that all in the population share, irrespective of their heterogeneity.

[16]The Cox model is another proportional hazard model. The difference between this model and the probit method I use is that the Cox model does not allow the specification of the baseline hazard since it fundamentally makes no assumption of its form - it cancels out in the conditional likelihood. I fit the data set to a Cox PH model without accounting for the grade repetitions to test the proportionality assumption of the migration event dummies. Although not reported, uncorrected Cox and probit models produce coefficients with similar estimates, where the precision of the Cox estimates are lower given the fewer observations of migrations because the method of preparing the data to fit this model does not easily allow the incorporation of grade repitions.

determines who enters the school system then this mechanism must be specified as duration analysis can not deal with this "left-censoring problem". That right-censoring occurs (that one observes people who survive to the end of the spell or survive beyond the survey time-frame) is not a problem as the nature of the analysis accounts for this. I assume that there is no initial condition problem and start all individuals at grade 1 for those who enter the school system.[17]

Fixed effects regressors are also problematic in probit models due to the issue of incidental parameters (or nuisance parameters).[18] However, this only concerns the nuisance parameter of the unobserved individual fixed effect, something I am not concerned with here as I am not conducting a fixed effect regression and merely control for time trends. As the main purpose of these treatments is to account for selection on unobserved heterogeneity, one solution for single spell data would be to fit a random effects estimator. The main worry here is that failing to control for the unobserved heterogeneity results in a bias toward "negative" duration dependence.[19] The resulting bias would work to generate an upward shift of the hazard function. However, as elaborated in Heckman and Singer (1984), the typical implementation of this random effects estimator is to make an assumption on the distributional form of the unobservables, which can become quite *ad hoc*. They present a robust method to nonparametrically estimate through maximum likelihood the parameters of interest and the distribution function of these unobservables in duration analysis. Currently, I do not implement this as even accounting for selection on unobservables the selection due to migration has not yet been accounted for. These techniques will prove useful in the next phase of the research to structurally estimate the mechanism of migration selection.

5.2.1 Duration Model

A baseline hazard consists of the following, which gives a dummy for each duration year of schooling:

$$\delta_\tau = \sum_{j=1}^{T^*} \delta_j \mathbb{1}(t = j) \tag{2}$$

T is censored according to $T^* = T - 1$, as the final period observed, $T = 12$ years

[17] I find that just slightly more than 2% of the children in my sample do not enter the schooling system.

[18] This problem was first studied by Neyman and Scott (1948) and elaborated by Lancaster (2000).

[19] These concepts developed in the analysis of unemployment spells, where negative duration dependence leads to lower probabilities of realizing the hazard of exiting unemployment. In the current context, negative duration dependence would increase the probability of realizing the hazard of dropping out.

of schooling, is not identified.[20] I interact equation (2) for each type of migration event (following the same notation as in equation (1) for event types) resulting in

$$\gamma_{i\tau}^{k} = \delta_{\tau} ParentMig_{is}^{k} \mathbb{1}(s = j) \tag{3}$$

In practice it is necessary to artificially censor dummies of equation (3) if a critical mass does not exist at one of the discreet time intervals.[21]

The conditional hazard rate that I estimate is then defined as

$$h_i(\tau, X_{it}; \theta) = P(t_i = \tau | t_i \geq \tau, X_i) = F\left(\delta_{\tau} + \sum_{k} \gamma_{i\tau}^{k} + \beta_k (ParentMig_{i\tau}^{k}) + \sum_{j} \eta_j X_{ij}\right) \tag{4}$$

In the above $h_i(.)$ is the hazard contribution of individual i. The function $F(.)$ is the normal CDF. The other variables are as previously described, only that now the controls X_{ij} do not contain any dummies controlling for time trends (although in principle I could). The parameter θ is the vector of parameters within the hazard that will be estimated. As discussed in footnote 20 the failure event (in the present context, dropping out of school) is defined as $y_{i\tau} = \mathbb{1}(t_i = \tau)$ and indicates the period the individual realizes the hazard specified in (4). I also define $a_{i\tau} = \mathbb{1}(t_i \geq \tau)$ as the dummy that identifies the periods the individual is alive in the sample. The log likelihood that is maximized to estimate the parameter vector θ is given by

$$\ln \mathcal{L}(\theta; X) = \sum_{i=1}^{N} \sum_{\tau=1}^{T^*} a_{i\tau} [y_{it} \ln h_i(\tau, X_{i\tau}; \theta) + (1 - y_{i\tau}) \ln(1 - h_i(\tau, X_{i\tau}; \theta))] \tag{5}$$

The left-hand portion of the bracketed equation is the hazard realized at i's failure event; while the right-hand portion is the survival of i when $t_i \geq \tau$ for current period τ.

[20] $T = 12$ is not identified because I define the failure event as dropping out of school prior to grade 12; that is, $y_{i\tau} = \mathbb{1}(t_i = \tau)$, where t_i is the termination time period of individual i (final school grade of an individual) and $t_i \leq T^*$. Anyone who survives to grade 12 would have a dummy vector full of 0s as they have $t_i = T > T^*$, and a probit would predict failure perfectly. The lack of mass points at the censored period does not identify it.

[21] Artificial censoring is necessary when there do not exist failures during a time period, as was discussed in footnote 20 for the final period. For example, the plots of the survival function depicted in appendix B.2.1 have artificially censored grades 4-5, grade 6-8, and grades 9-11. This is because there are no observed failures for grades 5, 6, 7, 8, 10, and 11 when the dummies of equation 3 are constructed discretely. Thus, I artificially censor them and instead create dummies for grades 1, 2, 3, 4-5, 6-8, 9-11.

5.3 Results

The results of estimating 1 are reported in figures B.1 - B.3. One feature of the results in general is that within a specification there is no difference between an inter-district migration event and an intra-district migration event (even for the specification in column 4 with the full set of currently available controls when a result is not significant at threshold levels). In magnitude, the mother and the father's migration seem to directionally increase children's attainment with inter-district migrations.

For migrations where fathers and mothers migrate (implying one of them stays behind, figures B.1 and B.2) specifications 2-4 are robust in the pure sense of the word, in that conditioning on current observables and de-trending leads to generally similar estimates. The same is not generally true for the case when both parents migrate. In comparison to specification 1 there is a decrease in the amount of years, indicating that selection on unobservables upwardly biases the estimates. Despite this the results generally fall in line with figure A.19. Interestingly, the conditioning reveals that mother's inter-district migration may be more important than closer movements. Further, the interaction seems to have little relevance, although this is more likely due to an imprecise estimation based on small occurrences in the data. I have not yet completed the cleaning of parental wages and will require the SUSENAS 2000 data set to help. I make note of this because one possible source of omitted variable bias is from excluding parental wages, as its exclusion may inflate the estimate of migration. But if parental education is a sufficient proxy for wages then its exclusion should not be problematic for the time-being. Overall the estimates reveal that a parent's migration status increases the average years of schooling a child attains by 2 months.[22]

The duration analysis reveals similar patterns to that described by the repeated cross-section. Figures B.4 - B.6 present the results for each parental migration type. The figures report the marginal effects from the probit, the response to a migration event in a given year. Although in general a migration event decreases the probability of dropping out of school (and as a consequence in the duration context, the hazard of realizing a dropout decreases), the results are sensitive to selection on omitted variables across specifications. Within the various specifications there is no difference between the type of migration; migrating in general seems to decrease the chances of dropping out. Although not reported in the output figures, those in urban locations are also more likely to stay in school; sex does not play much of a role in the estimations; and birth-order seems to also decrease the hazard of dropping out, indicating a

[22]An indonesian school year consists of about 10 months of education, from September to June. So a "school year" being 10 months this factor can be multiplied to the coefficients of figures reffig: FaMig - A.19 to obtain the equivalency.

life-cycle interpretation vs. the "quality-quantity" interpretation developed by Becker and Lewis (1973) and elaborated in Becker and Tomes (1976).[23]

From the results of the hazard analysis I construct the predicted survival plots of various groups of individuals based on specification (5) of the reported tables. I restrict these plots for inter-district migrations and because estimation results are similar across parents (likely attributable to the high correlation of fathers and mothers migrating). As a result I only report figures for the father's migration. These can be found, along with an explanation of the procedure, in appendix B.2.1. Migration tends to increase the proportion of those who survive into the next year, especially after completing the first level of schooling (after grade 6). The structure of the survival functions (and the predicted hazard, also reported) is a property of the data: there are very few dropouts within a level of schooling (grades 1-5, 7-8, 10-11). The end of a level is when I tend to observe a large proportion of people exit. Further, the primary school completion rate is high in Indonesia (as can be inferred from the descriptive statistics and now evident in the survival plots).[24]. A feature of these survival plots is that they do not go to 0, as one would presume. Although it is possible that the duration distribution is defective,[25] as figure A.17 reports some children eventually go to college. Since I have censored the data to the pre-tertiary cycle and do not consider graduation from secondary schooling a failure event the survivor function can not be defective, it is merely censored and will not reach 0 under the current paradigm.

