UNIVERSITÀ DEGLI STUDI DI MILANO

GRADUATE SCHOOL IN SOCIAL, ECONOMIC AND POLITICAL SCIENCES

DEPARTMENT OF ECONOMICS

PHD IN ECONOMICS

"EMPIRICAL ESSAYS IN DEVELOPMENT ECONOMICS"

VALERIA GROPPO

SUPERVISORS: GIORGIO BARBA NAVARETTI

MASSIMILIANO BRATTI

PHD COORDINATOR: MICHELE SANTONI

A.A.
2010/2011
SECS-P/01

ACKNOWLEDGMENTS

I am grateful to Massimiliano Bratti and Giorgio Barba Navaretti for careful and supportive supervision. I also express my gratitude to Mariapia Mendola and Sonia Bhalotra for precious suggestions and insightful discussions. I also thank Christine Peterson (RAND) for many important clarifications on the Indonesia Family Life Survey, and Juan Miguel Gallego, Francesca Modena and participants to the Third Italian Doctoral Workshop in Economics and Policy Analysis (Turin, 8-9 July 2010), for helpful comments.

CONTENTS

Chapter 1. Motivations and Thesis' Structure

Households in developing countries are characterized by low and volatile incomes. High income volatility is due, for instance, to the variability of weather conditions and commodity prices, as well as to households' exposure to economic shocks, such as illnesses, unemployment, and business failures. Developing settings are also characterized by a lack of formal insurance, savings and credit institutions, which would help households to increase their consumption and reduce risk. Households respond to these conditions by relying on various informal mechanisms, which increase their capacity to save and cope with risk.

As outlined by Morduch (1995), households can rely on informal mechanisms to cope with risk at various stages. They can indeed adopt *ex-ante* behaviors which help them to smooth income, protecting them from the occurrence of economic shocks. For instance, households' diversification into different economic activities represents an income smoothing strategy which is commonly used by households in developing countries. However, households can also undertake *ex-post* behaviors which would help them to smooth their consumption after the occurrence of negative economic shocks. Assets' sales, borrowing and savings are examples of *ex-post* behaviors used by households to smooth consumption in developing countries.

Given the important role that informal savings and insurance mechanisms play in developing settings, it is crucial to understand in which conditions individuals and households are more likely to rely on them, and which effects do these strategies have on key economic variables, such as education and expenditure.

This dissertation focuses on two behaviors, which can be used by individuals and households in poor settings to overcome market imperfections, thus increasing savings and reduce risk.

In particular, chapter 2 considers individuals' participation in rotating savings and credit associations (roscas), analyzing its determinants, and its consequences in terms of children's schooling. Roscas are frequently found in developing countries, and represent mostly informal associations, whose members meet regularly contributing money to a collective 'pot', which is assigned, usually in rotation, to each group member (Anderson and Baland, 2002). Roscas can ease the accumulation of savings and help their members to cope with adverse economic shocks, through the social capital they embed. In my analysis, I focus in particular on the Indonesian context, where participation in *arisans* (the Indonesian roscas) is relatively common. My study aims at evaluating the consequences of maternal participation in *arisans* on their children's secondary or tertiary school enrolment. Employing bivariate probit techniques in order to deal with the endogeneity problems associated with mothers' participation in *arisan*s, I estimate that maternal participation significantly increases the probability of children's school enrolment. The bivariate probit models, furthermore, provide evidence of a negative correlation between the unobservables determining rosca participation and the unobservables determining schooling. I interpret this result as evidence that in households that are more vulnerable to shocks, mothers invest less in their children's education, while increasing their partecipation in roscas, as an insurance and savings device.

Chapters 3 and 4, instead, focus on the role on individuals' and households' mobility as a strategy to overcome various market failures in developing settings.[1] In particular, these two chapters consider *internal* migration in Indonesia, highlighting its main determinants, and its consequences on consumption growth of household members who remain at origin.

In chapter 3 the determinants of internal migration in Indonesia are analyzed. As for chapter 2, also in this study I mainly follow an empirical approach, using the second and third waves of the Indonesia Family Life Survey dataset, corresponding to years 1997 and 2000 (IFLS2-3). I estimate individuals' migration between the two years, as a function of individual and household characteristics in 1997, before the considered migration episodes. The main contribution of the study is the distinction between 'individual migration' and 'family migration', where the first relates to movements which individuals undertake alone, while the latter indicates movements they undertake with their families. I estimate a Multinomial Probit Model of the determinants of individual migration, where the individuals' choice set entails three alternatives: not moving, moving alone, moving with the family. The estimates, performed on men and women samples separately, highlight that individual and family migration have very distinct determinants. For instance, an increase in household assets (proxied by the value of land owned) does not exert a significant effect on the probability of individual migration, while having a significant negative effect on the probability of family moves. Working status has a significant effect mainly on individual migration: individuals who are not working at the baseline are more likely to move alone, *ceteris paribus*. These results underline the importance of individual migration as a means to reach job and education opportunities that are not available at origin.

Moreover, for individual men's migration, we found evidence of lower mobility at intermediate levels of the household income distribution. As shown by Banerjee and Newman (1998), this might be related to the fact that individuals whose household has an intermediate income have a lower incentive to move, compared to individuals in households at the extremes of the income distribution. Indeed, intermediate levels of income at the village-level are likely to give access to a set of village-based informal insurance mechanisms, which would not be available in the city. Therefore, the poorest and the richest would be more likely to move, given that they are less affected by the absence of informal insurance mechanisms in urban areas. This effect might create poverty traps, for those individuals at intermediate household income levels who decide not to move, even when having economically advantageous job opportunities at destination. This effect underlines the importance of the presence of insurance mechanisms, both at origin and at destination, confirming the strong link between migration and insurance market imperfections, which is one of the key insights of the New Economics of Labour Migration.

The final chapter (joint with Mariapia Mendola, Università degli Studi di Milano Bicocca) studies the impact of internal migration on the consumption growth of origin households. We use IFLS2 and IFLS3 and we estimate whether households having a migrant in the period 1997-2000 have a different consumption

[1] The models of Todaro (1969) and Harris and Todaro (1970) explain migration by referring in particular to labour market imperfections, while the New Economics of Labour Migration mainly focus on the role of migration as a strategy to overcome capital and insurance markets' imperfections (Stark and Levhari, 1982; Stark and Bloom, 1985; Rosenzweig and Stark, 1989).

growth in the same period, compared to households not having migrants. The main empirical challenge of the analysis is the endogeneity of migration. We use a difference-in-difference (DID) specification, which removes all sources of time-constant unobserved heterogeneity. Yet, if unobserved characteristics shaping both the migration process and the living conditions of remaining household members vary in time and space (e.g. ambition), DID estimates will be still biased and inconsistent. Thus, we also use an instrumental variable method, as to take time-varying unobserved heterogeneity into account. The main contribution of the paper lies in the fact that we are able to disentangle the impact of current and return migration. Our results show that having a current migrant in the period 1997-2000 does not exert a statistically significant impact on origin households' consumption growth in the same time span, while having a return migrant significantly reduces per capita consumption growth. The data provide evidence that returning migrants have a significantly lower income in 2000 compared to non-migrants and current migrants. Thus, the negative effect of return migration is likely to be explained by a lower income potential of returning migrants with respect to current migrants and non-migrants.

Overall, while finding evidence of the importance of informal strategies used by households to overcome market imperfections, our results also point to the need to strengthen formal mechanisms which could be used by households in developing settings to reduce risk and increase their income and consumption opportunities. As suggested by Morduch (1999), some important policies in this regard are public programs for savings' promotion, micro-credit programs, crop insurance and employment guarantee schemes.

Bibliography

Anderson, S., and Baland, J.-M. (2002). The Economics of Roscas and Intra-Household Resource Allocation. *The Quarterly Journal of Economics*, 117(3), 963-995.

Banerjee, A., and Newman, A. (1998). Information, the Dual Economy, and Development. *Review of Economic Studies*, 65(4), 631-654.

Fafchamps, M., and Lund, S. (2003). Risk-sharing networks in rural Philippines. *Journal of Development Economics*, 71, 261-287.

Harris, J. R., and Todaro, M. P. (1970). Migration, Unemployment and Development: A Two-Sector Analysis. *The American Economic Review*, 60(1), 126-142.

Morduch, J. (1995). Income Smoothing and Consumption Smoothing. *The Journal of Economic Perspectives*, 9(3), 103-114.

Morduch, J. (1999). Between the State and the Market: Can Informal Insurance Patch the Safety Net? *The World Bank Research Observer*, 14(2), 187-207.

Rosenzweig, M. R., and Stark, O. (1989). Consumption Smoothing, Migration and Marriage: Evidence from Rural India. *The Journal of Political Economy*, 97(4), 905-926.

Stark, O., and Bloom, D. E. (1985). The New Economics of Labor Migration. *The American Economic Review*, 75(2), 173-178.

Stark, O., and Levhari, D. (1982). On Migration and Risk in LDCs. *Economic Development and Cultural Change*, 31(1), 191-196.

Todaro, M. P. (1969). A Model of Labor Migration and Urban Unemployment in Less Developed Countries. *The American Economic Review*, 59(1), 138-148.

