Universitat Autònoma de Barcelona

Essays on Migration, Altruism, and Intergenerational Mobiilty

Alberto Jose Ramirez

Essays on Migration, Altruism, and Intergenerational Mobiilty

Doctoral Thesis

International Doctorate in Economic Analysis
Departament d'Economia i d'Història Econòmica
Universitat Autonòma de Barcelona

Alberto Jose Ramirez
Candidate

Joan Llull
Supervisor

Inés Macho-Stadler
Tutor

Bellaterra, September 2018

Acknowledgements

I would first like to thank my supervisor Joan Llull and tutor Inés Macho-Stadler for their guidance and help throughout these years. Joan has been helping me since the beginning to guide me in the organization of research topics, the technical aspects related to my thesis, and pushing me along to realize this dissertation. You've helped me greatly when I've been stuck in both modeling and estimation, as well as giving me tips and techniques to improve my presentation skills during applied seminar presentations and the yearly evaluations. Inés has been supportive in both helping me improve my presentation skills during my second year, and in helping me decide upon what summer schools and conferences to participate in.

I would additionally like to thank Jordi Caballé for meeting with me over the years to advise me on the topics of intergenerational mobility and with regards to my career aspirations and prospects after my PhD. Lídia Farré, I would also like to thank you for providing me with the the codes to your and Franceso Fasani's paper that used the IFLS dataset. With your help I was able to shave off weeks of data cleaning and merging to quickly start exploring my thesis ideas. I would also like to thank Nezih Guner and Hannes Mueller for the insights you've shared regarding my thesis chapters in your respective student workshops.

Thanks to Àngels and Mercè for helping me along whenever I had a question about the program, my scholarship, and the paperwork that needed to be completed at the various stages of my doctoral career. But most of all, for being supportive when things seemed difficult.

I would also like to thank my colleagues Francesco, Alejandra, Andreii, Alexei, Matt, Tongbin, Alessandro, Lavinia, Anna, Analisa, Jesús, and José. Throughout the years we've bounced ideas off one another, helped each other on how communicating our ideas, and sought each other's help on how to implement them. And just as importantly, I thank you all for the emotional support that we've provided each other throughout the years in and out of our offices.

Moreover, without the assistance of the Agència de Gestió d'Ajuts Universitaris i Recerca I would not have been able to accomplish this thesis. So I gratefully acknowledge their financial support via the FI AGAUR 2015 FI_B00624 fellowship, which also financially supported my attendance at various summer schools and conferences.

Finally, I want to thank my family for their continuing support over the years. I want to dedicate this thesis to them, and specifically to my niece Elisa and new-born nephew Rafael. May your respective journeys in life lead you to fulfillment, however you choose to define it, and that you both realize the investments that we, your migrant ancestors, bestowed upon you.

Contents

Abstract

This thesis is a collection of essays focused on developmental economics. The first portion is dedicated to the analysis of transaction fees on remitting behavior and how the existence of transaction costs impact the canonical test of the altruistic remittance motive. The second portion is dedicated to the empirical analysis and estimation of the effect of family migration in Indonesia on the investment in their children's human capital. Our work contributes to the ongoing discussion in developmental economics on how migration and capital flows impact both those who stay behind and the future generations who might reap the fruits of the investments made by their progenitors.

The economic literature on remittances has not settled on what is the driving motive for the large flows of capital transferred by migrants to their home countries. In the first chapter we investigate the role of transfer fees and the cost incurred by migrant remitters on the altruistic remitting motive. We explore a theoretical treatment of the problem of transaction costs induced by the existence of a fee to send capital between two households, and how the distortion affects remitting behavior indirectly through inter-temporal effects on savings decisions. While we have not yet generalized our results to the class of convex functions for a transfer cost, we show in a simple two period model that remittances decrease as a function of increasing remittance fees due to the theoretical increase in savings that the existence of higher fees induce. We then ask how the existence of a transaction cost induced by a transfer fee affects the canonical test of the altruistic motive for inter-household transfers, first elucidated by Becker (1974). Here we find that the distortions from transfer costs negatively impact the inference of altruism from the theoretical limit based on Becker's test. Finally, we calibrate the model in the context of the Cuban migrant community in the U.S. We find, in a no cost model, that

the aggregate time series of remittance flows observed is mostly explained by altruism; and that the calibrated transaction fee assuming migrants are altruistic comes fairly close to the average fees reported in the literature for remittances to Cuba. We argue that this evidences the need to consider the scope of the remittance landscape when inferring the motive of this behavior, especially the altruistic one.

In the second chapter we shift focus to migration of households and the associated outcome on children's schooling attainment. This work is relevant given the increasing migration flows both within and between countries that can have disruptive effects on the family. As such the effects of migration on the household's children has been a consistent topic within the development literature because of the various dimensions through which migration can impact them. To explore this topic, we turn to the Indonesian Family Life Survey, a longitudinal panel data maintained by the RAND Corporation. We first analyze how internal migration in general affects wage premiums, given that the migration literature finds evidence that expected wage premiums between labor markets is the principal motivation for migration. Then we look at how family migration specifically affects schooling attainment in migrant children. Our analysis is descriptive in nature but points to positive associations in both cases, with family migration reducing the hazard of exiting higher schooling levels in a country where the government is still actively combating child labor, despite its illegality.

We take these two qualitative findings from the second chapter to develop in chapter three a simple intergenerational model of family migration and investment in a child's human capital. Investigating a plausible selection mechanism that plagued the endogeneity in the descriptive work allows us to not only comment on whether a causal effect exists but also on the magnitude and direction of effects. In conducting policy experiments we show that inducing family migration via full (or nearly full) subsidy, especially among low-skilled households, leads to higher average wages in the next generation than in the base case where families are tied to their home location. Finally, while relaxing the migration cost in our estimated model does lead to improvements, we also show that the cheapest policy outcome might be to relax the opportunity costs to educate that still exist in Indonesia, whose effects are greater and the costs arguably lower than a migration policy would accomplish. We also argue that this last result doesn't diminish the effects of migration, but enhances it as the apparent disparities between labor markets is empirically favoring migration and warrants a further look at regional investments in human capital. And while a full migration subsidy may seem extreme, we note that the Indonesian government has provided impoverished Jawanese the opportunity to move themselves and their families to other islands within Indonesia at no cost through transmigration programs, a topic only lightly touched upon in the literature.

Chapter 1

The Effect of Transfer Costs on Remittance Behavior in an Altruistic Framework

1.1 Introduction

This chapter builds on the branch of the remittance literature concerned with altruism while taking as a motivational example the Cuban diaspora in the U.S. Much has been said in the literature regarding remittances and the motivations for them. But seldom does one come across a study of how transfer fees may affect remitting behavior, and importantly on its effect in estimating altruism. In one of the few papers to study such this, Aycinena et al. (2010) conducted an experiment on migrants in the U.S. and find that decreasing fees results in total outflows of remittances, and larger average remittances. While their work establishes a causal relationship, it does not comment on the motivation for why agents were willing to remit more and unable to establish the mechanism.

We look at altruism as a motive to remit and the implications of remittance fees on the inference of altruism. A framework similar to Stark (1999) is used where our remitting agent is the welfare maximizer of his composite household. We find from the FOCs that the proportion of wage chosen to remit is hyperbolic, indicating that agents would remit less as a proportion of their wage as their wage grows. Put differently, they may not remit in proportion to the growth rate of their wage. We show that the main test researchers use to infer an altruistic motive from remittance data becomes distorted when remittance fees are incorporated into an altruistic framework. We solve a two period model to understand the dynamics of transfer fees on remittance behavior and the implications on the literature's method of testing for altruism. We find that under certain conditions the absolute transfer is decreasing in the transaction fee when the remitter assumes the cost of the transfer. This relationship is composed of both a direct effect on the transfer and an indirect effect resulting on how transaction fees affect inter-temporal substitution of consumption through the effect on the savings decision.

We then proceed to analyze in a one period case how the existence of a transaction fee affects the main test researchers have used to test altruism, first elaborated by Becker (1974). We find that in both assumption on the functional form the value of the transfer derivative deviates from its theoretical value. An implication of this finding is the introduction of bias when costs are not accounted for, which may result in the rejection of altruism when empirically tested. Which is to say, instead of estimating the effect of altruism one instead estimates a distortion of it. We then move to the two period specification and find that the indirect effect here of the transaction fee through the savings decision works to compensate the distortionary effect on what was found in the one period case. However, this inter-temporal, indirect effect on the transfer derivative decreases its compensatory contribution as the number of periods increases (this will depend on the way agents discount the future). In an experiment on remittance fees conducted by Aycinena et al. (2010) on El Salvadorian migrants in the U.S., the authors don't find conclusive evidence of an inter-temporal effect on remitting behavior due to fee changes, given the low power of the estimation, although the coefficients on the magnitude of savings increase as fees increase.

Finally we simulate the framework for illustrative purposes. Calibrating the altruism parameter to match aggregate remittances suggests that Cuban-Americans are arguably moderately altruistic. We then calibrate the transfer fee that matches the aggregate remittances at the upper bound of what we call rational altruism. This results in a transfer fee that is only slightly larger than that reported in Orozco (2009) and Gonzalez and Larson (2008). We regard these results as qualitative and not indicative that altruism is the complete motive given the limitation of the data used. It serves to show the potential of the model and to help understand the effects of transaction fees on remittances.

The rest of this chapter is organized as follows: in section 1.2 we review the literature on remittances and describe the transfer derivative (which is used as the main empirical test of altruism); section 1.3 documents the motivational diaspora in the U.S.; section 1.4 develops the framework and solves a 2 period model; section 1.5 analyzes the implications of transfer fees on transfer decisions and the transfer derivative; section 1.6 details the simulation, the data used, and the results; and section 1.7 concludes. The appendices contain proofs, comparative statics on the remaining model primitives, and figures.

1.2 Literature Review

The motives for remittances have been largely studied by economists and these concern such reasons as altruism (pure and selfish), mixed motives, strategic, insurance and moral hazard, and exchange of services. As Rapoport and Docquier (2005) point out in their survey, although models that describe these motivations for migrant remittances to their kin in their home country are largely compartmentalized to the individual framework, there is no reason to believe that they are mutually exclusive of each other. For this reason researchers have started to focus more on the mixed motives (only lightly touched upon in their survey).

The framework we deal with and discuss here deals with the altruistic motive. In his seminal paper, Becker (1974) proposed the altruism motive and presented a generalized way it could be modeled, along with implications on the price and income elasticity effects to the social income that agents transfer. Within this framework the agent who remits is conducting a joint maximization of his and the recipient's utility functions, the idea being that the remitter cares about the recipient's welfare. Stark (1999) introduced the notion of unilateral and mutual altruism by taking the weighted average of value functions, where the weighting parameter is the degree of altruism. Regardless of the modeling of the utilities, the main result to come from Becker's work is the transfer derivative, defined below:

$$\frac{\partial T}{\partial I^h} - \frac{\partial T}{\partial I^m} = -1 \tag{1.1}$$

where m is the remitter and h is the recipient. This equation captures the income substitution response between agents. Its interpretation boils down to the observation that in this model a one dollar increase in the migrant's income with a likewise decrease in the recipient's income should lead to a one dollar transfer. This constitutes the empirical test that researchers take to the data.

Empirically pure altruism has not fared well based on failure of tests of the transfer derivative; although it has been shown, for example in Lucas and Stark (1985) study of Botswana, that remittances increase with the remitter's income (as the altruistic model, and indeed most other models, predict). Moreover, Foster and Rosenzweig (2001) have empirically shown that under imperfect commitment full insurance is not attainable due to informational barriers, but altruism reduces (although doesn't eliminate) the commitment constraint. Raut and Tran (2005) look at the intergenerational transfers between parents and children in Indonesia and estimate a transfer derivative very close to -1. They use an interesting method to control for selection bias (the Altonji-Ichimura method), which accounts for the non separability between income and preferences on remitting behaviors. The authors here argue that the nonlinearity and heterogeneity inherent in the selection bias will generate transfer derivatives close to 0 if not accounted for.

Another branch of the literature looks at patterns of remittances through models of familial inter-temporal contracts. It is for this reason that remittances are sometimes viewed through the insurance lens. As pointed out above, these models have their limitations on the foundation of the commitment constraint. Altruism is variously used to try to relax that constraint, however it is still an issue since tests of full insurance are difficult to conduct. The limitation on the commitment constraint may be due to altruism, or due to a lack of the type of data needed to test for the full insurance that these models would predict, as has been pointed out by Kinnan and Townsend (2012).

1.3 The Cuban Migrant Community in the U.S.

We consider the Cuban diaspora for studying altruism as this group has idiosyncratic characteristics that may lend itself toward this motive. First, this group receives a preferential migration treatment from the US. In 1995 the US government amended the 1966 Cuban Adjustment Act that previously sanctioned the Cuban government over Cold War hostilities. This amendment led to the “Wet-foot, Dry-Foot” policy where, *balseros* who were caught in the waters of the Florida straights would be sent back to Cuba; but those who made it to American soil (regardless of the method) would have “dry feet”, qualifying them for legal permanent residency status and citizenship. The impetus for 1995 amendment was to prevent another Mariel situation that the Castro regime was threatening as a result of the continued US embargo during though economic conditions in Cuba after the collapse of the Soviet Union. This preferential treatment where Cuban migrants, regardless of how they enter the country, are not deemed illegal may reduce the incentive of reverse migration (especially since doing so would brand one a “counter revolutionary” in Cuba, with the stigma and jail time attached). This reduced likelihood for circular migration may call into question investment, profit, and loan/insurance motives of remittances.

The 1994 crisis (the “Special Period”) and the continued embargo led to a change in the official government attitude toward remittances in Cuba (Dilaz-

Briquets, 2013). This was arguably an economically and politically expedient move by the Cuban government given the shock discussed above due to the Soviet collapse. For despite the US embargo, Cuba's largest trading partner in the region became the US; however, due to the embargo, the US requires cash settlement on trade balances with Cuba regarding the food and medicines that can be traded - purchases on credit are forbidden. To thus capture the remittances entering the country the government opened dollar stores, offering staple products previously available through the *bodegas* that disbursed state-provided, rationed goods (and pricing these at market rates), as well as other higher goods (Blue, 2004). Thus, in a reversal of pre-Soviet policy, the Cuban government actively fostered remittances from abroad to fund itself. This has an obvious economic consequence as the purchasing power of the peso (CUP) decreased relative to the dollar; and the post 2004 change of using the convertible peso (CUC) has effectively maintained the system.

Citing the Bush administration's June 2004 implementation of remittance and travel restrictions (RTR), the Cuban government withdrew the American dollar from official circulation on November 8, 2004 (Gonzalez and Larson, 2008). As a consequence the state-owned dollar stores, at which Cubans were able to spend the remittance and other tourist dollars in their possession, were switched to the Cuban Convertible Peso, the CUC (the tourist currency - a form of Foreign Exchange Certificate - pegged to the American dollar, but in reality more expensive due to a tax placed on exchanging dollars to the CUC); its typical official exchange rate is 0.80CUC:1$. This made the CUC the second de facto currency on the island, along with the Cuban Peso (CUP), the currency used to pay wages, which has an exchange rate 25CUP:1CUC.

The RTRs restricted transfers through official forwarding agencies to a maximum of 300$/quarter and travel to the island to once every three years. Despite this remittances pre and post 2004 remained relatively stable. Yet the very conditions that favor Cuban migration to the US (the humanitarian crisis resulting from the embargo) are the very conditions that increase the cost to remit to the island. As Gonzalez and Larson (2008) notes in their survey of remitting agencies and documented in Orozco (2002), compared to the rest of the Latin American and Caribbean countries Cubans pay high transaction costs to remit.[1]

1.4 An Altruistic Remittance Framework with Transfer Costs

To study and then simulate the motivational context of Cuban remittances we develop in this section an altruistic remitting framework. We assume that the remitting agent transfers a proportion of his wage to his family member abroad (Agent 2). We further assume that the transfer fee is applied as a transfer cost when the remitting agent remits. To explore the effect of these transfer costs on the transfer decisions and on the transfer derivatives we make two functional form assumptions on how the fee is applied, which are inverses of each other.

1.4.1 The Framework

Suppose there are two agents each in two distinct economies: the migrant, agent 1, and the recipient, agent 2. These two agents share the same well-behaved, time separable utility function, with the same preference parameter (γ) and inter-temporal elasticity of substitution. Agent 1 is a representative agent of his diaspora community in the first economy who takes wages and interest rates as given. We therefore consider only a partial equilibrium. This agent supplies his labor inelastically and so leisure is not a concern. The same assumptions are symmetrically assumed for agent 2 in his economy. At time $t =$

[1] After 2009, the Obama administration reversed this policy allowing remittance transfers of up to 10,000\$/day to be sent to the island and unlimited mobility to Cuba for Cuban-American's. In the work that follows we average the sets of data to work with one data set; the results of averaging place the data within the mid 2000s.

0 agents in both economies draw from their wage distribution the full sequence of wages that they will receive in their lifetime, with perfect information on the other's wage process. Therefore there is no uncertainty in this framework. Agents end their lives with no resources, so bequests are not possible. Finally, transfer flows are such that agent 1 remits to agent 2. This implicitly assumes that the distribution of wages in agent 1's economy dominates the wages in agent 2's economy everywhere along the distribution. With this assumption a reverse flow can not be optimal in equilibrium. Shimada (2012)[2] assumes a similar expected wage process and shows that if the altruism parameter is above a certain threshold hen agent 1 will remit to agent 2.

With his wage sequence in hand, Agent 1 must decide on the triple (c_t, τ_t, a_{t+1}) given the current state vector $s = (a_t, y_t)$, where τ_t is the proportion of his wage he chooses to remit. Since we are working within a PE construct this ex-flux of capital is assumed to not have a general equilibrium effect. Further, agent 1 cares about agent 2 and so he incorporates the other agent's welfare in his utility function. Agent 1 starts life with no assets (since bequests have been excluded) and works until his death. Finally, agent 1 does not have access to financial markets and so is borrowing constrained - that is, $a_{t+1} \geq 0$.

Agent 2 receives a state wage with the relationship previously noted between the wages of agents 1 and 2. This would augment the state vector to $s = (a_t, y_t, \widetilde{y}_t)$, where $\widetilde{y}_t$ is the wage of agent 2.[3] Further, in this framework we do not afford agent 2 a savings mechanism. This can be relaxed by imposing a no-arbitrage condition between the two economies such that $r \geq \widetilde{r}$. In equilibrium we would expect this condition to be binding. However, we exclude this from the current modeling exercise. as it is more appropriate in a GE framework. Therefore, agent 2 is forced in each period to consume all his

[2]Shimada (2012) studies the remittance behavior of an agent before migrating; that is, under what levels of altruism will an agent migrate to provide remittances if the agent provides remittances at all.

[3]We further note that given the context of the motivational economy for agent 2, a PE for economy 2 is not implausible since the state centrally plans wages and interest rates.

available resources.

We take the perspective that the migrant, agent 1, is maximizing the utility over the composite household. Taking the perspective of the system as an entire household facilitates the analysis and, later, the simulation. Since agent 1 takes into consideration the welfare of agent 2 in his utility $\tau_t > 0$ in equilibrium.

We distinguish between agents 1 and 2 through tildes (where Agent 2 receives the tilde over all the similarly ascribed variables for Agent 1). So we write the value function as:

$$V_t(s_t) = \max_{c_t, \tau_t, a_{t+1}} (1-\alpha)U(c_t) + \alpha\widetilde{U}(\widetilde{c}_t) + \beta V_{t+1}(s_{t+1}) \tag{1.2}$$

subject to the following constraints:

$$c_t + a_{t+1} \leq (1 - \tau_t\varphi_i)y_t + (1+r)a_t \tag{1.3}$$

$$\widetilde{c}_t \leq \tau_t y_t + \widetilde{y}_t \tag{1.4}$$

$$c_t \geq 0 \qquad \widetilde{c}_t \geq 0 \qquad \tau_t \in [0,1] \qquad a_{t+1} \geq 0,\, \forall t \neq 1, T \tag{1.5}$$

where β is the discount factor as usually defined. Since U is well behaved the value function has the property that $V' > 0$ and $V'' < 0$. Concerning α, the altruism parameter, we follow Stark (1995). Whether the parameter is interpreted as one-sided altruism or two-sided altruism (in which case α is a reduced form parameter $\alpha(\alpha_1, \alpha_2)$, where α_i is the altruism the respective agent holds for the other) is innocuous. We follow Stark in the sense that Agent 1 is rational so that $\alpha \in (0, 0.5]$. Notice that folding (1.4) into (1.3) we have that Agent 1 would essentially be choosing the consumption stream $\widetilde{c}_t$ for agent 2. However, it seems more natural to model according to a sequence of transfers.

Definition 1.4.1. *A partial equilibrium is a sequence* $\{c(s,t), a'(s,t), \tau(s,t)\}$ *chosen by Agent 1 that maximizes* (1.2) *subject to* (1.3) - (1.5) *given the state* *vector* $s = (a(t), y(t), \widetilde{y}(t))$ *and parameter* r.

1.4.2 Two Period Model with CES Utility

In what follows we define the previous model in two periods. This is an arbitrary simplification that makes the analysis simple, where to keep ideas fixed one could take the first period as the working phase of an agent's life and the second period as the retirement phase. We therefore assume that inter-temporal wage structures are such that first period wages are at least as large as second period wages, and risk-aversion induces inter-temporal transfers to smooth consumption. We make use of these assumptions later.

We define below the utilities that we will work with in this two period framework. We assume that both agents share the same γ coefficient, and that risk aversion implies that $\gamma \geq 1$. This gives the following two period value function:

$$U(c, \widetilde{U}) = (1-\alpha)\frac{c_1^{1-\gamma}}{1-\gamma} + \alpha\widetilde{U}_1 + \beta\left[(1-\alpha)\frac{c_2^{1-\gamma}}{1-\gamma} + \alpha\widetilde{U}_2\right] \tag{1.6}$$

$$\widetilde{U}(\widetilde{c}_t) = \frac{\widetilde{c}_t^{1-\gamma}}{1-\gamma} \tag{1.7}$$

and the budget constraints:

$$t = 1: \quad c_1 = (1-\tau_1\varphi_i)y_1 - a_2 \qquad t = 2: \quad c_2 = (1-\tau_2\varphi_i)y_2 + (1+r)a_2 \tag{1.8}$$

$$\widetilde{c}_1 = \tau_1 y_1 + \widetilde{y}_1 \qquad \widetilde{c}_2 = \tau_2 y_2 + \widetilde{y}_2 \tag{1.9}$$

The parameter, φ_i, captures the transfer cost. When $i = 1$ this parameter has a functional form represented by $\varphi_1 = 1/(1-\overline{\varphi})$, where $\overline{\varphi} \in [0, 1)$ is the percentage representing the transfer fee. This is a convex function in the domain of $\overline{\varphi}$. This specification implies that agent 1 assumes the cost of the transfer and compensates his remittance for it. The assumption is reversed when $i = 2$, where $\varphi_2 = 1 - \overline{\varphi}$. In this linear case the fee would reduce the transfer agent 2 receives. This parameter enters the budget constraint as a proportional adjustments.

Solution of the Model

We solve for the first order conditions under the assumption of a partial equilibrium in wages, interest rates, and transfer fees. We leave the consumption decision as a residual since we are not explicitly concerned with analyzing this choice variable. Furthermore, we leave the proportional transfer decisions in each period as a function of the savings decision for analytical compactness.