A striking feature of the plots is that parental migration seems to have the most profound impact in the final years of the schooling cycle, where a predicted 12% more children are likely to survive into the next grade level. This is made all the more interesting given that upper secondary is non-compulsory. If at the margins, where a household may gain a productive unit of labor, parental migration really works to alleviate household constraints then parents may choose to forego substitution in favor of investing more in the child's human capital. This may be especially true if parents have an expectation that their child may be able to access the tertiary schooling system. Or as Raut and Tran (2005) point out, these Indonesian parents are likely motivated by the insurance such investment may yield them in their old

[23]Older parents may find themselves in a better economic situation as they age, and children born later may reap this benefit as compared to their older siblings. The Becker framework demonstrates that the income elasticity of demand of quality children is lower than the demand elasticity of child quantity. The lower quality elasticity raises the shadow price of increasing child quality. So more children should lead to lower investment in their quality. However, wealth and price effects are indeterminant in this framework, so predictions in this dimension are not elucidated.

[24]Duflo, Kuiper (2011), and Hertz and Jayasundera (2007) document similar results, indicating evidence of success of the INPRES program of the 1970s

[25]As explained by Heckman and Singer (1986), survivor functions should satisfy $S(t = \infty) = 0$. If this is not the case then the duration distribution is deemed defective, unless a plausible reason exists for this possibility.

age.

Moreover, as the patterns in figure A.21 demonstrate rural children on average have lower education than their urban counterparts. This can also be appreciated from the survival plot of figure B.10, where rural children have a higher hazard of dropping out conditional on having attained a current school level (and so lower proportions survive). When both parents migrate, the rural-born child attains almost as much education as an urban born child whose parent's don't migrate (and I can not discount the possibility that the parents migrated to an urban area) - also evident in figure A.21, which shows that those rural children whose parents migrate tend to have about one more year of schooling. Comparing the survival plots presented in figures B.9 and B.10 reveals that rural children whose parents migrate would survive in the school system at about the same proportions as an urban child whose parents did not migrate.

Because the omitted selection that sorts parents into a migration makes it *a priori* ambiguous whether the estimation is inflated or attenuated, the only way to interpret the results of this section is to note that so far the evidence points to a positive correlation and that parental migration may be reducing the hazard of exiting the school system prior to 12 years of completion. So in general, the extent to which parental migration may raise the quality of children by relaxing the budget constraint, and studying the labor-market outcome of these children, is a key question that the structural framework will help resolve. I caveat that although compulsory education is (by law) only for 9 years I have chosen to study the full cycle of pre-tertiary schooling as the general trend in Indonesia has been an increase in educational attainment.

6 Conclusion

Internal migration in Indonesia is quite extensive and evidence exists that an association exists between the schooling attainment of children and the parents that migrate. Linear estimates point to an increase of educational attainment equivalent to two more months of schooling. A duration analysis on the current data set to elucidate this pattern reveals that parental migration shifts the predicted hazard downward and thusly increases the survival of children within the school system. The impact is greatest in upper-secondary education, where around 12% more children are predicted to stay in the school system. The associated marginal probability of a child dropping out of school when a parent migrates also decreases.

This analysis was undertaken to document the motivational evidence for the next stage of the research analysis. There are two main sources of error: the selection on unobservables and the selection of parental migration. Evidence from the regression and the duration analysis suggests that not controlling for the first is upward biasing estimates of the migration coefficient, and overall may cause a negative duration dependence. Both of these issues can be adequately dealt with by generating a model to facilitate the estimation. Due to this, the evidence presented is qualitative in nature and works to guide the next phase of the research project.

A Appendix: IFLS Background and Descriptive Statistics

The first wave launched in 1993, covering 13 of the then 26 provinces on 6 islands in Indonesia.[26] As I elaborate in the next subsection, since the first wave the set of provinces and other regions have expanded. Figure A.1 illustrates a map of the provinces as they were sampled in the 1993 IFLS. These 13 provinces were chosen as they contain 83% of the population - that is, the survey in itself was not fully representative due to costs. As the IFLS did not survey the eastern provinces, SurveyMETER used the same techniques and surveys to conduct the IFLS-East 2012 survey of 2,547 households with 10,759 respondents in 9 eastern provinces in Indonesia (figure A.2). I plan on including this newly available data set to increase the source data, adding the variation afforded by these more distant provinces.

The survey sampling scheme for contacting households followed the Central Bureau of Statistics Indonesia's 1993 SUSENAS, a nationally representative socioeconomic survey conducted on 200,000 individuals nearly every year. These base dynasties were established in this wave, which totaled 7,224 contacted households yielding a sample size of 22,347 individuals. Subsequent waves in 1997, 2000, and 2007 sought to maintain high recontact rates with these dynasty households while also surveying the cadet households that were generated when a respondent in a dynasty household moved. The targeting of cadet households to survey followed certain rules to keep the sample, once weighted, closely representative of the original 1993 population in the 13 IFLS provinces. In this sense the sample size of the survey both in households and individuals grows with each wave. Recontact rates are quite high, reaching rates in the mid 90s. By the last wave 13,995 households were contacted, a little over half of which were dynasty households, comprising 44,103 individuals.

In the 1993 wave the head of household, spouse, other adults in the household (with a maximum interview of four adults), and two random children were targeted for interviews. Starting with the 1997 wave the procedure changed to interview all household members in the dynasty households (so that those in 1993 who were not interviewed now have interviews, conditional on still living in the household and being alive). Interviews within cadet households were restricted to the core family, of which at least one member must have been a dynasty household member in a previous wave.

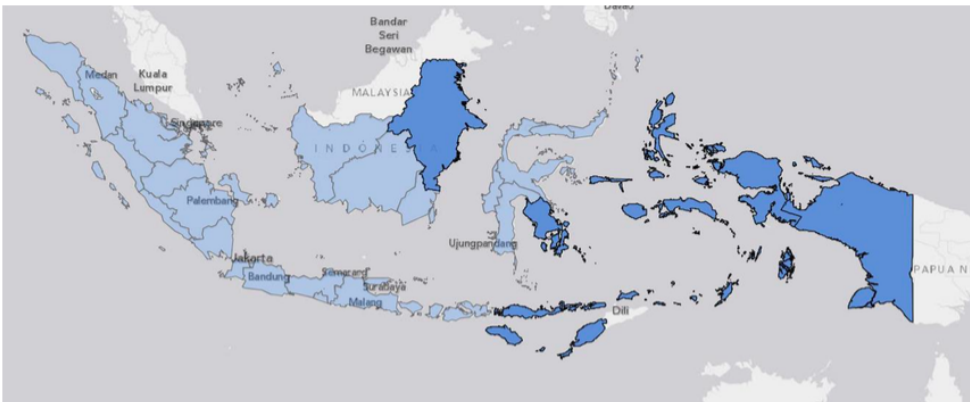
[26] Specifically the islands of Sumatra, Java, Bali, Kalimantan (Borneo), Sulawesi, Nusa Tenggara Barat.

A.1 Indonesia

Below are two maps of Indonesia, highlighting the areas where the IFLS surveys were administered (shaded regions). The IFLS-East 2012, a recently available data set, supplements the original RAND Corp. survey to account for the eastern provinces, which are not sampled in the parent survey. Indonesia currently consists of 34 provinces, 404 districts, 6543 sub-districts, and over 75,000 villages/towns/cities spread across 8 islands groups. The inclusion of these eastern provinces is important since they tend to be more rural and poorer than the western provinces covered in the original IFLS data set.

