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

Alberto José Ramírez*

June 10, 2016

Abstract

In my previous paper I documented evidence from the first four waves of the Indonesian Family Life Survey (IFLS) that parental migration has a correlated impact on the educational attainment of their children. Bivariate analysis shows that at the intensive margin these children are expected to complete more schooling years overall and complete more levels, finishing at least half of non-obligatory secondary education in Indonesia. Furthermore, a duration analysis estimated that parental migration early in the child's life was associated with about a 10% reduction in the hazard rate of exiting the school system at this level.

However, the previous analysis suffers from self-selection wherein parents sort themselves into migration states, coupled with the endogenous investment decision to educate a child. In this report I document my partial equilibrium dynastic model to account for this endogeneity. The start of a structural model is presented that captures the basic tradeoffs between choosing to migrate and investment in a child, but empirical specification is held off for the moment while a simulation study is undertaken. To this end a simple 2 period, 2 labor market Dynastic OLG model is distilled from the previous model to study if variation is generated by key economic parameters upon their perturbation in simulation. This analysis will facilitate guidance on how to specify the model for estimation and which data are necessary to identify the parameters.

[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; Nezih Guner and Jordi Caballe for their comments and input.

1 Introduction

While returns to education in the developing world are arguably higher than in the developed world (Montenegro and Patrinos, 2013), 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 it can be optimal to substitute away from education. Unless a sufficiently high wage is attained by parents to meet the sustenance level of household consumption, parents substitute the child's time toward labor (Basu and Van, 1998), even when schooling is considered socially optimal (Baland and Robinson, 2000). Notwithstanding, these theoretical results are not always predictive as other motives may influence parental decisions. For example, an exogenous increase in local economic activity may lead to a substitution even when conditioning on socio-economic status (Kruger et al., 2012).

Given the various mechanisms that may impact the decision to educate, a natural extension is to analyze whether parental migration is sufficient to relax the household budget constraint 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. As Kennan and Walker (2011) discuss, it is in the expectation of a permanent wage increase from differences in wages offered across locations (regardless of the source of difference) that induces the reallocation of labor supply through migration. But with migration comes uncertainty. So while 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 parent's migration into a different labor market would have on a child's educational attainment given the uncertainty associated with the choice through the income effect.

A propensity to migrate internally provides a country the mechanism to reallocate labor efficiently. And access to more local labor markets may have differential impacts to those children whose parents migrate. 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). What makes Indonesia an interesting study is its extensive internal migration where government policies have actively fostered transmigration. Given that the government has reduced barriers to internal migration high rates are prevalent, 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% per year. I find that migrant parents have higher education than their non-migrant counterparts, with roughly equal shares living in rural and urban areas.

The current research 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 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 while accounting for the tradeoffs associated with migration and education (the investment decisions in both), when parents are altruistically motivated toward their children and the future generations of their dynasty.

The previous, descriptive analysis conducted on the IFLS data set revealed a positive association between parental migration and childhood educational attainment. The data used is drawn from those children who have already completed their pre-tertiary education (and thus, fully observed). As Children in Indonesia have high rates of primary school completion, where only about 1/3 of children have up to primary schooling, the bivariate analysis shows that parental migration increases the expected schooling by up to 2 years (completing most 2/3 of the upper, non-obligatory secondary education). A duration analysis reveals that parental migration also decreases the hazard that children drop out of school, where the effect of migration is strongest at the secondary schooling level. 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.

In this current report I further document the results from reconstructing the wage history into a longitudinal panel data set, along with the dynastic structural model of parental investment, both in education and in the location choice. Regarding wages, descriptive analysis yields that a migration event in general is associated with a wage increase and that this increase is higher for those who move between provinces versus those who move within one. A fixed effect mincer regression estimates a 7% wage premium for an instantaneous move between provinces, while a move within a province yields a 3% wage premium. When the estimation is repeated on markets constructed by aggregating the provinces located within an island, the wage premium increases to 15% when migrating between islands vs. a 6.1% premium for migrating within an island. Importantly, when estimating all the models together in a seemingly unrelated regression I fail to reject the null hypothesis that the intraprovincial and intramarket migration premiums (of 3 and 6%) are statistically different, implying that markets close to each other seem to give in expectation the same boost as the market within a province. An update of these results to account for the post-migration premium is currently under analysis, as well as an analysis on the subset of adults whose parents migrated while they were young to understand if there is an impact to their labor market outcomes.

I present as well the skeleton of the structural model that captures the main tradeoffs that I wish to study, the investment in migration and the investment in education for households and considering only the first-born child.[1] Parents must decide whether moving to another location in anticipation of better labor market outcomes will increase their immediate welfare; and whether the investment in their child's education will likewise increase their welfare through the altruism they hold for the future outcome of their child. Moving implies a real, pecuniary cost. This decision to educate implies an ex-ante decrease in immediate welfare from the foregone income that the child could contribute to the household, captured through the cost of educating. The simple version of the model does not include a production function for educating. However, given that the survey includes administered cognitive tests, a measure of ability can be constructed. Coupled with the assumption that parental and child abilities are correlated due to heredity, this simple complication in the spirit of Cunha and Heckman (2008) and Cunha et al. (2010) can be easily incorporated.

Before empirically specifying the model, however, I first study in simulation a simpler version of the model to understand if the parameters capturing the economic tradeoffs are generating variation. Specifically, what is being tested upon solving the model is whether perturbing parameters (that will be chosen) yields differences in simulated outcomes to understand the extent of the variation needed to identify these parameters econometrically. To this end a 2 labor market, 2 period Dynastic OLG model is presented, along with the initial parameters selected. This portion is still in progress.

The rest of this report is organized as follows: section 2 elaborates on how my project fits within the related literature; section 3 relates the motivating descriptive evidence from the previous chapter along with the current results of the analysis of wages; section 4 contains the model; and appendices are presented for more detailed information on the education system in Indonesia, the nature of the dataset used, and the reduced form duration and mincer models and results from the analysis of the data.

2 Related Literature

My project is related to the current literature on migration and human capital development in its synthesis of the two strands; yet it departs from the current strand of literature on the effect of migration on the children of those migrants by the dynamic treatment of the topic. So while it concerns the movement of the whole household, this, and more broadly the topic of children-left-behind, is important and an open question in the migration literature. However, that I am studying the whole family within the migration event necessarily means that I am departing from this literature on the effect of missing parents in the household, where staggered migration or children-left-behind forms the

[1] The implication of this assumption is that the quality-quantity trade-off of multiple child households per Becker and Lewis (1973) and Becker and Tomes (1976) can not be studied.

main topic. The added disruption to the household of a single parent migration event may place the child at greater risk than a simple budget constraint issue may suggest, as these households may face varying consequences than just consumption allocation since the time investment in a child is completely lacking versus a migration where the whole family moves. So while the difference between an internal and external migration event is the type of labor market accessed, whether the family stays together or is temporarily disrupted adds a different complication that, for the moment, I am not considering.

This strand of the literature is resoundingly reduced form in the treatment of a typically dynamic problem of repeated migrations. 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 this problem (Borjas, 1999).

Other authors have reported results that give a more nuanced understanding of the role of parental migration in children's outcomes. 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. Hu (2013) finds a likewise negative effect for Chinese children left behind, though remittances here partially offset the parental absence. This may suggest that remittances are not capable of compensating the lack of time investment that parents could otherwise have made. China is one of the more interesting countries to study because every year it 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. By exploiting a policy reform in 1998 Pan (2012) estimated the effect on human capital generation under migration controls. He 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 migration into urban areas. Differences between countries are many, and relative skill sets and external networks that groups can tap may play important roles.

The strand of migration literature within labor economics that has had access to data allowing for a dynamic treatment of the topic focuses on the expected labor market outcomes of such events. This literature models decisions structurally to study the underlying effect of and quantify income shock responses and/or policy interventions from the shut down of the proposed mechanisms. One of the first papers in this literature to study the effect of different location choices on the labor market outcome of single males was due to Kennan and Walker (2011). They find that interstate migration is indeed explained by income prospects. Other authors have extended this framework to incorporate the more complex migration decisions of married couples. Gemici (2011) estimates the labor market effects and marital stability from the simultaneous migration decision of married couples in the U.S., modeling the decision problem between the husband and the wife through Nash Bargaining. She finds that family ties hinder both mobility and wage growth for men and women. While Lessem (2013) also models decisions for husband and wives she is more concerned with individual migration events under the constraint of border enforcement. Since her data on Mexican migrants reports staggered type migration events (with the husband moving first and the wife choosing to follow or not), she models a Markov-perfect equilibrium. The author reports a wage elasticity of US migration of -0.8 and a US migration elasticity of border enforcement of -0.42.

By contrast the literature on parental investment in human capital development is much more developed. Becker and Lewis (1973) and the followup Becker and Tomes (1976) developed the fundamental theories on the quantity-quality tradeoff in children and investments, where endogenization of the fertility decision is critical in the analysis of parental investments and the outcome of the child. More recently Cameron and Heckman (1998), Cunha and Heckman (2008), and Bernal and Keane (2010) have used Becker's initial framework to estimate the effect of investment in children's human capital development. This literature finds that parental time investment (and variably other other endowments) in the child directly impact the labor market outcomes later in life. The principal motivation for the investment in a child's human capital development is that parents are altruistic toward the well-being of their children. 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 activities. In this case the current well-being of the child is deemed more important - Basu and Van (1998) show this is perfectly rational. Taking these investment decision one step further, Gayle et al. (015a)[2] show that dynastic models can be estimated by considering that parents are altruistic towards their dynasty, foregoing the complication of how parents allocate welfare among their children. In this way the literature moves toward estimating the dynamic effects of investment decisions on social mobility. Here the authors consider parental inputs as time invested in children and de-

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

termine the dynastic discount factors associated with human capital transmission. Their paper reports that the different times invested by parents in white and black children A similar framework can be employed here to model the mechanism by which parental migration aids in the development of human capital investment.