Since agent 1 maximizes utility over the composite household, his decision problem is given by:

$$\tau_1^* = \frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}[y_1 - a_2^*] - \widetilde{y}_1}{y_1\left[1 + \varphi_i\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\right]} \tag{1.10}$$

$$\tau_2^* = \frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}[y_2 + (1+r)a_2^*] - \widetilde{y}_2}{y_2\left[1 + \varphi_i\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\right]} \tag{1.11}$$

$$a_2^* = \frac{[\beta(1+r)]^{1/\gamma}[y_1 + \varphi_i\widetilde{y}_1] - (y_2 + \varphi_i\widetilde{y}_2)}{(1+r) + [\beta(1+r)]^{1/\gamma}} \tag{1.12}$$

We note that this model nests the solution of one where there are no transfer frictions by setting the parameter $\overline{\varphi} = 0$ in the respective φ_i as previously defined.

It is also useful to define the linear transfer form of the previous model. To do so, we adjust equations (1.8) and (1.9) by removing the proportional transfer scheme and instead adding and subtracting, as appropriate, transfers T_t. The result yields optimal transfers T_t^* similar to (1.10) and (1.11) except the denominator removes the current period wage. So solving the model with a linear transfer gives us the relation

$$T_t^* = y_t \tau_t^* \tag{1.13}$$

With the assumption that y_t dominates $\widetilde{y}_t$ it can easily be shown that the following relationships hold: $\frac{\partial \tau_t^*}{\partial y_t} < 0$, $\frac{\partial^2 \tau_t^*}{\partial y_t^2} > 0$, $\frac{\partial \tau_t^*}{\partial y_j} > 0$ for $j \neq t$. We further note that just as in Rapoport and Docquier (2005) the following relationship is predicted once (1.13) is applied to (1.10) and (1.11): $\frac{\partial T_t^*}{\partial \widetilde{y}_t} < 0$. This is a natural result since we would expect agent 1 to remit less if agent 2's wage increases. However, as we will show shortly, the transfers T_t^* do not satisfy the transfer derivative in equation (1.1) unless $\overline{\varphi} = 0$.

1.5 Implications of Transfer Costs on Transfer Decisions and the Transfer Derivative

We now want to understand the relationship between transfers and the transfer fee parameter $\overline{\varphi}$. From the first order conditions, it is apparent that there are distortions to decisions in the presence of transfer costs, which have implications on both decisions and the measurement of the transfer derivative.

1.5.1 Transfer Decisions

In what follows we consider only the transfer in the first period, and to facilitate the calculations we use absolute transfers as shown in (1.13).[4]

Substituting equation (1.10) into the definition of a linear transfer (1.13) for the first period yields $T_1^*(a_2^*(\varphi_i(\overline{\varphi})), \varphi_i(\overline{\varphi}))$. By taking the total derivative of this first period transfer we can state the following claim on the effect of transfer fees on the transfer decision, based on the assumption of how the

[4]We consider only the first period since the second period would only consist of the direct effect of $\overline{\varphi}$ on T_2^*. This is simply because the savings decision made in the first period becomes the state variable in the second (and here, final) period.

transfer cost is applied to the remitting agent.

Proposition 1.5.1. *Assume that wages of agents are such that first period wages are at least as large as second period wages and that the economies are at steady state. Then, under the following conditions*

- $y_1 \geq \widetilde{y}_1 \left(\frac{\varphi_i(1-\alpha)}{\alpha}\right)^{1/\gamma} \left(\frac{\gamma-1}{\gamma}\right)$
- $a_2^* \leq y_1 \frac{\varphi_i}{\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma} \varphi_i \gamma + 1}$

the direction of a change in total transfers due to a change in the transfer fee in the first period is fully determined and leads to the following cases, given the two possible methods in which the transfer fee can be applied,:

1. $\frac{\partial T_1^*}{\partial \varphi} < 0$ *under convex transfer costs*
2. $\frac{\partial T_1^*}{\partial \varphi} > 0$ *under the inverse form.*

The proof is in appendix A.1. The first condition indicates that for the signs of the derivative to be determined it is not sufficient for agent 1's wage in the first period to simply be larger than agent 2's; it must be larger than some multiple of agent 2's wage. The same is true of the chosen amount to save for the second period, a_2^*, as it can't exceed some proportion of agent 1's period wage. When these conditions are satisfied, the sign of the derivative of the total transfer as a function of the transfer fee is fully determined; the opposite can not be guaranteed and in such a case the sign is ambiguous.

1.5.2 The Transfer Derivative

To analyze the effect of transfer fees on the transfer derivative in the literature, it helps to rewrite proportional transfers in absolute transfers, per equation (1.13).

Static Transfer Derivative

We define static income as the total available resources prior to transfers: $I^j \equiv I_t^j(y_t^j, a_{t+1})$ for $j = 1, 2$ at a constant t, and where $a_{t+1} \geq 0$ (depending on the particular agent and the assumption on their savings behavior).

We first start with the static transfer derivative, which is defined by placing the definition of I^j in τ^* of equation (**??**):

$$T^* = \frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma} I^1 - I^2}{1 + \varphi_i \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}. \tag{1.14}$$

Then the derivatives of this static transfer function with respect to each individual's income is given as:

$$\frac{\partial T^*}{\partial I^1} = \frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1 + \varphi_i \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}} \qquad \frac{\partial T^*}{\partial I^2} = \frac{-1}{1 + \varphi_i \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}. \tag{1.15}$$

From the above we obtain the following static transfer derivative at period t:

$$\frac{\partial T^*}{\partial I^2} - \frac{\partial T^*}{\partial I^1} = -\frac{1 + \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1 + \varphi_i \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}. \tag{1.16}$$

This result is a direct consequence of the distortionary nature that transfer fees have on remittances.[5] Gonzalez and Larson (2008) find in their survey of US Cuban remittance agencies that they constitute a percentage applied to the amount remitted, with fees decreasing in the amount remitted. As can be seen in (1.16) the above does not equal -1 unless the transfer fee is such that $\overline{\varphi} = 0$, the costless transfer. In this scenario, $\varphi_i = 1$ for both cases in which we

[5] The previous results were done with the first period transfer decision. However, the same result holds for the second period when income is redefined such that it is static.

consider that the remittance fee might be applied. Whether (1.16) is greater than or less than -1 will depend on the functional form of φ_i. This generates the following:

Proposition 1.5.2. *Suppose the transfer cost, $\overline{\varphi}$, is positive. Given the definition of static income the following cases hold:*

1. *In the case where transfer costs are convex in the transfer fee, so $\varphi_1 \geq 1$, the substitution response to a change in agent 2's income is such that agent 1 remits less than the change in agent 2's income (that is, $\frac{\partial T_t^*}{\partial I_t^2} - \frac{\partial T_t^*}{\partial I_t^1} \geq -1$).*
2. *Under the inverse case when transfer costs are linear in the transfer fee, so $\varphi_2 \leq 1$, the opposite effect occurs.*

Given the interval of the parameter $\overline{\varphi}$, the above result is straightforward. Nevertheless, to see these results it helps to note that the transfer derivative is a ratio (since transfers and income are both in logs). And as has been mentioned, the only case in which (1.16) is equal to -1 is when φ_i equals 1. This may help explain why the empirical remittance literature tends to reject the altruistic motive, as authors typically find estimates for the transfer derivative that are greater than -1. It suggests that when costs are positive an upward bias results if the costs are not controlled for (we refer the reader to Cox and Fafchamps (2008) for the survey on literature estimates); and that if altruism is truly the motive (or at least a contributive factor in the motive to remit) observing a transfer derivative larger than -1 may also be consistent with convex transaction costs assumed by the remitting agent.

The Two Period Transfer Derivatives

Here we derive the 2 period transfer derivative. We work directly with (1.10) and (1.11) applying (1.13), which incorporates the inter-temporal effect of savings. Instead of considering static income we analyze the transfer derivative with respect to wages in each period.

Taking the derivatives of (1.10) - (1.12) with respect to y_1, $\widetilde{y}_1$, y_2, and $\widetilde{y}_2$ we obtain the transfer derivatives for both periods:

$$\frac{\partial T^*}{\partial \widetilde{y}_1} - \frac{\partial T^*}{\partial y_1} = -\frac{1 + \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1 + \varphi_i \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}} - d_1(\varphi_i - 1) \tag{1.17}$$

$$\frac{\partial T^*}{\partial \widetilde{y}_2} - \frac{\partial T^*}{\partial y_2} = -\frac{1 + \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1 + \varphi_i \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}} - d_2(\varphi_i - 1) \tag{1.18}$$

where the values of d_1 and d_2 are strictly positive inter-temporal terms, defined below:

$$d_1 = \frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1 + \varphi_i \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}} \cdot \frac{[\beta(1+r)]^{1/\gamma}}{(1+r) + [\beta(1+r)]^{1/\gamma}} \tag{1.19}$$

$$d_2 = \frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1 + \varphi_i \left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}} \cdot \frac{(1+r)}{(1+r) + [\beta(1+r)]^{1/\gamma}}. \tag{1.20}$$

The transfer derivative now incorporates the direct effect of (1.16) and an indirect effect arising from the distortion's effect on inter-temporal substitution. This inter-temporal effect on the transfer derivative is determined by φ_i. So in the convex case, where $\varphi_i \equiv \varphi_1$, the transfer derivatives would move closer to -1 from above, as the inter-temporal term compensates the distortion. On the other hand, in the linear case, where $\varphi_i \equiv \varphi_2$, the inter-temporal effect compensates the distortion by moving closer to -1 from below. This indicates that in any one period, (1.16) is either an upper bound or a lower bound relative to φ_i. It is apparent, however, that this period compensation attributed to d_t decreases as $t \to \infty$ due to discounting and that $d_t \xrightarrow{t\to\infty} 0$. Now add

(1.17) to (1.18) to obtain

$$-\frac{2+\varphi_i\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1+\varphi_i\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}-\frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1+\varphi_i\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}<-1. \tag{1.21}$$

So over the two period life-cycle the cumulative transfer derivatives surpass -1, independently of the assumption regarding φ_i. Equation (1.21) generalizes to T periods by replacing the 2 with T and noting that

$$\sum_{t=1}^{T} d_t=\frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}}{1+\varphi_i\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}} \tag{1.22}$$

This shows that the cumulative effect on the transfer derivatives over the life-cycle is not only able to surpass -1, but is monotonically increasing in the number of periods, potentially unbounded. It suggests that when costs are real an upward bias results if the costs are not controlled for (we refer the reader to Cox and Fafchamps (2008) for the survey on literature estimates); and that if altruism is truly the motive (or at least a contributive factor in the motive to remit) observing a transfer derivative larger than -1 may also be consistent with Agent 1 assuming the transfer costs.

Cox and Fafchamps (2008) conjecture that "The pronounced transfer derivatives predicted by the altruism model could well be a good deal weaker once life-cycle considerations are taken into account." We've shown in a two period specification of the model that this in fact does occur (depending once more on φ_i). That is, we find that inter-temporal effects on the transfer derivative arise as a result of how φ_i affects agent 1's savings decision.

1.6 Model Calibration & Results

1.6.1 Data

The first requirement are the incomes of Cubans and Cuban-Americans. The incomes of the latter are obtained from the ACS and Census (when applicable) from 2000-2012. We define a Cuban-American as the head of household age 20 years or older, self-identifying as an ethnic Cuban, regardless of citizenship status. We take as income the total income reported by this population (including government transfers and pensions/retirement income). We use this income rather than labor income since we do not consider a government in our PE model. The wage profiles for each year are computed and then averaged, generating a representative income profile. Regarding Cubans, we have available only the average state-supplied income for 2012 (as previously reported in the Statistics subsection of the introduction). Blue (2004) and Orozco (2009) report similar figures for this average state wage indicating that it has been pretty stable over the years. They also report some survey results on informal income generated. However, the proportion of the surveyed population engaging in this is rather small. So, we will take as the working data the state supplied income. This point estimate is what we give agent 2 in the model simulation.

To aggregate the results of the representative agent, we assume that the head of household is the remitter. We pull from American Fact Finder the total population of Cuban-Americans and the total households for the above years. However, for the years 2001-2004 total households are no longer reported. So we impute these figures from the total population by using the average number of Cuban-Americans in a household reported in the 2000 census (2.7 persons/household) and take the average across all the years, denoted $\overline{H}$. From the above ACS and Census data sets we also pull the age distribution of the head of households for each year to help aggregate the results of the representative agent. As was done with incomes, we take the average of the age profiles to generate one distribution.[6]

Regarding remittances, aggregate figures are taken from the Havana Consulting Group, which publishes the yearly estimates on their website. These are cash remittances, and so do not include in-kind remittances. We average the realized remittances[7] to work with only one remittance figure, $\overline{R}$. An average cost of transferring a remittance is pulled from Orozco (2009), where the author reports a 17% transfer fee. Gonzalez and Larson (2008) report an average transfer fee of around 16% from a survey conducted in 2007 and 2008. These authors also found that the exchange fee in Cuba is typically assumed by the remitter (recall this is equivalent to a 20% tax imposed on the dollar).

1.6.2 Calibration

The calibration is conducted on the Cuban example from the perspective of the representative agent. We take the framework given by equations (1.2) - (1.5) and calibrate for the altruism parameter in a no cost model; and the transfer fees under each assumption of the functional form in which it may be applied as a cost, assuming the altruism parameter is at full reciprocity ($\alpha = 0.5$). This is done because we lack a degree of freedom to jointly calibrate both the altruism parameter and the transfer fee parameter for both cases. Further, we make an adjustment to the model to account for the exchange fee that is also

[6] Lacking information on the remitting population, we assume that the age profile conditional on remitting is the same as that of the underlying population, understanding that this assumption is rather strong.

[7] These figures are realized using the CPI to 2000 constant \$.

applied to transfers. We take this as a constant 20% given that the policy is to apply this amount as a tax on all US dollar denominated remittances.

Next, following Huggett (1996), we conduct a life-cycle analysis with 64 periods. Agents begin life at age 20, retire by age 65, and die with no bequests or assets at age 83. We operate on two grids for the choices of a_{t+1} and τ_t. Since agent 1 is borrowing constrained he is able to pick a savings from 0 up to his current period wage y_t. The proportion of income remitted, τ_t, is defined per (1.5). We constrain the endogenous parameters so that aggregate remittances match the average of the time series, according to the proportion of households, per the equation below:

$$\sum_{t=1}^{T}(\tau_t y_t)\phi_t \overline{H} = \overline{R}. \tag{1.23}$$

In the above, ϕ_t is the proportion of Cuban-Americans of age t.

The calibration of the other parameters are pulled from Huggett (1996) since Cuban-Americans in the US are subject to the same income process. Since we are not incorporating survival probabilities, we take the authors beta for mortality certainty as 0.99. We assume the wider economy is in steady state and set the interest rate as $r = 1/\beta - 1$ per the Euler condition. The preference smoothing parameter, γ, is set to 1.5 for both agents.[8] Havranek et al. (2013) find in a meta-analysis of the inter-temporal elasticity of substitution that the US falls in the interval $(0.5, 0.7]$. Since we assume CES utility this corresponds to a smoothing preference parameter of $\gamma \in (1.4, 2]$. So 1.5 seems reasonable.

[8]This can be relaxed, however, and figures A3 and A4 in Appendix section 1.4.2 show the results of the comparative statics of agent's smoothing parameter on altruism.

1.6.3 Results

The results of the calibration are reported in table A1. Figures are also located in the appendix in section A.3 for the comparative statics that were conducted on the model parameters.

The results in the first column indicates that Cuban-Americans are somewhat altruistic if 0.5 is taken as the benchmark. However, altruism may not be explaining remittance behaviors entirely; or frictions may suppress the observation of altruism if it is the true motive.

In the appendix we report the comparative statics conducted on the model where we allow the preference smoothing parameter of an agent to change while holding the other's constant. These are found in Figures A3 and A4. During the years 2000-2008 the model generates calibrated altruism parameters that are very close to each other, but after 2008 separation becomes visible.[9] This effect is capturing policy change that occurred between 2008 and 2009 (as discussed in the introduction), and the model is directly capturing the effect through the parameter. Policy, then, may contribute as a friction that dampens observed altruism.

We try to capture policy through a cost to transfer. As we previously noted the exchange tax is also incorporated into the cost structure so that agents account for this. The results in the second and third columns of table A1 pertain to the transfer fee specification. It shows that if we assume that Cubans are altruistic then under the first assumption on the form of φ_i the data is consistent with a transfer fee $\overline{\varphi}$ of 22%. Recall that the average cost has been found to be around 17%. If, however, we assume the functional form of φ_2 the resulting $\overline{\varphi}$ is 38%.

[9]Indeed the effect of averaging the data to conduct the analysis in Table 1 places the average around 2008.

In figure A1 we plot the per capita remittance profiles over the lifetime, net of the transfer fees that the model explains. These remittance profiles look like standard income profiles. Considering that remittances serve as an income supplement in any period, by construction we should expect them to be contingent on the ability to remit. Because there are no bequeath motive, we observe a slight uptick in end of life remittances. This residual may need further refinement based on ethnographic studies on remittance behaviors to understand how people remit over the life cycle.

1.7 Conclusion

In this chapter we've explored how transfers, when scaled by income, are decreasing in agent's wages and the affect that transfer costs impart on the remittance decision. We've shown how the distortion that arises from the existence of a fee to transmit a remittance to a kin abroad decreases the income substitution response between agents. Inter-temporal decisions have the effect of compensating the distortion from transfer fees; however, as an agent becomes more farsighted in his decision making process the compensatory effect diminishes. This has implications for researchers as failing to control or account for these costs in remitting leads to weaker estimates of the transfer derivative, and empirical rejection of altruism as a motive may result. This is not to say that altruism is the direct or principle reason that agents may send private transfers to their families in the developing world. After all, testing for altruism is quite difficult since this is inherently unobservable and so we must infer it. But the transfer derivative as currently used to test for altruism (cross-sectionally and without considering costs) does not seem adequate, especially since the model predicts that transfers are decreasing in transfer fees when the remitter assumes these costs. We argue that our results regarding the transfer derivative in such a case seems consistent with estimates in the literature.

We picked Cuban-Americans to help simulate our model because of differential policies that apply to this group of migrants in the US. We conjectured that they would likely be altruistically motivated when remitting to Cuba due to these policies. But also because of these policies it may be the case that testing for altruism would be difficult. Although the results are qualitative, the simulation illustrates the effect of ignoring transfer costs. In a model with no costs, the representative Cuban-American remitter seems somewhat altruistic. Accounting for costs, the transfer fee that fits the aggregate remittances when altruism is assumed comes close to the average reported by other authors. This does not, on its own, indicate that Cuban-Americans are in fact completely altruistic. But it does highlight the effect in failing to account for the real costs migrant populations face when remitting to their families, and future research should aim to account for these costs as they have a real effect on remittance behavior (see Freund and Spatafora (2008)).

Chapter 2

Descriptive Analysis of Internal Family Migration in Indonesia and the Human Capital Investment in Children

2.1 Introduction

In this chapter we document the descriptive analysis conducted on the Indonesian Family Life Survey (IFLS) to analyze the relationship between internal family migration and their children's human capital vis-à-vis schooling attainment.

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).

The IFLS data consists of 5 waves conducted between 1993 and 2014. The data we use in this and the proceeding chapter is drawn from the non-divorced families of adult children born between 1980 and 1997. We choose 1997 as the last cohort to analyze as by the 2014 wave this cohort should have finished their non-obligatory secondary education if they had entered that level. In this way, the pre-tertiary education is fully observed. As the idea of family is fluid across cultures, in the context of this analysis we consider the family unit to consist of the parents and their children. We reconstruct the family histories, education history of children, migration history, and wage history of the participants. When appropriate we augment the wage data with the data offered in IPUMS International of the 1995 Indonesian Census.

Internal migration affords families access to more local labor markets, which may have subsequent impact on children's educational attainment. Motivations to migrate may be driven by 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 fostered transmigration. In 2000 about 10% of the population lived in a province different from birth whereas only 1.5% lived abroad (Ducanes and Abella, 2009). In reconstructing of the full migration histories of adults in the IFLS, we find that internal provincial migration rates in Indonesia average about 1.5% per year. We also find that migrant parents have higher education than their non-migrant counterparts, with roughly equal shares living in rural and urban areas.

Our analysis reveals a positive association between internal family migration and childhood educational attainment, as well as a migration wage premium between Indonesian islands. We first find that children in Indonesia have high rates of primary school completion, where only about 1/3 of children complete up to primary schooling. Descriptive analysis shows that family migration increases the expected schooling by up to 1.5 years (completing at least half of the upper, non-obligatory secondary education). A duration analysis reveals that family migration is associated with a decrease in the hazard that children drop out of school, where the effect of migration is strongest after primary schooling. The analysis displays some sensitivity as a result of not yet accounting for the selection mechanism that sorts families 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.

Regarding wages, our mincer regressions show that a migration event in general is associated with a wage increase and that this increase is slightly higher for those who move between islands versus those who move between provinces within an island. A fixed effect mincer regression estimates a 12% wage premium for an instantaneous move between islands, while a move within an island yields a 5.6% wage premium, on average.

The rest of this chapter is organized as follows: section 2.2 documents the related literature of this chapter; section 2.3 describes the Indonesian Family Life Survey data along with the construction of the various datasets, as well as a discussion of the empirical evidence from the data that motivates the research question; section 2.4 analyzes the data to condition the means of interest on observables and the hazard associated with migration; section 2.6 concludes. The associated appendix elaborates on related concepts and contains the figures and tables of this chapter.

2.2 Literature Review

Our research concerns the movement of the whole household, the implications of which is an important and open question in the migration literature. However, that we are studying a family migration event necessarily means that we are departing from the bulk of the literature that places emphasis on analyzing the effect of missing parents in the household. The added disruption to the household of a single parent migration event may place the child at greater risk. The children in these households may face varying consequences than just consumption allocation since parental time investment may be 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, we are not considering.

This strand of the literature is also resoundingly reduced form in the treatment of a typically dynamic problem. 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 et al. (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, as well as 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 given 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.

Our analysis adds to this literature by investigating the role of migration of the family unit on children's educational attainment. This poses interesting intergenerational questions that we explore further in the subsequent chapter.

2.3 Data

The Indonesian Family Life Survey is an ongoing, longitudinal survey of Indonesian households conducted by the RAND Corporation in conjunction with SurveyMETER.[10] Five waves were fielded in the years 1993, 1997, 2000, 2007 and 2014. 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 B.1.

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. We use this rule to combine all five waves of household data. This procedure generates a database of 19,088 households comprising 83,766 individuals. We also note that the concurrency of Indonesian transmigration programs subsidized by the Indonesian government leads us to drop around 70 identifiable individuals from the dataset.

[10] A second, recently available data set, the IFLS-East, was conducted in 2012 by SurveyMETER. We make use of this data in a specific context in the next chapter. We refer the reader to Appendix B.1 for more information on this survey.