Chapter 2. Mothers' Participation in Roscas and Children's Schooling: Evidence from Indonesia

Valeria Groppo
Università degli Studi di Milano

1. Introduction and motivations

This paper analyses the effect of maternal participation in rotating savings and credit associations (roscas) on children's education in Indonesia. Roscas are one of the most common forms of informal financial institutions in developing countries and typically work as follows: a group of individuals meet regularly and, at each meeting, each individual contributes a certain amount into a collective 'pot'. The pot is then allocated to one member, by drawing lots or bidding. At future meetings, each member is still required to contribute, but past winners are excluded from receiving the pot. The meeting process is repeated until each participant receives the pot. Underlining that the main elements of roscas are *regularity* and *rotation*, Ardener (1964) defines roscas as 'associations formed upon a core of participants who agree to make regular contributions to a fund, which is given […] to each contributor in rotation' (1964: 201).[2]

Roscas are spread in developing settings: for instance, they are found in many African countries (Nigeria, Sierra Leone, Ghana, Cameroon and Uganda among others) and in Asian ones (India, Indonesia, Taiwan, Vietnam).[3] In this study, I examine the relationship between schooling and mothers' participation in *arisan*, the Indonesian rosca.

I focus on Indonesia mainly for two reasons: first of all, roscas represent the most common forms of group participation for Indonesian women: according to the Indonesia Family Life Survey (IFLS) 1997 and 2000, about 40% of Indonesian women participated in *arisans* during the year prior to the survey. Second, Indonesia is characterized by a historical tradition of mutual cooperation, which is at the basis of the formation of groups by individuals of a community (Geertz, 1962).

Women participate in *arisans* both for economic and social reasons. Among the economic reasons, the savings and the insurance motives are prevalent: roscas are often a way to purchase durable goods, or to pay for education or wedding expenses. These groups, furthermore, can help households to cope with the economic losses following illnessesses or natural disasters. Roscas may be also used because they provide higher returns compared to other savings devices. Besides economic motivations, *arisans* also constitute an important channel of social contact and information exchange. *Arisan* participation, hence, can be considered

[2] The actual functioning of roscas present a wide variety, not only across communities, but often within a single community: contributions can be fixed or can vary, they can be in cash or in kind. Sometimes interest rates are paid.
[3] For a qualitative comparative analysis of roscas in different countries, see Ardener (1964).

as a form of 'social capital', and we may say that this paper also contributes to the literature on the effects of social capital in developing countries.[4]

Maternal social capital can influence children's education in different ways: first of all, in a developing country, where credit and insurance markets are incomplete or nonexistent, participation in *arisans* may represent a risk-coping mechanism: hence, in the presence of unanticipated shocks, participating in a group may help households to smooth their consumption and thus also schooling expenditures.[5]

Second, roscas can ease the accumulation of savings, thanks to which households can afford to pay for durable goods, and for education expenditures as well.

Third, as highlighted by Baland and Robinson (2002), participation in roscas may be a way for women to gain resources, which they can decide how to spend, without being forced by their husbands or other family members in this decision. In a setting like Indonesia, in fact, women are often excluded or marginalized in the decision-making process about household expenditure, and *arisans* may provide them the means to have a 'say' in the household. These means can be either monetary or informational resources. Furthermore, women can also acquire confidence and independence from participation in roscas. Once additional resources are acquired, they can be spent for children's welfare. There is significant evidence, in fact, that the pattern of household resources' allocation differs depending on who is the income provider. For example, Pitt and Khandker (1998) showed that in rural Bangladesh credit extended to women increases household expenditure for food and working tools more than credit granted to men. Thomas (1990) studies intra-household resources allocation in Brazil and concludes that additional non-labour income in the hands of mothers has a greater impact on children's health than additional non-labour income under fathers' control.[6] Another channel through which maternal participation in roscas could influence children's education is the information exchange that takes place within community groups: through this 'informational channel', women could acquire more power in the household and/or could gain better acknowledgment of the importance of education. These mechanisms could imply an increase in the percentage of households' expenditures on education. Following the monetary and informational channels, thus, greater maternal participation in *arisans* would lead to better children's education.

However, involvement in community activities also requires a contribution of time and monetary resources from participating members: hence, when mothers participate in community activities, they could

[4] Throughout the paper, the expressions 'social capital', rosca and *arisan* participation will be used as synonimous. The expression 'social capital' is used in the literature to indicate that particular kind of asset which can be constituted and accumulated through social interaction. For this interaction to generate capital, a sufficient degree of trust among individuals is needed. Trust, in fact, eases the provision of private and public goods in the system, by lowering transaction costs. Lower transaction costs mean that resources have been freed in the system, which can be either accumulated or invested. Social capital is a multifaceted concept, which can be defined emphasizing different dimensions, e.g. Putnam *et al.* (1993) define social capital as features of social organizations such as trust, norms and network that can improve the efficiency of a society, while Granovetter (1985) focuses on the asset characterization, defining social capital as monetary or informational resources deriving from knowledge of others and interaction with them.

[5] Indeed, although there might be mismatches between the time in which an individual is supposed to receive the 'pot'and the time in which economic shocks happens, individuals may be given the 'pot' in advance with respect to the due timing, in case they are hit by unexpected shocks (Ardener and Burman, 1995).

[6] Strauss and Beegle (1996) provided a survey of the evidence on gender differences in resources' allocation.

reduce the amount of money the household can invest in education, and/or they might divert time inputs from child's supervision.

On the basis of the above reasoning, thus, it is *a priori* unclear whether a greater maternal involvement in roscas would increase children's education.

I investigate this issue through an econometric analysis of household survey data: in particular, I utilize the second and third waves of IFLS (IFLS2-3), referring to years 1997 and 2000 respectively, and containing information both on children's education and on adults' participation in community activities.

The main econometric issue encountered in the estimation is the potential endogeneity of maternal participation in *arisans*: in fact, omitted maternal characteristics influencing children's education may be correlated with participation in roscas, thus biasing the estimate of the coefficient of interest. Estimates could also be affected by reverse causality: mothers can decide whether to participate in *arisans* in part as a function of the amount of resources they desire to invest in children's education.

I take into account the possible endogeneity of participation in *arisans* through the following strategy: first, I include in the paper an analysis of the determinants of social capital formation, in order to understand which variables have to be included in the equation of interest, to minimize omitted variable bias. Second, I use bivariate probit techniques, taking into account the *simultaneity* of maternal decisions pertaining groups' participation and children's education.

Though with the limitations determined by possible omitted variable bias, my analysis suggests that maternal participation in *arisans* positively influences children's school enrolment.

The paper is structured as follows: Section 2 reviews selected studies relating social capital and economic outcomes; Section 3 describes the econometric issues and strategy. Section 4 contains a description of the dataset used in the analysis; in Section 5, the determinants of mothers' participation in *arisans* are explored, and in Section 6 the role of *arisans* in determining schooling is analyzed. Section 7 concludes.

2. Empirical literature on the effects of social capital

There is a wide literature analyzing the effects of social capital on economic outcomes, such as firm productivity (Fafchamps and Minten, 2002) or household income (Narayan and Pritchett, 1999).

Studies on the effects of social capital can be classified along many directions: first of all, micro-level analyses can be separated from macro-level ones (Knaack and Keefer, 1997). Looking at the context under study, we can distinguish analyses of social capital in OECD countries from analyses in developing countries.

Many studies of social capital effects measured social capital through membership in groups that provide some benefits, in terms of monetary and/or informational resources' sharing and public goods' production. This kind of measure is particularly frequent in studies of developing countries, while analyses of OECD contexts often measure social capital through family or friendship ties (Furstenberg and Hughes, 1995; Coleman, 1988).

Studies of developing and OECD countries also differ with respect to the investigated outcomes. Analyses of developing countries, in fact, generally focused on economic outcomes, such as per capita household expenditure (Narayan and Pritchett, 1999) and fertilizer adoption (Isham, 2002; Bandiera and Rasul, 2006). The range of outcomes analyzed by OECD studies is wider, going from strictly economic measures, such as returns to investments (Fernandez *et al.*, 2001), to outcomes with a more visible social aspect, such as criminal behavior (Hagan and McCarthy, 1995), or with a more intangible value, such as educational attainment (Furstenberg and Hughes, 1995; Coleman, 1988).

These differences in the social capital measures and outcomes reflect the different functions that social capital serves in the two contexts. In developing settings, in fact, social capital serves many functions that in developed countries are served by formal markets. These functions include above all the provision of insurance and credit. In developed countries, instead, where credit and insurance are provided by formal markets, the role of social capital is related to the functioning of existing markets and is impoverished from its economic aspect to the advantage of its social aspect.

For a technical evaluation of studies of social capital effects, the most relevant feature that distinguishes one study from another is the extent to which it addresses and solves the potential endogeneity of social capital, due to omitted variables, simultaneity or measurement error. Following Durlauf (2002), in the following sections, I first explain the main econometric issues faced by studies of social capital effects, and then I review selected studies, underlining to which degree they address and solve the highlighted issues. Given that my analysis is at the micro-level, I focus on micro-level studies of social capital effects.[7]

[7] However, an interesting macro-level study is that of Miguel *et al.* (2005), which focuses on Indonesia and asks whether social capital promoted industrialization in the country. The study utilizes both measures of *formal* social capital, such as number of state-led community groups, and *informal* social capital, such as the proportion of per capita expenditure on cerimonies. To identify the social capital effect, that study relies on a rich set of district controls and on the fact that the social capital measures employed predate the industrialization outcomes. However, the authors conclude that initial social capital does not predict subsequent industrial development in Indonesia.