The project is the synthesis of these various strands the literature. Its key contribution is to add to the current literature precisely by analyzing the mechanism through which family migration decisions affect human capital generation of subsequent generations in a dynasty.

3 Empirical Analysis

In appendix C I detail the main results from the previously conducted analysis. The key results from the bivariate analysis of means from the descriptive data is that parents who migrate have children who complete more years of schooling, as well as more levels. However, since these types of analyses may suffer from spurious aggregate group level effects a duration analysis is conducted to ascertain if parental migration affects the probability of exiting the schooling system. I find that parental migration mitigates the probability of a exiting, with a survivor plot (figure C.2) showing that this reduced probability occurs in the non-obligatory secondary schooling phase of the education cycle and increases the survival of children in the school system at this level by 9%.

Regarding wages the OLS mincer specification is conducted to obtain the premium regarding schooling. The estimation results show that the premium to education is in line with what others have found, not just in Indonesia but also in the developing world. Moreover there is a clear, immediate correlational premium to migrating. This is borne out in both an OLS and FE specification. As the FE specifications control for individual unobserved heterogeneity the migration premiums revealed in table C7 show that whether the migration is within a province or between a province there are consistent premiums associated with the move. However, the premium associated with an interprovincial migration is consistently higher (7%) and different than a within province migration (3%). I further study the premiums associated with what I call a "market" migration: migrating between 5 island groupings constructed from all the provinces constituting those groups. Here, the wage premium of migrating between islands is much higher (14.3%), underscoring the heterogeneity between agglomerated provincial groupings: provinces within an island may be more similar with regards to labor markets than provinces elsewhere and on other islands. Although I can not reject the null hypothesis that the premium associated with an intraprovincial migration (3%) and an intramarket migration (6%) are the same, as of yet it can not be claimed that this is the result of an equivalence between the labor market within a province and the labor market of its neighbors.[3] It may very well

[3]I accomplish the test by conducting a seemingly unrelated estimation of the two fixed-effect models

be the case that this is true; and it would imply that the similarity between labor markets of neighboring provinces brings down the estimated expected premium associated with an intraprovincial move.

I note that these wage premiums are instantaneous. Additionally I have sector information that I may exploit by adding as a fixed effect term for groups of related occupations. A future specification will augment this to reflect the expected wage premium following a provincial or market move. Migrants are not just migrating for an instantaneous wage boost but also on the expectation of some persistence. Another future specification will study the subset of the data corresponding to children to asses if a wage premium exists on the basis of previous parental migration (as well as an understanding of the intergenerational elasticity of wages to understand if social mobility is affected).

4 Model

The current project is concerned with the development of the dynastic model to structurally estimate the mechanism by which parental migration may mechanistically translate into greater investment in human capital of the subsequent generation. A structural model is elaborated first and then a simpler model is proposed to study if the key parameters of the model lead to variation in simulated outcomes upon perturbations. The model assumes stationarity over generations, and the state space of the initial generation, having no history of decisions, is considered to be exogenous to the problem.

4.1 A Structural Dynastic Model

The main features of the partial equilibrium model are that parents make decisions regarding the location of their family and whether to invest in their children's education. Parents are altruistic toward their dynasty and as such care about the labor market outcomes of their children and subsequent generations. These decisions become intertemporal between generations, wherein the household's decisions translate into the states of their children, and so on. Currently, the period in the model is one year (given that I have a longitudinal panel, this is feasible). However, it is likely that time periods will be aggregated into stages of life for ease of computation and at the moment this is not the concern. Agents are also not allowed to save, and as such there are no intertemporal assets to transfer. So agents must consume their period income; thus utility maximization is effectively lifetime income maximization.

to account for correlations between the models.

4.1.1 Model Setup

An agent is a household (husband and wife that makes decisions jointly) from a generation $g = (0, ..., \infty)$ alive for $t = 0...T$ periods. Further, the model will only consider the (genderless) first born child in a family. As such I forgo modelling the quantity-quality tradeoff that would also require modelling fertility decisions as elaborated in Becker and Lewis (1973) and Becker and Tomes (1976), and estimated by Gayle et al. (015a). Therefore, marriage decisions will not be considered; as only two generations are necessary to estimate the model, the initial old (the progenitors of their dynasty) start out married and with the child already born. Children can be educated for $s = 0, ..., S$ periods, where maximum $\#S = 13$ (if yearly, then 0 is kindergarten and 12 is the final year of non-obligatory schooling). If the model is alternatively defined as stages of life then S is defined as the level of schooling corresponding to the stage of the child's life.

One complication that will be considered for a future exercise is to model ability production in the spirit of Cunha and Heckman (2008) and Cunha et al. (2010) based on the cognitive tests gathered. To this end, when a child is born the household observes its ability, which is highly correlated with the household's own ability. So the educational investment works to affect the child's future ability.

Timing The household decides at the beginning of the period whether to move and to educate its child. The decisions are independent and commitment must be maintained: once the household moves to a new labor market they can not rescind on the educational investment decision based on observing the schooling quality in the area (that will be proxied by an amenity index on schooling). And as a household moves in search of an expected wage increase they must move to observe wages in the new location. This timing ensures that a reversal does not result in multiple equilibriums.

State Variables The state space includes the current location ℓ_t (which was the previous location choice in $t = t - 1$) and the previous location ℓ_{t-1}. As in Kennan and Walker (2011) and Lessem (2013) I employ a reduced state space memory on location history for computational feasibility, keeping only the previous location. Characteristics of the household are given by X_t (such as age of the household, labor market experience, and the agent's ability - which changes in each period), where $X_t = (X_t^h, X_t^w, X_t^c)$. In an extension that will consider abilities, agents also observe a_0, the child's initial ability; thereafter, child's ability a_t for $t > 0$ is observed and results from the educational investment parents made in the previous period. That is, the child's production of abilities generates a dynamic state space regarding a_t. The state vector for an agent is thus defined as $z_t = (\ell_{t-1}, \ell_t, a_0, a_t, X_t)$.

Production Technology of Child Abilities The model can be extended so that investments in education affects a child's ability (and in turn, whether to continue to educate). This will enter the model through a dynamic state space. Here it occurs through the propagation of abilities that will be subject to a (added-value) production technology, as described below:

$$a_t = f_{t-1}(a_{t-1}, z_t, \phi_t w_t^{hh}(j, z_t) \cdot e)$$

This technology will generate new abilities taking as inputs previous abilities and, if parents choose to educate, a portion of wages. (DEVELOP THIS SECTION FURTHER)

Decision Set A household makes two mutually exclusive decisions: choose a location j from the set of J locations (which also includes their current location); and whether to educate the child one more year (or school level, as the case may be), $e_t = 1$, or not, $e_t = 0$. The set J will be defined at most as the number of provinces (34), or at the least either groupings of provinces within geographical islands (5-8) or two markets (consisting of one of the main islands and the rest of Indonesia).

Following Gayle et al. (015b), this means that I can construct mutually exclusive indicators for these decisions, $I_{k,t}$, that takes value 1 for decision k at time t and 0 otherwise:

$$\begin{aligned}
&I_{1,t} = \mathbb{1}\{j = 1\}\mathbb{1}\{e = 0\} \\
&I_{2,t} = \mathbb{1}\{j = 1\}\mathbb{1}\{e = 1\} \\
&\ldots \\
&I_{K-1,t} = \mathbb{1}\{j = J\}\mathbb{1}\{e = 0\} \\
&I_{K,t} = \mathbb{1}\{j = J\}\mathbb{1}\{e = 1\}
\end{aligned}$$

These mutually exclusive decisions mean that $\sum_{k=1}^{K} I_{k,t} = 1$, for $K = \#e \cdot \#J$ decisions.

Utility Utility is defined as $u_t(c_t, z_t, \varepsilon_t)$. Since agents do not make a savings decision, the other assumptions in this partial equilibrium framework are that agents have a constant marginal utility of income and are able to borrow/lend at will according to whatever interest rate exists. I will also normalize prices in terms of the consumption good so that it is treated as a residual. These assumptions imply that the econometric specification of utility will be linear with risk-neutral agents.

For simplicity I also assume that agents are inelastic with respect to their labor supply and do not model this decision. As such I do not include, as is typically the case in the literature, the decision $I_{k,t}$ as a proxy for leisure. The main reason I do not currently do this is because of the possibly large amount of parameters that would need to be estimated if the decisions were included as a proxy for leisure (which would be $K - 1$

parameters, where K is obviously dependent on the size of the set of locations J) - it is not a difficult matter, however, to include the decisions as the proxy for leisure if the set of J locations is kept reasonably low.