2.3.1 Family Longitudinal Panel

We restrict our sample to households where the parents are not and have never been divorced (divorced households account for 12% of our data), are married to only one spouse, the marriage occurred by the time the first child is born, and both parents are living in the same household. We further restrict the sample to households where children were born between 1980 and 1997. The reason for this 18 year data range are the following: by the first wave of the IFLS study (1993 - conducted in 1992) the children of the 1980 cohort will have been observed as children and upon turning 13 will have already completed their primary schooling, and as such their EBTANAS school test scores should be available, as well as cognitive test scores as taken by the IFLS surveyors (which is given to those who are 24 years of age or younger); those born in the 1997 cohort should have finished the full, pre-tertiary schooling cycle had their parents made this education decision by the 2014 wave. Moreover, those born in this timeframe will have been fully exposed to the effects of the Sekolah Dasar INPRES Program[11] that Duflo (2001) and Hertz and Jayasundera (2007) analyzed to measure its impact on labor market outcomes and social mobility, respectively, in Indonesia. In this way, we do not have to account for differential government investment over time in the educational infrastructure of Indonesia.

We identify 9603 children corresponding to 5244 parents that meet the above criteria. The descriptives statistics of these children and their parents can be found, respectively, in tables C2 and C1. We expand this data into a longitudinal panel for a survival analysis and link the data to their parents, along with other household information.

[11] Please refer to appendix B.3 for more information on this program.

2.3.2 Migration Longitudinal Data

We reconstruct the migration histories of all individuals in the IFLS dataset across all five waves. As noted in appendix section B.2, since the IFLS started in 1993 the country's administrative regions have fragmented. Instead of using cross-walks to convert the codings of the administrative regions across the IFLS waves forward in time, we move them backwards to the 1993 coding. We do this as it is easier and the research question is preoccupied with provincial migration.

Table B12 documents the descriptive statistics of migration events for the entire IFLS dataset between 1980 and 2014 (the final wave year) at the Any Migration level, Inter-Provincial , and Inter-Island levels. The last two columns are what we have been calling "Provincial Migration". In determining shares, we only count an individual once for that type of migration event; for average moves per movers we consider the total moves up to the final observed wave at each level and relative to all individuals who have that type of event. For this reason we observe that average migration events increase in relation to the reported levels. Relative to the total migration sample, average lifetime moves per mover is decreasing in the reported levels: a back of the envelope calculation using the figures in table B12 reveals that relative to the entire migratory sample considered, average moves per movers is ~ 1.00 at the inter-provincial level; and ~ 0.38 moves per movers at the inter-island level.

Per the migration literature, we also split these migration events across 10 year cohorts starting with age 15[12]. These can be found in table B13, and graphically in figure B6. Consistent with the literature, average migration events is generally decreasing with age. However, we observe two curiosities in

[12]The IFLS considers adulthood to start at this age since non-obligatory upper-secondary schooling starts at this age.

the data: 1) among those who have ever engaged in an inter-island migration, there is an uptick in the average migration event for those aged 25-34; and 2) we observe an uptick in the average migration event at the 65+ age cohort. The first observation may be due to income effects. It may be more expensive to move between islands and so this type of migration may be out of reach for the youngest adults. The second could be older individuals moving into their children's (or other relative's) household.[13]

Finally, we report the average migration rates observed in the data starting from 1980 in table B14 and illustrated in figure B7; as well as the shares of repeat moves per cohort for those with more than one migration event in table B15. The average yearly provincial migration rate (the combination of the two time-series) is about 1.5%. This decomposes to an average yearly migration rate of 0.53% for inter-island migrations, and a rate of 0.90% for intra-island migrations. An interesting feature of the migration rates is the relative stability over time of inter-island migration rates relative to closer types of provincial migrations (intra-island migration rates). This is especially true prior to and after the 1997 Asian Financial Crisis.

2.3.3 Wage Longitudinal Panel

Wage data of all adults in the IFLS dataset are reconstructed from adult wage histories across all the waves, where we consider only the first reported job. An adult for the purpose of this analysis is anyone older than 15 years of age, and who has permanently left school. We then restrict ourselves to wage observations starting from 1980 until the end of the panel observation. We further use data from the World Bank to realize wages into 2000 constant INT dollars and the data on labor market participation to construct hourly

[13]We note that households in Indonesia contain several generations within them.

wage rates. Since the IFLS data contains the 2 digit ISCO-68 occupation codes, we can make use of these in wage analysis. We report the descriptive statistics of these wage histories for all adults in the reported timeframe in the IFLS dataset in table B8 located in the appendix. Additionally, we report the wages of the IPUMS 1995 census data in table B9. We notice that there are slight differences in the data. We also note that in decomposing table B8 by cohorts, as seen in table B10, there is a clear life-cycle pattern to hourly wages in both the mean and median wages.

To create estimates of median wages in each province, island, or island grouping we consider only the wages of individuals at the time the IFLS survey wave was conducted (where the person was living at the moment the survey was conducted). We also augment the IFLS dataset with data from the 1995 Indonesia census made available by IPUMS International, and the 2012 IFLS-East Survey conducted by Survey Meter (a survey conducted with the same instruments used by RAND, but conducted on the eastern provinces that were not considered in the original IFLS survey). These data are reported in table B11, and are used in chapter 3 to create the market wages used there as well as the figures discussed thusly.

Choropleths of geographical aggregations of wages are reported in figure B4. They highlight the heterogeneity in wages across the 27 provinces and 7 islands that constitute Indonesia in 1993. In the final classification (Market Group) we classify all provinces on the island of Jawa as Market 2, and “Everywhere Else” (all other provinces on all other islands) as Market 1.[14] The advantage of this demarcation is that, as shown in table B1, roughly half the population lives in Jawa and the other half everywhere else.

To understand if the augmented IFLS wages are consistent, we compare

[14]The wages according to this classification is used in chapter 3 as given in table C3.

with the 2015 Gross Regional Product (in 2000 constant dollars) for each province as obtained from the World Bank. Figure B5 graphs the spatial distribution of the median hourly wages, superimposed with bubbles that correspond to the size of the GRP in the respective province. For example, the province of Jakarta (the capital of Indonesia) has both the darkest blue and one of the largest bubbles (it is a relatively productive region and it contains a correspondingly high wage). The rank correlation between provincial wages and GRP is 0.41 and highly significant, indicating there is indeed a high correlation between productive regions and the wages offered in them.

2.3.4 Discussion of the Data

Internal migration is prevalent in Indonesia and has been fostered by government transmigration programs (initiated in 1950, through transmigration plans) to relieve pressures on densely populated islands. Between 1950 and 1968 some 90,000 households were relocated, and by 1997 6.5 million individuals (Farré and Fasani, 2012). The IFLS data also evidences substantial migration.[15] As reported in table B12, roughly half of the identified adults have engaged in some kind of migration, and nearly 18% of the sample has migrated across provinces.[16]

The selection associated with migration is quite evident. The two subsamples in table C1 are systematically different across many dimensions. For the combined sample of parents and adult children we find that adults have about as much education as reported by Duflo, of around 7.6 years; and the age of the sample is consistent with what Bryan and Morten (2015) finds using a different

[15] Although, as previously noted, only about 70 individuals self-report as transmigrants - we drop them and their families from the dataset.

[16] As table B1 elaborates, there are 4 types of geographical divisions: provinces, districts (*kabupatan*), sub-district (*kecamatan*), and at the lowest level villages (*kelurahan*, or the more frequent usage *desa*) and municipalities (*kotamadya*).

data set. Migrant parents also have higher educational attainment. Interestingly, those who don't have any migration events in their histories are no more older at indicated life stages. We also document the marriage contingency table for adult parents in table B3, which shows that given the assortative matching present in Indonesia, likes are marrying likes.

As reported in table B4, adult children whose parents migrate have generally higher education than those whose parents don't migrate and no more likely to repeat a grade. These children are also more likely to have attended Kindergarten and born in urban areas. We also note that children of migrant families have higher frequencies of having moved in their adult lives (where we observe their age 12 location to be different than their birth location). And that the age and the school grade when the family first migrates corresponds to just before entering the first level of school in grade 1.

As further evidence, we decompose the educational attainment of children across several relevant dimensions. Figures B6 and B7 report the cross-tabs of educational attainment across the urbanization of birth and family migration. Rural boys and girls obtain roughly the same level of education, while urban girls eek out, on average, slightly more than urban boys. Rural children who migrate obtain roughly 1.5 years more of schooling, while urbanites who migrate eek out approximately half a year more of education. Importantly, migration for the rural children looks to be associated with at least finishing one year of the non-obligatory secondary level (years 10-12).

2.4 Empirical Analysis

As the evidence of parental migration and children's educational attainment presented in the previous section was based on simple univariate and bivariate exploration, we now analyze the relationship of parent's migration and children's schooling. We first look at the migration premium in wages for all the adults in the sample from 1980 - 2014. We do so for both statistical power, and because in the next chapter we will consider the implications to the labor market outcomes of children of migrant families. Then we look at the subset of children and conduct a duration analysis to comment on the hazard of leaving school from a family migration event. This is done on an unbalanced panel of the longitudinal education history of the identified children, typing them to parental and household observables.

In all cases, a migration is considered a provincial migration. We may further partition the migration event into an inter-island and intra-island migration type, to take into consideration the geographical distribution of Indonesia. Both of these analysis are descriptive in nature as the selection mechanism is not yet accounted for, being left for chapter 3. Results are reported in the appendices.

2.4.1 Wage Analysis & the Migration Premium

In this section we report the results of both cross-sectional mincer regressions on IFLS wages starting from 1980 as well as fixed-effect regressions on the longitudinal data. We also report the cross-sectional mincer regressions for the the IPUMS data of the 1995 Indonesian census to compare the coefficients.

Cross-Sectional Mincer Regressions

We run the following models on both our reconstructed wage histories from IFLS as well as the 1995 IPUMS census.

$$\begin{aligned} ln(wage_{iopt}) = \alpha + \beta SchYrs_i + \gamma_1 age_{it} + \gamma_2 age_{it}^2 \\ + \sum_l \delta_l Mig_{it}^l + \sum_k \theta_k X_{it} + \sum_j \alpha_j + \varepsilon_{iopt} \end{aligned} \tag{2.1}$$

$$\begin{aligned} ln(wage_{iopt}) = \alpha + \sum_m \beta_m SchLvl_i^m + \gamma_1 age_{it} + \gamma_2 age_{it}^2 \\ + \sum_l \delta_l Mig_{it}^l + \sum_k \theta_k X_{it} + \sum_j \alpha_j + \varepsilon_{iopt} \end{aligned} \tag{2.2}$$

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 we 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}). X_{it} are the set of additional controls that may or may not time vary.[17] For the IFLS wage and migration histories we add the Mig_{it}^l sets of dummies for the type of migration event, and omit them for the IPUMS regressions. These migration events are defined as either movements between distinct islands (an interprovincial move that has occurred between two islands), or between provinces within an island.[18] We define these islands according to ISO 3166-2.[19]

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

[18] This serves two purposes: in chapter 3 we will only consider migration between two regions, so understanding the role of migration between aggregate regions seems more important; more practically and as we've previously discussed considering islands or provinces within them is easier to cross-walk back to the 1993 configuration.

[19] Island 1 is Sumatra; island 2 is Jawa; island 3 is Bali and all islands in the Nussa archipelago; island 4 is Kalimantan (Borneo); island 5 is Sulawesi; island 6 are the islands

Fixed-Effect Mincer Regressions

We then conduct fixed-effect mincer regressions on the full panel of available IFLS adult wage histories, as these regressions have the power of controlling for individual fixed effects. The models that we run are from the following equation:

$$ln(wage_{iopt}) = \alpha+\gamma_1 age_{it}+\gamma_2 age_{it}^2+\theta Urban_{it}+\sum_l \delta_l Mig_{it}^l+\sum_j \alpha_j+(\eta_i+\varepsilon_{iopt}) \tag{2.3}$$

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 can be augmented to account for marriage entrance and exits and birth of children.

2.4.2 Duration Analysis

The duration analysis of the data is conducted with a discrete-time probit estimation on our single spell data of children's educational histories.[20] This method for conducting the duration analysis is quite flexible and allows us to incorporate repeated grades into the analysis. It also allows us to estimate the baseline hazard[21] non-parametrically, generating a dummy for each grade level

of Maluku archipelago; and island 7 is New Guinea.

[20]This is essentially a proportional hazard model. We consider the data single spell because we assume that once an individual exits the school system they do not reenter, a reasonable assumption.

[21]The baseline hazard is the fundamental exit probability that all in the population share, irrespective of their heterogeneity.

and omitting the final year (grade 12) to anchor the estimation.

An issue that must be dealt with is the initial conditions problem - how individuals enter into the sample. For example, if there's a selection problem that determines who enters the school system then this mechanism must be specified as duration analysis can not deal with this "left-censoring problem". Right-censoring is not a problem (that is, when one observes survival to the end of the spell or survive beyond the survey time-frame). We assume that there is no initial condition problem and start all individuals at grade 1 for those who enter the school system.[22]

Fixed effects regressors are also problematic in probit models due to the issue of incidental parameters (or nuisance parameters).[23] However, this only concerns the nuisance parameter of the unobserved individual fixed effect, something we are not concerned with here as we are not conducting a fixed effect regression. As the main purpose of these treatments is to account for selection on unobserved heterogeneity, one solution for single spell data would be to fit a random effects estimator. The main worry here is that failing to control for the unobserved heterogeneity results in a bias toward "negative" duration dependence.[24] The resulting bias would work to generate an upward shift of the hazard function. However, as elaborated in Heckman and Singer (1984), the typical implementation of this random effects estimator is to make an assumption on the distributional form of the unobservables, which can become quite *ad hoc*. They present a robust method to non-parametrically estimate through maximum likelihood the parameters of interest and the dis-

[22]We find that around 2% of the children in our sample do not enter the schooling system.

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

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

tribution function of these unobservables in duration analysis. We currently do not implement this as even accounting for selection on unobservables there still remains the selection due to migration, which will be accounted for in the next chapter. This analysis is therefore descriptive and serves only to highlight the relationship between family migration and survival in the Indonesian schooling system.

Duration Model

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

$$\delta_\tau = \sum_{j=1}^{T^*} \delta_j \mathbb{1}(t = j) \tag{2.4}$$

T is censored according to $T^* = T - 1$, as the final period observed, $T = 12$ years of schooling, is not identified.[25] In practice it is necessary to artificially censor dummies of equation (2.4) if a critical mass does not exist at one of the discreet time intervals.[26]

The conditional hazard rate that we estimate is then defined as

$$h_i(\tau, X_{it}; \theta) = P(t_i = \tau | t_i \geq \tau, X_i) = F\left(\delta_\tau + \beta(FamilyMig_{i\tau}) + \sum_j \eta_j X_{ij}\right) \tag{2.5}$$

[25] $T = 12$ is not identified because we 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.

[26] Artificial censoring is necessary when there do not exist failures during a time period, as was discussed in footnote 25 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 (2.4) are constructed discretely. Thus, we artificially censor them and instead create dummies for grades 1, 2, 3, 4-5, 6-8, 9-11.

In the above $h_i(\cdot)$ is the hazard contribution of individual i. The function $F(\cdot)$ is the normal CDF. The controls X_{ij} are a vector that contain such observables as parent's education, birth urbanization, birth order, sex, and kindergarten participation. The parameter of interest, β, captures the correlation associated with family migration. The variable $FamilyMig_{i\tau}$ activates if a family migration event takes place during the child's schooling and stays activated thereafter to signal a state of "migrant". The vector of these parameters to be estimated are captured in the parameter θ. As discussed in footnote 25 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 (2.5). We 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

$$\ln \mathcal{L}(\theta; X) = \sum_{i=1}^{N} \sum_{\tau=1}^{T^*} a_{i\tau}[y_{it} \ln h_i(\tau, X_{i\tau}; \theta) + (1 - y_{i\tau}) \ln(1 - h_i(\tau, X_{i\tau}; \theta))] \quad (2.6)$$

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.

2.5 Results of Empirical Analysis

2.5.1 Migration Premium

The results of estimating the cross-sectional mincer models are presented in appendix section B.5, tables B16 and B17. The two tables display a concordance in the estimation of the education premium between the two datasets, and that it is in line with what has been found in the development literature.

The IFLS data further documents consistently large and significant migration premiums for a one-off migration event, where columns (6) and (12) in table B17 show an inter-provincial migration premium of 10%. The difference between an inter-island and intra-island migration premium are not significantly different.

The fixed-effect migration premiums in table B18 show that whether the migration is within an island or between islands there are consistent and significant premiums associated with the move. However, the premium associated with an island migration is consistently a little over twice that of a within island migration (specification 3, the preferred specification, results in 12% premium vs. 5.6% premium, respectively). The results are similar to what we find in the cross-section.

Since we are interested in the wage premium attributable to migration, the FE specifications have the advantage that we 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 we 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 we have accounted for unobserved heterogeneity (assuming it is constant over time) and whatever omitted variables are correlated with the other fixed effect terms included, but as of yet not the selection into migration, these specifications can not be interpreted causally.

Finally we presume that migrants are not just migrating for an instantaneous wage boost (especially if the whole family moves), but instead do so on the expectation of some persistence in the wage premium, future augmentations should include interacted time dummies to understand if these premiums decay.

2.5.2 Duration Results

We report the results of the estimation in table B19. The estimated coefficients are the raw outputs whose sign is the only thing interpretable. Marginal effects are not calculated as it is sufficient to use the sign of the coefficient to make a statement about the role of family migration on the hazard of exiting schooling. The indication of whether the hazard of exiting the school system increases is given by a positive coefficient, and a decrease in the hazard is given by a negative coefficient. We find that while there is sensitivity in the results of the analysis, the hazard of exiting school is significantly negative. This would indicate a drop in the hazard of exiting when a child enters a state of “migrant” when the family undergoes a provincial migration event. Although estimates of controls are not reported in table B19, those in urban locations are also more likely to stay in school; a child’s sex does not play much of a role; birth-order seems to also decrease the hazard of dropping out, and parental education is significant in reducing the hazard of dropping out of the school system. While the values of the coefficients can only be interpreted with a marginal effects analysis, these coefficients can be used to construct the survival plots of an individual with the average characteristics.

To construct the survival plot we generate the predicted data of a representative agent based on specification (5) of the estimation output in table B19. This representative agent has the mean values of the covariates.[27] To under-

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

stand the impact of provincial migration events we set the indicator of family migration to 1 (and 0 otherwise) to compare the plots against those whose parents do not migrate. As the average household's first migration event occurs while the child is around the age of 7 (just prior to entering school), we set the dummy equal to 1 for the migration case prior to first entering school.

In general we find that family migration mitigates the probability of exiting the school system. There are several takeaways in common between the unconditional plot (the K-M plot in figure B8) and the conditional plot estimated from the observables (the plot below it): within a schooling level (Grades 1-5, 7-8, 10-11) we do not see large drops in the amount of children surviving to the next grade; the largest drops are observed at the grade corresponding to the end of the schooling level (6 and 9); conditioning on observables we notice a larger drop in the amount of children that survive into the next grade vs. the unconditional plot. Given the push in the 1970s to increase the rates of primary education as explained in appendix B.3, that most children complete elementary school is to be expected.[28] It is in the subsequent schooling levels that family migrations serve as a mitigating factor against dropping out of school, with an average reduction of the hazard by 7%. Finally, it seems that the extensive margin (whether or not to invest in one more level of schooling) is more important than the intensive margin (how much schooling within a level) since children generally finish their schooling level, but face a hazard of not continuing.

Because the selection that sorts parents into migration is not taken into account, it is *a posteriori* ambiguous whether the estimation is inflated or attenuated, the only way to interpret the results of this section is to note that so far the evidence points to a positive correlation and that family migration

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

may be reducing the hazard of exiting the school system prior to 12 years of completion. So in general, the extent to which parental migration may raise the quality of children by relaxing the budget constraint, and studying the labor-market outcome of these children, is a key question that the structural framework in the next chapter will help to resolve. We caveat that although compulsory education is (by law) only for 9 years we have chosen to study the full cycle of pre-tertiary schooling as the general trend in Indonesia has been an increase in educational attainment.

2.6 Conclusion

Internal migration in Indonesia is quite extensive and an association exists between the schooling attainment of children and the parents that migrate. We documented wage migration premiums to adults that conduct a provincial migration event. We find that provincial migration is on average associated with a 10% wage premium, and that inter-island migrations tend to yield slightly larger premiums. A duration analysis on the educational history of children born between 1980 and 1997 elucidates that family migration shifts the predicted hazard of dropping out of school downward and thusly increases the survival of children within the school system by 7%. This is important as completion rates are lowest at the non-obligatory, upper-secondary level.

This analysis was undertaken to document the motivational evidence for the next chapter. There are two main sources of error: the selection on unobservables and the selection of parental migration. Evidence from the duration analysis suggests that not controlling for the first is upward biasing estimates of the migration coefficient, and overall may cause a negative duration dependence. In the next chapter we establish a causal link between family migration on an expected income basis and how this affects educational investment in their child.

Chapter 3

The Effect of Internal Family Migration in Indonesia on the Human Capital Investment in Children

3.1 Introduction

Given the various mechanisms that may impact the decision to educate, in this chapter we build on the previous analysis to understand if migration is sufficient to relax the household budget constraint so as to avail resources towards a child's education. Families may not necessarily be confined to their local labor market 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 based on regional wage differentials (regardless of the source of difference) that induces the reallocation of labor supply through migration.

The intuition in the theoretical framework of Basu and Van (1998) and Baland and Robinson (2000) lends itself to this question. These models comment on the substitution of children's time between labor and schooling in the face of constrained family resources. But with migration comes uncertainty, and so it is *a priori* ambiguous what the effect of family 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. Moreover, the implications for which labor market children end up in (and possibly educated in) given the family migration in a family event is not understood.

The question we address in this chapter is therefore twofold: is the family migration observed in the IFLS impacting the educational investment in children; and if so, what are the policy implications of affecting family migration decisions, and does this represent a cost effective means to increase the stock of high skilled children vis-à-vis a competing education subsidy?

We develop a structural model that captures the main tradeoffs regarding family migration and the household investment in education. We restrict the model to capture the more relevant, extensive margin of whether parents choose to educate their child. Our model is composed of two periods for each generation: a young period where agents are children and make no decisions; and an old/adult period wherein the household makes decisions for themselves and the child. To simplify the choice of location we consider a two markets model. Households 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 choice of location along with the possibility of educating their child in the new location will increase their welfare through the altruism they hold for the future outcome of their child and their dynasty. Moving implies a real, pecuniary cost. While the decision to educate represents an opportunity cost in immediate utility from the foregone income that the child could contribute to the household, captured through the cost of educating. Coupled with the possibility of choosing to educate the child in the more productive labor market, this implies that to offset costs expected discounted utility must be fairly high.

To estimate the model we filter all the reconstructed histories that were used for the empirical and descriptive analysis in the previous chapter, considering only the first-born child born between 1980 and 1997.[29] First, we consider a migration event in this collapsed dataset to be the last event to occur during the first 18 year period of childhood; for this reason, we can not in this framework consider return migration. Second, we partition the country of Indonesia into two regions: the island of Jawa being one labor market, and all other islands everywhere else in Indonesia being the other labor market. We further assume that an education decision is one that leads to the full 12 years of schooling. Regarding wages, we assign to families the median wage for their chosen location conditional on their skills; so we forego incorporating transient shocks to the expected wages of individuals as in our model they know perfectly what wages they will get. This produces a dataset of two periods per dynasty, where the first period overlaps the adult stage of parents and the child stage of their children. We then empirically specify the model assuming that households are borrowing constrained with log utility, while making an assumption on the joint migration and education decision via a nested logit. This nested logit is specified to account for the correlation between the unobservables in agent utilities across different groupings.