2.2 Main issues in the empirical analysis of social capital effects

The first problem faced by a researcher studying the effects of social capital is how to *define* the concept of social capital. This is a very important issue, since the precision with which social capital is defined strongly influences the empirical identification of the investigated relationships.

Durlauf (2002) distinguished *functional* notions of social capital from *causal* ones: the first describe social capital as norms or values easing cooperation, efficiency, and public goods' provision (Putnam *et al.*, 1993; Coleman, 1988). The second, instead, focus on the process by which a rational individual, given the cooperative behavior of others, cooperates guided by trust, expectations of reciprocity, and fear of sanctions. Fafchamps and Minten (2002) pointed out another conceptualization of social capital which is found in the literature: this notion sees social capital as an asset which benefits a single individual or a firm (Granovetter, 1985).

As highlighted by Durlauf (2002), many studies provided an imprecise definition of social capital, mixing functional and causal conceptions. This happened for example in Furstenberg and Hughes (1995), in which a failure in defining precisely social capital led to a lack of identification.

Identification problems also arise through the potential endogeneity of social capital, mainly related to omitted variables, reverse causality and measurement error. In the following paragraph, I assess to which degree previous empirical studies address and solve these issues.

2.3. Assessment of empirical studies

Following Durlauf (2002), the first study I consider is that of Furstenberg and Hughes (1995). This study examined the influence of social capital on seven dichotomous variables, among which high school graduation and college enrolment. Social capital variables are divided in two categories: *within family* social capital (e.g. how often the child does activities with parents, or whether the father lives in the home) and *community* social capital (e.g. a mother's religious involvement, or whether the mother has someone to turn to in hypothetical situations of difficulty).

Performing logistic regressions, the authors found statistical significant effects of many social capital measures on different outcomes and concluded that there exists a statistically significant association between most of their measures of social capital and socio-economic success in early adulthood.

As pointed out by Durlauf (2002), the set of controls used in the analysis of Furstenberg and Hughes (1995) is too small to rule out potential endogeneity of social capital and thus to maintain that social capital causally determines youths' outcomes.

To understand why, we could consider the regressor saying whether the family moved: this is a choice variable, and, as such, it is potentially related to the unobserved heterogeneity embedded in the error term. If, as suggested by Durlauf (2002), we consider the 'moves' variable as a proxy for parental investment in children, we see that the evidence that moves predict higher dropping out represents a link between parental investment and children's outcome, which has nothing to do with social capital. Given the specification used

by Furstenberg and Hughes (1995), it is hard to assert that a role for social capital exists, as opposed to parental investment in determining children's outcomes.

Furstenberg and Hughes (1995), furthermore, did not distinguish between social capital and other individual and contextual effects: a father's presence in the household, for instance, matters for child rearing in a number of ways, of which many can be completely unrelated to social influences. This identification failure mainly derives from the absence of a theory describing the mechanisms guiding social capital formation, and the lack of a consistent and precise definition of social capital: church attendance, for instance, is not necessarily associated with a strong social network which can help to recover from shocks and, instead, captures aspects of individual behavior which are not necessarily related to social capital (e.g. acceptance of strong constraints on behavior).

Narayan and Pritchett (1999), analyzing the effect of social capital on household expenditure in rural Tanzania, partially addressed issues of identification. They measured social capital trough a village-level index, constructed by combining village average group membership and groups' characteristics. Besides the social capital variable, their estimated equation included family- and village-level controls.

The authors acknowledged that social capital could be proxying for unobserved characteristics at the village-level, and, in order to assess the potential bias associated with omitted village variables, estimated the effect of village average variables (e.g. village mean education or assets) using two different datasets referring to different periods.[8] Estimates of the village effects across the two datasets are very weakly correlated: on the basis of this result, the authors maintained that the omitted variable bias is not significant, since unexplained village variation is almost exclusively associated with temporary random shocks or measurement error and is not related to time-persistent excluded variables. As a further robustness check, the authors included in the regressions all the village-level variables they could measure.[9] Adding these controls did not alter the significance of social capital. Hence, Narayan and Pritchett (1999) effectively limited omitted variable bias.

In order to address the potential endogeneity of social capital, the authors also used instrumental variables methods. However, the instruments they used do not fulfill all the requirements for identification. These instruments are, in fact, trust in strangers and in government officials. Considering that trust in strangers could depend on trustworthiness (i.e. that trust could reflect the actual behavior of others in a society), and that trustworthiness could influence expenditure, Durlauf (2002) casted doubts on the fact that the instruments are uncorrelated with the errors in the expenditure equation and concluded that the analysis of Narayan and Pritchett (1999) should be completed by finding other instruments that can be plausibly excluded *a priori* from the expenditure equation.

Besley and Levenson (1996), analyzing the relationship between rosca participation and durable goods' acquisition in Taiwan, addressed the problem of endogeneity of social capital via instrumental variables. The instrument they employed is the sector of occupation: more stable occupations, like the one in the public

[8] The two dataset used are the 1995 Social Capital and Poverty Survey (SCPS) and the 1993 Human Resource Development Survey (HRDS).

[9] In particular, they add measures of land quality, population density and banking facilities (which could all have a positive relationship with associational activity).

sectors, are positively associated with rosca participation. The use of this instrument, however, is based on the assumption that the employment sector is uncorrelated to the preference for durable goods, which is rather strong an assumption.

In the study of Fafchamps and Minten (2002), social capital effects are better identified with respect to Narayan and Pritchet (1999) and Besley and Levenson (1996). This study analysed the influence of social capital on sales' value among agricultural traders in Madagascar. The authors measured social capital using the number of traders known: the higher the number of traders known, the lower the transaction costs and thus the better firms' performance.

The identification strategy of this paper included three steps: first, the author used instrumental variables methods. Second, they added regressors which could capture omitted firms' and entrepreneurial characteristics potentially correlated with social capital. Third, they examined the channels through which social capital boosts firm productivity. As instrumental variables, the study used a wide range of entrepreneurial characteristics, such as human capital, family and personal background.[10] Since the instruments satisfied the requirements for identification, this analysis represents an empirical proof of the causal effect of social capital on firms' performance. [11]

Studies of *network effects* (such as Bandiera and Rasul, 2006; Isham, 2002; Betrand *et al.*, 2000) posed slightly different problems with respect to the endogeneity issues analysed so far. Network effects studies, indeed, focusing on the effect of *group* behavior on individual behavior, also face the so called *reflection problem* (Manski, 1993). In these analyses, researchers aim at distinguishing the role of group behavior ('social effects'), from the role of common shocks ('correlated effects') in influencing individual behavior.[12] Since I do not consider group effects, but I focus on the influence of individual maternal social capital on individual children's schooling, the problem posed by my research question is partially different from the one analysed in network effects studies, and it is more similar to the endogeneity problem faced by program evaluation studies. For this reason, I consider in this review also the study of Pitt and Khandker (1998), analyzing the impact of group-based credit programs on household behavior in Bangladesh.[13] This study faced endogeneity problems, similar to those arising when the social capital effects are analyzed. Moreover, group credit programs entail a strong component of trust, which is a fundamental dimension of social capital. However, the study of Pitt and Khandker (1998) distinguishes from analyses of social capital, because credit programs are interventions that are decided outside the population who participates in them. Differently, the groups considered in studies of social capital arise most often spontaneously among population members.

[10] Among personal background instruments, we find age and its square, whether the trader was born in the capital city, and its religion. Family backgroud instruments include whether the trader's father/mother was a farmer, whether the trader's father/mother attended primary school, and the number of father's/mother's years of experience in agricultural trade.

[11] However, some additional regressors included to further control for omitted variable bias (e.g. entrepreneurial traits such as entrepreneur's propensity to save and proxy for individualism and altruism) could be endogenous, since they could be themselves the result of the number of traders known, the measure of social capital employed by the authors.

[12] In studies of network effects, identification is achieved, for instance, using fixed effects (as in Bertrand *et al.*, 2004), or relying on nonlinear functional forms (as in Bandiera and Rasul, 2006).

[13] Grameen Bank, Bangladesh Rural Advancement Committee (BRAC), and Bangladesh Rural Development Board (BRDB).

This difference is what provides to studies of the impact of group-based programs (such as Pitt and Khandker, 1998) a further device to achieve identification, which is different from instrumental variables. While analyzing the impact of group-based programs, in fact, researchers can construct the sample survey so as to provide identification through a *quasi-experimental design*. This approach considers *eligibility* into the program as a quasi-experiment: the program's effect can be identified if the sample also includes households in villages with treatment choice ('program villages') that are non-eligible, i.e. they are still subject to treatment choice, but are non-eligible.[14] In this setting, a comparison of eligible and not eligible households, conditioning on village fixed effects and household characteristics, is sufficient to identify the treatment effect.

The analysis of social capital effects, if the social capital measures concerns groups that arise spontaneously from the population, cannot take advantage of the eligibility information, and has to rely on instrumental variables or on functional form to achieve identification.

In my analysis, furthermore, it is particularly difficult to find a credible instrument, since most of the factors influencing maternal participation in *arisan* are also likely to have a direct impact on children's schooling.

Give the difficulties in finding credible instruments for maternal participation in *arisans*, as I will explain in Section 3, I rely mainly on functional form assumptions to identify the effects of interest.

Table 2.1 summarizes the main characteristics of the studies on social capital effects here reviewed.

[14] The setting considered by Pitt and Khandker (1998), in particular, is characterized by the restriction that households owning more than one-half acre of land are precluded from joining any program.