The utility is, however, subject to a transitory, idiosyncratic payoff shock $\varepsilon_{k,t}$ from the decisions taken, which may represent a shock to either preferences, moving cost, or educational cost (with no way of knowing which one) and is associated with each discrete choice k. These shocks are essentially the aspect of the state space that is unobservable to the econometrician (but is observable to the agent) for each posible decision taken. This vector $\varepsilon_t = (\varepsilon_{0,t}, ..., \varepsilon_{K,t})$ of shocks is assumed i.i.d. across population and time and drawn from distribution function $G_\varepsilon(\varepsilon_t)$. And as is also typical in the literature I will also assume that the utility is additively separable in the observable and unobservable states of the decisions taken, so that $u_t(c_t, z_t, \varepsilon_t) = u_t(c_t, z_t) + \varepsilon_t$.[4]

Budget Constraint The agent's per period budget constraint is given by the following, which includes the costs associated with the decisions:

$$w_t^{hh}(j, z_t) = c_t + \delta(j, z_t)\mathbb{1}(j \neq \ell_t) + \phi_t \cdot e$$

where

$$w_t^{hh}(j, z_t) = w_t^h(j, X_t^h) + w_t^w(j, X_t^w) + w_t^c(j, a_t) \cdot (1 - e)$$

The parameter $\delta(j, z_t)$ is the moving cost associated with choice j, which activates as long as the choice is not to stay in the current location. The moving cost will be a function of parameters in the state space (which includes a fix cost, as well as costs based on distance between locations, population densities, province adjacency, and moving back to a previous location). ϕ_t is the investment in education. The agent's budget is the cumulative wage of all members in the household (which includes the child's contribution should it be decided that the child is not educated).

Wages are exogenous and will be taken as the median income of each of the j labor markets under the partial equilibrium in the labor market. While wages will be subject to transient fluctuations, migrants seek a permanent wage increase by picking a wage draw from a new location. It is this expected, permanent wage increase from location-based wage differentials that incentivise workers to reallocate their labor supply through migration.

Transition Probabilities A transition from the state z_t to possible states in z_{t+1} de-

[4]To include leisure I would have to include a second utility term that includes the utility for the decisions taken. Following Gayle et al. (015b) and Aguirregabiria and Mira (2010), this would lead to defining utility as $u_t(c_t, z_t, \varepsilon_t) = u_t^1(c_t, z_t) + \sum_k I_{k,t}^{\circ} \left(u_{k,t}^2(z_t) + \varepsilon_{k,t} \right)$, where the second parenthetical term is the portion of utility attributable to leisure under the separable utility assumption, composed of the systemic component and the transitory shocks of the decisions taken.

scribes the probability of entering the new state due to the uncertainty of the outcomes of the agent's investment decision. In the extension the agent does not observe the outcome of educating their child through ability until the next period, at which point they may choose to continue to educate or not. The result of the migration decision, having been made and the household having moved to the new location in period t, may or may not have been fruitful and the family may choose to migrate once more.

Since household decisions create a new generation, a second transition matrix defines the probabilities that the newly formed adult of generation $g+1$ obtains a set of characteristics given the state of the agent from the previous generation and the investment decisions this agent made.

4.1.2 Value Functions

I start deriving the value functions of the agent within his dynasty. The period $t = 0$ expected lifetime utility of the agent in generation g with the characteristics X is given by:

$$U_{gT}(X) = \mathrm{E}_0 \left[\sum_{t=0}^{T} \beta^t \left[u_t(c_t, z_t) + \sum_k I^{\circ}_{k,t} \varepsilon_{k,t} \right] \middle| X \right]$$

where $I^{\circ}_{k,t}$ represents the optimal sequence of decisions. And the dynastic aspect of the utility for an agent of generation g is:

$$U_g(X) = U_{gT}(X) + \alpha \beta^T \mathrm{E}_0 \left[U_{g+1}(X') \middle| X \right]$$

Here utility $U_{g+1}(X')$ given the characteristics X' (based on the parents investments) is the utility of generation $g+1$. Since agents are altruistic and derive utility from the wellbeing of their child they discount their child's expected lifetime utility according to the altruism parameter α. The recursiveness links generations within a dynasty.

Rearranging the budget constraint and inserting the consumption good into the deterministic component of the utility yields the choice specific utility (as now utility is a function of decisions):

$$u_{k,t}(z_t) = u_t(w_t^{hh}(j, z_t)(1 - \phi_t \cdot e) - \delta(j, z_t)\mathbb{1}(j \neq \ell_t), z_t)$$

Lifetime utility is now:

$$U_{gT}(X) = \mathrm{E}_0 \left[\sum_{t=0}^{T} \beta^t \left[\sum_k I^{\circ}_{k,t} (u_{k,t}(z_t) + \varepsilon_{k,t}) \right] \middle| X \right]$$

Optimality then means that the agent must choose a sequence of alternatives using decision rule $I(z_t, \varepsilon_t)$ given the vector of shocks. The optimal decision rule, as described by Gayle et al. (015b), is

$$I^{\circ}(z_t, \varepsilon_t) = \arg\max_{I_k \in I} \ \mathrm{E}_I \left[\sum_{t=0}^{T} \beta^t \left[\sum_k I^{\circ}_{k,t}(u_{k,t}(z_t) + \varepsilon_{k,t}) \right] + \alpha \beta^T U_{g+1}(X') \middle| X \right]$$

The value function associated with this is then given by:

$$V(z_{t+1}, \varepsilon_{t+1}) = \max_{I_k \in I} \ \mathrm{E}_I \left[\sum_{t'=t+1}^{T} \beta^{t'-t} \left[\sum_k I_{k,t'}(u_{k,t'}(z_{t'}) + \varepsilon_{k,t'}) \right] + \alpha \beta^T U_{g+1}(X') \middle| z_t, \varepsilon_t \right]$$

Bellman's optimality implies the recursive form is given by

$$V(z_t, \varepsilon_t) = \max_{I_k \in I} \ \sum_k I_{k,t} \left\{ u_{k,t}(z_t) + \varepsilon_{k,t} + \beta \mathrm{E} \left[V(z_{t+1}, \varepsilon_{t+1}) \middle| z_t, I_{k,t} = 1 \right] \right\}$$

The exante value function, the continuation value of being in state z_t prior to observing ε_t, is derived by marginalizing out the preference shocks. Following Aguirregabiria and Mira (2010) I define this is as:

$$\bar{V}(z_t) = \int V(z_t, \varepsilon_t) dG_{\varepsilon}(\varepsilon_t)$$

As previously defined, $G_{\varepsilon}(\varepsilon_t)$ is the CDF of the unobserved payoff shocks. I further define the transitions for states conditional on the previous state and the decision taken as $F(z_{t+1}|z_t, I_{k,t} = 1)$. Then the exante value function is given by:

$$\bar{V}(z_t) = \int \left[\max_{I_k \in I} \ \sum_k I_{k,t} \left\{ u_{k,t}(z_t) + \varepsilon_{k,t} + \beta \sum_{z_{t+1}} \bar{V}(z_{t+1}) F\left(z_{t+1} \middle| z_t, I_{k,t} = 1 \right) \right\} \right] dG_{\varepsilon}(\varepsilon_t)$$

Below I define the conditional value function (or the choice-specific value function, indexed by k for the k^{th} choice taken) as the discounted, present value of having taken the k^{th} choice, behaving optimally moving forward. I also explicitly account for the dynastic component by further defining a new transition matrix, $N(x|z_T, I_{k,T} = 1)$, describing how generation $g+1$ characteristics are determined based on the parental decisions taken.

$$\begin{aligned}
v_k(z_t) =& u_{k,t}(z_t) + \beta \sum_{z_{t+1}} \bar{V}(z_{t+1}) F\left(z_{t+1} \middle| z_t, I_{k,t} = 1 \right) \\
=& u_{k,t}(z_t) + \sum_{t'=t+1}^{T} \beta^{t'-t} \sum_{z_{t'}} \bar{V}(z_{t'}) F\left(z_{t'} \middle| z_t, I_{k,t} = 1 \right) \\
&+ \alpha \beta^{T-t} \sum_x \sum_{z_T} \bar{V}(x) N\left(x \middle| z_T, I_{k,T} = 1 \right) F\left(z_T \middle| z_t, I_{k,t} = 1 \right)
\end{aligned}$$

The above conditional value function contains the term $\bar{V}(x)$, the expected continuation value of the next generation of the dynasty after the agent's life. It is a function of the characteristics that have been passed to the next generation according to the transition function $N(\cdot)$. This explicitly presents the complication inherit in dynastic structural models: while termination value in period $T+1$ is such that $v_k(z_{T+1}) = 0$, at $t = T$ we must know $\bar{V}(x)$. So these models are infinitely lived from the dynastic perspective, but finite within generations. To solve the model, this portion of the terminal value function will have to be estimated so that the usual techniques of estimating the finite model can be applied.

As will be obvious later, this conditional value function is the key to obtaining the conditional choice probabilities. The exante value function now reads as

$$\bar{V}(z_t) = \int \left[\max_{I_k \in I} \sum_k I_{k,t}\{v_k(z_t) + \varepsilon_{k,t}\} \right] dG_\varepsilon(\varepsilon_t)$$

and the optimal decision rule becomes:

$$I^\circ(z_t, \varepsilon_t) = \arg\max_{I_k \in I} \sum_k I_{k,t}\{v_k(z_t) + \varepsilon_{k,t}\}$$

To obtain the solution, first I derive the conditional choice probabilities (CCPs) obtained by integrating the optimal decision rule over the unobservable states, $\varepsilon_{k,t}$:

$$p_k(z_t) = \int I^\circ(z_t, \varepsilon_t) dG_\varepsilon(\varepsilon_t) = \int \left[\prod_{k \neq k'} \mathbb{1}\left\{ v_k(z_t) - v_{k'}(z_t) > \varepsilon_{k,t} - \varepsilon_{k',t} \right\} \right] dG_\varepsilon(\varepsilon_t)$$

Assuming that the distribution $G_\varepsilon(\varepsilon_t)$ from which the components of vector ε_t are drawn is of type-1 extreme value[5] leads to the following closed-form results: the first is that the above CCPs take a logit form and is given by

$$p_k(z_t) = \mathrm{E}\left(I^\circ_{k,t} = 1 \middle| z_t \right) = \frac{e^{v_k(z_t)}}{\sum_s e^{v_s(z_t)}}$$

where $v_k(z_t)$ is the value when optimal I°_k is chosen, and $v_s(z_t)$ the conditional value function for any choice $s \in K$, including the optimal choice k; the other result of this assumption is that the expected payoff shock given the optimal decision choice and the state is

$$\mathrm{E}_\varepsilon\left(\varepsilon_{k,t} \middle| I^\circ_{k,t} = 1, z_t \right) = \gamma - \ln p_k(z_t)$$

where gamma is Euler's constant. And inserting the closed-form CCP into the above

[5]In the literature this is known as the CLOGIT - or conditional logit - assumption

equation implies the following result:

$$\underbrace{v_k(z_t) + \mathrm{E}_\varepsilon\left(\varepsilon_{k,t} \middle| I^\circ_{k,t} = 1, z_t\right)}_{=\bar{V}(z_t)} = \gamma + \ln\left(\sum_s e^{v_s(z_t)}\right)$$

So under this assumption on the CDF of the preference shocks and the assumption of separable utility in the observed and unobserved states, the exante value function also has a closed form solution, assisting in solving the dynamic programming problem. The other use of the CCPs is in defining the likelihood function in the empirical estimation. The solution to the model would follow McFadden (1973) and Rust (1987), and Gayle et al. (015b) for the dynastic estimation of the terminal value function, through backward induction.