Figure A.1: RAND Corp. IFLS Provinces

Figure A.2: SurveyMETER IFLS-East Provinces



A.1.1 Indonesian Geographical Units and Population

From the Indonesian Statistical Office (BPS) I obtained the aggregated data on the geographical subdivisions of Indonesia and the populations from the 2010 census, presented in the table below. A feature of Indonesia over the past several decades has been the increasing expansion in the set of geographical regions. In 1993, the first year of the IFLS, there were 26 provinces. By 2000 four more provinces were created from splits of the previous provinces; a further four more provinces resulted from splits by 2007 (during this period East-Timor was recognized as a sovereign country, breaking ties with Indonesia). Currently there are 34 provinces.[27] The master files of geographical codings available at Statistics Indonesia's website also indicates that districts and sub-districts have split, resulting in expanding sets at the finer levels of geographical codings. Because the geographical codes in the migration histories data set have not yet been cleaned, this indicates that my measures of migration are currently underestimated.[28]

Figure A.3: 2010 Census and Geographical Subdivisions in Indonesia

Province Name	Capital	IFLS or IFLS-East Province	Population at 2010 Census	Area (km²)	Population density	Island Group	Number of Districts (kabupaten)	Number of sub-Districts (kecamatan)	Villages (kelurahan)	Municipalities (kotamadya)
Special Region of Aceh	Banda Aceh		4494410	57956	77	Sumatra	18	275	6420	5
Bali	Denpasar	X	3890757	5780	621	Lesser Sunda Islands	8	57	698	1
Bangka–Belitung Islands	Pangkal Pinang		122296	16424	64	Sumatra	6	43	361	1
Banten	Serang		10632166	9662	909	Java	4	154	1530	4
Bengkulu	Bengkulu		1715518	19919	84	Sumatra	9	116	1442	1
Central Java	Semarang	X	32382657	40800	894	Java	29	573	8577	6
Central Kalimantan	Palangkaraya		221089	153.564	14	Kalimantan	13	120	1439	1
Central Sulawesi	Palu		2635009	61841	41	Sulawesi	10	147	1712	1
East Java	Surabaya	X	37476757	47799	828	Java	29	662	8502	9
East Kalimantan	Samarinda	O	3026060	139462	22	Kalimantan	6	89	1023	3
East Nusa Tenggara	Kupang	O	4683827	48718	92	Lesser Sunda Islands	20	286	2775	1
Gorontalo	Gorontalo		1040164	11257	94	Sulawesi	5	65	595	1
Jakarta (Special Capital Region)	Jakarta	X	9607787	664	12786	Java	1	44	267	5
Jambi	Jambi		3092265	50058	57	Sumatra	9	128	1319	2
Lampung	Bandar Lampung	X	760405	34623	226	Sumatra	12	206	2358	2
Maluku	Ambon	O	1533506	46914	32	Maluku Islands	9	76	898	2
North Kalimantan	Tanjung Selor		622350	71176	10	Kalimantan	4	47	381	1
North Maluku	Sofifi	O	103.087	31982	31	Maluku Islands	7	109	1041	2
North Sulawesi	Manado		2270596	13851	162	Sulawesi	11	150	1510	4
North Sumatra	Medan	X	12982204	72981	188	Sumatra	25	408	5649	8
Special Region of Papua	Jayapura	O	2833381	319036	8	Western New Guinea	28	330	3583	1
Riau	Pekanbaru		5538367	87023	52	Sumatra	10	153	1500	2
Riau Islands Province	Tanjung Pinang		1679163	8201	208	Sumatra	5	59	331	2
Southeast Sulawesi	Kendari	O	2232586	38067	51	Sulawesi	10	199	1843	2
South Kalimantan	Banjarmasin	X	3626616	38744	96	Kalimantan	11	151	1973	2
South Sulawesi	Makassar	X	8034776	46717	151	Sulawesi	26	301	2874	3
South Sumatra	Palembang	X	7450394	91592	86	Sumatra	11	217	2869	4
West Java	Bandung		43053732	35377	1176	Java	17	625	5827	9
West Kalimantan	Pontianak		4395983	147307	30	Kalimantan	12	175	1777	2
West Nusa Tenggara	Mataram	X	4500212	18572	234	Lesser Sunda Islands	8	116	913	2
Special Region of West Papua	Manokwari	O	760422	97024	8	Western New Guinea	10	149	1291	1
West Sulawesi	Mamuju		1158651	16787	73	Sulawesi	5	66	564	0
West Sumatra	Padang	X	5133989	42012	110	Sumatra	12	169	964	7
Special Region of Yogyakarta	Yogyakarta	X	3457491	3133	1138	Java	4	78	438	1
Total	**34**	**22**	**227045689**	**1771612.564**		**8**	**404**	**6543**	**75244**	**98**

Note: An X identifies an IFLS province, while a O identifies an IFLS-East province. Source data from the Indonesian Stasitical Office (BPS) 2010 Census

A.2 Adult Descriptives

In this section I present the descriptive statistics of the adults in the IFLS data set. Because I am concerned with the finished schooling of children whose parents have migrated while they were in school, I look for those individuals who have finished their

[27]As reported by Statistics Indonesia `http://www.bps.go.id/website/fileMenu/Perka-BPS-No--151-Tahun-2014--Kode-dan-Wlayah-Kerja-Statistik-Tahun-2014.pdf`

[28]Lidia Farré was kind enough to provide me a file that will be used to clean the regional codes, developed by Benjamin Olken at MIT. This file can currently clean codes up to 2002. Given the previous discussion it will have to be updated to account for geographical changes.

pre-tertiary schooling. As the survey only explicitly asks the question if an individual has left school to those who are younger than 24 years of age, for those who are older than 24 I take their schooling as given and consider it complete. Table A.4 shows some descriptive statistics for adults I am able to identify. I also stratify the sample according to those who have ever migrated and those who have not. In total I identify 47,159 adults

Figure A.4: Adult Descriptive Statistics

Characteristics		All Adults (age>14)		Migrants (Any Migration)		Non-Migrants	
		mean	stdev	mean	stdev	mean	stdev
Sex (Male)		48.4%	50.0%	49.91%	50.00%	47.23%	49.92%
Age (last survey)		39.0	17.1	39.11	16.12	38.91	0.11
Education	No School	10.9%		6.38%		13.08%	
	Primary	34.9%		30.75%		36.96%	
	Lower Secondary	16.6%		17.47%		16.17%	
	Upper Secondary	24.3%		28.30%		22.31%	
	College	10.1%		14.79%		7.77%	
	Obs	47159		21270		25889	
Years of Schooling		7.55	4.30	8.08	4.17	7.11	4.35
	Obs	45626		20764		24862	
Marriage	Unmarried	18.72%		10.59%		25.40%	
	Married	68.88%		77.80%		61.55%	
	Sep./Est./Div.	3.08%		3.41%		2.80%	
	Widow	9.04%		8.01%		9.88%	
	Don't Know/Miss	0.28%		0.18%		0.37%	
	Obs	47153		21268		25885	

A.2.1 Marriage

Below is a marriage contingency table describing the assortative matching present in Indonesia. This table encompasses the the parents of the identified children who have finished their schooling and constructed following Greenwood et al. (2014). This table will be useful in the structural estimation to exogenously assign men and women according to the nonparametrically estimated probability of marriage.[29] The contingency table displays the observed frequencies in the left-hand side of a column and the expected frequencies based on random matching on the right-hand side. I further highlight in red the highest observed frequencies (and second highest when it seems large). Perfect assortative matching is described by the diagonal of the column. As the figure shows, there is a fair amount of assortative matching in Indonesia.

[29]As I do not explicitly intend to explain marriage decisions. A similar exercise will be conducted for female fertility.

Figure A.5: Marriage Contingency Table

Husband / Wife	No School		Primary		Lower Secondary		Upper Secondary		College	
No School	8.8%	2.8%	4.0%	7.0%	0.1%	1.5%	0.1%	1.4%	0.0%	0.5%
Primary	11.2%	11.0%	36.4%	27.5%	3.1%	6.0%	1.1%	5.5%	0.3%	2.1%
Lower Secondary	0.8%	2.8%	7.0%	6.9%	3.4%	1.5%	1.5%	1.4%	0.3%	0.5%
Upper Secondary	0.2%	3.0%	4.3%	7.6%	3.9%	1.7%	4.9%	1.5%	1.1%	0.6%
College	0.0%	1.5%	1.0%	3.8%	0.9%	0.8%	2.9%	0.8%	2.4%	0.3%
Marginal	21.1%		52.8%		11.5%		10.5%		4.1%	
Total Couples	7440									
Pearson Chi2(16)	5.2e+03									

A.2.2 Migration

I present here the described data of adult migration across all the IFLS waves. When observations are not reported I report the standard error of the mean to conserve space. I also stratify the sample according to different types of regional migrations that individuals may have ever engaged, where movements toward the right of the table can be generalized as migrations that could constitute increasing distance.[30]

Figure A.6: Adult Migration Events and Shares

Migration Characteristics		Any Migration mean	Any Migration stdev	Inter-District mean	Inter-District stdev	Inter-Provincial mean	Inter-Provincial stdev	Island Hoppers (Interprov) mean	Island Hoppers (Interprov) stdev
Migrant Shares	Whole Sample	45.1%	49.8%	32.0%	46.7%	17.9%	38.3%	4.4%	20.5%
	Among Migrants			71.0%	45.4%	39.6%	48.9%	9.8%	29.8%
Age		39.11	16.12	38.71	15.82	38.60	15.59	40.17	15.46
	Observations	21270		15096		8430		2093	
Years of Shooling		8.08	4.17	8.53	4.05	8.61	3.99	8.95	3.98
	Observations	20764		14748		8242		2058	
			std err		std err		std err		std err
Moves per Movers	(Wt. Avg):	1.73	0.01	1.94	0.01	2.12	0.01	2.73	0.03
	Cohorts								
	15-24	1.89	0.01	2.10	0.01	2.26	0.02	2.72	0.04
	25-34	1.67	0.01	1.87	0.02	2.04	0.02	2.87	0.06
	35-44	1.45	0.01	1.58	0.02	1.73	0.03	2.51	0.11
	45-54	1.34	0.02	1.47	0.03	1.56	0.05	2.23	0.14
	55-64	1.28	0.02	1.33	0.03	1.41	0.06	2.21	0.21
	65+	1.34	0.03	1.47	0.05	1.58	0.09	2.40	0.30
	Observations	30302		19959		10739		2410	

The figures below graph the migration events according to cohorts. As is usually the case in migration literature, I observe a decreasing average of migration events with older cohorts. There is an interesting reversal in this trend for the oldest cohort. This may represent retirement-type migrations as I observe parents living in the households of their older children. The second figure reports rural and urban migrations, which shows that there is little difference between the two types across cohorts.