Our results demonstrate the moving cost to be 72% of the median income of high skilled households. Moving cost must necessarily be this large given that we only observe a total of 63 migration events in a data of 3,645 households when considering a migration event to be this aggregated. Regarding education costs, the opportunity cost is slightly lower in Jawa than in the rest of the country, but statistically the two costs are equivalent and likely driven by the fact that a little over 70% of our dataset constitutes low skilled households. We find education costs to be around 60% of the wage of low skilled households and

[29]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.

23% for high skilled households. Simulating the model based on the estimated parameters results in a relatively good fit of the data across both generations of the data, with the exception that low skilled households in the model do not move due to the borrowing constraint.[30] We also find that the model is robust to the specification of the altruism parameter α. Moreover, conditional moments are captured fairly well on various measures.

We then conduct a counterfactual analysis by giving agents a migration or an education subsidy, and comment on the welfare implication of the two. We find that not only do migration subsidies induce low skilled households to take the joint decision of migrating and educating the child in the new location, there is also an increase in the overall relative proportion of high skilled children in the economy.[31] This effect kicks in starting with a subsidy that compensates the household for around 90% of the migration event. On the other hand, the migration subsidy applied to high skilled households doesn't markedly increase the overall relative proportion of high skilled children produced by these households; rather, the subsidy serves to induce a family migration event, choosing to substitute away from educating in the current region and educating in the new region. For these high skilled households located outside of Jawa, we observe a 26% jump in the relative proportion of educated children resulting from families migrating to Jawa and educating their children there; The effect of the subsidy flattens at 70% of the costs. The difference in behavioral responses between these two household types is likely the result of relaxing a borrowing constraint (for low skilled households) vs. the role of wage differentials between regions (which is the relevant margin for high

[30]We do observe these households move in the data; however, they do so with low probability. Since low skilled households in the first generation account for a little over 70% of the households in our data, they do contribute weight to the migration rate accounting for a little under half the events. The effect of the borrowing constraint is that we underestimate the overall migration rate

[31]Relative to the initial state of the household.

skilled households). We conclude that based on marginal responses, if a migration subsidy were enacted by policy makers perfectly observing household types, the subsidies of 70% for high skilled households and 90% for low types would suffice.

Analyzing the effect of educational subsidies we find that they provide a mechanism for low skilled households to produce relatively more high skilled children. Here, we find that fully subsidizing education for the low skilled households increases the relative proportion of educated children elsewhere in Indonesia by 19% vs. the 9% increase observed by fully subsidizing migration; and an increase in the relative proportion of 25% in Jawa vs. 12% with full subsidy to migration.

We conclude with a welfare comparison, which shows that a 90% migration subsidy produces overall the same absolute proportion of high skilled children and is equivalent in welfare terms to a 25% education subsidy; and the benefits of a full migration subsidy can be obtained with a 30% education subsidy. The overall proportion of educated and thus high skilled children produced by families given an education subsidy is a little over twice that produced by giving them solely a migration subsidy. Regressing whether a child is educated on whether the parents are educated in our simulated datasets further shows that an education subsidy increases the base probability that a child will be educated, and lowers the "premium" probability of being educated afforded to those children from high skilled households; large migration subsidies also increase the base probability, but not to the extent an education subsidy does and does not affect the premium probability. Applying an education subsidy may represent a more cost effective way to increase both the stock of productive individuals and lower intergenerational correlations as the cost of migration is several times larger than the cost to educate. Moreover, the full benefits to the economy may come from fully subsidizing education across the board.

The rest of this paper is organized as follows: section 3.2 elaborates on how this paper fits within the related literature; section 3.3 details the data and the descriptive statistics of the households; section 3.4 documents the model and its econometric specification; section 3.5 the results of the estimation and of the simulation exercise; section 3.6 analyzes the role of migration and education subsidies, along with the welfare impact; we conclude in section 3.7. The appendices contain a generalized version of the model as well as tables and figures.

3.2 Literature Review

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 to conduct policy experiments on the proposed mechanism. 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 expected 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 staggard migration events under the constraint of US-Mexico border enforcement.

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 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) 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.[32] 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 determine the dynastic discount factors associated with human capital transmission. A similar framework can be employed here to model the mechanism by which parental migration impacts human capital

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

investment.

This chapter is a synthesis of these various strands the literature. Its key contribution is an analysis of how family migration decisions play into human capital investment of subsequent generations and the role of policy decisions to affect social mobility.

3.3 Data

The data used in this chapter is subset of the data from the previous chapter, where we consider only families where the first-born child was born between 1980 and 1997. We consider only the first-born child as the child of interest so as to avoid the complications of the quantity quality tradeoff as identified by Becker and Lewis (1973), understanding that this may impact some results. In total we identify 3,645 households (which is to say, 3,645 children) that satisfy our requirements. The distribution of the households across Indonesia are shown in figure C1. The associated migration events of these families is shown in table C4.

As we partition the country into labor markets consisting of the the island region of Jawa as one market, and the rest of the country as the other labor market, we present in table C3 the median wages that will be used in the model section. These wages are taken from the data discussed in the previous chapter's section 2.3.3. We further document the descriptive statistics of the labor histories of the fathers and mothers that constitute the parents in table C1, and of the adult children once they enter the labor market in table C2.

For skills we create a measure based on educational attainment, a dummy based on whether the individual has completed at least the full cycle of pre-tertiary schooling (high skilled), low skilled otherwise.[33] We use these skills to create the transition function in the empirical specification, where education maps 1:1 to skill attainment in our model.

3.4 Model

The motivating analysis in chapter 2 suffers from endogeneity as the parental choice to migrate is inherently a selection problem. Here, we develop a simple 2 period, 2 market dynastic model to structurally estimate the mechanism by which parents select themselves into migration and how this selection may mechanistically translate into greater investment in human capital in the subsequent generation. It is a simple search and match model, and Appendix C.1 details a general model that incorporates intragenerational dynamics. 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. We then specify the econometric model to estimate via the nested logit configuration.

3.4.1 A 2 Period, 2 Markets Model

Our model is a simplified version of the one found in appendix C.1. We reduce the country of Indonesia into two regions based on the distribution of the population in the 2010 census, as detailed in table B1. We also compress time periods into a young and old period. Decisions are made by parents (who we interchangeably call households) and children in their young period receive

[33] RAND includes occupation codes at the two digit level based on ISCO-68. So we also make use of the O*NET database on occupational skills and abilities to create a secondary skill measure for future use. First we crosswalk the ISCO-68 codes to ISCO-88 codes, collapse the three digit occupation codes to a two digit code, sum across all ability measures, and create a dummy based on the median skill (where high skilled (1) are all those occupations that are above the median, low skilled otherwise).

the investments, either passively (being located in their adult life where their parents chose to migrate to) or actively through the educational decision the parents make. This model will be used to understand the intergenerational consequences of parental decisions.

Model Setup

Timing Given a dynasty D call the initial old $g = 0$. Agents live for only two periods, so old is defined as $t = 2$ and young as $t = 1$. We synonymously refer to agents as a household. 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.[34] 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. When agents make decisions they do so for their child; their child realizes those investments in their state space when they become old and through the labor market outcomes those endowments imply. Since we will not consider population growth for simplification, a "child" in this context is the household of the next generation, g'.

Decisions Agents in period $t = 2$ of life decide whether to move to a new location, of which there are two (the home location and the other market), or stay. They simultaneously decide if they will educate the $g + 1$ agents in their $t = 1$ period of life; agents in $t = 1$ make no decisions at this stage. If agents choose to move or stay, they obtain a wage according to their state space. The wages in the market are exogenous and thus we solve a partial

[34] Which is quite strong, but reasonable given that in Indonesia we observe assortative matching across education levels with the typical lower off-diagonal, as evidenced in table B3.

equilibrium model. Agents making decisions must commit to the decision to educate after moving (so that the location does not induce a switch in this decision). We define a mutually exclusive decision in this model as $I_{k,t}$, the four possible being:

$$I_{1,t} = \mathbb{1}\{j = 1\}\mathbb{1}\{e = 0\}$$
$$I_{2,t} = \mathbb{1}\{j = 1\}\mathbb{1}\{e = 1\}$$
$$I_{3,t} = \mathbb{1}\{j = 2\}\mathbb{1}\{e = 0\}$$
$$I_{4,t} = \mathbb{1}\{j = 2\}\mathbb{1}\{e = 1\}$$

State Space and Transition Matrix 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.[35] Skill takes a value of $h = 0$ if the agent has less than full, pre-tertiary schooling, and $h = 1$ otherwise. So parental investment e in the young period becomes h in the old period of life and enters the state space. Since education investment e evolves deterministically into skill level h, the transition function, which we define as $N(\cdot)$, is essentially a matrix of 1s and 0s. In this framework then, education is not a risky investment.

Utility Utility is defined as $u(c, z, \varepsilon)$. We assume that agents are subject to incomplete markets, and as we do not allow a savings decision we make them borrowing constrained in this partial equilibrium model. We also normalize prices in terms of the consumption good so that it is treated as a residual.

[35] 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.

These assumptions imply that agents are risk averse and the utility will reflect this. Particularly, since agents do not make a savings decision, identification of a risk-aversion parameter will be problematic. So we make the simple CRRA assumption with the relative risk aversion parameter equal to 1 to obtain log utility. We also assume that agents are inelastic with respect to their labor supply and do not model this decision.

The utility is, however, subject to a transitory, idiosyncratic payoff shock ε_k from the decisions taken, which may represent a shock to either preferences, moving cost, or education 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 possible decision taken. This vector $\varepsilon = (\varepsilon_1, ..., \varepsilon_4)$ of shocks is assumed i.i.d. across population and time and drawn from distribution function $G_\varepsilon(\varepsilon)$.[36] As is also typical in the literature we will also assume that the utility is additively separable in the observable and unobservable states of the decisions taken, so that $u(c, z, \varepsilon_t) = u(c, z) + \varepsilon$.

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

$$\mathcal{B}(k, z) = c - w^{hh}(j, h) - \delta \mathbb{1}(j \neq \ell) - \phi_j \mathbb{1}(e = 1) \geq 0, \text{ for } j = 1, 2 \qquad (3.1)$$

The parameter δ 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 is bilateral, being symmetric whether the agent moves from market 1 to market 2 or vice versa. ϕ_j is the location-specific investment in education. The agent's

[36]In the nested logit specification we relax this assumption of i.i.d unobserved shocks.

budget is the cumulative wage of the household.

Value Function

We 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):

$$V_{d,g,t=2}(z, \varepsilon) = \max_{I_k \in I} \sum_k I_k \{v_{d,g,t=2}(z, k) + \varepsilon_k\} \tag{3.2}$$

where

$$v_{d,g,t=2}(z, k) = u(c) + \alpha \mathrm{E}\left[V_{d,g',t=2}(z, \varepsilon)|z, I_k = 1\right] \tag{3.3}$$

and utility, being the same across generations, is $u(c) = \log(c)$, where c is given in the budget constraint above.[37] 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 their fruits are realised in the next generation through the labor market outcome. As agents care about the next generation they discount the value of the next generation's utility, $V_{d,g',t=2}(z)$, by their altruism towards it, α.

Defining the ex ante value function as the value of generation g being in state z prior to observing their ε so as to marginalize out the shock, gives

$$\bar{V}(z) = \int V(z, \varepsilon) dG_\varepsilon(\varepsilon). \tag{3.5}$$

[37] When we estimate the current model to generate the starting values for the nested logit configuration we present in section 3.4.2, we will add an additional parameter, σ. This parameter scales utility to understand the degree of importance of the preference shocks in this model. The closer σ is to 1, the smaller the role that preference shocks play in the decisions that agents take. Conversely, the farther the scale parameter is from 1 (and at the limit, if it approaches ∞) the greater the role the shocks play and the lower the role of the model itself. The choice specific function as related in equation (3.3) would then be

$$v_{d,g,t=2}(z, k) = \frac{u(c)}{\sigma} + \alpha \mathrm{E}\left[V_{d,g',t=2}(z, \varepsilon)|z, I_k = 1\right]. \tag{3.4}$$

Using the above we can define the expected value of the next generation, who as children in period $t = 1$ received direct investment e and indirect investment j along with the intergenerational transition function $N(\cdot)$, as

$$\mathrm{E}\left[V_{d,g',t=2}(z,\varepsilon)|z, I_k = 1\right] = \sum_z \bar{V}_{d,g',t=2}(x) N\left(x|z, I_k = 1\right) \tag{3.6}$$

where x is the state of the g' agent (the child) in their $t = 2$ period of life. Plugging (3.6) above into equation (3.3), then (3.3) into equation (3.4), and finally (3.4) into equation (3.5) rewrites the system into the following recursive form:

$$\bar{V}_{d,g,t=2}(z) = \int \left[\max_{I_k \in I} \sum_k I_k \{v_{d,g,t=2}(z,k) + \varepsilon_k\}\right] dG_\varepsilon(\varepsilon_t). \tag{3.7}$$

Assuming that the shocks follow a Generalised Extreme Value Type 1 distribution, we can write the the above as

$$\bar{V}(z) = \gamma + \ln\left(\sum_s e^{v(z,s)}\right) \tag{3.8}$$

where γ is the Euler-Mascheroni constant. The policy function of the above model yields the conditional choice specific probabilities (CCPs), as outlined in Appendix C.1. These CCPs take the form

$$p(k|z) = \frac{e^{v(z,k)}}{\sum_s e^{v(z,s)}} \tag{3.9}$$

and enter the log-likelihood function.

3.4.2 Econometric Specification

In this section we econometrically specify the equations in the previous section. We partition the country into the two markets: market 1 consists of all the islands that is not Jawa; market 2 is the island of Jawa. We then partition the decision set and specify the nested logit version of the model in section 3.4.1 to estimate the parameters, accounting for the possibility of correlation in unobserved ε_k for some of the choices.

Wages We do not estimate a mincer equation for wages. Rather, we specify the wages the agent receives as a function of the skill in their state space. Upon choosing the location and paying the moving cost the agents receive their corresponding wage. The wages are the median wage offered in each location according to the agents' skill. We take these wages directly from the data and normalize them to the low skilled wage on the island of Jawa. These wages correspond to the final column in appendix table C3. We note that giving only one wage as opposed to two wages (one per parent in the household) seems fairly justified given the low female labor participation in Indonesia, as evidenced in table B10.

Terminal Value Functions Terminal value functions are the permanent incomes of the g' generation based on the state that a household in this generation finds itself in given the investment decisions made by the previous generation in each dynasty. Alternatively, this can be the period income and not necessarily the presumed permanent income.

Utility and Consumption We specify below the consumption of the normalized good, which is unchanged from equation (3.1)

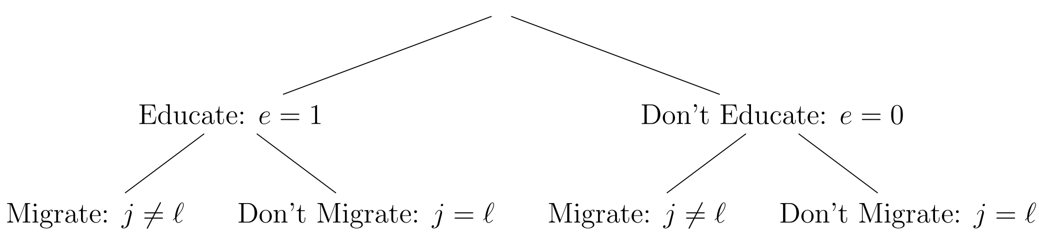
$$c(k, z) = w^{hh}(j, h) - \delta \mathbb{1}(j \neq \ell) - \phi_j \cdot \mathbb{1}(e = 1), \text{ for } j = 1, 2. \tag{3.10}$$

Utility is then simply $u(c(k, z)) = \log(c(k, z))$. Given log utility we set $c(k, z) > 0$ by giving agents whose value of $c(k, z) \leq 0$ a very small number.

Nested Logit Specification

Because there are varying degrees of variances around the components of the decisions that constitute the joint decision of educate and/or migrate, we make a further assumption on the unobserved, transitory preference shocks ε. Specifically, we assume that the joint decision can be partitioned into subsets according to the tree below, where the education level form the branches (we denote these as $b = 1, ..., B$, where in the present case $B = 2$), and the limbs within a branch form the migration level (and these we denote as $m = 1, ..., M_b$, where $M_b = 2$ in the present case given there are only two location choices). This partition implies that the independence of irrelevant alternatives (IIA) needed in estimation of the model in the previous section is relaxed to hold only within branches, but not across branches. Put differently, there is proportional substitution across limbs in a branch, but not across branches. The parameter ρ_b that will be estimated along with the other parameters of the model is an indication of the degree of independence in unobserved utility among the alternatives in nest b.[38]

Figure 3.4.1: Nesting Structure of Joint Decisions



[38]These ρ_bs also give an indication of the correlation between preference shocks $\varepsilon_{b,m}$ that exist within branches. McFadden (1977) shows that this relationship is given by $\rho_b = \sqrt{1 - \text{corr}(\varepsilon_{b,m}, \varepsilon_{b,-m})}$ for the present case of two limb alternatives.

This tree should not be interpreted as sequential decision making. Rather, the partitioning structure that the tree affords is a way of linking different but interdependent choices and decomposing a single decision to minimise the possibly restrictive condition of cross-alternative substitution (Hensher, 1986). Further, the structure of the tree is not arbitrary. From Tables C4 and C2 the variance of the migration decision is lower than the variance of the education decision. As the nested logit requires that the component of the within-limb correlation that the branch decisions contribute be smaller than the component the limb decision contributes, the above tree is implied by the data.[39] Further, as Hensher (1986) elaborates, the nested logit also gives us a test of the consistency of the assumed structure with utility maximisation. Namely, that the individually estimated ρ_bs must be between 0 and 1.

Specification We present below the specification of the model for estimation with the nested logit. First, the conditional choice probabilities presented in equation (3.9) can be decomposed into the branch ($p(b|z)$) and limb ($p(m|b,z)$)) probabilities according to the equation

$$p(k|z) \equiv p(b,m|z;\theta) = p(b|z;\theta) \cdot p(m|b,z;\theta), \tag{3.11}$$

where we have decomposed the k^{th} decision as the joint b, m combination according to the nest in figure 3.4.1. Since in our model there are no branch varying regressors, we can write the following closed form equations for the

[39] We also invert the nesting structure, but statistically reject this model where migration are branches and education forms the limbs

calculation of the branch probabilities as

$$p(b|z;\theta) = \frac{\exp(\rho_b \cdot \text{IV}_b)}{\sum_{s=1}^{B} \exp(\rho_s \cdot \text{IV}_s)}, \quad \text{where } \text{IV}_b = \log\left(\sum_{r=1}^{M_b} \exp\left(\frac{v(b,r,z;\theta)}{\rho_b}\right)\right). \tag{3.12}$$

The choice-specific value function in the above equation, $v(b,r,z;\theta)$, is the same as those presented in equation (3.3), except that we are now expressly including the parametrisation vector θ that was previously implied. The limb probabilities are calculated as

$$p(m|b,z;\theta) = \frac{\exp\left(\frac{v(b,m,z;\theta)}{\rho_b}\right)}{\sum_{s=1}^{M_b} \exp\left(\frac{v(b,s,z;\theta)}{\rho_b}\right)}. \tag{3.13}$$

We will therefore estimate two additional parameters along with those in the budget constraint equation (3.10).

Log-Likelihood Function The likelihood function of the model is simply the joint probability that we observe an agent's decision as being the maximal choice given the parameters. In this simple framework, the log-likelihood function is given by

$$\mathcal{L}(\theta) = \sum_{i \in HH} \log p_i(b,m|z;\theta) \tag{3.14}$$

where $\theta = (\delta, \phi_{1=\text{Everywhere Else}}, \phi_{2=\text{Jawa}}, \rho_{1=\text{Educate Branch}}, \rho_{2=\text{Don't Educate Branch}})$, and HH is the total number of households in the data.

3.5 Results and Simulations

3.5.1 Identification

The nested logit is solved with full information maximum likelihood (FIML). It is well known that FIML estimation of the log-likelihood function for nested logits are not globally concave. To avoid conducting sequential estimation, we first estimate the non-nested model in section 3.4.1 along with the specification in section 3.4.2 for the utility, consumption, and wages. As suggested by Hensher (1986), we then use the results of this estimation as the starting value for the nested logit specification (and make starting guesses for the ρ_bs of 0.5).

The model parameters ϕ_1 and ϕ_2 are identified due to the variation in educational decisions of those who stay as well as those who enter from the other region. The moving parameter cost δ is identified only for bilateral movements of equal cost; so we can not, in this framework, identify a location dependant cost of moving from one region to the other. The only parameter that can not be identified is the altruism parameter α. We instead calibrate it accordingly: Raut and Tran (2005) show that Indonesian families reach Beckerian levels of altruism according to their estimates. So we take a sufficiently high discount factors (0.99) and we elevate it to 18 for the full possible years of the $g + 1$ generation in their $t = 1$ period of life. This sets the value to $\alpha = 0.834$. A separate sensitivity analysis is conducted to understand the role that α plays in the estimation of the parameters of the model.

3.5.2 Estimation Results

The results of the previous section yield the parameter estimates presented in table C5. We present both the results of the nested logit (column 2) as well as the non-nested model (column 1, used as starting values for the nested logit). The likelihood ratio test between the two models rejects the non-nested model in favor of the nested model. Furthermore, a test of whether the nested model is consistent with random utility maximization can be conducted by testing if the null hypothesis that the parameter $\rho_b = 1$ is rejected. An F test on the joint parameters both reject the null in favour of consistency with random utility at the 95% level.

What is evident, is that between the non-nested model and the nested logit model the moving cost parameter decreases by about 90%. This may be due to the relative degree of independence of the migration decision within the educate branch. In the data, the high skilled households tend to educate proportionately more than the low skilled households, and it is mostly these high skilled households that move and/or educate. Regardless, whether the household stays or moves choosing to educate confers a larger increase in expected future utility given that the next generation (the child) will be in a higher state. So the moving cost can be much smaller relative to a model that applies IIA across all decisions.

Regarding the moving costs, it is difficult to provide a monetary value to give a clear cost due to the utility specification. However, we can give a monetary interpretation in proportion to wages. With the weighted average of the median wages being INT$1.33/hr[40], using the parameter estimate from the estimation results in table C5 we see that a family migration cost is a little over 135% of this averaged wage. Put differently, the average household would need to borrow at least 35%, or obtain a likewise pay increase. If the average household wishes to further educate their child in the new location

[40]We calculate this using the information in column 4 of table C3 and the distribution of skills in the dataset to obtain the weighted average obtained from the final column in table C10

this would require either a minimum 80% rise in wage or an equivalent sum borrowed. These are substantial sums from a development context. However, the model restricts migration to those in the high states given the borrowing constraint. So a counterfactual would need to include the relaxation of such a constraint. If we restrict ourselves to those households in the high skill state where the weighted average of median wages is INT\$2.48/hr, then the moving cost represents about 72% of their wage; choosing to also educate in the new region would imply a cost representative of 96% of their wage. To understand these figures we can compare them to the moving costs estimated by Bryan and Morten (2015). In their study they find that the average moving cost in terms of fractional wages for individuals moving between Indonesian districts are 0.59 in 1976 and 0.42 in 1995. For the average high skilled households in our data spanning observations from 1980 to 1997 we get 0.72. A bit higher, but understandably so given we are estimating family migration costs between island groupings as opposed to individual migration.