4.2 A 2 Market Model for Simulation

The previous section described a structural dynastic model to implement in the estimation of the research question. What is lacking yet is the econometric specification of the model. However, prior to specification I will study if the model parameters generate variation in equilibrium outcomes when the parameters are perturbed. I distill the model into a dynastic OLG model where decisions made by parents (who constitute the initial old) become the state space of the children (who are essentially the household) in the next generation.

4.2.1 Model Setup

Given a dynasty D the initial old are $g = 0$. Agents live for only two periods, so old is defined as $t = 2$ and young as $t = 1$. Those agents in the second and final period of their life are exogenously bestowed with the child that constitute the new young, who are in their first period of their life. Similar to before, when agents make decisions they do so for their child; their child realises those investments in their state space when they become old and through the labor market outcomes those endowments imply. Since I will not consider population growth for simplification, a "child" in this context is the household of the next generation. The wages in the market are exogenous and thus I solve a partial equilibrium model where I once again assume stationarity across generations.

Agents in $t = 2$ period of life decide to move to a new location, of which there are two, and educating those agents in $t = 1$; agents in $t = 1$ make no decisions. Furthermore there is no marriage pairing that creates the household, either in g or $g + 1$ - in a sense, households create new households composed of agents that are perfectly matched in terms of skill h.[6] Once more, agents making decisions must commit to the decision to educate

[6]Which is quite strong, but reasonable given that in Indonesia I observe assortative matching across

after moving (so that the location does not induce a switch in the decision). Since I forgo a leisure decision, it is not valued and the agents supply all their time to the labor market. They also draw utility from the expected outcome of those agents in their care when they enter their $t = 2$ period of life and do so altruistically. The state space of the agent in their second period of life is defined as the vector $z = (\ell, h)$ where ℓ is the location of the agent and h is their skill. Since agents can only decide to move once in their lifetime there is no need to retain previous location histories.[7] Skill takes a value of $h = 0$ if they were never educated, and $h = 1$ if the agent is educated. So parental investment e in the young period (similarly defined as in the structural model setup) becomes h in the old period of life and enters the state space. In one scenario this investment is not subject to uncertainty (an educational investment necessarily generates skills in the next generation); in another scenario I relax this assumption to make the agent's investment in education uncertain (investment $e = 1$ does not yield $h = 1$ with some probability).

4.2.2 Value Function

In the previous subsection I made certain assumptions that lead to linear utility. However, the assumption on the distribution of the unobserved preference shocks lead to concavity of the expected value function. For this exercise I could make a similar assumption and follow the literature in solving the model to then simulate it.[8] I write the following value function from the point of view of the old-age agent, indexed by $(d, g, t = 2)$ (dynasty d, generation g, and period of life t):

$$V_{d,g,t=2}(z, \varepsilon) = \max_{I_k \in I} \sum_k I_k \left\{ v_{d,g,t=2}(z, k) + \varepsilon_k \right\}$$

where

$$v_{d,g,t=2}(z, k) = u(c) + \alpha \mathrm{E} \left[V_{d,g',t=2}(z', \varepsilon') \middle| z, I_k = 1 \right]$$

and utility, being linear and the same across generations, is $u(c) = c$. The interpretation of the above problem and the solution is the same as previously described. Since the generation ends at period 2 there is no discounting of the future through β. However, decisions can be viewed as inter-temporal investments as the fruits of at least one of them (education) is realized in the next generation through the labor market outcome. And as agent does care about the next generation they discount the value of the next

education levels with the typical lower off-diagonal, as evidenced in table B1.

[7] Of course if intergenerational mobility were to be studied, keeping at least the previous generations location may be helpful to understand if there is a persistence in location and how it may affect future generations.

[8] Otherwise I would have to make an assumption regarding either one or both of the costs associated with the decision (specifically a convexity assumption, which is not obvious why and in what way they should be convex); or make a curvature assumption on utility (for example, assume that utility is CRRA for ease of simulation since decisions have an inter-temporal interpretation).

generation's utility, $V_{d,g',t=2}(z')$, by their altruism towards it, α.

The budget constraint is given by:

$$c = w^{hh}(h,j) - \delta\mathbb{1}(\ell \neq j) - \phi \cdot e$$

and

$$w^{hh}(h,j) = w^{g}_{t=2}(h,j) + w^{g+1}_{t=1}(h,j) \cdot (1-e)$$

where the cost of moving and to educate have the same interpretation as before, only that with two markets the moving cost is a constant as it is symmetrical (at least if the cost is just a measure of distance); and instead of an investment in education derived from a portion of wages, it is simply a cost (since no mechanism is elaborated on how the investment would generate skill). I am breaking up the household wage into the components contributed by the members of the household, except that in this simplified framework I adjust it so that an agent (the household) begets a new agent (the next household) and these are the contributors. Further, if what I call a child here is not educated (so $e = 0$), it necessarily posses skill $h = 0$ and can only contribute such a wage to the household in its young period.

4.2.3 Parameter Selection

For the exercise I pick parameters to both solve the model and for the simulation exercise. To that end I use the data to guide the exercise. The altruism parameter is selected to match complete, reciprocal altruism ($\alpha = 0.5$). The fixed cost for moving is selected to be $\delta = 0.4$ (or 40% of period wages); and the cost for educating a child (taken to be location independent) is set at $\phi = 0.2$ (or 20% of period wages).

The two markets are selected by aggregating the provinces located in the island of Sumatra as well as Riau Island province into Market 1, and all other provinces on the other islands of Indonesia into Market 2. The median of log hourly wages are then tabulated according to skill, h, and normalized based on the non-educated wages in Market 1. The table below lists the full parameter selection for solving the value functions in the model:

Parameters			Value
Altruism	α		{0,0.95}
Moving Cost	δ		[0,2]
Educ. Cost	$\phi_{\ell=2}$		[1,2]
Wages	Market 1	$w^1_{h=0}$	1
		$w^1_{h=1}$	2.40
	Market 2	$w^2_{h=0}$	1.06
		$w^2_{h=1}$	2.34

The last parameters to define are the probabilities that constitute the 16×4 matrix describing the discretised transition function. This transition function is equivalent to the one described in the previous subsection and defined as $N(x|z_T, I_{k,T} = 1)$, only now it is defined as $N(z'|z, I_{k,t=2} = 1)$ and z' is the state of the $g + 1$ agent in their old period of life. Under the case where educating leads to the certain outcome of the next generations' skill attainment, the following describes the degenerate empirical transition probabilities:[9]

$$
N(z'|z, I_{k,t=2} = 1) = \begin{array}{c c}
z, I_{k,t=2} = 1 \downarrow \ \backslash \ z' \rightarrow & \begin{array}{cccc} h=0,l=1 & h=0,l=2 & h=1,l=1 & h=1,l=2 \end{array} \\
\begin{array}{c}
h=0,l=1,j=1,e=0 \\
h=0,l=1,j=1,e=1 \\
h=0,l=1,j=2,e=0 \\
h=0,l=1,j=2,e=1 \\
h=0,l=2,j=2,e=0 \\
h=0,l=2,j=2,e=1 \\
h=0,l=2,j=1,e=0 \\
h=0,l=2,j=1,e=1 \\
h=1,l=1,j=1,e=0 \\
h=1,l=1,j=1,e=1 \\
h=1,l=1,j=2,e=0 \\
h=1,l=1,j=2,e=1 \\
h=1,l=2,j=2,e=0 \\
h=1,l=2,j=2,e=1 \\
h=1,l=2,j=1,e=0 \\
h=1,l=2,j=1,e=1
\end{array} &
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{pmatrix}
\end{array}
$$

4.2.4 Model Results and Simulation

TO BE COMPLETED

[9] In appendix D I consider other transition functions based on an uncertain outcome regarding educational investment.

A IFLS Background

The Indonesian Family Life Survey is an ongoing, longitudinal survey of Indonesian households conducted by the RAND corporation in conjunction with SurveyMETER.[10] Four waves were fielded between 1993, 1997, 2000, and 2007, with the most recent wave 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 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.[11]

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

[10]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.

[11]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.

cross-sections I collapse the panel dimension.

The first wave launched in 1993, covering 13 of the then 26 provinces on 6 islands in Indonesia.[12] 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.