Table 2.1. Micro-Level Empirical Studies on Social Capital (SC) Effects

	Outcome	Social Capital Measures	Data	Econometric Strategy	Conclusions	Potential Sources of Bias
Furstenberg and Hughes (J. of Marriage and the Family 1995)	- Educational attainment - Labour force status - Avoiding criminal behavior - Avoiding births before age 19	- Within family SC[15] (e.g. support to/from mother, n. of child's friends mother knows) - Outside family SC (e.g. mother's religious involvement, mother's help network[16], child changed school due to move)	1984-87 Baltimore Study, US (longitudinal)	OLS Logistic Regressions. Identification via addition of the following controls: - Parents' human capital - Youth's outcomes 3 years before	The following SC measures are significantly related to outcomes: - Support to/from mother - N. of child's friends mother knows - Mother's help network	- SC assumed exogenous - SC not precisely defined - Small set of individual controls - Small sample size
Besley and Levenson (The Economic J. 1996)	Rate of durable goods' ownership by households	Household participation in roscas (dummy=1 if in the hh there is at least one rosca participant)	1977-91 Personal Income Distribution Survey, Taiwan (repeated cross-sections)	OLS and IV. IVs used: - Occupation - Twice-lagged rosca participation	Participation in roscas is significantly positively linked to durables' accumulation	IVs may not satisfy the exclusion restriction
Narayan and Pritchett (Ec. Development and Cultural Change 1999)	Per capita household expenditure	Village-level social capital index (combining membership and group characteristics)	1995 Social Capital and Poverty Survey, rural Tanzania	OLS and IV. IVs used: - Trust in strangers - Trust in gov't officials	SC is an important determinant of household expenditure	- Data aggregation - IVs may not satisfy the exclusion restriction
Fafchamps and Minten (Oxford Ec. Papers 2002)	Firms' value added and total annual sales	Number of (non-family) traders known	1997 Survey of Agricultural Traders, Madagascar	OLS and IVs. IVs used: - Trader's human capital - Location - Trader's personal/family background - Trader's informal borrowing history	SC has a strong positive effect on traders' performance	- IVs may not satisfy exclusion restriction

[15] 'SC' indicates Social Capital.
[16] This variable is a dummy which is equal to one if repondent reported having someone to turn to in all of four hypothetical situations.

3. Empirical strategy

The equation of interest is in the form:[17]

$$School_{chv} = \beta_0 + \beta_1 \mathbf{X}_{chv} + \beta_2 \mathbf{H}_{hv} + \beta_3 \mathbf{V}_v + \gamma M\_Arisan_{chv} + u_{chv} \quad (1)$$

where *c* refers to the child, while *h* and *v* indicate the child's household and village of residence, respectively. As dependent variable, I consider a dummy which equals one if the child is enrolled in secondary or tertiary education, and zero otherwise. I consider secondary or tertiary education because in Indonesia primary education is almost universal, and thus secondary and tertiary levels represent a proper target to which policy measures should be addressed in order to improve the overall educational outcomes of the population. $\mathbf{X}_{chv}$ is a set of individual-level characteristics, while $\mathbf{H}_{hv}$ and $\mathbf{V}_v$ are vectors of household- and village-level variables, respectively. M_Arisan_{chv} is a dummy which equals one if the child's mother participated in *arisans* during the year previous to the survey and zero otherwise. γ and β_j (j=0, …, 3) are parameters, among which γ is the parameter of interest. u_{chv} represents the error term.

As anticipated in the introduction, the main challenge of the analysis is to deal with the potential endogeneity of social capital, which has two main sources. First of all, there could be *omitted variables* which are correlated both with the decision to participate in *arisans* and with the choice of the amount to invest in education. For instance, mothers that are more extroverted and thus more involved in community activities could also have a greater interest in their children's education: if this is the case and I do not control for maternal extroversion, I obtain upward biased estimates of the parameter of interest. In fact, the estimated parameter would be capturing not only the effect of *arisans*, but also the influence of maternal attitude towards education. A second example of omitted variable bias can be made with reference to household vulnerability to shocks. More vulnerable households could have greater participation rates in *arisans*, utilized as a way to recover faster from shocks. At the same time, in this kind of households children may have lower school enrolment rates, since parents, given the higher perceived risk of shocks, could invest less in education, compared to parents in less vulnerable households, and use child labour as a source of household finance. As a result, the omission of household vulnerability to economic shocks among control variables could bias downward the social capital coefficient estimate.

Omitted variables could also imply a failure to distinguish between the effect of *arisans*, from that of legal factors or political institutions.

Reverse causality could also be an important source of bias: when a mother decides to send her child to school, she has to acquire monetary resources to invest in education and may try to obtain them by participating in *arisans*. Hence, participation in *arisans* may be partly determined as a function of a mother's desired level of education for her child.

In order to correct for these potential biases, first of all I analyze the determinants of social capital formation. As pointed out by Durlauf (2002), indeed, an important reason for omitted variable bias in studies

[17] For notational simplicity, the time index is omitted.

of social capital effects is their lack of analysis of the determinants of social capital formation. Since my measure of social capital is a dichotomous one, I estimate a probit model, which is outlined as follows:

$$\text{Arisan}^{*}_{ahv} = \alpha_0 + \alpha_1 \mathbf{X}_{ahv} + \alpha_2 \mathbf{H}_{hv} + \alpha_3 \mathbf{V}_v + u_{ahv} \quad (2)$$

where *a* refers to the adult, while *h* and *v* indicate his or her household and village of residence. Arisan^{*}_{ahv} is not observed, while we observe $\text{Arisan}_{ahv} = 1$ if $\text{Arisan}^{*}_{ahv} > 0$, 0 otherwise. $\mathbf{X}_{ahv}$ is a set of individual level characteristics, while $\mathbf{H}_{hv}$ and $\mathbf{V}_v$ are vectors of household- and community-level variables, respectively. u_{ahv} is an i.i.d. standard normally distributed error. I estimate this model across the samples of all adults, men, women, and mothers.

After having understood which are the variables influencing *arisan* participation by mothers, I include in equation (1) those variables which could influence both maternal social capital and education, and estimate the following probit model:

$$\text{School}^{*}_{chv} = \beta_0 + \beta_1 \mathbf{X}_{chv} + \beta_2 \mathbf{H}_{hv} + \beta_3 \mathbf{V}_v + \gamma \text{M\_Arisan}_{chv} + u_{chv} \quad (3)$$

where School^{*}_{chv} is unobserved, while we observe $\text{School}_{chv} = 1$ if School^{*}_{chv}, > 0, 0 otherwise. M_Arisan_{chv} is a dummy equal to one if child's mother participates in *arisans* and zero otherwise, while $\mathbf{X}_{chv}$, $\mathbf{H}_{hv}$, and $\mathbf{V}_v$ are the same vectors considered in (2). u_{chv} is an i.i.d. standard normally distributed error.

In order to take into account the endogeneity of social capital, I use a bivariate probit regression, which models jointly mothers' decisions regarding *arisan* participation and children's schooling.

Two stage least squares represent an alternative strategy which has been used in the literature (see Section 2). However, I could not find a credible instrument for maternal participation in *arisans*. Given that group activities are highly linked to community characteristics, an ideal instrument would have been a community characteristic unrelated to children's schooling and influencing maternal participation in *arisans*. In particular, since the regressor of interest is at the individual level (mother's participation), the instrument could have been constructed by interacting a community variable and an individual-level maternal characteristic, both influencing *arisan* participation.

The IFLS community level questionnaire, administered to community leaders, entails a section regarding communities' traditional law, in which the following question is asked: 'According to traditional law, are there groups in the community which utilize the principle of mutual cooperation?'. This *historical* information is a particularly good instrument candidate, mainly for two reasons: first of all, it pre-dates children's schooling, and, therefore, it is unaffected by village-level random shocks. Second, it represents a key determinant of social capital formation. As underlined by Glaeser *at al.* (2002), indeed, 'small differences in initial conditions can generate large divergence in long-run levels of social capital' (2002; F442). This is because social capital has strong *interpersonal complementarities*, which imply that an

increase in individual-level social capital translates into a more than proportional increase in aggregate social capital (*aggregate* or *social multiplier effects*).[18]

IFLS data show that the historical presence of groups based on the ethic of mutual cooperation is significantly and positively associated with mothers' participation in *arisans*. However, when controlling for other village-level characteristics proxying for village economic development, this positive association is no longer significant. This could be due to the fact that social capital may contribute to a village's development, affecting many of its characteristics. Interacting the village-level variable above described with maternal characteristics determining *arisan* participation could have solved the 'relevance of the instrument' problem. Unfortunately, it was impossible to find a maternal characteristic which could credibly satisfy the exclusion restriction. I considered for instance mother's literacy, which, in most settings represents a necessary conditions for participation in *arisans*. The information about literacy satisfies the relevance requirement, but the exclusion restriction is violated, since mother's literacy has a direct impact on children's schooling. In other words, we cannot maintain that participation in *arisan* is the only channel through which mothers' literacy influences children's schooling.[19]

Given both the difficulties in finding a proper instrument, and the binary nature of both my dependent and independent variables of interest, I consider a bivariate probit model. In particular, I estimate a *recursive* bivariate probit model (Greene, 2003).