Although it seems that the utility cost of educating a child in the first region all the way through to 12 years of schooling is consistently more expensive, the two parameter values are statistically equivalent. This would indicate that on aggregate, the provinces on the islands other than those in Jawa have reached fairly similar level of costs to educate. This doesn't mean that they're necessarily equivalent in quality; simply that given the low relative difference observed between the high skilled median wages in the two markets, the difference in education costs between the two markets isn't going to be markedly different from each other. However, the costs are quite substantial and significant. Low skilled households must give up a substantial sum to educate their child and move them to a higher wage state. The weighted average of their median wages across both locations is INT\$0.98/hr. Taking as the cost to educate the average of the education cost parameters we get that choosing to educate costs about 60% of their wages (compare with high skilled households for whom the cost represents about 23%). High skilled households who choose to both move and educate must give up about 95% of their wages.

The closed-form solutions for the conditional choice probabilities allows us to generate this matrix for the solution (column 2 of table C5). As we see in tables C6 and C7, the conditional choice probabilities based on the model parameter estimates fit fairly well the choice probabilities we observe in the data. However, the model restricts the decision of agents in low skilled states from migrating. While low skilled households make a migration decision with very low probability in the data, in the present model the borrowing constraint prevents an agent from taking this decision. The nested logit does a fairly better job of matching the migration decisions of high skilled migrants, but does a poorer job properly matching the educational decisions of high skilled households who don't migrate. This is especially true among the high skilled households not located in Jawa (row 3), for which we observe a discrepancy of 0.12 between the "No Education No Migration" and "Education No Migration" decisions - the model under predicts their education decision when they choose to stay. The education decisions of the high skilled in Jawa are fairly well captured.

Given that the present model lacks other sources of agent heterogeneity, the dual assumptions of risk aversion and imperfect markets via a borrowing constraint help to fit the first two decisions. Although not presented here, a linear utility specification fails to capture the differences in the second decision (column 2) between low and high state households: it uniformly specifies across all states essentially the same distribution of decisions taken. A CCP matrix generated by the non-nested model (column 1 of table C5) results in similarly estimated probabilities for the first two columns in table C7, but poorly matches the probability of making a joint migration decision in columns 3 or 4. Future work to better match the high skilled households located elsewhere in Indonesia would look to incorporate another source of heterogeneity, namely the child's ability.

3.5.3 Sensitivity Analysis

We make two assumptions in the model that may make estimates sensitive to the assumed parameter settings.[41] The first is that we have set the parameter $\alpha = 0.834$, based on taking a high level of altruism (0.99) and discounting it for the first 18 years of a child's life. The second, implicit assumption, is that the relative risk aversion parameter is set to $\gamma = 1$ by using log utility. Given the added complexity of respecifying the utility function to account for different relative risk aversion parameters and no clear guidance from the literature[42] as to its calibration, we do not analyse the sensitivity of parameter estimates to different γs.

Therefore, we test the sensitivity of the estimates for different levels of α. Namely, Raut and Tran (2005) identify a level of altruism for Indonesian parents in their study that implies an $\alpha = 0.95$. We set the model to this level

[41]The creation of the skill measure to classify households is another assumption that could be tested to understand how our estimates change if we instead consider, for example, using the O*NET data to create a skills measure. However, this leads to several implications that would require a new model: first, educational investment is now risky, for which at least the child's ability would need to be incorporated; second, given the previous, our transition function is no longer a 1:1 mapping of education to skills and would require us to estimate it as well (assuming parents know the distribution associated with the risk of education and its realisation of a high skilled child, conditional on the child's ability and the parent's own state). We would expect in such a case that education costs would increase and leave this for future research.

[42]The literature on Indonesian estimates of relative risk aversion is sparse, and perhaps with the IFLS dataset better estimates can be obtained. Currently, the only study that we were able to find is Gandelman and Hernández-Mutillo (2014), who find an estimate of $\gamma = 1.24$. However, they report in their tables that they fail to reject the null at the 90% level that this parameter is equal to 1.

of altruism at the high end, and at the low end an $\alpha = 0.75$ and $\alpha = 0.5$ (the latter being more or less reciprocal altruism).

We find that the parameters are generally insensitive to the higher levels of α, specifically those above 0.75. In general we notice that the moving cost parameter decreases at a much steeper rate than the slight increase we observe in the education cost parameters; however even these changes are not substantial and all are significant. The parameters ρ_b are much more stable across specifications. We take this as a sign of the robustness of our results.

3.5.4 Goodness of Fit

We simulate the model with the estimated parameters on the same number of households found in the data. These households are further distributed across the two regions and the two skill sets according to the distribution in the data. We first look at the unconditional rates of migration and educational decisions taken in the simulated dataset and compare against those in the data.

We see that the model underestimates the migration rates, being about two times less than the rate in the data. This is expected given the CCP results discussed in table C7. As the model restricts borrowing, and the utility specification further restricts consumption to be positive, it assigns 0 probability to decisions involving migration for the low state households. However, in the data we notice that although the migration rate of low skilled households in the two regions is very low, the fact that these households make up a little over 70% of the data (last column in table C10) means that they contribute to the absolute number of migration events, and therefore add weight to the unconditional mean - in the data they are responsible for a little under half the number of events (67, as reported in row 2 of table C4). We can condition the mean on the household states to understand them better. Per table C10 we can see that the model does a well enough job of capturing the migration decisions of the high skilled, and does a fairly good job of capturing the education decisions.

We can also look at the outcomes for the next generation to understand how the distribution of their states fare against that which we find in the data. In general we find that the model does an adequate job of matching the unconditional proportions as seen in table C11. Not withstanding, we also report the transitional matrix and compare the probabilities of moving between states as generated by the model and compare with the data, as shown in table C12.

We can see from table C12 that low skilled households in both regions choose to educate their children in roughly equal proportions, and that the model captures this well. High skilled households overwhelmingly choose to educate their child, however, as we would expect given the literature on persistence in educational attainment across generations. Further, in the data, high skilled households that choose to migrate tend to also choose to educate their child. This occurs roughly 4 times more often for those high skilled households that move to the more productive Jawa, and 2 times as often for high skilled Jawanese households that move elsewhere in Indonesia. Our simulated dataset underestimates these rates by roughly half in both cases.

Overall, the current model does a fairly good job of fitting the data and we consider it a good candidate to describe the DGP of the original data. The nested logic specification helps us achieve some of the fitting, considering there is only one source of heterogeneity in the model (the initial household skill level). By construction the model doesn't capture the migration probabilities of low skilled states, and slightly underestimates the probability that the high skilled households located elsewhere in Indonesia educate their child.

Other sources of heterogeneity can be incorporated to improve the fit, which may allow for the relaxation of the borrowing constraint to better capture the migration of the low skilled households, which is left to future research.

3.6 Effect of Migration and Education Subsidies

We conduct two counterfactuals on the model consisting of the effect of subsidising migration or education. As these costs are quite substantial and may impact the low skilled differently, a policy analysis should determine whether one of these dominate. We apply these subsidies as follows: we generate the new parameter(s) by applying a reduction via a subsidy τ, $\hat{\theta}_{counter} = \hat{\theta}(1-\tau)$, where $\hat{\theta}$ is the particular parameter (or parameters in the case of education subsidies) we are affecting, *ceteris paribus*. Since our CCPs are closed formed solutions as given in equation (3.11), we generate a new CCP matrix based on the new parameter(s). These CCPs are then used to create simulation datasets composed of the same number of initial households with the same skill proportions as found in the data. For the time being we forego analysis on a mixture of policies. We then analyse the effect of the subsidy we applied on the distribution of the next generation's states conditional on the previous generation's starting states. We finally conduct a welfare analysis of the effect of these subsidies by comparing the overall change in the proportion of high skilled children produced and the labor market outcomes for the next generation against the baseline results of the model, as well as the implications for social mobility.

3.6.1 Migration Subsidies

From the results of estimating the model we know that in terms of fractional wages low skilled households would need to be compensated about 80% of the migration cost for them to consider moving. We use this starting percentage and also consider a 70% compensation (given that high skilled households have to forego about 72% of their income for a migration event) along with a 90% and 100% (fully subsidised) migration compensation for all households in the simulated datasets. The baseline case of no subsidy corresponds to the results from the original model. While the subsidy could be means tested to apply for only the low skilled, high skilled households may also benefit from the subsidy and may also generate an impact on welfare. For example, a high skilled household located elsewhere in Indonesia may gain by switching to the island of Jawa, where the median income is higher for their current generation and for the next generation once the child is educated there and enters the labor market. Other authors discuss what policies regarding migration subsidies might look like. Bryan and Morten (2015) note that these may take the role of better language training, infrastructure development that facilitates mobility, or migrant welcome centers to facilitate ethnic integration into the new region. Bazzi et al. (2016) suggest that based on Indonesia's transmigration program, which moved poor, low skilled households out of Jawa and randomly into the outer islands at no cost, benefits of migration derived from regional productivity differences may be overstated.[43] Here too, a migration subsidy might help in acclimating households to their new location. We remain agnostic to the methods of a policy application and instead analyse the overall effect of it.

[43] This had a differential impact on households who were not matched agroclimatically to their new locations (as most were farmers, those poorly matched had to learn new skills and were less productive on average).

The results of applying the migration subsidies and simulating the model is presented in figure C2, one for each type of household and their initial location. Starting with the 80% migration subsidy, we observe that low skilled households not only start to migrate, but invest in their child's education in the new location; however, the relative proportion of high skilled children produced by the low skilled households located elsewhere dips below baseline. As the subsidies increase, we observe a corresponding increase in the substitution response between locations and a relative increase in the proportion of high skilled children. For low skilled households located in Jawa, we can see that a full migration subsidy increases the relative proportion of high skilled children by 12% from a baseline of 44%, and by 9% for children originally born into low skilled households elsewhere in Indonesia (from a baseline of 48%); the application of a 90% subsidy, however, seems to produce the optimal marginal gain.

Contrast this with the effect of the subsidy on high skilled households, where we observe them substituting between locations rather than substantially increasing the relative stock of high skilled children (the overall relative stock of high skilled children produced by high skilled households in both regions is relatively flat throughout the subsidy schedule, owing to intergenerational persistencies). The largest increase in the marginal substitution of locations for these high skilled households occurs with a subsidy of 70%.[44] At this level of subsidy the high skilled located elsewhere in Indonesia increase their substitution response, moving to Jawa and educate their children there (a jump from about 3% to 29%). This has obvious welfare implications just as subsidizing the migration of low skilled households does (instead, perhaps, constrained by their preference shocks). The substitution response by the

[44] This is not surprising, as we found in the results section that the migration cost for the average high skilled household represents 72% of their income.

high skilled households is large because we are probably over incentivizing them given they are not necessarily subject to borrowing constraints. The low skilled households have to overcome their borrowing constraint and so we don't see such dramatic substitution in location, although we see clear increases overall in the proportion of high skilled children these households produce once the threshold is crossed.

In short, we observe that low skilled households engage in both migration and educational investment in the new location in response to a migration subsidy. Furthermore, these households increase the relative proportion of educated children above the baseline model. High skilled households mainly use the subsidy to facilitate a move to a new location and educate in that location, while only slightly increasing the relative proportion of educated children from baseline. Given that we observe different marginal responses to the subsidies based on the initial household skill level, this analysis would imply that on its own, the optimal migration subsidy for low skilled households is roughly 90% of the cost and 70% for high skilled households (conditional on any policy maker able to perfectly identify such households).

3.6.2 Education Subsidies

We next ask what an education subsidy would look like and how this would impact households. Similar to the previous exercise, we construct a subsidy schedule ranging from no subsidy (the baseline case) to full subsidy, increasing the subsidy in 25% increments. We also subsidize the education costs in both regions; as such we look at the effect of decreasing the ϕ_j in table C5. In this case we notice that since the borrowing constraint remains binding for the low skilled households we do not observe family migration events in the face of education subsidies, implying children become adults in the location of their original household. We provide the results of this exercise in figure C3.

We notice overall that low skilled households substantially increase their investment in education as we reduce their cost to educate, especially among those in Jawa (who are much more incentivized to educate given their children will receive a larger wage once they enter the labor market compared to the low skilled located elsewhere in the country). Those low skilled households in Jawa produce roughly 25% more educated children when given a full subsidy compared to an increase of 19% in the relative proportion of educated children produced by low skilled households located elsewhere. The reason the baseline case for Jawa is lower is simply that low skilled labor in this region has a lower return than low skilled labor outside Jawa. So a subsidy produces greater results in this region.

The effect of an education subsidy on high skilled households is much less dramatic. While increasing subsidies does incentivize high skilled households located outside of Jawa to migrate there (in the figure we observe an increase in the substitution response among these households), the overall gain of 4% in the relative proportion of educated children they produce is much more modest in comparison to the gains seen among their low skilled counterparts. The high skilled households in Jawa, meanwhile, have a dampened substitution response between locations, perhaps as they take into account that switching locations yields less pay for theirs and the next generation and that it costs slightly more to educate outside of Jawa. The substitution response increases slightly for sure as we increase the subsidy, but this small margin is probably due to some households being subject to a preference shock for that location. The overall relative proportion of educated children these households produce increases by almost 9% compared to baseline when education is fully subsidized. Among these high skilled households, we note that the overall gains from the increasing education subsidy really originates from the substitution of locations.

In short, an educational subsidy works to increase the proportion of children coming from low skilled households that become educated children, with a minimal effect on the high skilled households. The low skilled households in Jawa are more incentivized to educate with such a policy as the return to education in this region is much higher. An education subsidy does induce a small substitution response among high skilled households elsewhere in Indonesia, choosing to move to Jawa and educate their child there.

3.6.3 Welfare Implications on the Next Generation

The natural question to ask if we don't consider a mixture of subsidies is, which of the two subsidies is best. For this analysis we consider the labor market outcome of the next generation given the household decisions along the subsidy schedules. We use the median wages in table C3 and assume that in our partial equilibrium environment these wages remain unchanged for the next generation, despite changes in the spatial distribution and educational attainment of the next generation.[45] We aggregate the total income of the economy, generate the per capita income and create the welfare index by normalizing against the baseline model. In this way we can compare the results of both subsidies. We note that in the baseline model, as we have previously reported, the weighted average of median wages is INT$1.33/hr for the first generation. The baseline model produces a weighted average of median wages of INT$1.76/hr for the next generation, a 32% increase from the original generation.

We first report the overall composition of the next generation's states based on parental investment decisions, and the corresponding changes in those pro-

[45] Any real wage growth between generations is also inconsequential in both the model and the simulations as this would just scale the results, given our homothetic preferences.

portions based on the parent's responses to the subsidy in table C13. While these proportions may look different to those in figures C2 and C3, they are in fact related. We must simply take into consideration that each of those figures was generated relative to the initial state of the parents, so adding and subtracting shaded areas between graphs gives the overall composition for a location.

The subsidies to migration induce large decreases in the low skilled located in Jawa in the next generation. These shares are balanced with larger increases in the high skilled located elsewhere. We posit that the reason this is the case is that for the low skilled in Jawa, the migration subsidy leads families to migrate because the income for low skilled labor elsewhere in Indonesia is greater. Recall that figure C2 implied that low skilled households in both locations started switching locations in response to the subsidy. The net effect of these substitution responses is what we are capturing here: the substitution of low skilled households out of Jawa is larger than the substitution of families located elsewhere into Jawa.

In response to an education subsidy, however, we notice much larger differentials in the composition. Low skilled households overwhelmingly choose to educate in response to this subsidy. Moreover, as these households' borrowing constraints are binding given that we have not helped them move, the next generation originating from these households are now high skilled in the same location. A full accounting of how many of these stayers compose the high skilled households in their origin location, and how many are new entrants from high skilled families that moved to different regions, would need to take into account the origin location of the households. For example, when we fully subsidise education there is a commensurate drop of 6.4% in the proportion of the next generation composing low skilled households located outside of Jawa.

This 6.4% drop must mean that a portion of this share came from low skilled households outside of Jawa that educated their child there, and is now within the 8.6% increase we observe in high skilled in the same location.

Overall, we note that both subsidies generate an overall increase in the amount of educated (and thus, highly skilled) children in the next generation. However, the extent of the overall gains produced by either subsidy are different considering the costs associated with both migration and education - which in turn leads to different tradeoffs dominating the resulting outcome. We confirm this with the overall labor market outcomes for the economy, as presented in figure C4. It demonstrates the results of the welfare analysis of the subsidy schedules on the weighted average of median earnings in the economy for the next generation.[46]

The education subsidy dominates the migration subsidy along the whole schedule. The effect of a migration subsidy doesn't kick in for the whole economy until the 90% level is reached. From this graph (and table C15 in appendix C.3) we can infer the equivalent subsidies that produce the same results in the economy. A 90% migration subsidy (the amount that induced the largest marginal response in low skilled households) is equivalent in welfare terms to a roughly 25% educational subsidy. If we want the same welfare outcome associated with fully subsidizing migration events, we could achieve this with a policy that subsidizes education to the tune of 30% of the current cost across both regions. That being said, if the benefit to the overall economy is greater than the cost of fully subsidizing education, this would produce the greatest welfare gains (a 15% increase from the baseline model among the next generation, and a 52% intergenerational increase in earnings).

Finally, instead of generating intergenerational matrices for each subsidy,

[46]The data used to generate this figure is located in table C15.

we conduct some basic OLS regressions on the simulated data sets to better visualize how the subsidies affect social mobility with regards to education. As our education variables are dummies, we are estimating a linear probability model of the form:

$$EducChild_h = \alpha + \beta EducParent_h + u_h. \tag{3.15}$$

We present the results in table C14 for each specification estimated based on the type of subsidy and the corresponding schedules. The constant, α, is the base probability of being educated, corresponding the the probability that the average low skilled household educates their child; while the coefficient β estimates the "premium" above the base probability associated for a child whose parents are educated.

First we note that for the baseline model we estimated, the overall probability that a child is educated conditional on the parent's education is given as the weighted average of the diagonal probabilities in table C12 (Simulation column) or the weighted average of the coefficients (summing when necessary) of column 1 (Baseline) in table C14. In both cases we get the overall average probability of being educated as 0.557. This is similar to the result of Hertz and Jayasundera (2007). These authors find that the probability an Indonesian child is educated in 1978 (the last year of their data) is around 0.56. As the children in our data are born after 1978 and benefit from the lasting effects of the INPRES Program these authors use as instruments in their analysis, we take this as further confidence in the results of our model.

While both subsidies significantly increase the base probability that the average low skilled household educates their child, the education subsidy increases this probability to a greater extent than the migration subsidy alone.

This effect of an education subsidy also kicks in at lower subsidies. We also notice that the education subsidy does a much better job of nullifying the premium probability of obtaining education based on having educated parents (breaking the intergenerational correlation, also shown in figure C5). The migration subsidy does not accomplish this. It increases the base probability so everyone has a higher chance of getting educated, but children from high skilled households still retain an advantage that the education subsidy markedly reduces. And as we can observe in figure C6, the overall intergenerational correlation in education is reduced with an education subsidy.

We conclude that both subsidies increase the overall proportion of children that enter into the high skilled states and that their labor market outcomes are better than the current baseline. However, education subsidies dominate in their ability to better impact the outcome for the next generation, and would be generally more cost effective to implement. In the context of Indonesia it may perhaps be better for low skilled households to stay put in their current location and receive a subsidy to incentivize them to educate their children, who are then set up to enter other labor markets as skilled workers when they become adults.

3.7 Conclusion

In this chapter we have proposed a causative mechanism to understand the tradeoffs between the dual investments of migration and education, and its impact on the subsequent generation. We show through the mechanism that describes the DGP of our Indonesian households that family migration does indeed cause households to invest in the education of their children in the new location. The model itself makes a strong assumption on the inability of low skilled households to migrate; however, it fits the data fairly well considering that the only source of heterogeneity is a household's skill. It is also robust to the specification of the altruism parameter. Counterfactuals were conducted by giving all households either a migration subsidy or an education subsidy. Here we find that while reducing the cost barrier to migration causes low skilled households to both migrate and increase educational investment in the new location (thereby increasing overall production of high skilled children), a better policy is to subsidise education - possibly fully to obtain the greatest overall benefit. This would place the next generation at a higher skill state to then make their own location choices with a lower proportion of the population in states where a borrowing constraint is more likely to bind. We show that in the case of Indonesia policy mechanisms exist to increase intergenerational mobility among children in low skilled households.

Future research would focus on adding more dimensions of heterogeneity in both moving types for parent households and ability types for children. These two conditions may allow for the relaxation of a budget constraint and the assumption of CRRA utility. The current model also collapses the initial generation's 18 years of child rearing into one period. This allows us to capture the extensive margin of educational investment as impacted by migration quite easily, but foregoes the ability to comment on migration's impact on the intensive margin of educational investment. A future model that considers life stages of children (to capture the relative extensive/intensive margin of one more schooling levels), as well as regarding education as a risky investment (by incorporating child ability type and different measures of skill attainment), might allow us to comment on the relative importance of the intensive margin. We note that in such a case, what we have estimated in this paper are then upper bounds for the effect of educational subsidies, as risky educational investments imply larger opportunity costs to educate. Unless there is a crossing with the effect of the migration subsidy, we would expect to find dampened effects to education subsidies under risky investment relative to what we document here; yet it would still be cost effective relative to inducing family migration through a migration subsidy. Furthermore, including more dynamics within the parental generation would allow us to "give back" households their yearly wage they actually earn instead of assigning them a median wage based on their ascribed state. Another role is to understand if a mixture of migration and education subsidies does better than applying either one alone; or if a tax on the high skilled can be used to finance one of these subsidies.

A Chapter 1 Appendices

A.1 Proof of Proposition 1.5.1

Proof. To start the proof we take the total derivative of the transfers in period 1, leaving the functional form of the transfer cost unspecified so as to flexibly adjust for each assumption as to how the transfer fee may be applied.

Recall that the total transfers of period 1 have the following form:

$$T_1^*(a_2^*(\varphi_i(\overline{\varphi})), \varphi_i(\overline{\varphi})).$$

So the total derivative is given by

$$\left.\frac{dT_1^*}{d\overline{\varphi}}\right|_i = \frac{\partial T_1^*}{\partial a_2^*}\frac{\partial a_2^*}{\partial \varphi_i}\frac{d\varphi_i}{d\overline{\varphi}} + \frac{\partial T_1^*}{\partial \varphi_i}\frac{d\varphi_i}{d\overline{\varphi}}. \tag{A.1}$$

This relation highlights the indirect and direct effect that a change in transfer fees has on the remittance decision, where the indirect effect is generated from the distortion's effect on the savings decision.

It is useful to restate equation (1.10) in terms of total transfers, per equation (1.13), to obtain

$$T_1^* = \frac{\left[\frac{\alpha}{\varphi_i(1-\alpha)}\right]^{1/\gamma}[y_1 - a_2^*] - \widetilde{y}_1}{1 + \varphi_i\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}}, \tag{A.2}$$

from which it is obvious that $\frac{\partial T_1^*}{\partial a_2^*} < 0$.