[30]In reading the tables I point out that movements toward the right-hand side columns (the larger geographical subdivisions) are not strict subsets, as individuals over their reported histories may have engaged in several types of migration events. Shares, however, still have a per capita interpretation.

In these figures I note that an inter-district move is a migration event between districts within the same province.

Figure A.7: Average Migration Events across Cohort

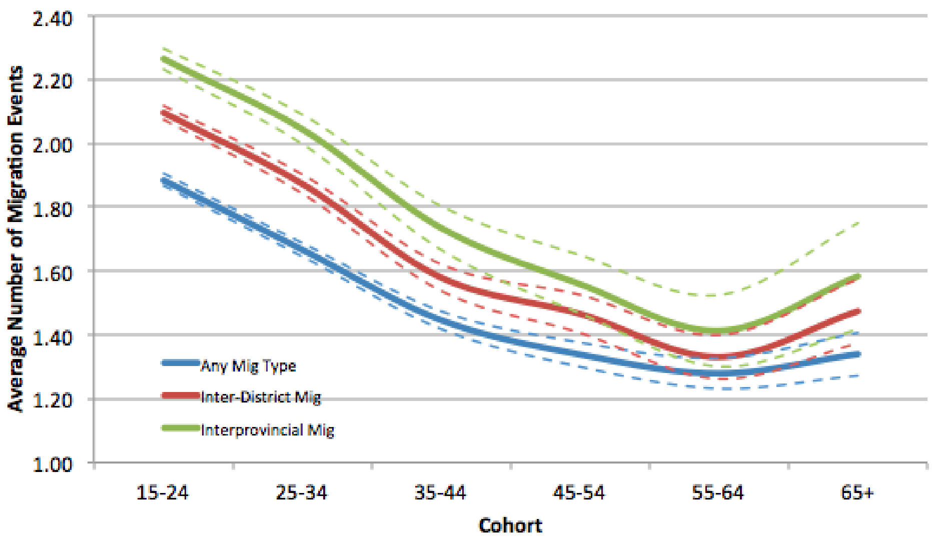
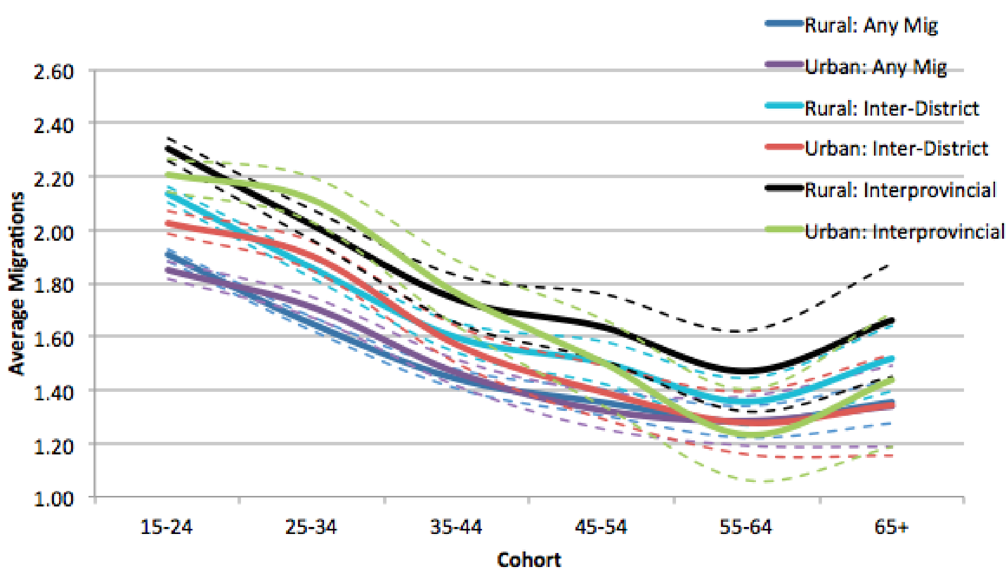


Figure A.8: Urban Rural Migration Events across Cohorts



As I ultimately aim to structurally estimate the research question, given the foreseeable computational limitations associated with the expanded state space of estimating the finest level of migration locations, I "zoom" into this subset of the data with inter-district migration events.[31]

[31] I define as an inter-district migration one in which the person crossed districts, not to be confused with the definition used to construct the figures above. Since a migration to another district can include a district in another province, this categorization includes inter-provincial migration. In the analysis section I make use of this definition to create the migration dummies.

Figure A.9: Inter-District Migration Shares Across Cohorts

	Inter-District Migration Events					
	Migrant Shares (Whole Sample)		Sex (Male - Movers)		Birth Urbanization	
Cohorts	mean	std err	mean	std err	mean	std err
15-24	38.81%	0.28%	50.90%	0.47%	31.22%	0.44%
25-34	24.79%	0.28%	58.80%	0.64%	35.26%	0.64%
35-44	12.64%	0.27%	59.90%	1.14%	34.81%	1.15%
45-54	7.27%	0.26%	56.29%	1.87%	31.74%	1.83%
55-64	5.12%	0.27%	51.46%	2.71%	27.24%	2.52%
65+	3.61%	0.26%	44.26%	3.68%	18.71%	2.99%
Obs	87361		20395		19398	

Figure A.9 reports the shares of migrations among the different cohorts with this type of migration. Interestingly it describes an upside-down U shape to the sex of migrants, with the oldest cohort predominantly composed of women. Migrants further differ on the educational attainment dimension. Figure A.10 below details the average migration events per cohort according to the levels of education completed, along with the share of individuals. On average, those with higher education generally have more events.

Figure A.10: Average Migration Events per Cohort and Education Levels

Migrant Characteristics		Inter-District Migration Events									
		No Schooling		Up to Full Primary		Up to Full Secondary		Up to Full Tertiary		Up To Completion of College	
Cohorts		mean	std err	mean	std err	mean	std err	mean	std err	mean	std err
15-24	# moves/mover	1.75	0.05	2.06	0.02	2.14	0.03	2.10	0.02	2.19	0.03
	share	40.45%		35.50%		35.11%		39.15%		52.60%	
25-34	# moves/mover	1.55	0.05	1.86	0.03	1.80	0.03	1.87	0.03	2.03	0.04
	share	25.83%		19.93%		22.55%		26.00%		37.28%	
35-44	# moves/mover	1.58	0.08	1.63	0.04	1.58	0.06	1.46	0.04	1.63	0.06
	share	14.06%		10.78%		12.70%		13.48%		17.32%	
45-54	# moves/mover	1.42	0.08	1.45	0.04	1.38	0.08	1.64	0.10	1.47	0.10
	share	7.44%		6.70%		6.65%		7.93%		10.69%	
55-64	# moves/mover	1.26	0.05	1.36	0.06	1.36	0.10	1.43	0.13	1.29	0.14
	share	4.91%		4.48%		7.73%		5.87%		7.17%	
65+	# moves/mover	1.56	0.08	1.45	0.08	1.31	0.13	1.17	0.17	1.33	0.33
	share	3.29%		3.93%		4.71%		2.63%		3.90%	

Migrants also tend to engage in repeat migrations, indicating a fair amount of dynamics. Figure A.11 reports the total shares of migrants with more than one move.