Defining $\text{School}_{chv} \equiv y_{1chv}$, $\text{M\_Arisan}_{chv} \equiv y_{2chv}$, and $\mathbf{W} \equiv [\mathbf{X}_{chv}\ \mathbf{H}_{hv}\ \mathbf{V}_{v}]$, and assuming that ε_{1chv}, ε_{2chv} are standard normally i.i.d. with correlation ρ, the bivariate probit model is written as follows:

$$y^{*}_{1chv} = \theta_1\mathbf{W} + \gamma y_{2chv} + \varepsilon_{1chv} \quad (4)$$

$$y^{*}_{2chv} = \theta_2\mathbf{W} + \varepsilon_{2chv} \quad (5)$$

Where y^{*}_{lchv} (l = 1, 2) are 'latent' variables for child *chv*'s schooling and participation in *arisan* by his or her mother respectively. We only observe $y_{l,chv} = 1$ if $y^{*}_{l,chv} > 0$, and zero otherwise (l = 1, 2). Equations (4) and (5) contain the same set of covariates, with the exception that (4) includes also the dependent variable of (5), y_{2chv}. The vector $\mathbf{W}$ includes the same variables of models (2) and (3).

Defining $q_{lchv} \equiv 1$ if $y_{lchv} > 0$, -1 otherwise (l = 1,2), and $\rho^{*}_{chv} \equiv q_{1chv}q_{2chv}\rho$, the log likelihood is:

$$\text{Log L} = \Sigma_{chv}\log\Phi_2[q_{1chv}(\theta_1\mathbf{W} + \gamma y_{2chv}),\ q_{2chv}\theta_2\mathbf{W},\ \rho^{*}_{chv}] \quad (6)$$

Where Φ_2 is the cumulative distribution function of the standard bivariate normal.

[18] The information about the historical presence of groups using an ethic of mutual cooperation in the village is also utilised as a measure of informal social capital by Miguel *et al.* (2005).

[19] An analogous reasoning can be made with reference to mothers' occupation. It has been showed (Levenson and Besley, 1996) that individual participation in roscas is positively associated with individual occupation in 'stable', sectors (e.g. the public sector, as opposed to the primary one). This relationship is verified also by IFLS data. However, mothers' sector of occupation is not a credible instrument for her participation in *arisans*, since the stability of a mother's occupation matters for children's schooling for many reasons, not only because it favours *arisan* participation.

Wilde (2000) shows that exclusion restrictions are not necessary for the theoretical identification of the model, if there exists a varying exogenous regressor. Hence, given also that I did not find a credible instrument for social capital, I first estimate the model without exclusion restrictions, only relying on functional form to achieve identification. As a robustness check, I perform two bivariate probits without exclusion restrictions, including different sets of covariates. The estimation of these two models provides consistent results.

However, Monfardini and Radice (2008) show that exclusion restrictions, though unnecessary to formal identification, are important because 'inference on the estimated correlation coefficient through the Wald test requires a great deal of sample information and deteriorates sharply in the absence of exclusion restrictions' (2008: 281). Hence, I also perform a bivariate probit model in which each equation contains one exclusion restriction. These restrictions, though, are not justified theoretically, but only empirically: I excluded from (4) one variable which was insignificant in that equation, while being strongly significant in (5), and viceversa. Results from the model including 'restrictions' are consistent with those from the models without them.

As a further robustness check, I estimate the three bivariate probits (two without restrictions and one including restrictions) on subsamples of the observations, obtaining a confirmation of the results.

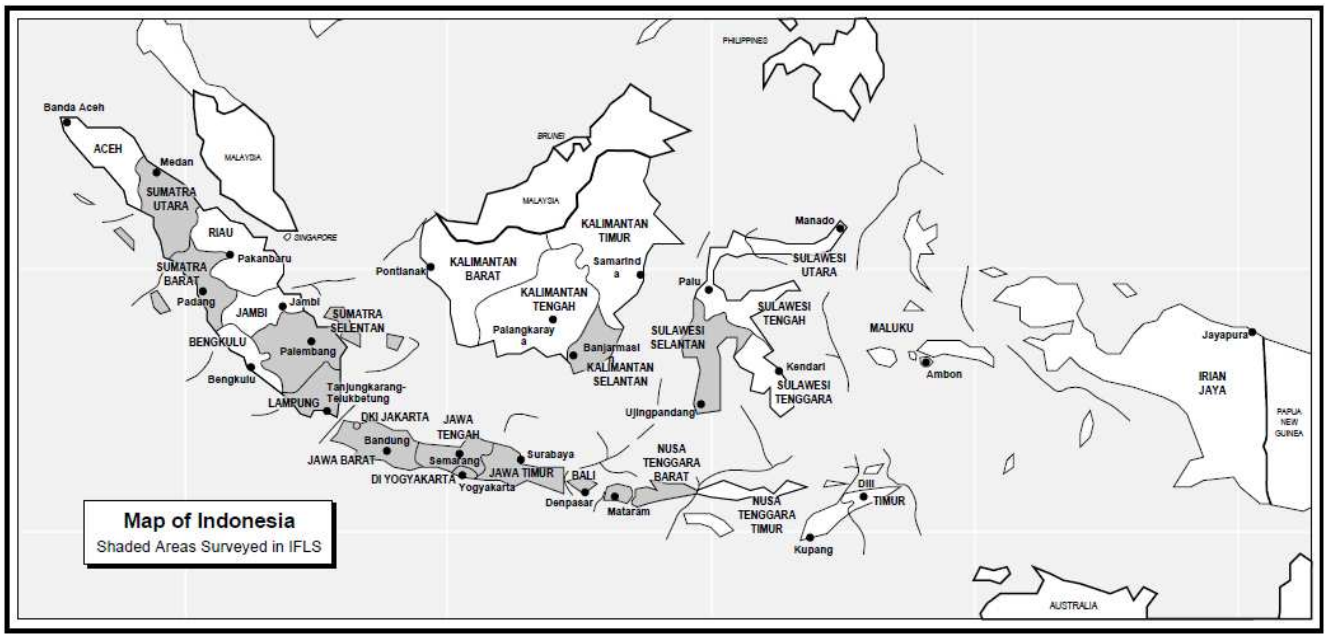
All models are estimated using the data from waves 1997 and 2000 of the Indonesia Family Life Survey (IFLS2-3) in a pooled sample. Furthermore, all regressions have robust standard errors, clustered at the district-level.

4. Data

I use the data originating from the second and third waves of the Indonesia Family Life Survey (IFLS2-3), administered in years 1997 and 2000 respectively. Because IFLS is a longitudinal survey, IFLS2 and IFLS3 draw their sample from IFLS1 and IFLS1-2 respectively. The ILFS1 sample is stratified by province and rural/urban location, and includes 13 out of the then 27 Indonesian provinces.[20] The resulting sample contains 83% of the population. In the following map, shaded areas represent the surveyed provinces.

[20] IFLS sample includes: four provinces on Sumatra (North Sumatra, West Sumatra, South Sumatra, and Lampung), all five Javanese provinces (DKI Jakarta, West Java, DI Yogyakarta, and East Java), and four provinces among the remaining major islands (Bali, West Nusa Tenggara, South Kalimantan, and South Sulawesi).

Figure 4.1. IFLS Sampled Provinces



Source: Frankenberg and Karoly (1995).

Within the 13 provinces, enumeration areas (EAs) were randomly chosen, over sampling urban EAs and EAs in smaller provinces to ease urban-rural and Javanese-non-Javanese comparisons. Within each EA, households were randomly selected. As a result of this sampling scheme, described by Frankenberg and Karoly (1995) and outlined in more datails in the next chapter, IFLS1 interviewed 7,224 households. In IFLS2, 7,698 households were contacted, meaning they were interviewed, had all members died since last survey, or had joined another IFLS household, which had been previously interviewed. Of the IFLS2 contacted households, 6,821 were original IFLS1 households and 877 were split-off households.[21] Thus, in IFLS2 a re-contact rate of 94.4% of the IFLS1 was reached. A total of 10,574 households were then contacted in IFLS3, of which 6,800 are IFLS1 households and 3,774 are split-off since IFLS1. The re-contact rate in IFLS3 amounts to 95.3% of the IFLS1.[22] 6,564 households were interviewed in all three waves of IFLS: 1, 2 and 3.

Information about *arisan* participation is collected starting from IFLS2: hence, the sample I use in the analysis only includes individuals who were either in IFLS2 or in IFLS3, or in both.

IFLS also gathered community-level information, such as village infrastructures, education and health facilities, credit opportunities.

The following sections provide decriptive statistics and econometric results for the analysis of the determinants of *arisan* participation by adults (Section 5), and for the analysis of the relationship between maternal participation in *arisans* and children's schooling (Section 6).

[21] *Split-off households* are new households interviewed in IFLS2 or IFLS3 because they contained a *target respondent*, who is an IFLS1 household member selected for IFLS2 or 3 either because he/she had provided detailed individual level information in IFLS1, or had been age 26 or older in IFLS1, or meet other criteria (for further details, see Strauss *et al.*, 2004).

[22] Percentage is over IFLS1 households with at least some members living in the last survey.

5. Determinants of participation in *arisans*

As explained in Section 3, the first part of the econometric analysis explores the determinants of adults' participation in *arisans*, by estimating model (2). This preliminary analysis has mainly the purpose to understand which control variables ought to be included in the schooling equation, to minimize omitted variable bias.