A.1 Indonesia

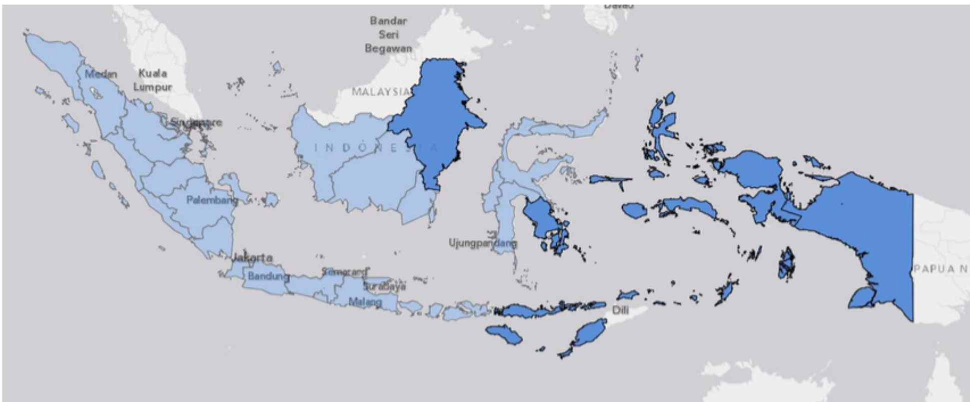
Below are two maps of Indonesia, highlighting the areas where the IFLS surveys were administered (shaded regions). The IFLS-East 2012, a recently available data set, supplements the original RAND Corp. survey to account for the eastern provinces, which

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

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 A1. 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.[13] The master files of geographical codings

[13] 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`

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.[14]

B 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).[15]

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) issued in 2003 declared compulsory up to age 15 (the first two levels). 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), which are not currently publicly funded. 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 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 that must be completed to enter the next level. Decentralized examining of students oc-

[14]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.

[15]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.

Table A1: 2010 Census and Geographical Subdivisions in Indonesia

Province Name	ISO[4]	Capital	IFLS or IFLS-East Province	Population at 2010 Census	Area (sq.km)	Population density	Island Group	Number of Districts (kabupaten)	Number of sub-Districts (kecamatan)	Villages (kelurahan)	Municipalities (kotamadya)
Special Region of Aceh	ID-AC	Banda Aceh		4494410	57956	77	Sumatra	18	275	6420	5
Bali	ID-BA	Denpasar	X	3890757	5780	621	Lesser Sunda Islands	8	57	698	1
Bangka-Belitung Islands	ID-BB	Pangkal Pinang		122296	16424	64	Sumatra	6	43	361	1
Banten	ID-BT	Serang		10632166	9662	909	Java	4	154	1530	4
Bengkulu	ID-BE	Bengkulu		1715518	19919	84	Sumatra	9	116	1442	1
Central Java	ID-JT	Semarang	X	32382657	40800	894	Java	29	573	8577	6
Central Kalimantan	ID-KT	Palangkaraya		221089	153.564	14	Kalimantan	13	120	1439	1
Central Sulawesi	ID-ST	Palu		2635009	61841	41	Sulawesi	10	147	1712	1
East Java	ID-JI	Surabaya	X	37476757	47799	828	Java	29	662	8502	9
East Kalimantan	ID-KI	Samarinda	O	3026060	139462	22	Kalimantan	6	89	1023	3
East Nusa Tenggara	ID-NT	Kupang	O	4683827	48718	92	Lesser Sunda Islands	20	286	2775	1
Gorontalo	ID-GO	Gorontalo		1040164	11257	94	Sulawesi	5	65	595	1
Jakarta (Special Capital Region)	ID-JK	Jakarta	X	9607787	664	12786	Java	1	44	267	5
Jambi	ID-JA	Jambi		3092265	50058	57	Sumatra	9	128	1319	2
Lampung	ID-LA	Bandar Lampung	X	760405	34623	226	Sumatra	12	206	2358	2
Maluku	ID-MA	Ambon	O	1533506	46914	32	Maluku Islands	9	76	898	2
North Kalimantan	ID-KU	Tanjung Selor		622350	71176	10	Kalimantan	4	47	381	1
North Maluku	ID-MU	Sofifi	O	103.087	31982	31	Maluku Islands	7	109	1041	2
North Sulawesi	ID-SA	Manado		2270596	13851	162	Sulawesi	11	150	1510	4
North Sumatra	ID-SU	Medan	X	12982204	72981	188	Sumatra	25	408	5649	8
Special Region of Papua	ID-PA	Jayapura	O	2833381	319036	8	Western New Guinea	28	330	3583	1
Riau	ID-RI	Pekanbaru		5538367	87023	52	Sumatra	10	153	1500	2
Riau Islands Province	ID-KR	Tanjung Pinang		1679163	8201	208	Sumatra	5	59	331	2
Southeast Sulawesi	ID-SG	Kendari	O	2232586	38067	51	Sulawesi	10	199	1843	2
South Kalimantan	ID-KS	Banjarmasin	X	3626616	38744	96	Kalimantan	11	151	1973	2
South Sulawesi	ID-SN	Makassar	X	8034776	46717	151	Sulawesi	26	301	2874	3
South Sumatra	ID-SS	Palembang	X	7450394	91592	86	Sumatra	11	217	2869	4
West Java	ID-JB	Bandung		43053732	35377	1176	Java	17	625	5827	9
West Kalimantan	ID-KB	Pontianak		4395983	147307	30	Kalimantan	12	175	1777	2
West Nusa Tenggara	ID-NB	Mataram	X	4500212	18572	234	Lesser Sunda Islands	8	116	913	2
Special Region of West Papua	ID-PB[6]	Manokwari	O	760422	97024	8	Western New Guinea	10	149	1291	1
West Sulawesi	ID-SR	Mamuju		1158651	16787	73	Sulawesi	5	66	564	0
West Sumatra	ID-SB	Padang	X	5133989	42012	110	Sumatra	12	169	964	7
Special Region of Yogyakarta	ID-YO	Yogyakarta	X	3457491	3133	1138	Java	4	78	438	1
Total		**34**	**22**	**227045689.1**	**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 curred from 1965 through 1980 via the *Ujian Negara* (State Exam), when a switch was made to a more centralized exam structure, the *Evaluasi Belajar Tahap Akhir National* (National Final Learning Evaluation - commonly abbreviated to EBTANAS). 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 and for the moment they will not be considered in the research.

B.1 Marriage

Below is a marriage contingency table describing the assortative matching present in Indonesia. This table encompasses the marriage pairings of those couples who have never divorced (to control for possible selection issues associated with termination of marriage) and constructed following Greenwood et al. (2014). The contingency table displays the observed frequencies in the left-hand side of a column and the expected frequencies based on random matching. I further highlight in red the highest observed frequencies (and second highest when it looks large). Perfect assortative matching is described by the diagonal across columns. As the figure shows, there is a fair amount of assortative matching in Indonesia, along with the typical off-diagonal pairing of educated men with lower educated women.

Table B1: Marriage Contingency Table

Husband / Wife	No School		Primary		Obl. Secondary		Non-Obl. Secondary		College	
No School	**2.70%**	**0.30%**	1.40%	1.60%	0.10%	0.90%	0.00%	1.10%	0.00%	0.40%
Primary	3.90%	2.60%	**24.50%**	**13.70%**	5.50%	7.20%	2.00%	9.10%	0.20%	3.60%
Obl. Secondary	0.40%	1.30%	6.40%	6.60%	**6.10%**	**3.50%**	4.00%	4.40%	0.50%	1.80%
Non-Obl. Secondary	0.20%	2.10%	4.70%	11.10%	**7.10%**	**5.90%**	**14.50%**	**7.40%**	3.00%	2.90%
College	0.00%	0.90%	0.60%	4.70%	1.10%	2.50%	4.50%	3.10%	**6.30%**	**1.30%**
Marginal	7.20%		37.70%		20.00%		25.10%		10.00%	
Total Couples	9,608									
Pearson Chi2(16)	7.90E+03									

C Results of Empirical Analysis

C.1 Bivariate Analysis: Children's Education

Here I report bivariate results from the data regarding children's education relative to parental migration. These migration events are separated into intradistrict and interdistrictl migrations. Intradistrict are also intraprovincial since districts are a subset of provinces; while inter-district migration events can be intraprovincial or inter-provincial. A district corresponds to a Kecamatan (as stated in table A1). I am currently working on regenerating this according to provincial migrations. Figure C.1 presents the highest level of schooling completed of children whose parents may or may not have migrated while the child was young and still in school. This unconditional frequency plot shows that those children whose parents migrate obtain more schooling over their schooling careers; while the children of parents who self-select into not migrating complete less schooling. The tables below further report that parental migration is associated with an increase in the average years of completed education.