Figure A.11: Share of Repeat Movers

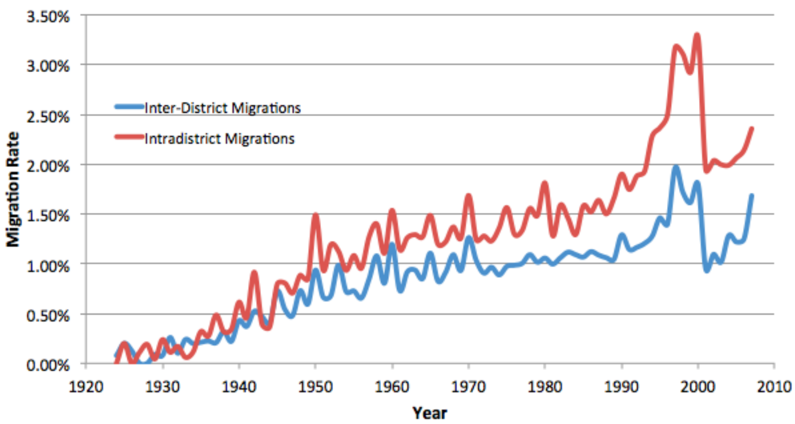
	Repeat Moves (>1 move)			
	Any Migration		Inter-District Migration	
Cohorts	mean	std err	mean	std err
15-24	51.60%	0.40%	36.11%	0.45%
25-34	40.94%	0.50%	26.58%	0.57%
35-44	31.33%	0.81%	20.29%	1.38%
45-54	24.77%	1.21%	18.95%	1.48%
55-64	22.51%	1.75%	17.25%	2.05%
65+	26.55%	2.40%	16.94%	2.78%
Total (Wt.Avg.)	44.26%	0.29%	30.81%	0.32%

I report the yearly migration rates I observe in the data in the graph below.[32] The average yearly migration rates (the combination of the two time-series) is about

[32]These two categorizations in a given year are not exactly mutually exclusive since the survey

2.09%, with an estimated growth rate of 0.04%/year. This decomposes to an average migration rate of 1.25% with a growth of 0.03%/year for intradistrict migrations, and a rate of 0.85% with a growth of 0.02%/year for inter-district migrations. An interesting feature of the migration rate is the quick growth prior to the 1997 Asian Financial Crisis, the sudden drop in 2000, and the subsequent steeper growth of the inter-district migration rate vis-à-vis the intradistrict rate after 2001.

Figure A.12: Yearly Migration Rates



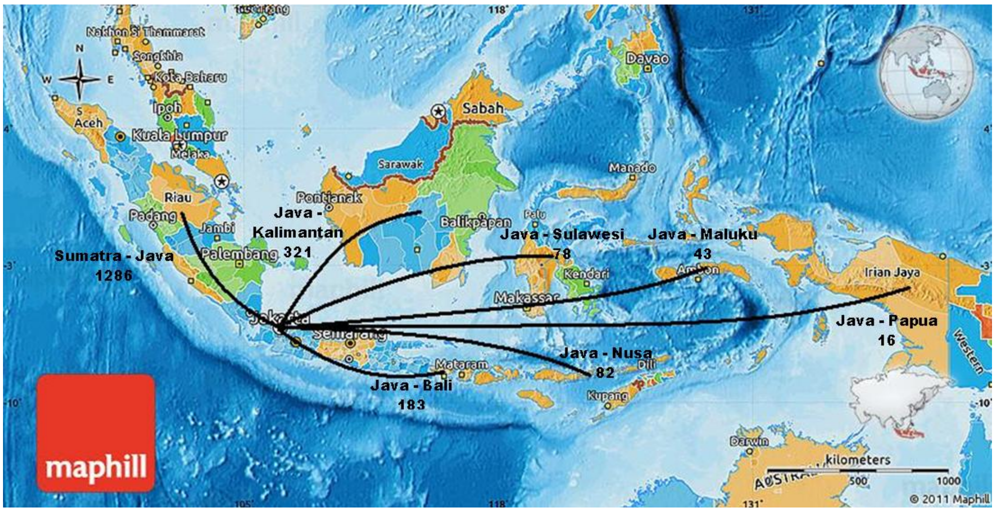
In figure A.6 I reported the island hoppers in the IFLS data set (the special interprovincial move between island groups). I notice in the distribution of island hops that migrants tend to hop to their nearby islands. The figures below report the frequency of island hops I observe in the data between the 8 major island groups as distinguished in the census data presented in figure A.3, and the movements to-and-from the island of Java superimposed on a map of Indonesia. It is not unusual that this island would receive the largest share of movements, as it is the most populous and contains the capital of Jakarta.

Figure A.13: Distribution of Island Hops in the IFLS

Island Hopper Freq	Sumatra	Java	Bali	Kalimantan	Nusa (East-West)	Sulawesi	Maluku (North-South)
Java	1286	-	-	-	-	-	-
Bali	9	183	-	-	-	-	-
Kalimantan	39	321	4	-	-	-	-
Nusa (East-West)	0	82	57	20	-	-	-
Sulawesi	0	78	17	84	28	-	-
Maluku (North-South)	2	43	0	4	0	20	-
Papua (West-East)	6	16	3	1	3	15	3
Total	**1342**	**723**	**81**	**109**	**31**	**35**	**3**

asks for migration events that lasted longer than 6 months, leaving open the possibility of observing at most 2 types of events in the same year. I do not find any instance of this in the data, however.

Figure A.14: Java Island Hops in IFLS



A.3 Adult Children

As the research question concerns itself with educational attainment, I restrict myself to the analysis of those who have completed schooling. In general, these children are a sub sample of the adults from the previous section. However, I also include children younger than 15 who were coded as completing their schooling.

To find and link children to their parents I utilize the relationship coding and the variable that identifies the mother and father in the survey. I restrict this sample to those who are coded as the biological children of the head of household. Therefore, I leave out other potential children (such as grandchildren, nieces/nephews, cousins) for the time being. Descriptive statistics of these adult children and their parents are located below in figure A.17. I do not stratify the sample according to inter-district or intradistrict migration events of the parents to get an overall picture of the pattern. The figures below detail characteristics of these families and a correlation of the parents migration.

Figure A.15: Descriptive Statistics of Indonesian Families

Family	All Valid Children		In School: Parents Don't Mig		In School: Father Any Mig		In School: Mother Any Mig		In School: Parents Any Mig		Kwallis Test
	mean	std err	mean	std err	mean	std err	mean	std err	mean	std err	of Diff
Num of Children (Avg of Birth Order)	2.23	0.02	2.24	0.02	2.22	0.08	2.14	0.07	2.18	0.06	p=0.0986
Avg Educ. Parents (years)	5.24	0.04	4.95	0.05	6.42	0.23	6.38	0.22	7.50	0.17	p=0.0001
Parents Married	88.09%	0.37%	87.07%	0.41%	97.58%	0.91%	80.82%	2.21%	100.00%	0.00%	p=0.0061
Average Size (persons)	4.04	0.02	4.04	0.02	4.11	0.08	3.86	0.08	4.18	0.06	p=0.1642
Obs	7672		6553		289		318		512		

Figure A.16: Correlation of Parental Migration

Correlation	Mother Migrates
Father Migrates	0.6407***

Figure A.17: Descriptive Statistics of Indonesian Children

Children		All Valid Children		In School: Parents Don't Mig		In School: Father Any Mig		In School: Mother Any Mig		In School: Parents Any Mig	
		mean	std err	mean	std err	mean	std err	mean	std err	mean	std err
Sex (Male)		51.6%	0.4%	51.3%	0.4%	52.9%	1.6%	50.0%	1.6%	54.2%	1.4%
Age (Finish School)		15.38	0.03	15.20	0.08	15.57	0.11	16.14	0.09	16.55	0.08
Birth Year		1979.3	0.1	1979.2	0.1	1980.1	0.2	1978.5	0.3	1979.6	0.2
Years of Schooling		9.30	0.03	9.12	0.03	10.18	0.09	10.51	0.08	10.81	0.08
Grade Repitions (at least 1 year)		19.5%	0.3%	19.3%	0.3%	19.3%	1.3%	20.2%	1.3%	18.1%	1.1%
Education	No School	2.5%		2.2%		0.0%		0.0%		0.0%	
	Primary (or less)	29.6%		29.6%		18.7%		14.0%		10.9%	
	Lower Secondary	20.7%		21.0%		21.9%		20.5%		16.8%	
	Upper Secondary	33.1%		32.9%		35.6%		41.9%		40.5%	
	College	16.6%		14.3%		23.7%		23.6%		31.7%	
	Obs	17189		13970		918		1012		1289	
Attended Kindergarten?		21.2%	0.4%	20.0%	0.4%	26.2%	1.8%	22.3%	1.6%	29.1%	1.3%
	Obs	13540		11022		626		663		1229	
School Administration (Public)		92.6%	0.2%	92.6%	0.3%	94.5%	0.9%	92.7%	1.0%	92.5%	0.8%
	Obs	12936		10513		595		633		1195	
Indonesian Proficiency	Speak	47.9%	0.4%	48.1%	0.5%	48.8%	2.0%	48.1%	1.9%	45.6%	1.4%
	Obs	13617		11085		631		672		1229	
	Read	96.9%	0.1%	96.6%	0.2%	97.8%	0.6%	98.7%	0.4%	98.9%	0.3%
	Obs	13633		11097		631		673		1232	
	Write	96.5%	0.2%	96.2%	0.2%	97.6%	0.6%	98.4%	0.5%	98.4%	0.4%
	Obs	13633		11097		631		673		1232	
Urbanization Birth Location: Urban		37.1%	0.4%	33.8%	0.5%	43.7%	2.0%	51.9%	1.9%	55.5%	1.4%
	Obs	13544		11033		622		663		1226	
Age 12 location different from birth:		13.5%	0.3%	8.6%	0.3%	19.5%	1.6%	28.3%	1.7%	47.1%	1.4%
	Obs	13617		11088		630		668		1231	