Previous studies analyzed the role of the following determinants of individual's participation in roscas:

i) *Gender*: there is consistent evidence in the literature about the fact that being female significantly increases the probability of joining roscas. This result has been verified in different countries, such as Taiwan (Levenson and Besley, 1996), and Kenya (Anderson and Baland, 2002). As pointed out by Levenson and Besley (1996), gender differentials in rosca participation could be related to two main factors: on the demand side, women could have a greater benefit in joining roscas, compared to men, given that women have smaller opportunities to obtain credit from the formal financial sector. On the supply side, women's social network could be stronger than men's, allowing for the presence of more credible sanction mechanisms and thus the formation of more effectively functioning roscas.

ii) *Position in the lifecycle*: age and its square significantly influence social capital formation in general (Glaeser *et al.*, 2002), and, in particular, rosca participation (Levenson and Besley, 1996). As highlighted by the authors, this suggests that, in earlier stages of the lifecycle, joining roscas may be a savings device in order to buy indivisible goods. When individuals become older, it is plausible that they do not need roscas' services anymore, either because they have enough savings, or because they are less interested in social interaction.

iii) *Household income*: this variable is expected to significantly increase rosca participation, mainly for two reasons: first, only individuals having a certain income flow can contribute regularly a part of it to roscas; second, households having higher incomes are likely to have higher assets, and thus are viewed as more reliable rosca members (Levenson and Besley, 1996). However, past literature reported contrasting results about the influence of household income on rosca participation. Levenson and Besley (1996), for instance, found a positive association between household income and rosca participation in Taiwan. Anderson and Baland (2002), instead, concluded that household income does not significantly influence rosca participation among the inhabitants of the Kenyan slum of Kibera (Nairobi). In this context, in fact, participation appears to be primarily determined by a woman's bargaining power in the household, as proxied by the her share of the couple's income. In particular, there appears to be an inverted-U shaped relationship between rosca participation and a woman's income share in the household. Thus, when women's weight in the household is very low or very high, women do not join roscas. With low weight, they do not have access to roscas, while with higher weight they have a smaller need to join.[23]

[23] This result can be related to that of Varadharajan (2003), who finds an inverted U-shaped relationship between *arisan* participation and per capita household expenditure. As explained by the author, this suggests that, at lower levels of income, individuals are not able to participate in *arisan*, because they lack the resources to contribute. When income increases, they are able to join the group and to take advantage of its services, which they presumably need less when income reaches a certain threshold. Hence, *arisan* participation would be highest at intermediate levels of income.

Hence, the relevance of household income as a determinant of participation seems to depend on the context of analysis: in settings where women are more autonomous from their husbands and have a greater weight in household's decision-making, for example, women's share of couple income could play a smaller role in determining rosca participation, while household income could be more important. Differently, in contexts in which women are more marginalized in the household's decision-making, women's share of income could be the main determinant of the decision to join roscas.

iv) *Occupation*: according to Levenson and Besley (1996), rosca participation is related to the stability of occupation. As a result, individuals who are employed in the private sector are less likely to join roscas than individuals working in the public sector. Furthermore, individuals working in agriculture, fishing or forestry have smaller participation rates compared to those employed in industry or services.

v) *Human capital*: there is consistent evidence that more educated individuals have higher levels of social capital (Helliwell and Putnam, 2007; Glaeser *et al.*, 2002). This mainly depends on three factors: first, more educated individuals are more likely to have higher income, thus affording to pay for regular roscas contributions. Secondly, educated individuals satisfy the literacy requirement needed to participate. Third, increased education may imply an greater awareness of the benefits of rosca participation.[24]

vi) *Community-level factors*: the importance of community characteristics in determining social capital has been showed, among others, by Glaeser *at al.* (2002), who underlined above all the importance of historical factors in influencing investment in social capital.

Given the above analysis, I include in equation (2) the following regressors:

1. Individual (adult) variables:
    a. Demographic: sex, age and its square, dummies for education level (elementary and secondary/tertiary, with no education being the omitted category), a dummy indicating literacy, and a dummy for Islamic religion.
    b. Occupation: I include a dummy indicating whether the individual worked during the year previous to the survey.
2. Household variables:
    a. Proxies for household wealth: dummies indicating house ownership and availability of electricity in the household.
    b. Household size and composition: dummy indicating whether the household head is a female, number of children in the household, and their average age.

[24] Varadharajan (2003) also finds a strong positive influence of education on the probability of joining *arisans*. The author also plotted *arisan* participation and borrowing against education, showing that less educated individuals are more likely to borrow than to join *arisans* to meet their financial constraints. After a certain education threshold, instead, *arisan* participation is more frequent than borrowing. As suggested by the author, this also might indicate that individuals join *arisans* not only for economic reasons.

3. Community variables:
    a. Dummy for rural/urban location.
    b. Proxies for economic development: dummies indicating whether inside the village a post office and a formal credit institution are present; percentage of households using electricity.
    c. Other variables proxying for community's status: dummies indicating whether during the year previous to the survey there was an improvement (a distaster) in the village, which benefited (damaged) at least 30% of the population.

All specifications also include time dummies and province fixed effects, and have robust standard errors, clustered at the district-level. I estimate the probit models on the following samples: all adults, men, women, and mothers. Table 5.1. shows mean characteristics for the adult samples (variable description is given in the Appendix). As it can be seen from the table, *arisan* participation is highest for the subsample of mothers.

Table 5.1. Summary Statistics for the Adult Samples
(Mean and Standard Deviation)

	ADULT	MEN	WOMEN	MOTHERS[A]
Individual (Adult) Variables				
Female	.522	0	1	1
	(.499)	(0)	(0)	(0)
Age	36.672	36.352	36.963	41.857
	(16.521)	(16.318)	(16.698)	(14.520)
Muslim	.921	.921	.921	.925
	(.269)	(.269)	(.269)	(.263)
Elementary	.437	.426	.445	.529
	(.496)	(.494)	(.497)	(.499)
High School/University	.458	.515	.405	.294
	(.498)	(.499)	(.490)	(.456)
Literacy	.838	.899	.782	.738
	(.368)	(.302)	(.412)	(.439)
Worked_Last_Year	.612	.762	.474	.516
	(.487)	(.425)	(.499)	(.499)
Arisan	.318	.206	.421	.480
	(.466)	(.404)	(.493)	(.499)
Household Variables				
House	.876	.875	.876	.874
	(.330)	(.330)	(.329)	(.331)
H_Electricity	.933	.932	.934	.926
	(.249)	(.250)	(.247)	(.260)
H_Head_Female	.131	.088	.170	.159
	(.338)	(.283)	(.376)	(.365)
H_Size	6.495	6.498	6.492	6.351
	(2.639)	(2.608)	(2.666)	(2.564)
N_Children	3.136	3.149	3.124	3.012
	(1.786)	(1.755)	(1.813)	(1.669)
Age_Children	13.347	13.509	13.199	12.115
	(5.054)	(5.045)	(5.057)	(5.06)

Table 5.1. (Continued)

	ADULT	MEN	WOMEN	MOTHERS[A]
Village Variables				
Rural	.465	.471	.458	.496
	(.499)	(.499)	(.498)	(.500)
Post	.246	.245	.246	.228
	(.430)	(.430)	(.431)	(.419)
Credit_Formal	.632	.627	.637	.625
	(.482)	(.483)	(.480)	(.483)
V_Electricity	85.849	85.708	85.977	84.947
	(20.348)	(20.489)	(20.218)	(21.015)
Disaster	.134	.138	.130	.138
	(.341)	(.345)	(.336)	(.344)
Improvement	.241	.338	.244	.247
	(.428)	(.426)	(.429)	(.431)
N.	22,606	10,793	11,813	8,250

Notes:
[A.] This column considers all mothers in the sample, not only mothers of children aged 13-24.
Descriptives for mothers of children aged 13-24 are reported in Table 6.1, which considers as unit of analysis the child (observations for mothers having more than one child aged 13-24 are repeated). However, considering only one observation per mother, the descriptives do not change substantially.

Table 5.2. shows the marginal effects for the probit models of *arisan* participation:

Table 5.2 Probit Models of *Arisan* Participation
(Marginal Effects at Average Characteristics)

Dependent Variable:
Arisan = 1 if the individual participated in *arisans* during the year previous to the survey

	ALL	MEN	WOMEN	MOTHERSA
Individual (Adult) Variables				
Female	.288***			
	(.0102)	-	-	-
Age	.026***	.009***	.039***	.026***
	(.001)	(.002)	(.002)	(.003)
Age_sq	-.000***	-.000***	-.000***	-.000***
	(.000)	(.000)	(.000)	(.000)
Muslim	.093***	.049**	.135***	.127***
	(.023)	(.024)	(.034)	(.048)
Literacy	. 116***	.066***	.154***	.145***
	(.014)	(.016)	(.018)	(.021)
Elementary	.113***	.065**	.119***	.132***
	(.022)	(.027)	(.026)	(.029)
High School/University	.192***	.123***	.229***	.274***
	(.025)	(.026)	(.032)	(.034)
Worked_Last_Year	.086***	.107***	.091***	.086***
	(.010)	(.011)	(.014)	(.016)
Household Variables				
House	.020	.034**	.031	.026
	(.014)	(.015)	(.018)	(.022)
H_Electricity	.011	-.001	.024	.033
	(.032)	(.029)	(.041)	(.045)
H_Head_Female	-.038***	-.028	-.043**	-.038*
	(014)	(.018)	(.018)	(.021)

Table 5.2 (Continued)