Figure C.1: Schooling Completion Frequencies When Migration Occurs While In School

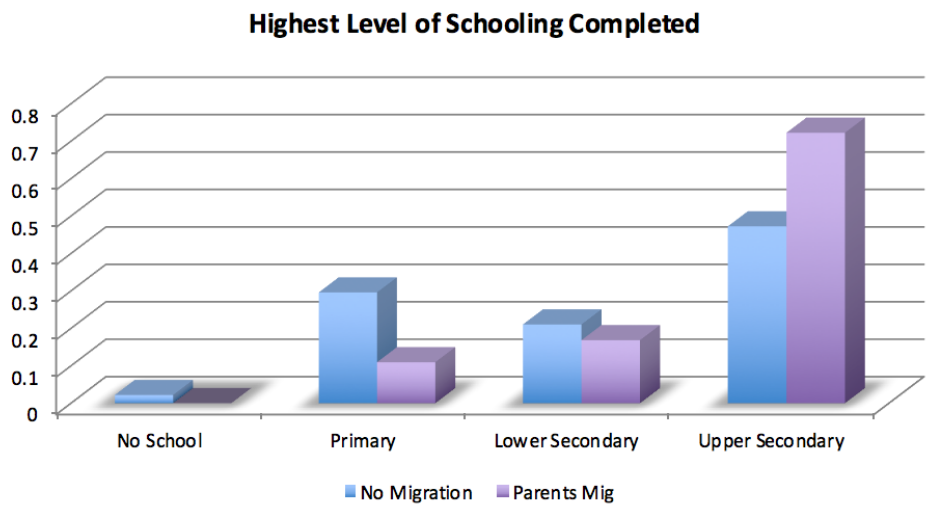


Table C1: Child's Education (years) by Sex and Urbanization of Birth

Sex	Rural	Urban
Girls	8.49	10.63
	(0.06)	(0.06)
Boys	8.74	10.61
	(0.05)	(0.06)

Table C2: Child's Education (years) and Parental Migration Type

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

Table C3: Child's Education (years) by Urbanization of Birth and Parental Migration Type

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

C.2 Duration Analysis

This section reports the results of estimating the discrete time probit hazard models. I first report the model that was used to estimate the coefficients of the duration of children in the Indonesian school system and the results of the estimation.

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

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

T is censored according to $T^* = T - 1$, as the final period observed, $T = 12$ years of schooling, is not identified.[16] I interact equation (1) for each type of migration event (following the same notation as in equation (**??**) for event types) resulting in

$$\gamma_{i\tau}^{k} = \delta_\tau ParentMig_{is}^{k}\mathbb{1}(s = j) \tag{2}$$

In practice it is necessary to artificially censor dummies of equation (2) if a critical mass does not exist at one of the discreet time intervals.[17] The k is the index for the parental migration type.

The conditional hazard rate that I estimate is then defined as

$$h_i(\tau, X_{it}; \theta) = P(t_i = \tau | t_i \geq \tau, X_i) = F\left(\delta_\tau + \sum_k \gamma_{i\tau}^{k} + \beta_k(ParentMig_{i\tau}^{k}) + \sum_j \eta_j X_{ij}\right) \tag{3}$$

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 16 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 (3). 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

$$\ln \mathcal{L}(\theta; X) = \sum_{i=1}^{N} \sum_{\tau=1}^{T^*} a_{i\tau}[y_{it} \ln h_i(\tau, X_{i\tau}; \theta) + (1 - y_{i\tau}) \ln(1 - h_i(\tau, X_{i\tau}; \theta))] \tag{4}$$

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.

I report below the results of parental migration in table C4. The estimated coefficients are the raw outputs whose sign is the only thing interpretable. Marginal effects are not calculated as I care only for the sign (the indication of whether the hazard of exiting the school system increases (a positive coefficient) or decreases (a negative coefficient)).

[16] $T = 12$ 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.

[17] Artificial censoring is necessary when there do not exist failures during a time period, as was discussed in footnote 16 for the final period. For example, the plots of the survival function depicted in this appendix 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 2 are constructed discretely. Thus, I artificially censor them and instead create dummies for grades 1, 2, 3, 4-5, 6-8, 9-11.

These coefficients can, however, be used to construct the survival plots of an individual with the average characteristics.

Table C4: Duration Results

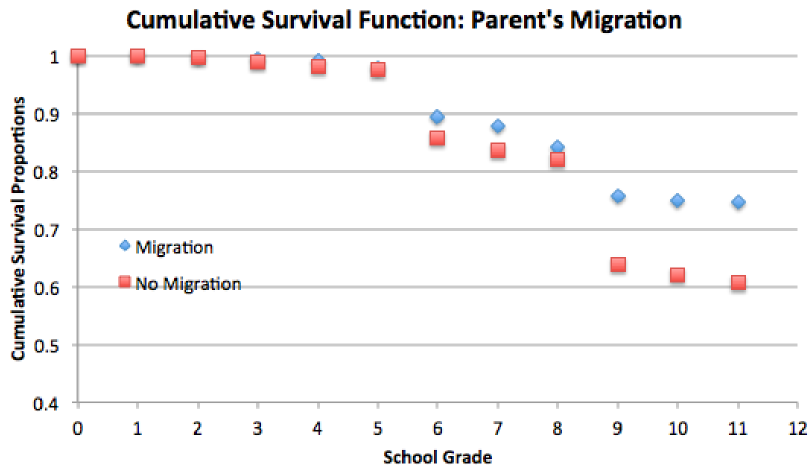
PROBIT: School Exit	(1)	(2)	(3)	(4)	(5)
Father's Migration	-5.806***	-5.106***	-4.767***	-2.345***	-2.371***
	(0.0483)	(0.0675)	(0.173)	(0.0919)	(0.188)
Mother's Migration	-5.538***	-4.947***	-4.943***	-2.465***	-2.271***
	(0.0349)	(0.129)	(0.148)	(0.0857)	(0.118)
Parent's Migration	-5.584***	-4.806***	-5.241***	-3.181***	-2.579***
	(0.0289)	(0.146)	(0.0853)	(0.195)	(0.126)
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	61,574	53,010	59,900	53,068	51,465

Notes: Robust standard errors in parentheses, clustered at the family level. The output reports the probit coefficients. The baseline hazard is non-parametrically specified as a sequence of dummies for each school grade. To correct for the non proportionality of the migration event, the event is interacted with the baseline hazard. Controls: individual controls consist of the birth urbanization, sex, and interactions of sex with the other two controls; parent's 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 significance are at: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

To construct the survival plot I generate the predicted data of a representative agent based on specification (5) of the estimation output in table C4, and only report the plot for the case when both parents migrate (as this has the largest and significantly different result from the other two cases). In general this representative agent has the mean values of the covariates.[18] 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 parents do not migrate. As figure C.2 demonstrates, the probability of exiting the schooling system is reduced by around 9-10%. Although the coefficient of parental migration is only slightly significantly different from the others, the interaction with the baseline hazard generates the deviation in the survival plot.

[18]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 C.2: Survival Function: Parent's Migration



C.3 Wage Analysis & the Migration Premium

In this section I report the results of mincer regressions on the IFLS wages of adults and children who are no longer in school. I also report the mincer regressions for the the IPUMS data of the 1995 Indonesian census to compare the coefficients.

IFLS asked its participants whether they worked and, if so, to report the documented history starting from the first job up to the last one, as well as interruptions. Monthly wages in Rupiahs are reported as well as the labor participation in terms of hours per week and weeks per year. Given this information I am able to obtain the hourly wage of all individuals who reported working and their wage. Included in the information is the ISCO 2 digit occupation code and the 2 digit sector code for the industry the person was employed in. That this information was collected and being a harmonized classification means that it is easy to compare across surveys (IPUMS and IFLS). It is also useful as a means to control for them as fixed effects.

The wages have to be realized and there are two ways to do this. The first is to use the CPI as made available by the World Bank to realize wages to a chosen base year in Rupiahs. The disadvantage of this method is that the CPI within Indonesia is not constant given the disparity between urban and rural regions, even within and between provinces. The statistics office of Indonesia does not provide an adequate time series to realize wages at a finer level (for example, at the province-urbanization level) leaving only that of the national level provided by the World Bank as the viable option. The other option is to use the GDP deflators and PPP data from the U.S. and Indonesia provided by the World Bank to realize the wages and convert them into international dollars. This process suffers from the same problem as CPI, however it has an advantage that in converting to INT$ wages are more readily understood than leaving them in Rupiah. Another draw back with either of the methods (and one out of my control) is that both the CPI and the GDP deflator data from the World Bank is available starting from 1960 and 1961, respectively. Thus, all wage observations prior to these years will be censored.

I use several models to estimate the mincer regressions, using both OLS and FE specifications. Since I am interested in the wage premium attributable to migration, the FE specifications have the advantage that I can control for individual characteristics through time. The implicit assumption in the FE specification is that these individual characteristics are the fixed, natural component of ability that I do not observe; but which is biasing the OLS specification since it is both captured within the error term and correlated with the other observables (in this sense the OLS specifications suffer from unobserved heterogeneity). As only time varying variables can be in these specifications and migration events are time varying, the results should yield a better estimate. Since I have accounted for unobserved heterogeneity (assuming it is constant over time) and omitted variable bias captured within the other fixed effect terms included, but as of yet not the selection into migration these, specifications can not be interpreted causally.

C.3.1 Descriptive Results on Wages

FILL IN WITH THE DESCRIPTIVE TABLE OF WAGES

C.3.2 Specifications and Estimation Results

The OLS specifications that I estimate are the following:

$$ln(wage_{iopt}) = \alpha + \beta SchYrs_i + \gamma_1 age_{it} + \gamma_2 age_{it}^2 + \sum_l \delta_l Mig_{it}^l + \sum_k \theta_k X_{it} + \sum_j \alpha_j + \varepsilon_{iopt}$$

$$ln(wage_{iopt}) = \alpha + \sum_m \beta_m SchLvl_i^m + \gamma_1 age_{it} + \gamma_2 age_{it}^2 + \sum_l \delta_l Mig_{it}^l + \sum_k \theta_k X_{it} + \sum_j \alpha_j + \varepsilon_{iopt}$$

In the above equation hourly wages are observed for each individual i, in occupation o, in province p, in year t. Years of schooling, $SchYrs_i$, is used in one set of specifications; and in another I substitute for this a set of dummies for attained schooling levels (omitting the case of no schooling as the base case). Since there is a constant in the model, α, the set of fixed effects, α_j, exclude one of the levels as a base case. The fixed effects included in the above regression are year fixed effects (α_t), occupation fixed effects (α_o), province fixed effects (α_p), or province-year fixed effects (α_{pt}). Mig_{it}^l are sets of dummies for the type of migration event.[19] X_{it} are the set of additional controls that may or may not

[19]Here I include dummies for interprovincial and intraprovincial migration events, and market and intramarket migration events (these are aggregated provinces across island groups, which generates less error given that provincial expansions occur within, and not between, islands and have not yet been corrected). Future augmentations will include dummies for continuation premiums will be added to augment the specification.

time vary.[20] I also run the above specification on the IPUMS census data, adjusting for observables that are available.