Next, we look at equation (1.12). The sign of $\frac{\partial a_2^*}{\partial \varphi_i}$ is determined by agent 2's inter-temporal wage structure:

$$\Delta\widetilde{y} \equiv [\beta(1+r)]^{1/\gamma}\widetilde{y}_1 - \widetilde{y}_2. \tag{A.3}$$

If the economy is at steady state, then the Euler condition derived from maximizing equation (1.2) yields $r = \frac{1}{\beta} - 1$, and so (A.3) simplifies to $\Delta\widetilde{y}_{SS} \equiv \widetilde{y}_1 - \widetilde{y}_2$. As we have assumed that first period wages are at least as large as second period wages, it follows that $\frac{\partial a_2^*}{\partial \varphi_i} > 0$. This would indicate that if migrant remitters (agent 1) faced lower fees, they would save less as they transfer remittances forward in time.

The sign of $\frac{\partial \varphi_i}{\partial \overline{\varphi}}$ is dependent on the functional form. Recall that under the first functional form assumption $\varphi_1 = \frac{1}{1-\overline{\varphi}}$; and its inverse is $\varphi_2 = 1 - \overline{\varphi}$. This implies that $\frac{\partial \varphi_1}{\partial \overline{\varphi}} > 0$ and $\frac{\partial \varphi_2}{\partial \overline{\varphi}} < 0$.

Finally, we need to analyze $\frac{\partial T_1^*}{\partial \varphi_i}$, as it is this derivative that fully determines the sign of (A.1). Taking the derivative of (A.2) with respect to φ_i we obtain

$$\frac{\partial T_1^*}{\partial \varphi_i} = -\frac{\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}}{\gamma\left[1+\varphi_i\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\right]^2}\left(\begin{array}{c}\frac{\partial a_2^*}{\partial \varphi_i}\varphi_i\gamma\left[1+\varphi_i\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\right]+ \\ \varphi_i\left[\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\gamma y_1-(\gamma-1)\widetilde{y}_1\right]- \\ \varphi_i y_1 + a_2^*\left[\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\varphi_i\gamma+1\right]\end{array}\right). \tag{A.4}$$

The sign of the above can only be determined if the following conditions hold (derived from the above bracket terms) :

$$y_1 \geq \widetilde{y}_1\left(\frac{\varphi_i(1-\alpha)}{\alpha}\right)^{1/\gamma}\left(\frac{\gamma-1}{\gamma}\right) \tag{A.5}$$

$$a_2^* \leq y_1\left(\frac{\varphi_i}{\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\varphi_i\gamma+1}\right). \tag{A.6}$$

If these conditions hold, then the derivative is $\frac{\partial T_1^*}{\partial \varphi_i} < 0$.

Analyzing the terms multiplying $\widetilde{y}_1$ and y_1, respectively, in the above conditions it becomes clear that under the first functional form assumption, where the transfer cost is convex in $\overline{\varphi}$, the multiplier in both cases is larger than the inverse case of assumption 2. [47] This indicates that if transfer costs are convex then agent 1's wage would have to be larger than in the case when transfer costs are linear in $\overline{\varphi}$; but it also means that he would likely be able to save more as his income would be larger.

With this derivative established it follows from equation (A.1) that $\frac{\partial T_1^*}{\partial \overline{\varphi}} < 0$ when transfer costs are convex in the transfer fee; and $\frac{\partial T_1^*}{\partial \overline{\varphi}} > 0$ when transfer costs are linear in the transfer fee, establishing the claim.

□

A.2 Comparative Statics of of the Model Primitives

In this section we consider the relationship between the two remaining model primitives, α and γ. We elucidate this relationship as changes in γ may have a perceived effect on transfers, even if it is expected that altruism (if it is fundamental) should have no correlation with an agent's desire to smooth consumption. Although this is the expected result, it is likely the case that one would infer a change in altruism as an agent's smoothing parameter changes. To show this, we first find the relationship between a_2^* and γ:

$$\frac{\partial a_2^*}{\partial \gamma} = -\frac{[\beta(1+r)]^{1/\gamma} \ln[\beta(1+r)]\{[(1+r)y_1 + y_2] + \varphi_i[(1+r)\widetilde{y}_1 + \widetilde{y}_2]\}}{\gamma^2\left[(1+r) + [\beta(1+r)]^{1/\gamma}\right]^2} \geq 0 \tag{A.7}$$

where the above is non-negative since the term $\ln[\beta(1+r)] \leq 0$ when r is below or at the steady state.[48] Using this relationship, we compute the derivative of

[47]Plug in the respective transfer costs for each method in which the transfer cost may be applied to both (A.5) and (A.6), and the result becomes immediate when analyzing the effect as $\overline{\varphi}$ or γ reach their upper bounds.

[48]If the economy is at steady state, then $\frac{\partial a_2^*}{\partial \gamma} = 0$.

the transfers T^* w.r.t. γ to obtain:

$$\frac{\partial T^*}{\partial \gamma} = \tag{A.8}$$

$$-\frac{\left[\gamma^2 \frac{\partial a_2^*}{\partial \gamma}\left\{1+\varphi_i\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\right\}+\left[(y_1-a_2^*)+\varphi_i\widetilde{y}_1\right]\ln\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)\right]\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}}{\gamma^2\left[1+\varphi_i\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\right]^2}$$

The sign of the above is partially dependent on the term $\ln\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)$. With convex transfer costs this term will be negative and with linear transfer costs it will be positive. Further, per footnote 48, if the economy is at steady state the sign of (A.8) is fully determined: under convex costs, $\frac{\partial T^*}{\partial \gamma} > 0$; and under linear costs, $\frac{\partial T^*}{\partial \gamma} < 0$. Then taking the derivative of the transfers with respect to the altruism parameter one obtains the following relationship:

$$\frac{\partial T^*}{\partial \alpha} = \frac{\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\left[y_1 - a_2^* + \varphi_i\widetilde{y}_1\right]}{\gamma\alpha(1-\alpha)\left[1+\varphi_i\left(\frac{\alpha}{\varphi_i(1-\alpha)}\right)^{1/\gamma}\right]^2} > 0 \tag{A.9}$$

This last relationship shows that transfers are increasing in altruism, a result similar to Shimada (2011) and that reported in Rapoport & Docquier (2005) via Stark (1995). With (A.9) We can now use the implicit function theorem to show the relationship between the parameters γ and α at a constant transfer stream $\overline{T}^*$:

$$\left.\frac{\partial \alpha}{\partial \gamma}\right|_{i=1} = -\left.\frac{\partial T^*/\partial \gamma}{\partial T^*/\partial \alpha}\right|_{\overline{T}^*} < 0 \quad , \quad \left.\frac{\partial \alpha}{\partial \gamma}\right|_{i=2} = -\left.\frac{\partial T^*/\partial \gamma}{\partial T^*/\partial \alpha}\right|_{\overline{T}^*} > 0 \tag{A.10}$$

We would expect then to infer at steady state and under convex costs that as an agent's smoothing parameter increases the observed altruism decreases as he returns to preoccupy himself with his own consumption needs before concerning himself with the second agent, and the reverse when costs are linear.

Figures A2 - A4 are the simulation results of the comparative statics of gamma on altruism at constant transfers and holding one agent's preference parameter constant while varying the other's. In the case when agents have the same preference parameter the model yields the same result as the prediction. For the cases that could not be analyzed analytically two outcomes arise (Figures A3 and A4). The first result has that as agent 1's preference parameter increases (holding agent 2's parameter constant), agent 1's altruism must decrease to support the constant stream of remittances. In the second case, as agent 2's preference parameter increases while agent 1's parameter remains unchanged, agent 1's altruism must increase to support a constant stream of transfers. We graph these for each year to show the break that occurs after 2008.

A.3 Tables and Figures

Table A1: Calibration Results

Parameter	No Costs (1)	Convex Cost (2)	Linear Cost (3)
$\beta^{\dagger}$	0.99	0.99	0.99
$\gamma^{\dagger}$	1.5	1.5	1.5
r[†]	$\frac{1}{\beta}-1$	$\frac{1}{\beta}-1$	$\frac{1}{\beta}-1$
α^{*}	0.3615	-	-
$\overline{\varphi}_1$	-	0.2196	-
$\overline{\varphi}_2$	-	-	0.3804
$\overline{H}^{\ddagger}$		562,000	
$\overline{R}^{\ddagger}$		\$1,452	

[†] Sources mentioned in section 1.6.2.
[‡] Sources mentioned in section 1.6.1: $\overline{R}$ in millions (2000 constant \$).
[*] Calibrated under No Costs; assumed value of 0.5 under each transfer cost functional form assumption.

Figure A1: Per Capita Remittance Profiles from the Model (Net of Remittance Costs)

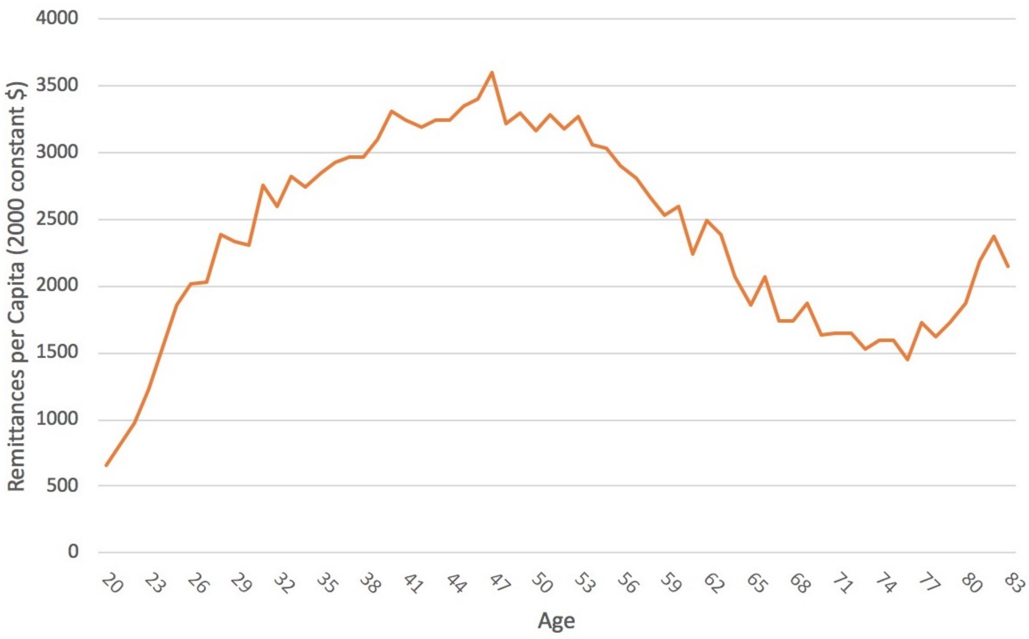


Figure A2: Comparative Statics: Effect of γ on α. Both agents share the same γ and aggregate transfers are held constant.

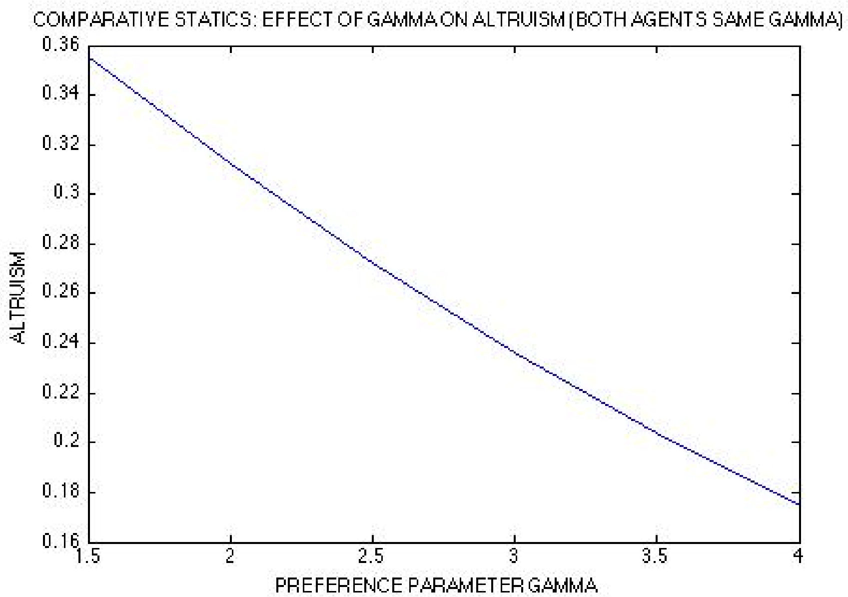


Figure A3: Comparative Statics: Effect of Agent 1's γ on α. Agent 2's $\gamma = 2$ held constant and aggregate transfers are held constant.

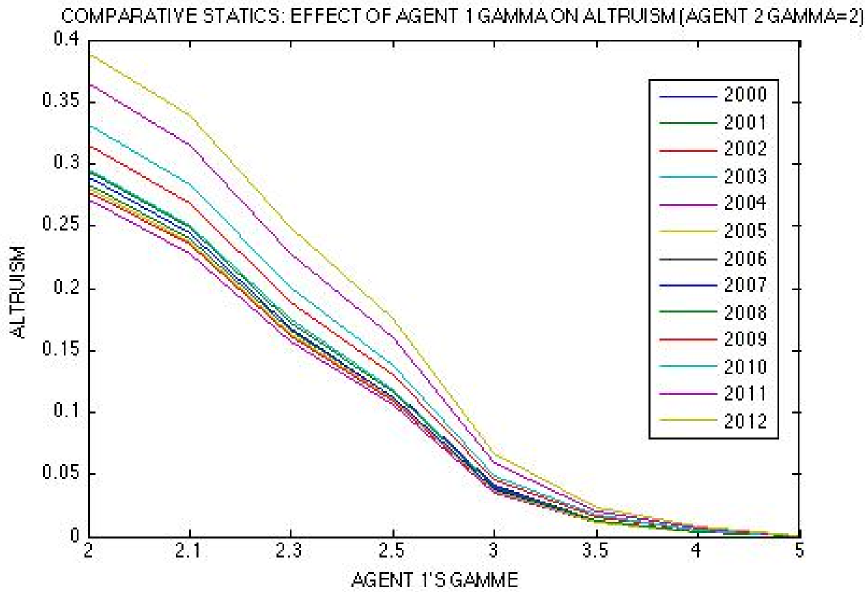
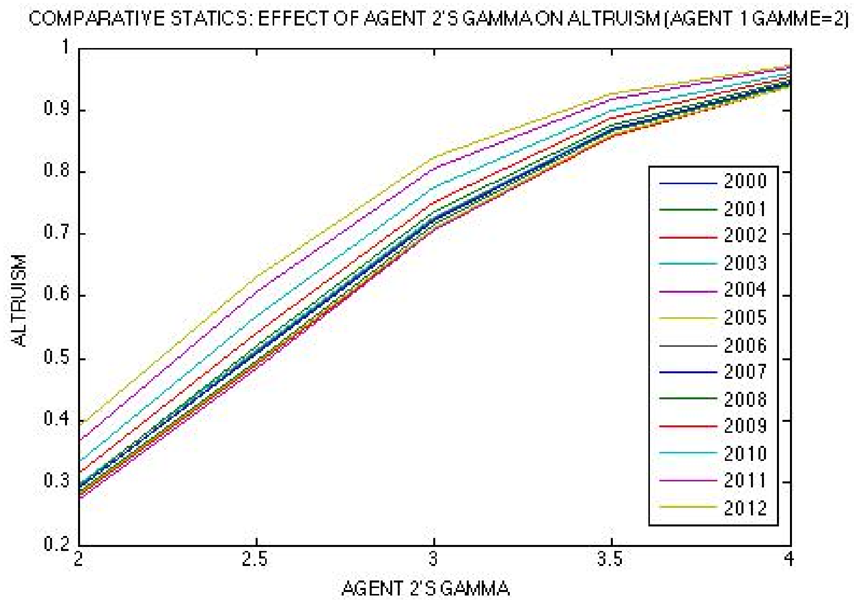


Figure A4: Agent 1's $\gamma = 2$ held constant and aggregate transfers are held constant.



B Chapter 2 Appendices

B.1 Further Background on IFLS

The first wave launched in 1993, covering 13 of the then 26 provinces on 6 islands in Indonesia.[49] As we elaborate in the next subsection, since the first wave the set of provinces and other regions have expanded. Figure B1 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 B2). We include this newly available data set to increase our observations when conducting analysis on IFLS wages, 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 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 rep-

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

resentative 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 90%. By the last wave in 2014 19,088 households were contacted, a little over half of which were dynasty households, comprising 83,766 individuals.

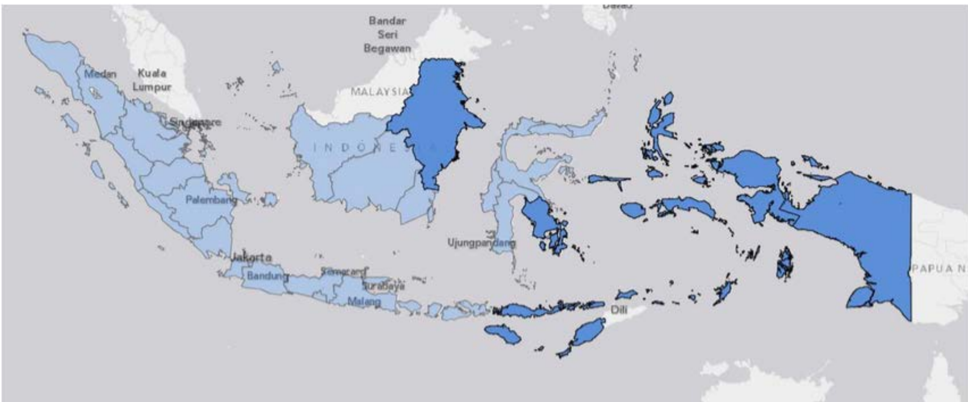
In the 1993 wave the head of household, spouse, other adults in the household (with a maximum interview of four adults), and two random children were targeted for interviews. Starting with the 1997 wave the procedure changed to interview all household members in the dynasty households (so that those 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.

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

Figure B1: RAND Corp. IFLS Provinces

Source: IFLS Frankenberg and Karoly (1995)

Figure B2: SurveyMETER IFLS-East Provinces



Source: IFLS-East Sikoki et al. (2013)

B.2 Indonesian Geographical Units and Population

From the Indonesian Statistical Office (BPS) we obtained the aggregated data on the geographical subdivisions of Indonesia and the populations from the 2010 census, presented in the table B1. 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 recognised as a sovereign country, breaking ties with Indonesia). Currently there are 34 provinces.[50] The master files of geographical codings available at Statistics Indonesia's website also indicates that districts and sub-districts have split, resulting in expanding sets at the finer levels of geographical codings.

B.3 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).[51]

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

[50] 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`

[51] 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 our data set of Indonesia we find evidence that may be consistent with this via an uptick in migration for those aged 65+, who are overwhelmingly female in composition.

Table B1: 2010 Census and Geographical Subdivisions in Indonesia

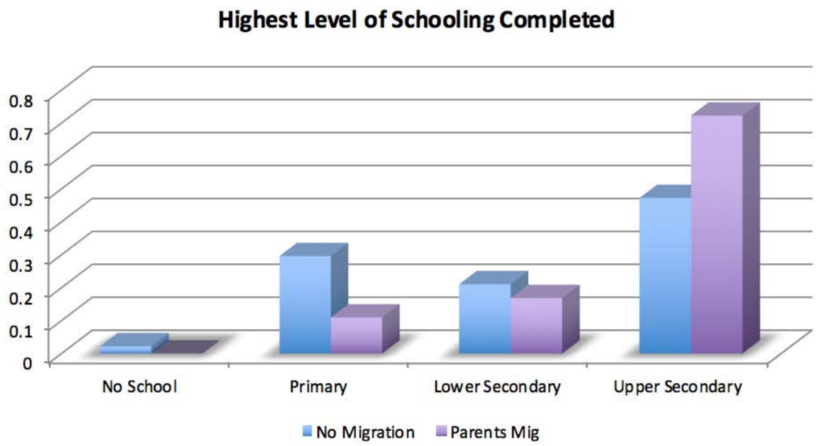
Province Name	ISO[4]	Capital	IFLS or IFLS-East Province	Population 2010 Census	Area (sq. km)	Population Density	Island Group	Number of Districts (kabupaten)	Number of sub-Districts (kecamatan)	Villages (kelurahan)	Municipalities (kotamadya)
Special Region of Aceh	ID-AC	Banda Aceh		4494410	57956	77	Sumatra	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 (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 our research we do not distinguish between the secular, religious, or vocational routes of educational attainment and treat them all equally, as our 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 occurred 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 our IFLS data set we 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.

In our data we find the following schooling completion rates conditional on the parents migration (interprovincial).

Figure B3: Schooling Completion Frequencies When Migration Occurs While In School



B.4 Descriptive Statistics

Family Data

Table B2: Descriptive Statistics of Parents (Children born between 1980 and 1997)

Parents	Don't Migrate		Migrate Provincial	
	mean	std	mean	std
Father				
Urbanization (Birth location)	0.32	0.46	0.46	0.49
Schooling Years	6.29	4.14	8.07	4.02
Age of Marriage	23.95	4.31	24.64	4.71
Age when first child born	26.73	5.17	26.59	4.33
Age of First Family Migration	-	-	32.07	7.02
Mother				
Urbanization (Birth location)	0.34	0.47	0.48	0.5
Schooling Years	5.11	4.02	6.84	4.08
Age Marriage	19.63	3.98	20.11	4.51
Age when first child born	22.41	4.62	22.1	3.84
Age of First Family Migration	-	-	27.75	6.3
Observations*	5004		240	

Observations are the sets of parents

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. We 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 B3: 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									

Table B5: Family Composition (Children Born 1980 - 1997)

	mean	std
Num of Children (Avg of Birth Order)	2.23	1.36
Avg Educ. Parents (years)	5.79	3.79
Average Size (persons in HH)	4.04	1.47
Observations (HH)	4,973	

Children's educational attainment is decomposed by urbanization and sex and by whether the family engaged in a provincial migration event. Substantial heterogeneity is apparent along urbanization and sex.

Table B4: Descriptive Statistics of Children (born 1980 - 1997)

Children	**Family Don't Migrate**		**Family Migrates (Provincial)**	
	mean	std	mean	std
Age Enter School	6.59	1.08	6.58	1.23
Age of child when family first migrates	-	-	6.32	4.87
School grade when family first migrates	-	-	0.77	4.89
Schooling Years	9.45	2.86	10.78	2.64
Proportion start school age >7	0.09	0.29	0.08	0.27
Proportion repeat a school year	0.01	0.11	0.01	0.08
Attended Kindergarten	0.31	0.46	0.43	0.49
Age Exit School	16.35	2.88	17.06	2.5
Proficiensies in Indonsian				
Write	0.99	0.09	0.99	0.08
Speak*	0.46	0.49	0.49	0.5
Read	0.99	0.08	0.99	0.08
Sex (Male)	0.51	0.49	0.5	0.5
Urbanization (Birth location)	0.32	0.47	0.47	0.49
Religion (Islam=1)	0.08	0.27	0.08	0.27
Observations	9226		377	

*Speak: this variable is measuring the proportion of children who speak Bahasa Indonesian at home, not those who are able to speak Bahasa Indonesian.