Dependent Variable: Arisan = 1 if the individual participated in arisan during the year previous to the survey	ALL	MEN	WOMEN	MOTHERS[A]
Household Variables				
H_Size	-.005**	-.004	-.007**	.000
	(.002)	(.003)	(.003)	(.004)
H_N_Children	-.000	.001	-.001	-.009
	(.003)	(.003)	(.005)	(.006)
H_Age_Children	-.001*	-.000	-.004***	.002
	(.001)	(.000)	(.001)	(.001)
Village Variables				
Rural	-.041*	-.021	-.061**	-.096***
	(.024)	(.020)	(.029)	(.035)
Post	.033	.017	.048*	.072**
	(.023)	(.023)	(.020)	(.035)
Credit_Formal	..039**	.031*	.047**	.053**
	(.017)	(.015)	(.020)	(.022)
V_Electricity	.000	.000	.000	.000
	(.000)	()	(.000)	(.002)
Disaster	-.011	-.024	.003	.007
	(.022)	(.021)	(.028)	(.031)
Improvement	-.024	-.022	-.023	-.024
	(.017)	(.015)	(.022)	(.026)
N.	22,606	10,793	11,813	8,250
Pseudo-R2	.17	.12	.15	.17

Notes:
***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.
Robust standard errors clustered at district level are reported in parentheses.
Province fixed effects and year dummy are included.
[A.] This column considers all mothers in the sample, not only mothers of children aged 13-24.Estimations of the determinants of arisan participation for mothers of children aged 13-24 are reported in Table 5.3, which considers as unit of analysis the child: hence, observations for mothers having more than one child aged 13-24 are repeated. However, when we consider only one observation per mother, results do not change substantially.

Results in the first column of Table 5.2 confirm that *arisans* are mainly female's associations: being female increases the probability of joining *arisans* by 29 percentage points. Life-cycle effects are also significant and consistent with expectations: the probability of investing in social capital has an inverted-U shaped relationship with age. As expected, more educated individuals are more likely to participate in roscas. The influence of education is particularly important for women and mothers, while it is less important for men: having secondary or tertiary education, is associated with an increase in mothers' probability of joining *arisans* by 27 percentage points, compared to having no education, while the impact is of 12 percentage points in the male sample. Also in the case of literacy, the effect referred to men is much lower than the one referred to women or mothers. Having worked during the year previous to the survey also increases the probability of joining *arisans*. The impact effect is about .09, and it is slightly lower for mothers, than for women and men. This is consistent with mothers being less likely to work than men and childless women in

general, and could suggest that, when working, mothers can devote less time to *arisan* participation, compared to men and childless women, because of their domestic duties.

Among household characteristics, we can consider house ownership and female headship as proxies for household wealth, given that, on average, households owning the house they occupy are wealthier than non-owner households, and female headed households are poorer than male headed ones. Column 2 shows a positive statistically significant association of wealth with *arisan* participation by men. The influence of wealth on participation declines, when considering women and mothers, for whom female headship is the only wealth proxy which is significantly associated with *arisan* participation. Hence, we see that education is more relevant than wealth for women's participation, while for men wealth is more relevant. If we think of women's education as being positively related to their power in the household, these results may suggest that, for women's participation, weight in the household decision making is more relevant than wealth. These results could also suggest that women participate in *arisans* not only for financial reasons. In other words, the social component of *arisans* could be more relevant for women than it is for men.

The number of children in the household does not seem to influence *arisan* participation, while average children's age increases participation only for women: in particular, one year increase in average children's age decreases women's probability of joining *arisans* by .04 percentage points. This negative marginal effect could suggest that, when children are older, they contribute to household expenses through some working activities, and could thus reduce women's need of *arisans'* financial services.

Among village-level characteristics, the presence of a post office and of a formal credit institution in the village are significantly positively associated with *arisan* participation, indicating that a sufficient level of economic development is required for *arisans* to be formed.[25] The positive association between *arisans* and formal credit institutions again suggests that *arisans* are not a perfect substitute for formal finance and that people may join them not only for financial reasons.

Table 5.3 reports the estimation of probit models of *arisan* participation for mothers of children aged 13-24. In this estimations, the unit of analysis is the child: hence, observations relative to mothers having more than one child aged 13-24 are repeated. Considering only one observation per mother does not change substantially the results. Regressions in Table 5.3 includes all variables which are included in the probit models having as dependent variable children's schooling. Hence, father's characteristics are also included (see Section 6). Furthermore, three models are estimated, and indicated as ARISAN PROBIT 1-2-3: ARISAN PROBIT 1 includes individual, household, and village controls, ARISAN PROBIT 2 differs from ARISAN PROBIT 1 since it excludes two village variables which were not significant in ARISAN PROBIT 1 (percentage of households having electricity in the village and dummy indicating whether there was a village improvement during the year previous to the survey), and includes a new regressor, the number of mother's children aged 0-6. ARISAN PROBIT 3 excludes from the regressors' set of ARISAN PROBIT 2 another insignificant regressor, the number of children in the household. The three probit models correspond to the

[25] This interpretation is consistent with the qualitative analysis of Geertz (1962), according to which *arisan* formation can be related to the achievement of a certain degree of monetization in the economy and can be viewed as a product of the 'shift from a traditionalistic agrarian society to an increasingly fluid commercial one' (1962; p.260).

probit models of the determinants of school enrolment (Table 6.1.1), and can be compared with the results from the bivariate probit models (Tables 6.2.1 and 6.2.2).

Estimations in Table 5.3 confirm the results of Table 5.2. I underline the following differences: mother's age and its square are not significant, while, from Table 5.2, they resulted to be significant determinants of mother's participation in *arisans*. Moreover, regressions in Table 5.3 indicate that mother's participation in *arisans* is significantly negatively associated with the number of children in the household, which is instead insignificant in the estimation referred to all mothers (Table 5.2).

Table 5.3. Probit Models of *Arisan* Participation for Mothers of Children Aged 13-24
(Marginal Effects at Average Characteristics)

Dependent Variable: M_Arisan = 1 if the mother participated in *arisans* in the year prior to the survey	ARISAN PROBIT 1	ARISAN PROBIT 2	ARISAN PROBIT 3
Individual (Child) Variables			
Female	.026*	.025*	.025*
	(.014)	(.014)	(.014)
Age	-.003	-.003	-
	(.002)	(.002)	
Muslim	.225***	.222***	.221***
	(.059)	(.058)	(.058)
M_Age	.013	.012	.012
	(.015)	(.015)	(.015)
M_Age_sq	-.000	-.000	-.000
	(.000)	(.000)	(.000)
M_Literacy	.068*	.068*	.068*
	(.039)	(.039)	(.039)
M_Elementary	.027	.028	.028
	(.055)	(.055)	(.055)
M_High	.137**	.142**	.142**
School/University	(.060)	(.060)	(.060)
M_Worked_Last_Year	.089***	.086***	.087***
	(.022)	(.022)	(.022)
Individual Variables			
F_Literacy	.110***	.113***	.113***
	(.036)	(.036)	(.036)
F_Elementary	.049	.048	.048
	(.054)	(.054)	(.054)
F_High School/University	.173***	.173***	.172***
	(.063)	(.063)	(.063)
F_Worked_Last_Year	-.017	-.021	-.019
	(.038)	(.042)	(.042)

Table 5.3. (Continued)

Dependent Variable: M_Arisan = 1 if the mother participated in *arisans* in the year prior to the survey	ARISAN PROBIT 1	ARISAN PROBIT 2	ARISAN PROBIT 3
Household Variables			
House	.065*	.060	.060*
	(.038)	(.037)	(.037)
H_Electricity	.010	.038	.038
	(.065)	(.058)	(.058)
H_Head_Female	-.124	-.124	-.123
	(.118)	(.118)	(.118)
H_Size	.019**	.018**	.018**
	(.008)	(.008)	(.008)
H_N_Children	-.038***	-.034***	-.034***
	(.011)	(.012)	(.012)
H_Age_Children	.002	.001	.000
	(.003)	(.003)	(.003)
M_N_Children_0-6	-	-.020	-.021
		(.026)	(.026)
Village Variables			
Rural	-.114***	-.126***	-.126***
	(.040)	(.037)	(.037)
Post	.084*	.088*	.088*
	(.045)	(.045)	(.045)
Credit_Formal	.060*	.062*	.062*
	(.032)	(.032)	(.032)
V_Electricity	.001	-	-
	(.001)		
Disaster	-.044	-.053	-.053
	(.043)	(.042)	(.042)
Improvement	-.050	-	-
	(.034)		
N	5,349	5,349	5,349
Pseudo-R2	.18	.18	.18

Notes:
***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.
Robust standard errors clustered at district-level are reported in parentheses.
Province fixed effects and year dummy included.

After having seen which variables determine mother's *arisan* participation, the following section analyzes the relationship between children's schooling and mothers' participation in *arisans*.

6. Relationship between maternal social capital and children's school enrolment

I analyze the relationship between mothers' *arisan* status and children's enrolment by using IFLS2 and IFLS3 data, in a pooled sample. I consider youths aged 13-24, which is the age range of secondary and tertiary education, and restrict the sample to the two major islands, Sumatra and Java, thus including 9 provinces. Considering Java and Sumatra only, identification is improved, not only because Java and Sumatra contains more similar provinces, and so unobserved heterogeneity is reduced, but also because these two islands have the most important tradition of community participation compared to other Indonesian islands (Geertz, 1962; Ardener, 1964). The final sample comprises 5,349 youths, whose characteristics are summarized in Table 6.1.