The parents of these children can be identified via the parental coding. In order to be able to find these children I have to observe their parents in the same household in at least one wave. Since the average child within the groupings in the above table was born, at the earliest, in 1978 this would have made them 15 years of age in the 1993 wave of the IFLS. A little over 1/3 of my sample falls below this birth year. It will be necessary to look at the pregnancy histories of women in the household, and specifically the parents of these children, to ascertain the degree of possible selection that is induced in the survey. For the moment I ignore this possibility as multi-generation households seems to be a feature of Indonesia, and adult children tend to live with their parents (Indonesia, 2013). The below figure describes some statistics of these parents.

Figure A.18: Descriptive Statistics of Indonesian Parents

Parents		All Valid Children		In School: Parents Don't Mig		In School: Father Any Mig		In School: Mother Any Mig		In School: Parents Any Mig	
		mean	std err	mean	std err	mean	std err	mean	std err	mean	std err
Father											
Age (First Child Born)		28.52	0.12	28.77	0.14	27.09	0.46	27.69	0.48	27.71	0.27
	Obs	6545		5225		351		273		696	
Years of Schooling		6.07	0.05	5.77	0.06	6.80	0.24	7.42	0.27	8.29	0.18
	Obs	6399		5388		283		227		501	
Urbanization of Birth		21.9%	0.6%	20.1%	0.6%	25.6%	2.6%	25.9%	3.2%	33.3%	2.2%
	Obs	5084		4151		277		185		471	
Mother											
Age (First Child Born)		22.65	0.12	22.93	0.12	20.53	0.48	21.38	0.43	21.94	0.24
	Obs	7331		5939		312		384		696	
Years of Schooling		4.98	0.05	4.72	0.05	6.11	0.27	6.09	0.23	6.97	0.19
	Obs	7189		6116		254		313		506	
Urbanization of Birth		23.9%	0.6%	21.7%	0.6%	26.9%	2.9%	33.6%	2.7%	38.7%	2.2%
	Obs	5935		4907		238		301		489	

Children's educational attainment is decomposed by urbanization and sex and by inter-district and intradistrict parental migration. Substantial heterogeneity is apparent along urbanization and sex, while parental migration by itself does not display much change.

Figure A.19: Inter- & Intradistrict Parental Migration and Education

In School: Parents Don't Mig	In School: Father Intradistrict Mig	In School: Father Inter-District Mig
9.02	10.12	10.21
(0.03)	(0.133)	(0.134)
In School: Parents Don't Mig	**In School: Mother Intradistrict Mig**	**In School: Mother Inter-District Mig**
9.02	10.60	10.44
(0.03)	(0.129)	(0.109)
In School: Parents Don't Mig	**In School: Parents Intradistrict Mig**	**In School: Parents Inter-District Mig**
9.02	11.14	10.76
(0.03)	(0.132)	(0.135)

Figure A.20: Education by Urbanization

Cross Tab: Children's Education (years)	Rural	Urban
Girls	8.49 (0,06)	10.63 (0,06)
Boys	8.74 (0,05)	10.61 (0,06)

Figure A.21: Education by Urbanization and Parental Migration

Cross Tab: Children's Education (years)	In School: Parents Don't Mig	In School: Father Intradistrict Mig	In School: Father Inter-District Mig
Rural Child	8.47 (0.041)	9.51 (0.195)	9.67 (0.199)
Urban Child	10.44082 (0.049)	11.15054 (0.170)	11.40 (0.180)
Cross Tab: Children's Education (years)	In School: Parents Don't Mig	In School: Mother Intradistrict Mig	In School: Mother Inter-District Mig
Rural Child	8.47 (0.041)	10.03 (0.179)	9.79 (0.231)
Urban Child	10.44082 (0.049)	11.40304 (0.132)	11.03 (0.172)
Cross Tab: Children's Education (years)	In School: Parents Don't Mig	In School: Parents Intradistrict Mig	In School: Parents Inter-District Mig
Rural Child	8.47 (0.041)	10.44 (0.244)	9.73 (0.245)
Urban Child	10.44082 (0.049)	11.87879 (0.138)	11.71 (0.142)

B Appendix: Results of Empirical Analysis

B.1 Cross-Sectional Analysis

This section reports the regression results from the cross-sectional analysis discussed in section 5.1. The three figures below report the results of estimating equation (1).

Figure B.1: Cross-Sectional Analysis: Father Ever Migrated - Child in School

OLS: School Years	(1)	(2)	(3)	(4)	(5)
Father's Migration	0.778***	0.577***	0.309***	0.512***	0.215*
Inter-district	(0.151)	(0.138)	(0.116)	(0.135)	(0.111)
Father's Migration	0.775***	0.459***	0.418***	0.418***	0.212*
Intradistrict	(0.144)	(0.136)	(0.122)	(0.141)	(0.121)
Interaction	-0.397	-0.211	-0.232	-0.0404	-0.0492
	(0.303)	(0.287)	(0.255)	(0.302)	(0.242)
Individual Controls		X			X
Father's Schooling			X		X
Child Schooling Controls				X	X
Observations	6,006	5,069	5,905	5,073	4,967
R-squared	0.011	0.136	0.328	0.212	0.434

Robust standard errors in parentheses, clustered at the family level. The output reports the levels of schooling attainment. The dependant variable is the total attained schooling years. Controls: individual controls consist of the birth urbanization, birth order, sex, and interactions of sex with the other two controls, as well as birth year dummies; father's schooling is the years of father's education and birth year dummies interacted with the father's education; child schooling consists of kindergarten participation, a dummy for ever repeating a grade, proficiency in Bhasa Indonesian (excluding linguistic proficiency). Observations are total individuals. Reported signifcance are at: *** p<0.01, ** p<0.05, * p<0.1

Figure B.2: Cross-Sectional Analysis: Mother Ever Migrated - Child in School

OLS: School Years	(1)	(2)	(3)	(4)	(5)
Mother Ever Migrates	0.726***	0.512***	0.361***	0.446***	0.258**
Inter-district	(0.147)	(0.138)	(0.129)	(0.145)	(0.125)
Mother Ever Migrates	0.734***	0.478***	0.321***	0.463***	0.186
Intradistrict	(0.140)	(0.133)	(0.118)	(0.134)	(0.117)
Interaction	-0.249	-0.174	-0.392	-0.0614	-0.190
	(0.311)	(0.284)	(0.248)	(0.268)	(0.229)
Individual Controls		X			X
Mother's Schooling			X		X
Child Schooling Controls				X	X
Observations	6,030	5,090	5,944	5,095	5,004
R-squared	0.011	0.137	0.293	0.211	0.408

Robust standard errors in parentheses, clustered at the family level. The output reports the levels of schooling attainment. The dependant variable is the total attained schooling years. Controls: individual controls consist of the birth urbanization, birth order, sex, and interactions of sex with the other two controls, as well as birth year dummies; mother's schooling is the years of mother's education and birth year dummies interacted with the mother's education; child schooling consists of kindergarten participation, a dummy for ever repeating a grade, proficiency in Bhasa Indonesian (excluding linguistic proficiency). Observations are total individuals. Reported signifcance are at: *** p<0.01, ** p<0.05, * p<0.1

Figure B.3: Cross-Sectional Analysis: Parents Ever Migrated - Child in School

OLS: School Years	(1)	(2)	(3)	(4)	(5)
Parents Ever Migrate	0.931***	0.698***	0.266**	0.597***	0.200*
Inter-district	(0.135)	(0.126)	(0.110)	(0.133)	(0.111)
Parents Ever Migrate	1.101***	0.836***	0.427***	0.757***	0.316***
Intradistrict	(0.130)	(0.129)	(0.109)	(0.124)	(0.111)
Interaction	-0.216	-0.274	-0.0719	0.0324	-0.0317
	(0.240)	(0.245)	(0.254)	(0.260)	(0.242)
Individual Controls		X			X
Mother's Schooling			X		X
Child Schooling Controls				X	X
Observations	6,121	5,162	6,034	5,166	5,075
R-squared	0.020	0.143	0.350	0.216	0.445

Robust standard errors in parentheses, clustered at the family level. The output reports the levels of schooling attainment. The dependant variable is the total attained schooling years. Controls: individual controls consist of the birth urbanization, birth order, sex, and interactions of sex with the other two controls, as well as birth year dummies; parents schooling is the average of parental education and birth year dummies interacted with the parents education; child schooling consists of kindergarten participation, a dummy for ever repeating a grade, proficiency in Bhasa Indonesian (excluding linguistic proficiency). Observations are total individuals. Reported signifcance are at: *** p<0.01, ** p<0.05, * p<0.1

B.2 Survival Analysis

This section reports the results of estimating the discrete time probit hazard described in maximizing eq (5). The results in the figures below report the predicted marginal effects at the mean of the covariates.