I define the fixed effect specification as follows, maintaining only the time varying parameters from the previous equation:

$$ln(wage_{iopt}) = \alpha + \gamma_1 age_{it} + \gamma_2 age_{it}^2 + \theta Urban_{it} + \sum_l \delta_l Mig_{it}^l + \sum_j \alpha_j + (\eta_i + \varepsilon_{iopt})$$

The above specification contains the following adjustments to control for unobserved heterogeneity: the unobserved individual fixed affect, η_i, is included and placed in parenthesis with the error term to highlight that its exclusion from the OLS model is biasing the wage premium attributable to migration; and the fixed effect terms α_j included are either year, or province-year and occupation, or occupation-province and year-province fixed effects. These specifications will be augmented in the future to account for marriage entrance and exits and birth of children. Finally, to compare the results across the fixed effect specification I conduct seemingly unrelated post-estimation analysis. I then conduct multi-equation restriction tests of regressors. I find that the estimated coefficients of an intraprovincial migration wage premium and an intramarket wage premium are not statistically different when comparing specifications 1-3 and 4-6 in table C7; while the estimated coefficients for interprovincial wage premiums and intermarket wage premiums are statistically different.[21] Enriching the models will further control for individual heterogeneity and a cleaner measure of the wage premiums to migration.

The results of estimating these models are presented in the subsequent tables.

[20] The only time varying control is the urbanization of the location - whether rural or urban - of the location of the individual.

[21] At the standard 95% confidence level.

Table C5: Mincer Regressions of IPUMS Wages - 1995 Indondesia

log(hourly wage)	(1)	(2)	(3)	(4)	(5)	(6)
School Years	0.0969***	0.0581***	0.0587***			
	(0.000581)	(0.000726)	(0.000719)			
Up to Primary				0.268***	0.200***	0.191***
				(0.0107)	(0.0103)	(0.0102)
Up to Lower Secondary				0.530***	0.348***	0.346***
				(0.0119)	(0.0117)	(0.0116)
Up to Upper Secondary				0.887***	0.533***	0.533***
				(0.0113)	(0.0117)	(0.0116)
Up to College				1.422***	0.914***	0.915***
				(0.0123)	(0.0134)	(0.0133)
Age	0.0615***	0.0445***	0.0457***	0.0590***	0.0435***	0.0447***
	(0.000967)	(0.000935)	(0.000918)	(0.000967)	(0.000935)	(0.000918)
Age sq	-0.000545***	-0.000386***	-0.000394***	-0.000551***	-0.000392***	-0.000401***
	(1.28e-05)	(1.22e-05)	(1.20e-05)	(1.28e-05)	(1.23e-05)	(1.20e-05)
Sex	-0.229***	-0.152***	-0.141***	-0.268***	-0.167***	-0.156***
	(0.00476)	(0.00515)	(0.00507)	(0.00480)	(0.00517)	(0.00509)
Urbanization	-0.0459***	0.0216***	0.00690	-0.0249***	0.0239***	0.00919*
	(0.00500)	(0.00503)	(0.00506)	(0.00497)	(0.00502)	(0.00505)
Religion	0.122***	0.0956***	0.114***	0.117***	0.0941***	0.109***
	(0.00624)	(0.00596)	(0.00742)	(0.00622)	(0.00595)	(0.00741)
Language	-0.197***	-0.224***	-0.0793***	-0.201***	-0.227***	-0.0767***
	(0.00520)	(0.00503)	(0.00665)	(0.00518)	(0.00502)	(0.00664)
Occupational FE		X	X		X	X
Province FE			X			X
Observations	105,749	105,749	105,749	105,755	105,755	105,755
R-squared	0.372	0.440	0.461	0.377	0.442	0.463

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table C6: OLS Mincer Regressions of IFLS Wages

log(hourly wage)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Provincial Migration				0.134***	0.129***	0.0968***	0.0925***				0.137***	0.129***	0.0947***	0.0905***
				(0.0136)	(0.0136)	(0.0138)	(0.0140)				(0.0137)	(0.0136)	(0.0138)	(0.0140)
Within Province Migration				0.0673***	0.0645***	0.0453***	0.0381***				0.0648***	0.0641***	0.0416***	0.0344***
				(0.0119)	(0.0118)	(0.0118)	(0.0123)				(0.0119)	(0.0118)	(0.0118)	(0.0123)
School Years	0.102***	0.0885***	0.0868***	0.102***	0.0883***	0.0867***	0.0868***							
	(0.000506)	(0.000579)	(0.000586)	(0.000506)	(0.000579)	(0.000586)	(0.000586)							
Primary								0.379***	0.341***	0.303***	0.378***	0.341***	0.303***	0.304***
								(0.00608)	(0.00610)	(0.00617)	(0.00608)	(0.00610)	(0.00617)	(0.00618)
Obligatory Secondary								0.764***	0.682***	0.640***	0.763***	0.681***	0.640***	0.641***
								(0.00774)	(0.00789)	(0.00799)	(0.00774)	(0.00789)	(0.00799)	(0.00800)
Non-Obligatory Secondary								1.085***	0.955***	0.909***	1.083***	0.953***	0.909***	0.909***
								(0.00762)	(0.00805)	(0.00817)	(0.00762)	(0.00805)	(0.00817)	(0.00818)
Tertiary								1.450***	1.235***	1.202***	1.447***	1.232***	1.200***	1.201***
								(0.00903)	(0.0105)	(0.0106)	(0.00903)	(0.0105)	(0.0106)	(0.0106)
Age	0.0871***	0.0846***	0.0843***	0.0872***	0.0848***	0.0844***	0.0844***	0.0844***	0.0828***	0.0825***	0.0846***	0.0830***	0.0826***	0.0825***
	(0.000608)	(0.000610)	(0.000609)	(0.000608)	(0.000610)	(0.000609)	(0.000609)	(0.000611)	(0.000611)	(0.000610)	(0.000611)	(0.000611)	(0.000611)	(0.000611)
Age sq	-0.000870***	-0.000844***	-0.000838***	-0.000872***	-0.000846***	-0.000839***	-0.000838***	-0.000861***	-0.000839***	-0.000834***	-0.000862***	-0.000840***	-0.000835***	-0.000834***
	(7.72e-06)	(7.71e-06)	(7.70e-06)	(7.72e-06)	(7.71e-06)	(7.70e-06)	(7.71e-06)	(7.76e-06)	(7.75e-06)	(7.73e-06)	(7.76e-06)	(7.75e-06)	(7.73e-06)	(7.74e-06)
Sex	-0.261***	-0.254***	-0.254***	-0.260***	-0.253***	-0.254***	-0.254***	-0.278***	-0.268***	-0.271***	-0.277***	-0.267***	-0.271***	-0.271***
	(0.00369)	(0.00407)	(0.00407)	(0.00369)	(0.00407)	(0.00407)	(0.00407)	(0.00372)	(0.00410)	(0.00410)	(0.00373)	(0.00410)	(0.00410)	(0.00410)
Urbananization	0.193***	0.121***	0.112***	0.192***	0.121***	0.111***	0.112***	0.209***	0.128***	0.116***	0.208***	0.127***	0.116***	0.117***
	(0.00392)	(0.00421)	(0.00436)	(0.00392)	(0.00421)	(0.00436)	(0.00437)	(0.00392)	(0.00422)	(0.00437)	(0.00392)	(0.00422)	(0.00437)	(0.00438)
Religion	-0.0207***	-0.0140**	-0.0207***	-0.0206***	-0.0139**	-0.0210***	-0.0209***	-0.0132**	-0.00409	-0.0135*	-0.0130**	-0.00401	-0.0137*	-0.0148**
	(0.00591)	(0.00590)	(0.00745)	(0.00591)	(0.00590)	(0.00745)	(0.00747)	(0.00592)	(0.00591)	(0.00747)	(0.00592)	(0.00591)	(0.00747)	(0.00749)
Ethnicity	0.0980***	0.0913***	0.00596	0.0986***	0.0917***	0.00675	0.00353	0.0751***	0.0743***	-0.00938*	0.0758***	0.0748***	-0.00857	-0.0115**
	(0.00381)	(0.00382)	(0.00531)	(0.00381)	(0.00382)	(0.00531)	(0.00532)	(0.00382)	(0.00382)	(0.00533)	(0.00382)	(0.00382)	(0.00533)	(0.00534)
Constant	-2.414	-2.622***	-2.453***	-2.418	-2.626***	-2.509***		-2.201	-2.601	0.107	-2.207	-2.606	0.121	
	(3,417)	(0.165)	(0.266)	(3,417)	(0.165)	(0.266)		(4,244)	(2,639)	(11,916)	(4,243)	(2,638)	(11,918)	
Year FE	X	X	X	X	X	X		X	X	X	X	X	X	
Occupation FE		X	X		X	X	X		X	X		X	X	X
Province FE			X			X				X			X	
Year*Province FE							X							X
Observations	549,138	549,129	549,129	549,138	549,129	549,129	549,129	550,053	550,044	550,044	550,053	550,044	550,044	550,044
R-squared	0.367	0.386	0.397	0.367	0.387	0.397	0.406	0.373	0.388	0.399	0.373	0.388	0.399	0.408