Table B6: Education by Urbanization (Children born 1980 - 1997)

Cross Tab: Children's Education (years)	Rural		Urban	
	mean	std	mean	std
Girls	9.51	3.10	11.02	2.57
Boys	9.45	3.12	10.81	2.81

Table B7: Education by Urbanization and Family Migration (Children born 1980 - 1997)

Cross Tab: Children's Education (years)	Family Doesn't Mig		Family Prov Mig	
	mean	std	mean	std
Rural Child	9.16	3.12	10.02	2.76
Urban Child	10.76	2.76	11.2	2.52

Wage Data

Table B8: IFLS Wage Data

	mean	std	min	25%	50%	75%	max	count
Hours/Week	42.22	21.86	1	28	42	54	168	235457
Weeks/Year	42.08	14.36	0	40	50	52	52	235456
Wage/Hr (INT$)	5.13	334.37	0	0.45	1.09	2.47	135595	235457
School Years	6.93	4.43	0	3	6	12	13	226181
Skill Level*	0.29	0.45	0	0	0	1	1	226100
Age	37.36	13.8	15	27	35	46	98	235457
Sex	0.35	0.48	0	0	0	1	1	235447
Urbanization	0.37	0.48	0	0	0	1	1	234440

*Skill level of occupation determined by using the O*NET database

Table B9: 1995 IPUMS Census Wage Data

	mean	std	min	25%	50%	75%	max	count
Hours/Week	45.14	14	1	37	45	52	98	106890
Wage/Hr (INT$)	2.85	7.47	0.06	1.04	1.79	3.2	768	106890
School Years	8.47	3.83	0	6	9	12	13	106122
Skill Level*	0.4	0.49	0	0	0	1	1	105971
Age	33.44	11.55	15	25	32	41	91	106890
Sex	0.3	0.46	0	0	0	1	1	106890
Urbanization	0.58	0.49	0	0	1	1	1	106890

*Skill level of occupation determined by using the O*NET database

Table B10: IFLS Wages by Age Cohorts

		Cohort								Cohort					
	Stats	15 - 24	25 - 34	35 - 44	45 - 54	55 - 64	65+		Stats	15 - 24	25 - 34	35 - 44	45 - 54	55 - 64	65+
Hours/week	count	45802	66364	55786	36739	21136	9630	Skill Level	count	43967	63909	53493	35120	20300	9311
	mean	46.1	43.16	42.24	40.36	37.85	33.91		mean	0.38	0.32	0.29	0.24	0.17	0.14
	std	21.26	21.76	22.21	21.61	21.42	20.64		std	0.49	0.47	0.45	0.43	0.38	0.35
	min	1	1	1	1	1	1		min	0	0	0	0	0	0
	25%	35	29	28	25.5	21	18		25%	0	0	0	0	0	0
	50%	48	42	42	40	35	35		50%	0	0	0	0	0	0
	75%	56	56	56	51	49	48		75%	1	1	1	0	0	0
	max	168	168	168	168	168	168		max	1	1	1	1	1	1
Weeks/year	count	45801	66364	55786	36739	21136	9630	Age	count	45802	66364	55786	36739	21136	9630
	mean	40.3	42.44	42.77	42.45	42.02	42.65		mean	20.26	29.53	39.18	49.14	58.83	70.18
	std	15.73	14.11	13.79	13.94	14.2	13.76		std	2.67	2.84	2.85	2.86	2.83	4.87
	min	0	0	0	0	0	0		min	15	25	35	45	55	65
	25%	31	40	40	40	36	40		25%	18	27	37	47	56	66
	50%	50	50	50	50	50	50		50%	20	30	39	49	58	69
	75%	52	52	52	52	52	52		75%	23	32	42	52	61	73
	max	52	52	52	52	52	52		max	24	34	44	54	64	98
Wage/hr (INT$)	count	45802	66364	55786	36739	21136	9630	Sex	count	45793	66363	55786	36739	21136	9630
	mean	2.44	3.94	8.88	5.14	5.59	3.35		mean	0.4	0.34	0.34	0.34	0.35	0.32
	std	32.16	66.69	674.81	47.94	150.68	38.95		std	0.49	0.47	0.47	0.48	0.48	0.47
	min	0	0	0	0	0	0		min	0	0	0	0	0	0
	25%	0.41	0.55	0.52	0.45	0.32	0.26		25%	0	0	0	0	0	0
	50%	0.94	1.24	1.27	1.14	0.81	0.63		50%	0	0	0	0	0	0
	75%	1.87	2.64	2.94	2.93	2.05	1.55		75%	1	1	1	1	1	1
	max	4981.9	14156.38	135595.86	3319.84	16479.2	3295.84		max	1	1	1	1	1	1
School Years	count	44086	63925	53671	35266	20108	9125	Urban	count	45648	66218	55642	36533	20898	9501
	mean	8.61	8.14	6.82	5.69	3.85	2.55		mean	0.4	0.39	0.38	0.36	0.31	0.29
	std	3.6	4.2	4.34	4.38	3.99	3.25		std	0.49	0.49	0.49	0.48	0.46	0.45
	min	0	0	0	0	0	0		min	0	0	0	0	0	0
	25%	6	6	3	2	0	0		25%	0	0	0	0	0	0
	50%	9	9	6	6	3	1		50%	0	0	0	0	0	0
	75%	12	12	12	9	6	6		75%	1	1	1	1	1	1
	max	13	13	13	13	13	13		max	1	1	1	1	1	1

Table B11: Provincial Median Wages and Average Schooling Years

		Total Sample		Low Skilled*		High Skilled*	
Island Code	Provincial Code	Median Wage/hr (INT$)	School Years Average	Median Wage/hr (INT$)	School Years Average	Median Wage/hr (INT$)	School Years Average
1	11	1.89	8.81	1.57	6.17	3.05	12
	12	1.43	7.67	1.25	5.77	2.27	12
	13	1.47	7.66	1.23	5.71	2.57	12
	14	2.21	9.15	1.81	6.49	3.07	12
	15	1.87	8.00	1.63	5.97	2.91	12
	16	1.18	6.83	0.98	5.28	2.55	12
	17	1.87	8.86	1.28	6.19	3.18	12
	18	0.88	5.85	0.78	4.73	1.89	12
3	31	1.99	8.87	1.51	6.36	3.10	12
	32	1.33	6.59	1.10	4.92	2.65	12
	33	0.96	6.01	0.85	4.77	2.05	12
	34	0.82	6.66	0.67	4.80	1.84	12
	35	1.02	6.06	0.90	4.64	2.05	12
5	51	1.15	5.99	0.96	4.03	2.32	12
	52	0.76	5.00	0.68	3.71	2.09	12
	53	1.49	8.57	0.79	5.98	3.11	12
6	61	1.28	7.16	1.12	4.30	1.58	12
	62	1.71	7.53	1.07	5.85	3.11	12
	63	1.21	5.71	1.08	4.34	2.17	12
	64	1.79	8.90	1.38	6.18	2.83	12
7	71	1.39	10.83	0.91	6.75	1.45	12
	72	1.27	6.60	1.24	6.00	3.49	12
	73	0.88	5.71	0.77	4.33	2.00	12
8	81	1.05	7.73	0.87	5.89	2.14	12
9	91	1.91	7.05	1.79	5.10	2.55	12

*This partitioning is useful for Chapter 3, where the data in table C3 comes from the above. Low Skilled are those with less than 12 years of pre-tertiary education; High skilled are those with 12 years.

The data used to construct the median wages comes from all 5 waves of the IFLS using only wages of the survey wave year; the IFLS-East data to capture the under or un-sampled Eastern provinces in the original IFLS study; and the IPUMS 1995 Indonesian Census Data. In total we have 341,409 wage observations.

Figure B4: Median Wages Across Provincial/Island Groupings and Skills

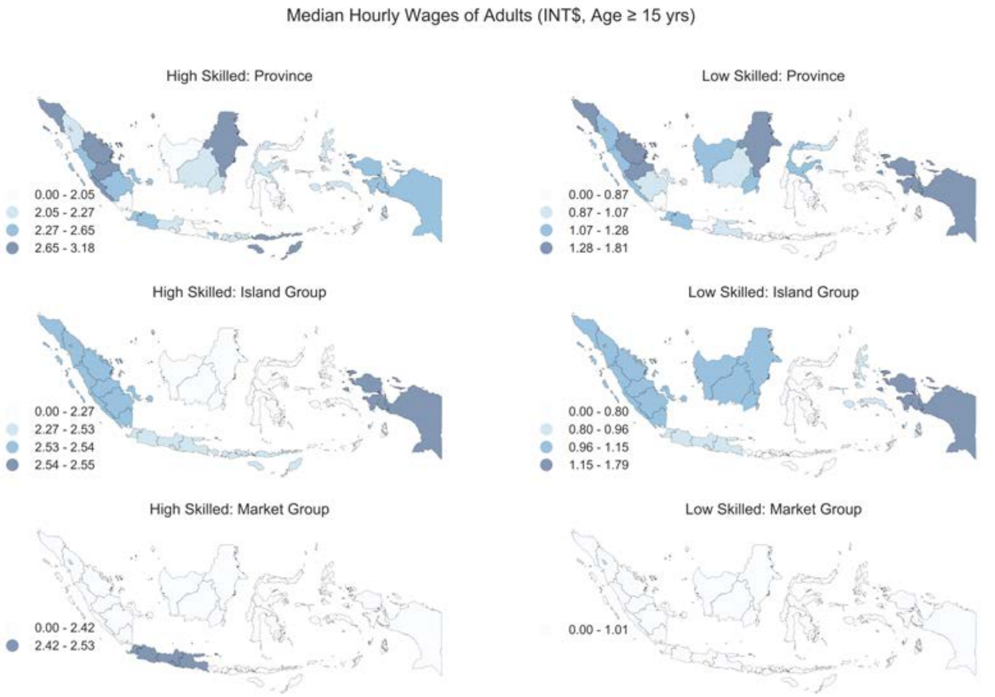


Figure B5: Provincial Wages and Gross Regional Product (as Bubbles)

Migration

Table B12: Adult Migration Events and Shares in IFLS (1980 - 2014)

Migration Characteristics		Any Migration		Inter-Provincial Migration		Inter-Island Migration	
		mean	std	mean	std	mean	std
Migrant Shares	Whole Sample	0.45	0.50	0.18	0.38	0.04	0.21
	Among Migrants			0.36	0.48	0.12	0.30
Moves per movers		1.82	1.03	2.28	1.15	2.90	1.43
	Observations	24,452		10,816		2,980	

Table B13: Adult Cohort Migration Events in IFLS (1980 - 2014)

	Any Migration		Inter-Provincial Migration		Inter-Island Migration	
Cohorts	mean	std	mean	std	mean	std
15-24	1.90	1.13	2.29	1.28	2.65	1.33
25-34	1.67	1.02	2.07	1.28	2.82	1.40
35-44	1.44	0.81	1.74	1.06	2.54	1.43
45-54	1.36	0.76	1.66	0.94	2.36	1.19
55-64	1.28	0.59	1.38	0.68	2.25	0.85
65 +	1.36	0.66	1.62	0.81	2.50	0.85
Wt. Avg:	1.73	1.05	2.13	1.26	2.67	1.35
Observations	30,302		11,610		3,201	

Observations may contain repeats of individuals across (but not within) cohorts, as individuals may migrate at different stages in their life; and individuals may repeat across migration types.

Figure B6: Average Migration Events across Cohorts

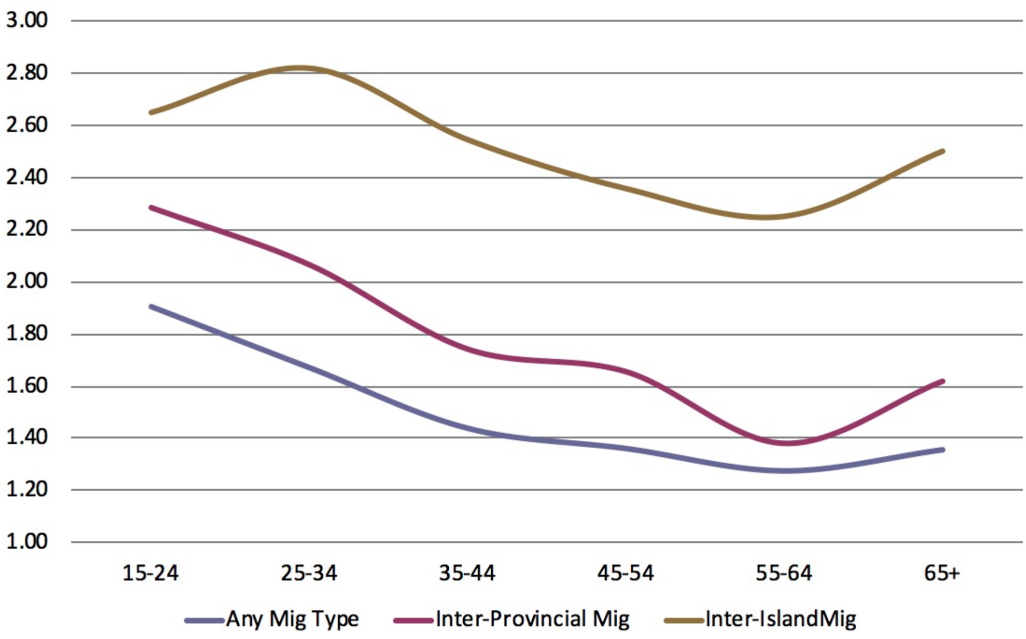


Table B14: Average Yearly Migration Rates: All IFLS (1980 - 2014)

Migration Type	mean	std	total moves
Provincial	0.0151	0.1219	24,660
Inter-Island	0.0053	0.0728	8,655
Intra-Island	0.0090	0.0945	14,698

Means are proportions

Figure B7: Yearly Provincial Migration Rates 1980 - 2014

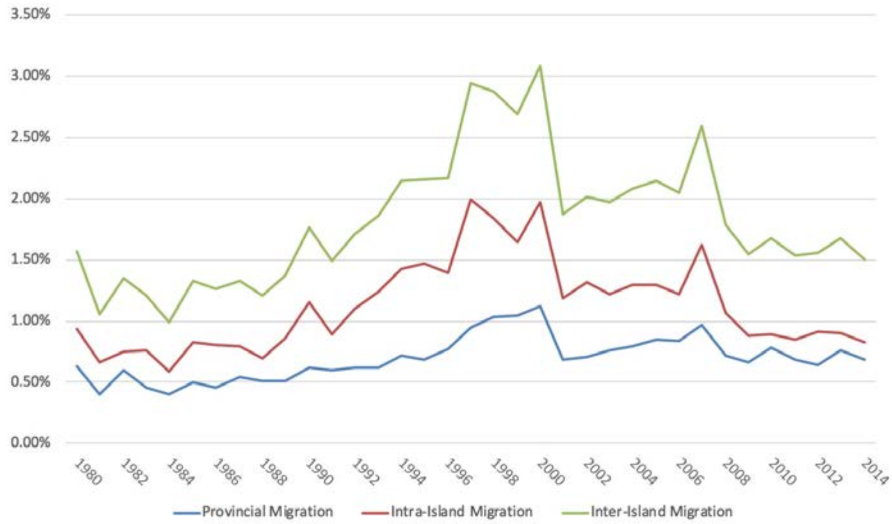


Table B15: Share of Repeat Moves (>1 move) Among All Migrants in IFLS (1980 - 2014)

	Any Migration		Inter-Provincial Migration		Inter-Island Migration	
Cohorts	mean	std	mean	std	mean	std
15-24	0.53	0.50	0.48	0.50	0.43	0.50
25-34	0.42	0.49	0.36	0.48	0.38	0.48
35-44	0.31	0.46	0.31	0.46	0.36	0.48
45-54	0.25	0.43	0.30	0.46	0.39	0.49
55-64	0.21	0.41	0.17	0.38	0.22	0.42
65+	0.27	0.44	0.22	0.41	0.21	0.41
Wt.Avg.	0.45	0.50	0.42	0.49	0.40	0.49
Observations	40,829		11,610		4,665	

Observations may contain repeats of individuals across (but not within) cohorts, as individuals may migrate at different stages in their life; and individuals may repeat across migration types.

B.5 Results of Empirical Analysis

Cross-Sectional Wage Analysis

Table B16: Mincer Regressions of IPUMS Wages - 1995 Indonesian Census

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.

Table B17: Mincer Regression of Adult IFLS Wages

ln (hourly wage)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Island Migration			0.160	0.149	0.116	0.109				0.162	0.145	0.115
			(0.039)	(0.039)	(0.039)	(0.039)				(0.039)	(0.039)	(0.039)
Within Island Migration			0.113	0.091	0.094	0.094				0.133	0.099	0.101
			(0.023)	(0.023)	(0.023)	(0.023)				(0.023)	(0.023)	(0.023)
Education	0.100	0.076	0.100	0.076	0.075	0.075						
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)						
SchLvl = 1, Primary							0.309	0.261	0.248	0.309	0.261	0.248
							(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
SchLvl = 2, Obl Secondary							0.650	0.530	0.510	0.648	0.529	0.509
							(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
SchLvl = 3, Non-Obl Secondary							0.982	0.797	0.777	0.980	0.795	0.775
							(0.018)	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)
SchLvl = 4, College							1.507	1.221	1.206	1.505	1.220	1.205
							(0.019)	(0.022)	(0.022)	(0.019)	(0.022)	(0.022)
Age	0.040	0.035	0.041	0.036	0.036	0.036	0.034	0.032	0.032	0.035	0.033	0.033
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age sq.	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sex	-0.256	-0.261	-0.255	-0.260	-0.260	-0.260	-0.287	-0.276	-0.277	-0.287	-0.276	-0.277
	(0.008)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)
Urbanization	0.180	0.115	0.181	0.116	0.119	0.121	0.172	0.103	0.106	0.174	0.105	0.107
	(0.007)	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)
Religion	0.092	0.091	0.092	0.091	0.087	0.088	0.074	0.082	0.085	0.073	0.081	0.084
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)
Ethnicity	0.007	0.006	0.007	0.006	0.005	0.005	0.006	0.005	0.004	0.006	0.005	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Year FE	YES	YES	YES	YES	YES		YES	YES	YES	YES	YES	YES
Occupation FE		YES		YES	YES	YES		YES	YES		YES	YES
Island FE					YES				YES			YES
Year*Island FE						YES						
Observations	201,530	201,429	201,530	201,429	201,429	201,429	201,530	201,429	201,429	201,530	201,429	201,429
R-squared	0.258	0.275	0.258	0.275	0.276	0.278	0.262	0.277	0.278	0.263	0.277	0.278

Robust standard errors in parentheses

Fixed-Effect Wage Analysis

Table B18: FE Mincer Regressions of IFLS Wages

ln (hourly wage)	(1)	(2)	(3)
Island Migration	0.146	0.119	0.122
	(0.042)	(0.042)	(0.042)
Within Island Migration	0.059	0.055	0.056
	(0.026)	(0.026)	(0.026)
Age	0.103	0.103	0.132
	(0.006)	(0.013)	(0.017)
Age sq.	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Urbanization	0.031	0.017	0.015
	(0.022)	(0.022)	(0.022)
Year FE	YES		
Occupation FE		YES	
Year*Island FE		YES	YES
Occupation*Island FE			YES
Observations	232,201	232,201	232,201
R-squared	0.223	0.230	0.232

Robust standard errors in parentheses.

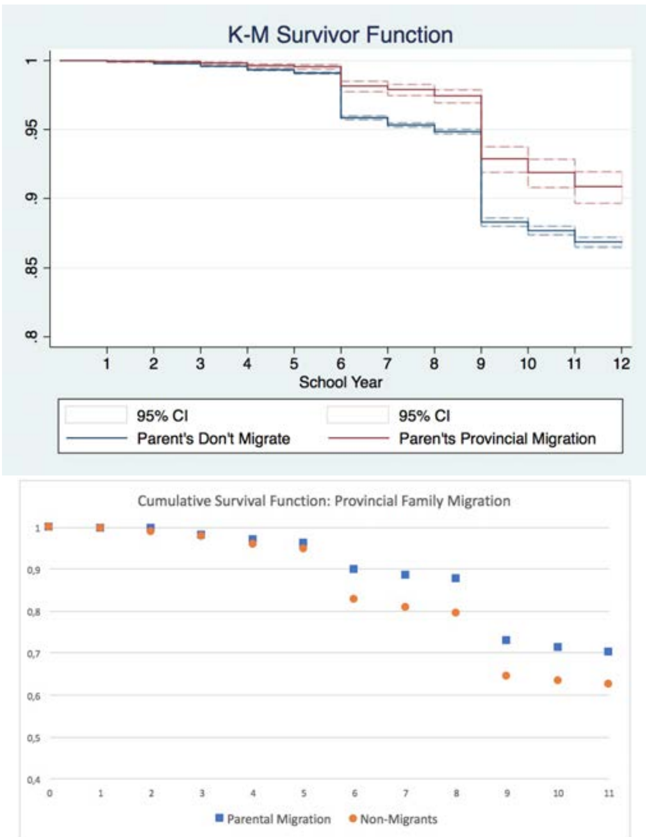
Duration Analysis

Table B19: Duration Results

PROBIT: School Exit	(1)	(2)	(3)	(4)	(5)
Family Migration	-5.573	-4.992	-5.012	-3.422	-2.148
	(0.0221)	(0.092)	(0.0751)	(0.178)	(0.118)
Baseline Hazard	X	X	X	X	X
Individual Controls		X			X
Parent's Schooling			X		X
Child Schooling Controls				X	X
Observations	96,450	87,886	95,776	87,944	86,341

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. 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

Figure B8: Survival Function: Parent's Migration



C Chapter 3 Appendices

C.1 Structural Dynastic Model of Migration and Educational Investment

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 inter-temporal 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 we 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.

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 we forgo modeling the quantity-quality tradeoff that would also require modeling 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) We 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 $z_t = (\ell_{t-1}, \ell_t, a_0, a_t, X_t)$.

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 we 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 $u_t(c_t, z_t, \varepsilon_t)$. For simplicity we assume that agents are inelastic with respect to their labor supply and do not model this decision. As such we do not include, as is typically the case in the literature, the decision $I_{k,t}$ as a proxy for leisure. The main reason we 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 education 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 we will also assume that the utility is additively separable in the observable and unobservable states of the decisions taken, so that $u_t(c_t, z_t, \varepsilon_t) = u_t(c_t, z_t) + \varepsilon_t$.[52]

Budget Constraint The agent's per period budget constraint is given by

[52] To include leisure we 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.

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 incentivize workers to reallocate their labor supply through migration.

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

Value Functions

We 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 *ex ante* 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) We 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. We 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:

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

Below we 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. We 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.