Figure B.4: Survival Analysis: Hazard Contribution of Father's Migration on School Exit

PROBIT: School Exit	(1)	(2)	(3)	(4)	(5)
Father's Migration	-0.0317***	-0.0250***	-0.0208***	-0.0198***	-0.0122***
Inter-district	(0.00102)	(0.000948)	(0.000857)	(0.000781)	(0.000669)
Father's Migration	-0.0318***	-0.0250***	-0.0211***	-0.0198***	-0.0121***
Intradistrict	(0.00102)	(0.000959)	(0.000875)	(0.000776)	(0.000669)
Baseline Hazard	X	X	X	X	X
Baseline Hazard*Mig	X	X	X	X	X
Individual Controls		X			X
Father's Schooling			X		X
Child Schooling Controls				X	X
Observations	85,685	74,012	84,257	74,025	72,497
Pseudo R-squared	0.310	0.310	0.310	0.310	0.310

Robust standard errors in parentheses, clustered at the family level. The output reports the marginal effects of the probit at the means of the regressors. The baseline hazard is nonparametrically specified as a sequence of dummies for each school grade. To correct for the non proportionality of the migation event, the event is interacted with the baseline hazard. Controls: individual controls consist of the birth urbanization, birth order, sex, and interactions of sex with the other two controls; father's schooling is the years of father's education; child schooling consists of kindergarten participation, a lag on schooling grade, proficiency in Bhasa Indonesian (excluding linguistic proficiency). Grade repeats are accounted for in the data. Observations are person-years. Reported signifcance are at: *** p<0.01, ** p<0.05, * p<0.1

Figure B.5: Survival Analysis: Hazard Contribution of Mother's Migration on School Exit

PROBIT: School Exit	(1)	(2)	(3)	(4)	(5)
Mother's Migration	-0.0315***	-0.0259***	-0.0214***	-0.0202***	-0.0126***
Inter-district	(0.00102)	(0.000989)	(0.000883)	(0.000802)	(0.000685)
Mother's Migration	-0.0318***	-0.0263***	-0.0218***	-0.0203***	-0.0128***
Intradistrict	(0.00102)	(0.00101)	(0.000894)	(0.000820)	(0.000694)
Baseline Hazard	X	X	X	X	X
Baseline Hazard*Mig	X	X	X	X	X
Individual Controls		X			X
Mother's Schooling			X		X
Child Schooling Controls				X	X
Observations	86,427	74,107	85,096	74,103	72,791
Pseudo R-squared	0.303	0.303	0.303	0.303	0.303

Robust standard errors in parentheses, clustered at the family level. The output reports the marginal effects of the probit at the means of the regressors. The baseline hazard is nonparametrically specified as a sequence of dummies for each school grade. To correct for the non proportionality of the migation event, the event is interacted with the baseline hazard. Controls: individual controls consist of the birth urbanization, birth order, sex, and interactions of sex with the other two controls; mother's schooling is the years of mother's education; child schooling consists of kindergarten participation, a lag on schooling grade, proficiency in Bhasa Indonesian (excluding linguistic proficiency). Grade repeats are accounted for in the data. Observations are person-years. Reported signifcance are at: *** p<0.01, ** p<0.05, * p<0.1

Figure B.6: Survival Analysis: Hazard Contribution of Parent's Migration on School Exit

PROBIT: School Exit	(1)	(2)	(3)	(4)	(5)
Parent's Migration	-0.0290***	-0.0222***	-0.0186***	-0.0179***	-0.0102***
Inter-district	(0.00117)	(0.00107)	(0.000946)	(0.000889)	(0.000721)
Parent's Migration	-0.0288***	-0.0222***	-0.0180***	-0.0175***	-0.00995***
Intradistrict	(0.00115)	(0.00108)	(0.000930)	(0.000861)	(0.000702)
Baseline Hazard	X	X	X	X	X
Baseline Hazard*Mig	X	X	X	X	X
Individual Controls		X			X
Parent's Schooling			X		X
Child Schooling Controls				X	X
Observations	61,222	52,645	60,338	52,627	51,685
Pseudo R-squared	0.302	0.302	0.302	0.302	0.302

Robust standard errors in parentheses, clustered at the family level. The output reports the marginal effects of the probit at the means of the regressors. The baseline hazard is nonparametrically specified as a sequence of dummies for each school grade. To correct for the non proportionality of the migation event, the event is interacted with the baseline hazard. Controls: individual controls consist of the birth urbanization, birth order, sex, and interactions of sex with the other two controls; parents schooling is the average of parental education; child schooling consists of kindergarten participation, a lag on schooling grade, proficiency in Bhasa Indonesian (excluding linguistic proficiency). Grade repeats are accounted for in the data. Observations are person-years. Reported signifcance are at: *** p<0.01, ** p<0.05, * p<0.1

B.2.1 Survival Plots

To construct the survival plots I generate the predicted data of a representative agent based on specification (5) of the estimation outputs in the previous section. In general this representative agent has the mean values of the covariates.[33] To understand the impact of inter-district migration events I set the indicator of the parent's migration to 1 (and 0 otherwise) to compare the plots against those whose parent does not migrate. For simplicity I report the graphs for the father's inter-district migration, as the coefficient estimates of the other parent(s) migration events are not substantially different.

Figure B.7: Survival Function: Father's Inter-district Migration

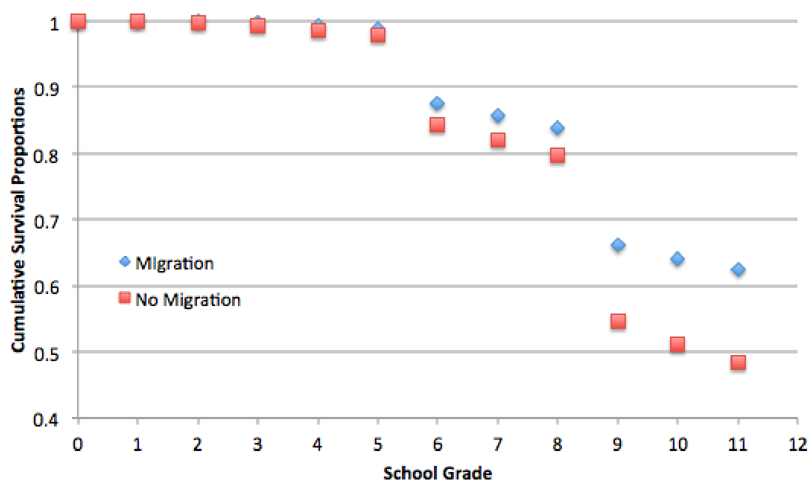
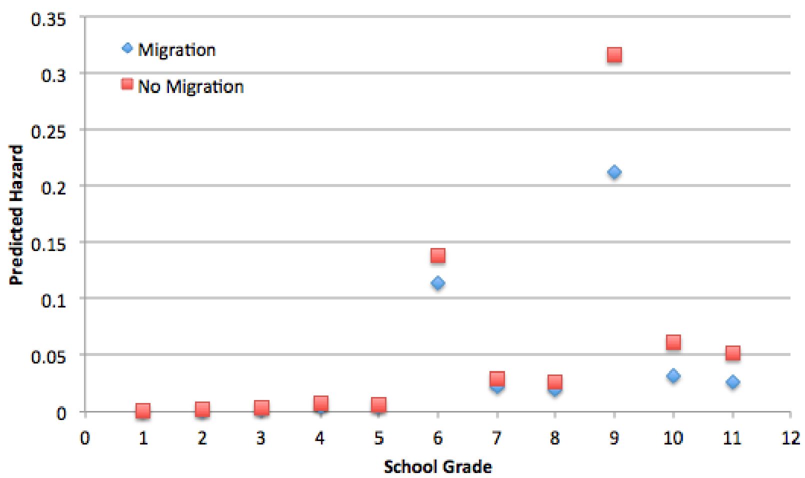


Figure B.8: Predicted Hazard Function: Father's Inter-district Migration



[33]For the proficiency in reading and writing Bhasa Indonesian I give a value of 1 for the indicator as the mean values reported in the descriptive tables is close to 100%.