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table C7: FE Mincer Regressions of IFLS Wages

log(hourly wage)	(1)	(2)	(3)	(4)	(5)	(6)
Provincial Migration	0.0973***	0.0714***	0.0735***			
	(0.0136)	(0.0140)	(0.0140)			
Within Province Migration	0.0365***	0.0300**	0.0308**			
	(0.0116)	(0.0121)	(0.0121)			
Market Migration				0.164***	0.148***	0.143***
				(0.0235)	(0.0236)	(0.0236)
Within Market Migration				0.0628***	0.0614***	0.0606***
				(0.0164)	(0.0164)	(0.0164)
Age	0.0879***	0.0837*	0.0812	0.0945***	0.101***	0.105***
	(0.00325)	(0.04311)	(0.0513)	(0.00333)	(0.0132)	(0.0329)
Age sq	-0.000871***	-0.000855***	-0.000935***	-0.000942***	-0.000938***	-0.000936***
	(8.38e-06)	(8.46e-06)	(7.462e-11)	(8.58e-06)	(8.61e-06)	(8.63e-06)
Urbananization	0.0449***	0.0348***	0.0887***	0.103***	0.0971***	0.0971***
	(0.0074)	(0.0077)	(0.0079)	(0.00762)	(0.00764)	(0.00766)
Year FE	X			X		
Occupation FE		X			X	
Occupation*Province FE			X			
Year*Province FE		X	X			
Occupation*Market FE						X
Year*Market FE					X	X
Observations	630,404	630,384	630,404	630,404	630,384	630,384
Number of individuals	31,488	31,482	31,487	31,488	31,482	31,482
R-squared	0.430	0.445	0.437	0.430	0.445	0.447

Notes: Markets are defined as island groupings, such that all provinces in the island of Sumatra along with the island provinces of Riau are market 1; all provinces in the island of Jawa are market 2; Bali and all islands in the Nussa archipelago are market 3; all provinces in Kalimantan (Borneo) are market 4; and all provinces in the Sulawesi island, along with all the islands of Maluku and provinces in the island of New Guinea are Market 5. The continued premium indicator takes a value of one upon the first inter-Market migration event and forever after. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Comparisons of the IPUMS results in table C5 and those from the IFLS results in tables C6 and C7 portray similar results for the coefficients, although the estimates in the IFLS are slightly higher. This may result not from the larger dataset per se, rather the time dimension associated with the IFLS data results in less observations within a time period vis-à-vis the IPUMS data being cross-sectional in year 1995.

D Transition Function for Uncertain Investment

In this appendix I document the transition functions obtained from the data when the agent's investment in education is uncertain. This would mean that with some probability the investment, if undertaken, fails and the child is subject to wage $w^{\ell}_{h=0}$ in their adult life. Further, it has practical implications for the investing agent in their lifetime: the uncertain outcome regarding the educational investment means they are necessarily foregoing the child labor that would otherwise be available. Since agents in my simpler model do not observe the ability of the child prior to making the decision, the cut-offs that I use are a rough proxy for transition probabilities that might be estimated in the full model when parents make this uncertain decision observing their child's ability (for example if they invest in a certain level of schooling but the child is unable to finish it). Generating an uncertain outcome to educational investment will allow me to ascertain how such uncertainty affects agent's decisions.

There are two ways to accomplish this. One is to assume that anything below 6 years of schooling (that is, not having finished primary education) yields $h = 0$ (equivalent to not having been educated at all). Such a transition function is shown below, where the sub-index on the function N indicates the cutoff of 6^{th} grade:

$$
N_6(z'|z, I_{k,t=2} = 1) = \begin{array}{r} z, I_{k,t=2} = 1 \downarrow \ \backslash \ z' \rightarrow \\ h=0,l=1,j=1,e=0 \\ h=0,l=1,j=1,e=1 \\ h=0,l=1,j=2,e=0 \\ h=0,l=1,j=2,e=1 \\ h=0,l=2,j=2,e=0 \\ h=0,l=2,j=2,e=1 \\ h=0,l=2,j=1,e=0 \\ h=0,l=2,j=1,e=1 \\ h=1,l=1,j=1,e=0 \\ h=1,l=1,j=1,e=1 \\ h=1,l=1,j=2,e=0 \\ h=1,l=1,j=2,e=1 \\ h=1,l=2,j=2,e=0 \\ h=1,l=2,j=2,e=1 \\ h=1,l=2,j=1,e=0 \\ h=1,l=2,j=1,e=1 \end{array}
\begin{array}{c} \begin{array}{cccc} h=0,l=1 & h=0,l=2 & h=1,l=1 & h=1,l=2 \end{array} \\
\left(\begin{array}{cccc}
1 & 0 & 0 & 0 \\
0.1416 & 0 & 0.8584 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0.1545 & 0 & 0.8455 \\
0 & 1 & 0 & 0 \\
0 & 0.1518 & 0 & 0.8482 \\
1 & 0 & 0 & 0 \\
0.1171 & 0 & 0.8829 & 0 \\
1 & 0 & 0 & 0 \\
0.0871 & 0 & 0.9129 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0.0205 & 0 & 0.9795 \\
0 & 1 & 0 & 0 \\
0 & 0.0396 & 0 & 0.9604 \\
1 & 0 & 0 & 0 \\
0.0262 & 0 & 0.9738 & 0
\end{array}\right) \end{array}
$$

The second method is to use Indonesia's law to guide this outcome. As appendix B discussed, compulsory education in Indonesia is up to the 9^{th} grade for those children in later generations. So in this case I consider anything below 9 years of schooling as uneducated, such that any investment $e = 1$ that leads to less than 9 years of schooling results in skill $h = 0$ for the child (once again, the sub-index on N highlights this cutoff):

$$N_9(z'|z, I_{k,t=2} = 1) = \begin{array}{r} z, I_{k,t=2} = 1 \downarrow \ \backslash \ z' \rightarrow \\ h=0,l=1,j=1,e=0 \\ h=0,l=1,j=1,e=1 \\ h=0,l=1,j=2,e=0 \\ h=0,l=1,j=2,e=1 \\ h=0,l=2,j=2,e=0 \\ h=0,l=2,j=2,e=1 \\ h=0,l=2,j=1,e=0 \\ h=0,l=2,j=1,e=1 \\ h=1,l=1,j=1,e=0 \\ h=1,l=1,j=1,e=1 \\ h=1,l=1,j=2,e=0 \\ h=1,l=1,j=2,e=1 \\ h=1,l=2,j=2,e=0 \\ h=1,l=2,j=2,e=1 \\ h=1,l=2,j=1,e=0 \\ h=1,l=2,j=1,e=1 \end{array} \begin{array}{c} \begin{array}{cccc} h=0,l=1 & h=0,l=2 & h=1,l=1 & h=1,l=2 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.3799 & 0 & 0.6201 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0.3496 & 0 & 0.6504 \\ 0 & 1 & 0 & 0 \\ 0 & 0.5125 & 0 & 0.4875 \\ 1 & 0 & 0 & 0 \\ 0.3964 & 0 & 0.6036 & 0 \\ 1 & 0 & 0 & 0 \\ 0.2635 & 0 & 0.7365 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0.1365 & 0 & 0.8635 \\ 0 & 1 & 0 & 0 \\ 0 & 0.2078 & 0 & 0.7922 \\ 1 & 0 & 0 & 0 \\ 0.1429 & 0 & 0.8571 & 0 \end{pmatrix} \end{array}$$

Obviously these are strong assumptions on the progression of skill attainment due to investment. Particularly because the mincer regressions do not justify these cutoffs (even one year of schooling is associated with non-trivial returns to wages). However I do this purely as an exercise. It is interesting to note that the probability that a child obtains skill when an investment occurs increases when parents themselves are skilled. It also seems that the probability that a child obtains skill increases conditional on the investing agent being skilled and choosing to migrate to the other labor market.

References

Aguirregabiria, V. and P. Mira (2010, January). Dynamic discrete choice structural models: A survey. *Journal of Econometrics 156*(1), 38–67.

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.

Bernal, R. and M. P. Keane (2010, May). Quasi-structural estimation of a model of child care choices and child cognitive ability production. *Journal of Econometrics 156*(1), 164–189.

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.

Cameron, S. V. and J. J. Heckman (1998, April). Life cycle schooling and dynamic selection bias: models and evidence for five cohorts of american males. *The Journal of Political Economy 106*(2), 262–333.

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

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

Cunha, F. and J. J. Heckman (2008, September). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, May). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

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 (2015a, February). What accounts for the racial gap in time allocation and intergenerational transmission of human capital? Washington Universtiy in St. Louis.

Gayle, G.-L., L. Golan, and M. A. Soytas (2015b, August). Estimation of dynastic life-cycle discrete choice models. Washington Universtiy in St. Louis.

Gemici, A. (2011, July). Family migration and labor market outcomes. Royal Holloway, University of London.

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.

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.

Kennan, J. and J. R. Walker (2011, January). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

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.

Lessem, R. (2013, July). *Mexico-U.S. immigration: effects of wages and border enforcement*. Ph. D. thesis, University of Wisconsin-Madison. Tepper School of Business, Carnegie Mellon University, Pittsburgh.

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.

McFadden, D. L. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. Academic Press: New York.

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.

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.

Rust, J. (1987, May). Optimal replacement of gmc bus engines: an empirical model of harold zurcher. *Econometrica 55*(5), 999–1033.

Strauss, J., K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herawati, and F. Witoelar (2004, March). The third wave of the indonesia family life survey (ifls): overview and field report. Technical report, RAND WR-144/1-NIA/NICHD.

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.