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

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 we 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[53] leads to the following closed-form results: the first is that the above CCPs take a logit form and is given by

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

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

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

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

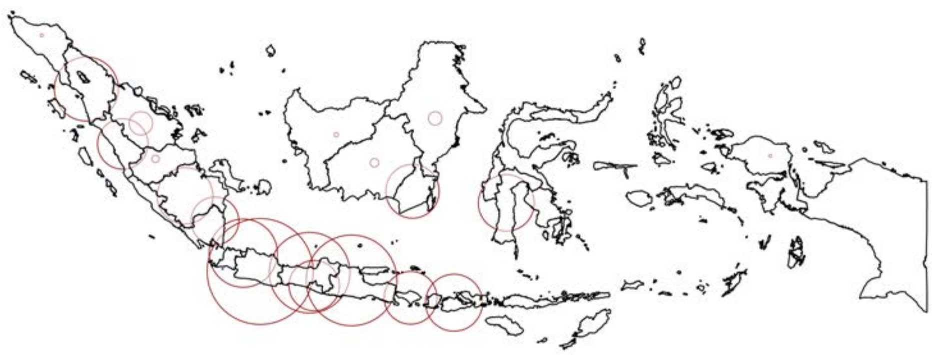
where gamma is Euler's constant. And inserting the closed-form CCP into the above equation implies the following result:

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

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 follows McFadden (1973) and Rust (1987), and Gayle et al. (015b) for the dynastic estimation of the terminal value function, through backward induction.

C.2 Descriptive Statistics and Figures

Figure C1: Locations of 3645 Identified Households



The location of the households as of their last observed wave in the IFLS. The size of the bubble indicates the relative size of the density of households in the province.

Table C1: Descriptive Statistics of Parents' Labor History (First-Born Child born between 1980 and 1997)

Parents	Don't Migrate		Migrate (Provincial)	
	mean	std	mean	std
Father				
Age Enter First Job	13.43	2.72	15.14	2.43
Labor Market Participation				
hrs/week	46.47	22.22	48.5	22.34
weeks/year	40.91	14.86	42.34	13.81
Average wage/hr (2000 INT$)				
Skill=Low	1.11	3.25	1.56	3.16
Skill=High	1.21	2.81	1.65	2.75
% Entering High Skilled Job*	0.32	0.47	0.35	0.48
Mother				
Age Enter First Job	12.88	2.21	14.26	2.16
Labor Market Participation				
hrs/week	46.92	22.84	48.51	23.072
weeks/year	39.14	15.84	39.77	15.44
Average wage/hr (2000 INT$)				
Skill=Low	1.07	3.12	1.36	2.84
Skill=High	1.11	2.91	1.58	3.31
% Entering High Skilled Job*	0.33	0.47	0.34	0.47

*This variable is measuring the proportion whose first job is high skilled, as determined by using the O*NET database and described in footnote 33.

Table C2: Descriptive Statistics of First-Born Children's Labor History (born 1980 - 1997)

Children	**Family Don't Migrate**		**Family Migrate (Provincial)**	
	mean	std	mean	std
Parental Rate of Schooling Investment (Up to Grade 12=1)	0.52	0.49	0.66	0.48
Age of First Job	16.62	3.46	17.16	3.08
% Entering High Skilled Job*	0.26	0.44	0.41	0.46
Average wage/hr (2000 INT$) First Job				
Skill=Low	1.77	2.9	1.99	2.38
Skill=High	1.82	3.26	2.12	3.05
Labor Market Participation				
hrs/week	46.14	22.47	49.05	22.42
weeks/year	39.6	15.58	40.62	14.92

*This variable is measuring the proportion of children whose first job is high skilled, as determined by using the O*NET database and described in footnote 33.

Table C3: Market-Grouped Median Hourly Wage

Market	Name	Skill Level	Wage/hr (INT$)	Normalized Wage
1	Everywhere Else	Low Skilled	1.01	1.047
		High Skilled	2.42	2.513
2	Jawa	Low Skilled	0.96	1.000
		High Skilled	2.53	2.626

Table C4: Family Migration Events

	Number of Events		Rate		Total Events
Family Migration Type	mean	std	mean	std	
Provincial	1.51	0.91	0.065	0.24	237
Jawa <–>Everywhere Else	1.66	0.93	0.0173	0.13	63

C.3 Tables and Figures

Estimation Results

Table C5: Estimates of the Structural Parameters

Parameter	Non-Nested Model	Nested Logit Model
δ (moving cost)	3.436	1.802
	(0.329)	(0.072)
$\phi_{1=\text{Everywhere Else}}$	0.629	0.598
	(0.031)	(0.029)
$\phi_{2=\text{Jawa}}$	0.619	0.562
	(0.025)	(0.021)
$\rho_{1=\text{Educate Branch}}$		0.723
		(0.122)
$\rho_{2=\text{Don't Educate Branch}}$		0.317
		(0.127)
σ (scale)	1.043	
	(0.052)	
Likelihood	-3151.208	-2779.833
$\chi^2(1)$		742.75

CCP Results

Table C6: CCPs Implied by the Data

State↓ / Decision →	No Education No Migration	Education No Migration	No Education Migration	Education Migration
Low Skilled, Everywhere Else	0.512	0.474	0.006	0.008
Low Skilled, Jawa	0.536	0.455	0.003	0.006
High Skilled, Everywhere Else	0.248	0.706	0.009	0.036
High Skilled, Jawa	0.296	0.675	0.01	0.02

Table C7: CCPs Generated by the Nested Logit

State↓ / Decision →	No Education No Migration	Education No Migration	No Education Migration	Education Migration
Low Skilled, Everywhere Else	0.529	0.471	0	0
Low Skilled, Jawa	0.505	0.495	0	0
High Skilled, Everywhere Else	0.368	0.583	0.01	0.039
High Skilled, Jawa	0.354	0.628	0.006	0.011

Sensitivity Results

Table C8: Alternative Specifications Varying α

Parameter	Baseline $\alpha = 0.834$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 0.95^*$
δ (moving cost)	1.802	1.899	1.885	1.714
	(0.072)	(0.081)	(0.081)	(0.069)
$\phi_{1=\text{Everywhere Else}}$	0.598	0.455	0.566	0.638
	(0.028)	(0.037)	(0.027)	(0.025)
$\phi_{2=\text{Jawa}}$	0.562	0.399	0.525	0.637
	(0.021)	(0.028)	(0.021)	(0.018)
$\rho_{1=\text{Educate Branch}}$	0.723	0.749	0.729	0.714
	(0.122)	(0.125)	(0.124)	(0.119)
$\rho_{2=\text{Don't Educate Branch}}$	0.317	0.315	0.316	0.318
	(0.127)	(0.119)	(0.120)	(0.126)
Likelihood	-2779.833	-2799.708	-2783.976	-2775.031

*based on the altruism parameter implied by Raut and Tran (2005)

Goodness of Fit Results

Table C9: Simulation Results of Decision Rates

	Simulation	Data
Migration Rate	0.009	0.017
Educational Investment	0.512	0.526

Table C10: Simulation Results of Decision Rates - By State

State	Migration		Education		Proportion in Datasets
	Simulation	Data	Simulation	Data	
Low Skilled, Everywhere Else	0.000	0.014	0.479	0.482	0.343
Low Skilled, Jawa	0.000	0.010	0.477	0.461	0.427
High Skilled, Everywhere Else	0.045	0.045	0.642	0.742	0.091
High Skilled, Jawa	0.021	0.029	0.638	0.694	0.139

Table C11: Distribution of Next Generation's States

States of Next Generation	Simulation	Data
Low-Skilled, Everywhere Else	0.207	0.201
Low-Skilled, Java	0.278	0.273
High-Skilled, Everywhere Else	0.218	0.232
High-Skilled, Java	0.297	0.294

Table C12: Intergenerational Transition Matrix

Household State ↓ / Child State →	Low Skilled Everywhere Else		Low Skilled Jawa		High Skilled Everywhere Else		High Skilled Jawa	
	Simulation	Data	Simulation	Data	Simulation	Data	Simulation	Data
Low Skilled, Everywhere Else	0.501	0.512	0.000	0.006	0.499	0.474	0.000	0.008
Low Skilled, Jawa	0.000	0.003	0.475	0.536	0.000	0.006	0.525	0.455
High Skilled, Everywhere Else	0.367	0.248	0.015	0.009	0.588	0.706	0.03	0.036
High Skilled, Jawa	0.01	0.01	0.326	0.296	0.012	0.02	0.652	0.675

Counterfactual Results

Figure C2: Effect of Migration Subsidies on the Next Generation (Relative to Initial State of Parents)

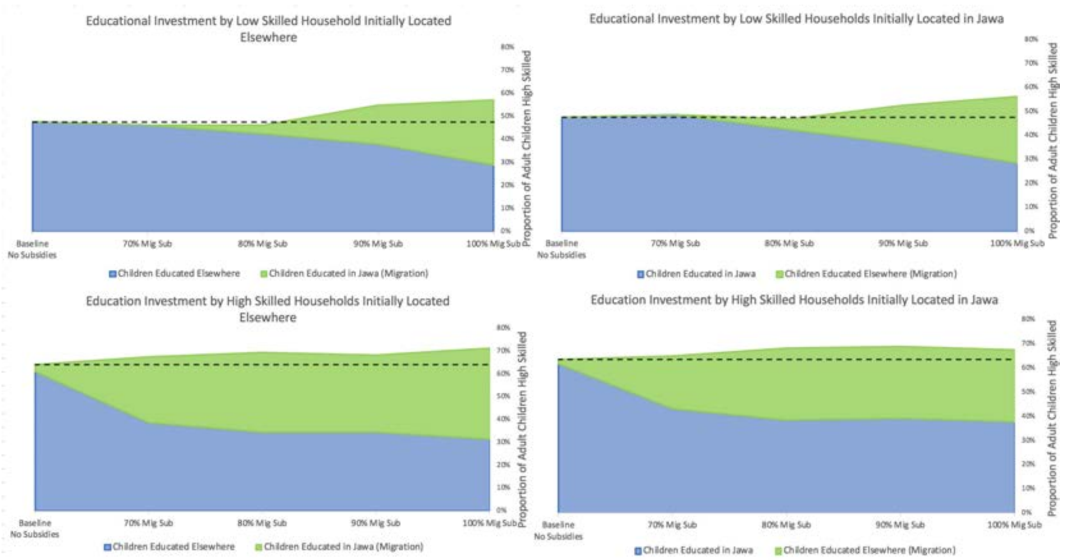


Figure C3: Effect of Education Subsidies on the Next Generation (Relative to Initial State of Parents)

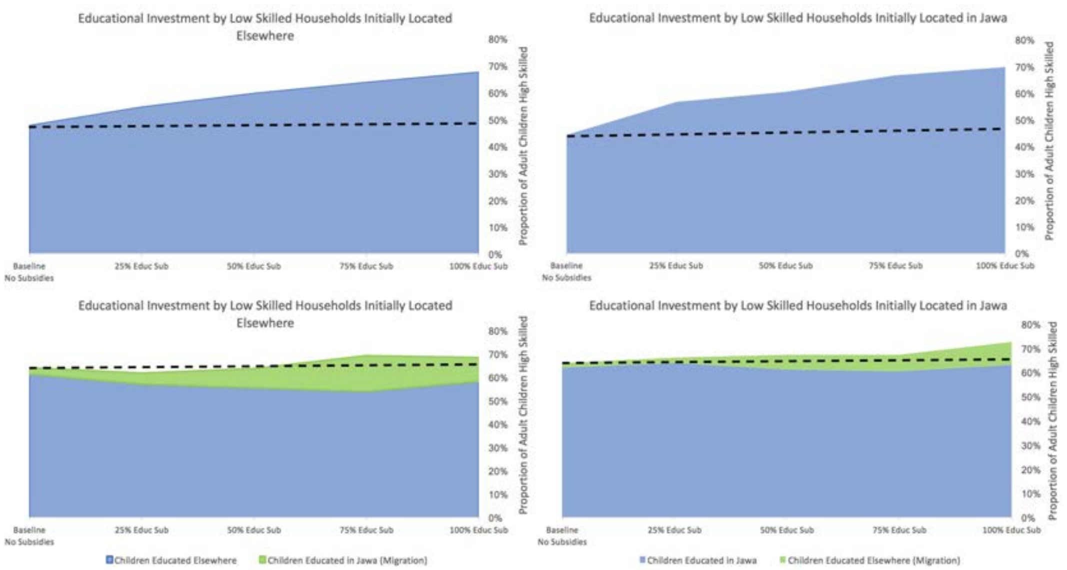


Table C13: Proportional Changes in the States of the Next Generation in Response to Subsidies

Subsidy Schedule →	Baseline No Subsidies	Changes Due to Migration Subsidy 70%	80%	90%	100%	Changes Due to Education Subsidy 25%	50%	75%	100%
Low Skilled Everywhere Else	20.7%	2.9%	4.0%	1.5%	1.9%	-1.9%	-3.8%	-6.0%	-6.4%
Low Skilled Jawa	27.8%	-3.5%	-4.2%	-8.1%	-10.0%	-4.3%	-6.1%	-8.6%	-11.3%
High Skilled Everywhere Else	21.8%	0.7%	2.2%	6.1%	7.1%	2.1%	4.2%	5.4%	8.6%
High Skilled Jawa	29.7%	0.0%	-1.9%	0.5%	1.0%	4.1%	5.7%	9.1%	9.1%
Overall High Skilled Gain		0.6%	0.3%	6.6%	8.1%	6.2%	9.9%	14.5%	17.7%

Figure C4: Welfare Impact of Subsidies on the Labor Market Outcome of Next Generation

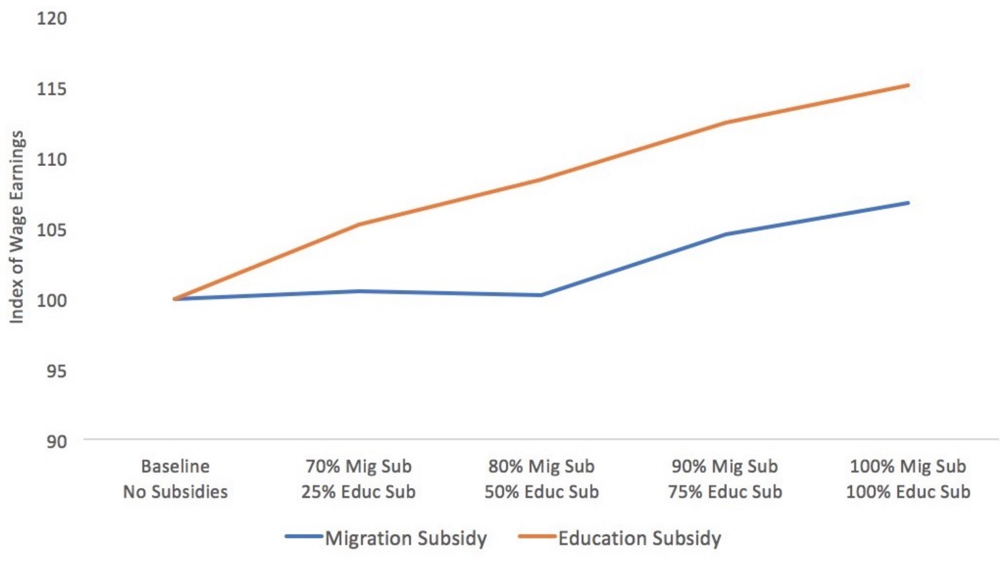


Table C14: Estimation Results: Social Mobility Regressions

Dep. Var Child Educated	Baseline	Migration Subsidy 70%	80%	90%	100%	Education Subsidy 25%	50%	75%	100%
Constant	0.482	0.478	0.482	0.521	0.555	0.555	0.611	0.647	0.683
	(0.016)	(0.009)	(0.007)	(0.003)	(0.001)	(0.010)	(0.009)	(0.006)	(0.007)
Parent Educated	0.144	0.201	0.199	0.167	0.139	0.096	0.064	0.041	0.031
	(0.002)	(0.011)	(0.005)	(0.001)	(0.002)	(0.008)	(0.006)	(0.007)	(0.001)

Standard errors in parentheses clustered at the location level.

Figure C5: Coefficients of Social Mobility Regressions

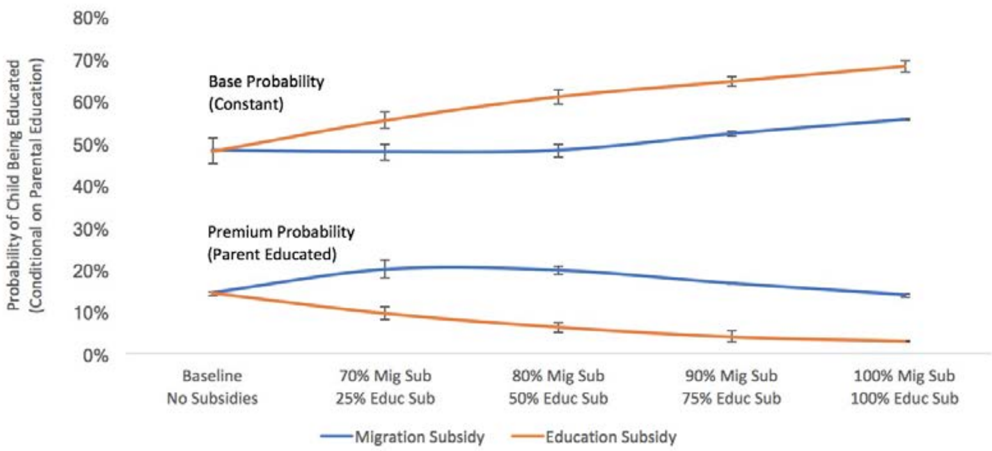
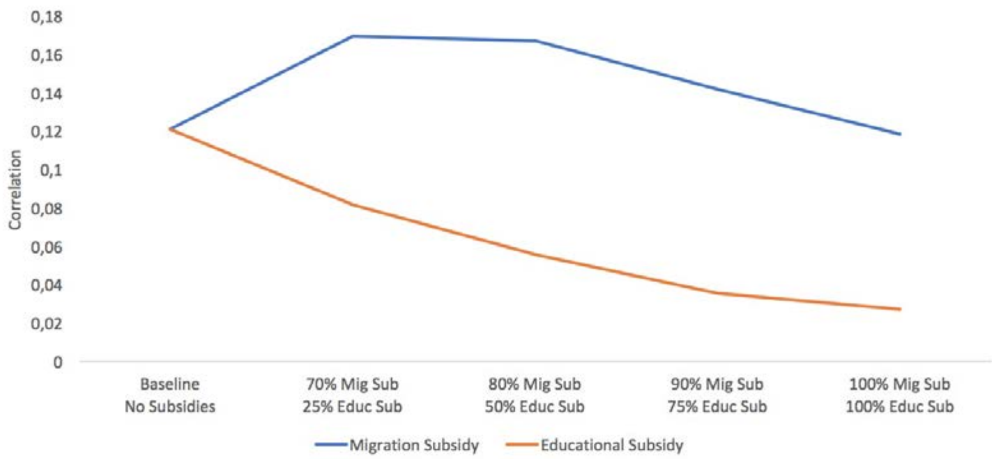


Figure C6: Effect of Subsidies on Intergenerational Correlations



The table below documents the effect on the economy of the two subsidies for the next generation (used to construct figure C4).

Table C15: Impact of the Subsidy Schedule on the Next Generation's Aggregate Median hourly Income

	Migration Subsidy INT$/hr	Education Subsidy INT$/hr	Migration Subsidy Indexed	Education Subsidy Indexed
Baseline No Subsidies	1.76	1.76	100.00	100.00
70% Mig Sub 25% Educ Sub	1.77	1.85	100.61	105.34
80% Mig Sub 50% Educ Sub	1.76	1.91	100.23	108.50
90% Mig Sub 75% Educ Sub	1.84	1.98	104.54	112.49
100% Mig Sub 100% Educ Sub	1.88	2.02	106.86	115.11

Bibliography

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.

Aycinena, D., C. Martinez, and D. Yang (2010, October). The impact of transaction fees on migrant remittances: Evidence from a field experiment among migrants from el salvador.

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.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2016, September). Skill transferability, migration, and development: Evidence from population resettlement in indonesia. *American Economic Review 106*(9), 2658 – 2698.

Becker, G. S. (1974). A theory of social interaction. *The Journal of Political Economy 82*(6), 1063–93.

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.

Blue, S. (2004). State policy, economic crisis, gender, and family ties: Determinants of family remittances to cuba. *Economic Geography 80*(1), 63–82.

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 O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (1 ed.), Volume 3, Part A, Chapter 28, pp. 1697–1760. Elsevier.

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

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.

Cox, D. and M. Fafchamps (2008). Extended family and kinship networks: Economic insights and evolutionary directions. In *Handbook of Development Economics*, Volume 4, Chapter 58. North Holland.

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.

Dilaz-Briquets, S. (2013). Remittances to cuba: An update. In *Papers and Proceedings of the Eighteenth Annual Meeting of the Association for the Study of the Cuban Economy (ASCE)*, Volume 18, pp. 154–159.

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.

Foster, A. and M. Rosenzweig (2001). Imperfect commitment, altruism, and the family: Evidence from transfer behavior in low-income rural areas. *Review of Economics and Statistics LXXXIII*(3), 389–407.

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.

Freund, C. and N. Spatafora (2008). Remittances, transaction costs, and informality. *Journal of Development Economics 86*(2), 356–366.

Gandelman, N. and R. Hernández-Mutillo (2014, October). Risk aversion at the country level. Technical Report 2014-005B, Federal Reserve Bank of St. Lous Working Paper Series.

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.

Gonzalez, M. and S. Larson (2008). Survey of cuban remittance forwarding agencies in the united states: Preliminary findings. *Cuba in Transition 18*, 293–299.

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.

Havranek, T., R. Horvath, Z. Irsova, and M. Rusnak (2013). Cross-country heterogeneity in intertemporal substitution. Working Paper 156, William Davidson Institute.

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

Hensher, D. A. (1986, April). Sequential and full information maximum likelihood estimation of a nested logit model. *The Review of Economics and Statistics 68*(4), 657–667.

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., D. J. McKenzie, G. Esquivel, and E. Schargrodsky (2005). The effects of migration on child health in mexico. *Economia 6*(1), 257–289.

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

Huggett, M. (1996). Wealth distribution in life-cycle economies. *Journal of Monetary Economics 38*(3), 469–494.

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

Kinnan, C. and R. Townsend (2012). Kinship and financial networks, formal financial access, and risk reduction. *American Economic Review 102*(3), 289–293.

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

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

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

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

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. and O. Stark (1985). Motivations to remit: Evidence from botswana. *Journal of Political Economy 93*(5), 901–918.

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, 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.

McFadden, D. L. (1977). Modeling the choice of residential location. Cowles Foundation Discussion Papers 477, Cowles Foundation for Research in Economics, Yale University.

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

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

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

Orozco, M. (2002). Worker remittances: The human face of globalization. Working paper, Inter-American Development Bank.

Orozco, M. (2009). The cuban condition: Migration, remittances, and its diaspora. Technical report, Inter-American Dialogue.

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.

Rapoport, H. and F. Docquier (2005). The economics of migrants' remittances. Discussion Paper 1531, Institute for the Study of Labor.

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.

Shimada, A. (2012). Migration decisions, expected remittances, and altruism. *International review of Economics 59*(3), 285–296.

Sikoki, B., F. Witoelar, J. Strauss, E. Meijer, and W. Suriastini (2013). Ifls east user's guide and field report. Technical report, SurveyMETER and University of Southern California.

Stark, O. (1999, April). *Altruism and Beyond*. Number 9780521663731. Cambridge University Press.

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.

Strauss, J., F. Witoelar, B. Sikoki, and A. Wattie (2016, March). The fifth wave of the indonesian family life survey (ifls4): overview and field report. Technical report, RAND WR-1143/2-NIA/NICHD.