Figure B.9: Survival Function: Father's Inter-district Migration and Rural Children

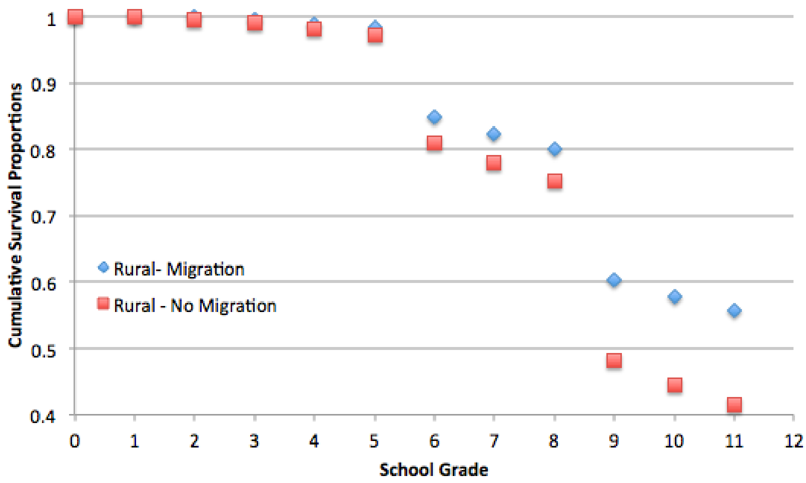
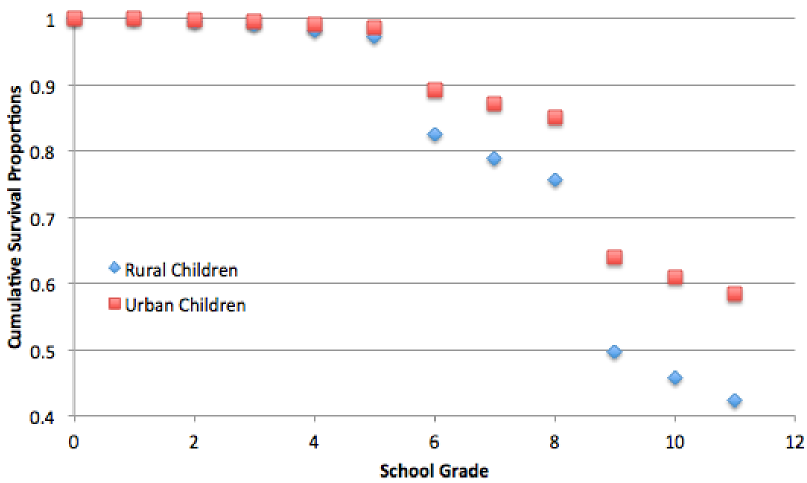


Figure B.10: Survival Function: Rural vs. Urban Children



References

Antman, F. M. (2012, October). Gender, educational attainment, and the impact of parental migration on children left behind. *Journal of Population Economics 25*(4), 1187–1214.

Baland, J.-M. and J. A. Robinson (2000, August). Is child labor inefficient? *The Journal of Political Economy 108*(4), 663–679.

Basu, K. and P. H. Van (1998, June). The economics of child labor. *American Economic Review 88*(3), 412–427.

Becker, G. S. and H. G. Lewis (1973, March). On the interaction between the quantity and quality of children. *Journal of Political Economy 81*(2), 279–288.

Becker, G. S. and N. Tomes (1976, August). Child endowments and the quantity and quality of children. *Journal of Political Economy 84*(4), S143–62.

Borjas, G. J. (1987, September). Self-selection and the earnings of immigrants. *The American Economic Review 77*(4), 531–553.

Borjas, G. J. (1999). The economic analysis of immigration. In A. O. and C. D. (Eds.), *Handbook of Labor Economics* (1 ed.), Volume 3, Part A, Chapter 28, pp. 1697–1760. Elsevier.

Bryan, G. and M. Morten (2015, February). Economic development and the spatial allocation of labor: evidence from indonesia. Stanford University.

Clifton-Sprigg, J. (2014, October). Out of sight, out of mind? educational outcomes of children with parents working abroad. University of Edinburgh.

Clifton-Sprigg, J. (2015, June). Educational spillovers and parental migration. *Labour Economics 34*, 64–75.

Cox, A. and M. Ureta (2003, December). International migration, remittances, and schooling: evidence from el salvador. *Journal of Development Economics 72*(2), 429–462.

Ducanes, G. and M. Abella (2009). Prospects for future outward flows: China and southeast asia. Technical report, ILO Asian Regional Programme on Governance of Labour Migration WP No. 24.

Duflo, E. (2001, September). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review 91*(4), 795–813.

Farré, L. and F. Fasani. (2012, November). Media exposure and internal migration - evidence from indonesia. *Journal of Development Economics 102*(C), 48–61.

Ferrone, L. and G. C. Giannelli (2015, March). Household migration and child educational attainment: the case of uganda. IZA Discussion Papers 8927.

Frankenberg, E. and L. Karoly (1995, November). The 1993 indonesian family life survey: overview and field report. Technical report, RAND, Santa Monica, CA.

Frankenberg, E. and D. Thomas (2000, March). The indonesia family life survey (ifls): study design and results from waves 1 and 2. Technical report, RAND DRU-2238/1-NIA/NICHD, Santa Monica, CA.

Gayle, G.-L., L. Golan, and M. A. Soytas (2015, February). What accounts for the racial gap in time allocation and intergenerational transmission of human capital? Washington Universtiy in St. Louis.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2014, May). Marry your like: assortative mating and income inequality. *American Economic Review 104*(5), 348–353.

Hanson, G. H. and C. Woodruff (2003, April). Emigration and educational attainment in mexico. University of California-San Diego.

Heckman, J. J. and B. Singer (1984, March). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica 52*(2), 271–320.

Heckman, J. J. and B. Singer (1986). The econometric analysis of longitudinal data. In Z. Griliches and M. lntriligator (Eds.), *Handbook of Econometrics*, Volume 3, pp. 1690–1761. Elsevier.

Hertz, T. and T. Jayasundera (2007, August). School construction and intergenerational mobility in indonesia. American University Working Paper Series No. 2007-18, Washington, D.C.

Hildebrandt, N. and D. J. McKenzie (2005, April). The effects of migration on child health in mexico. *Journal of LACEA Economia*.

Hu, F. (2013, July). Does migration benefit the schooling of children left behind? evidence from rural northwest china. *Demographic Research 29*(2), 33–70.

Hunt, J. (2012, October). The impact of immigration on the educational attainment of natives. NBER Working Papers 18047, National Bureau of Economic Research, Inc.

Indonesia, S. (2013). Indonesia demographic and health survey 2012. Technical report, National Population and Family Planning Board and Ministry of Health.

Keele, L. (2010, January). Proportionally difficult: testing for nonproportional zazards in cox models. *Political Analysis 18*, 189–205.

Klemp, M., C. Minns, P. Wallis, and J. Weisdorf (2013, May). Picking winners? the effect of birth order and migration on parental human capital investments in pre-modern england. *European Review of Economic History 17*(2), 210–232.

Kruger, D., R. R. Soares, and M. Berthelon (2012, January). Household choices of child labor and schooling: a simple model with application to brazil. *Journal of Human Resources 47*(1), 1–31.

Kuiper, J. C. (2011). Indonesia: a country study. Technical report, Library of Congress Federal Research Division.

Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics 95*, 391–413.

Lucas, R. E. B. (1997). Internal migration in developing countries. In M. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, Volume 1, Part B, pp. 721–798. Elsevier.

McKenzie, D. and H. Rapoport (2006, June). Can migration reduce educational attainments? depressing evidence from mexico. CReAM Discussion Paper Series 0601, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.

Montenegro, C. E. and H. A. Patrinos (2013). Returns to schooling around the world. Background paper for the World Development Report 2013.

Neyman, J. and E. Scott (1948, January). Consistent estimates based on partially consistent observations. *Econometrics 16*(1), 1–32.

Pan, Y. (2012, February). The effect of labor mobility restrictions on human capital accumulation in china. Working Papers 2012-5, The George Washington University, Institute for International Economic Policy.

Rahmi, U. (2011, February). An evaluation of indonesian national examination. Deakin University, Melbourne.

Raut, L. and L. H. Tran (2005, August). Parental human capital investment and old-age transfers from children: is it a loan contract or reciprocity for indonesian families. *Journal of Development Economics 77*(2), 389–414.

Strauss, J., F. Witoelar, B. Sikoki, and A. Wattie (2009, April). The fourth wave of the indonesian family life survey (ifls4): overview and field report. Technical report, RAND WR-675/1-NIA/NICHD.