Table 6.1. Summary statistics for the children's sample

	Mean	Standard Deviation
Individual (Child) Variables		
Female	.472	.499
Age	17.336	3.218
Muslim	.918	.274
M_Age	43.326	7.264
M_Age_sq	1929.941	661.276
M_Arisan	.539	.498
M_Elementary	.630	.482
M_High School_University	.243	.429
M_Literacy	.758	.429
M_Worked_Last_Year	.571	.494
F_Elementary	.547	.497
F_High School_University	.373	.483
F_Literacy	.868	.338
F _Worked_Last_Year	.937	.242
School	.533	.498
Household Variables		
House	.889	.313
H_Electricity	.944	.228
H_Head_Female	.010	.100
H_Size	6.937	2.326
N_Children	3.834	1.729
Age_Children	14.983	3.517
Village Variables		
Rural	.432	.495
Post	.267	.442
Credit_Formal	.618	.485
V_Electricity	87.118	19.193
Disaster	.145	.352
Improvement	.231	.421
N.	5,349	

Table 6.2 shows average children's school enrolment rates by mothers' *arisan* status: we see that both in 1997 and in 2000, and both in urban and in rural areas average enrolment rates of children whose mother participates in *arisans* are greater than those of children whose mother does not participate. The differences in average enrolment are statistically significant at the 1% level.

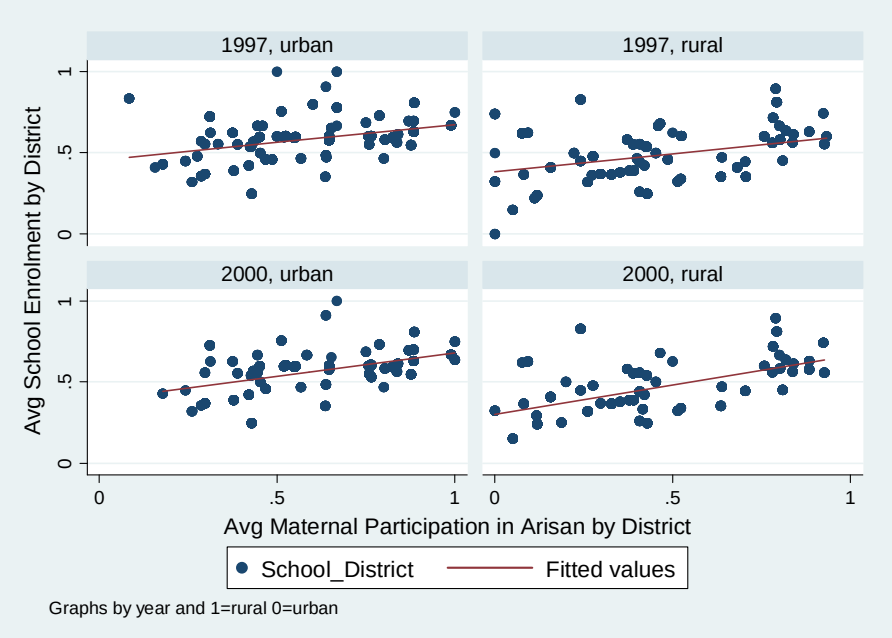
Table 6.2. Average children's enrolment rates by mothers' *arisan* status

			School
1997	Urban	M_Arisan=1	.665
		M_Arisan=0	.539
		Diff	*.126***
	Rural	M_Arisan=1	.585
		M_Arisan=0	.440
		Diff	*.145****
2000	Urban	M_Arisan=1	.597
		M_Arisan=0	.466
		Diff	*.131****
	Rural	M_Arisan=1	.517
		M_Arisan=0	.338
		Diff	*.179****

Notes:
School is a dummy which is equal to 1 if the child is enrolled in secondary or tertiary school, and zero otherwise.
*** indicates that the difference between means is significant at the 1% level.

The positive association between children's enrolment and *arisan* participation is also showed in Figure 6.1, plotting average enrolment rates by district against average *arisan* participation by district. In both years and areas there is a positive association between mothers' participation and children's enrolment.

Figure 6.1 Average Children's Enrolment Rates by *Arisan* Status of Mothers



Notes: School_District indicates average school enrolment by district.

6.1 Univariate probit model of children's enrolment

The first model I estimate to understand whether *arisan* participation has an influence on children's schooling is the univariate probit, indicated as model (3) in Section 3. Estimation results are showed in Table 6.1.1: SCHOOL PROBIT A only includes the regressor of interest (mothers' participation in *arisans*), SCHOOL PROBIT B adds other individual and household characteristics, while in SCHOOL PROBIT 1 village controls are included.

SCHOOL PROBIT 1, 2, and 3 of Table 6.1.1 correspond to the probit models of maternal *arisan* participation of Table 5.3: SCHOOL PROBIT 2 differs from SCHOOL PROBIT 1, since it excludes two village variables which were not significant in SCHOOL PROBIT 1 (percentage of households having electricity in the village and dummy indicating whether there was a village improvement during the year previous to the survey), and includes a new regressor, the number of mother's children aged 0-6. SCHOOL PROBIT 3 excludes another insignificant regressor, the number of children in the household. SCHOOL PROBIT 1, 2, and 3 can be compared with bivariate probit's results of Section 6.2, Tables 6.2.1 and 6.2.2.

Table 6.1.1. Probit Models of School Enrolment
(Marginal Effects at Average Characteristics)

Dependent Variable:
School = 1 if the child was in enrolled in secondary or tertiary education, and 0 otherwise.

	SCHOOL PROBIT A	SCHOOL PROBIT B	SCHOOL PROBIT 1	SCHOOL PROBIT 2	SCHOOL PROBIT 3
Individual (Child) Variables					
Female		-.014	-.014	-.014	-.014
		(.021)	(.021)	(.021)	(.021)
Age		-.151***	-.152***	-.152***	-.152***
		(.005)	(.005)	(.005)	(.005)
Muslim		-.064	-.070	-.067	-.066
		(.046)	(.048)	(.050)	(.050)
M_Age		.015	.013	.010	.008
		(.014)	(.014)	(.014)	(.013)
M_Age_sq		-.000	-.000	-.000	-.000
		(.000)	(.000)	(.000)	(.000)
M_Literacy		.135***	.128***	.130***	.131***
		(.025)	(.025)	(.025)	(.025)
M_Elementary		.043	.044	.042	.042
		(.032)	(.033)	(.033)	(.032)
M_High		.208***	.197***	.202***	.203***
School/University		(.040)	(.041)	(.042)	(.042)
M_Worked_Last_Year		.000	.008	.005	.006
		(.020)	(.020)	(.020)	(.020)
M_Arisan	.148***	.083***	.068***	.067***	.068***
	(.020)	(.026)	(.026)	(.025)	(.025)

Table 6.1.1. (Continue) Probit Models of School Enrolment

(Marginal Effects at Average Characteristics)

Dependent Variable:

School = 1 if the child was in enrolled in secondary or tertiary education, and 0 otherwise.

	SCHOOL PROBIT A	SCHOOL PROBIT B	SCHOOL PROBIT 1	SCHOOL PROBIT 2	SCHOOL PROBIT 3
Individual Variables					
F_Literacy		.127***	.119***	.119***	.119***
		(.033)	(.033)	(.033)	(.033)
F_Elementary		.006	-.002	.000	.001
		(.038)	(.038)	(.038)	(.038)
F_High		.188***	.172***	.176***	.176***
School/University		(.042)	(.041)	(.042)	(.042)
F_Worked_Last_Year		.013	.021	.020	.018
		(.033)	(.032)	(.032)	(.033)
Household Variables					
House		-.008	.023	.020	.021
		(.035)	(.038)	(.037)	(.037)
H_Electricity		.192***	.148**	.167***	.165***
		(.055)	(.063)	(.055)	(.055)
H_Head_Female		.174**	.168**	.174**	.179**
		(.073)	(.075)	(.080)	(.079)
H_Size		.007	.005	.004	-.001
		(.007)	(.007)	(.007)	(.005)
H_N_Children		-.014	-.014	-.010	-
		(.010)	(.010)	(.010)	
H_Age_Children		.017***	.016***	.013***	.013***
		(.040)	(.039)	(.004)	(.004)
M_N_Children_0-6		-	-	-.028	-.034*
				(.020)	(.020)
Village Variables					
Rural			-.054**	-.065**	-.065**
			(.027)	(.028)	(.028)
Post			.089***	.089***	.088***
			(.031)	(.031)	(.031)
Credit_Formal			.008	.007	.008
			(.024)	(.025)	(.025)
V_Electricity			.000	-	-
			(.000)		
Disaster			-.026	-.027	-.028
			(.029)	(.028)	(.028)
Improvement			.027	-	-
			(.025)		
N	5,349	5,349	5,349	5,349	5,349
Pseudo-R2	.03	.41	.41	.41	.41

Notes:

***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively. Robust standard errors clustered at district level are reported in parentheses. Province fixed effects and year dummy included.

From Table 6.1.1 we see that the positive association between maternal participation in *arisan* and enrolment is robust to the inclusion of individual and household characteristics (SCHOOL PROBIT B), and of village controls (SCHOOL PROBIT 1): controlling for individual, household and village characteristics, children whose mother participates in *arisans* are 6.8 percentage points more likely to be in school than children of non-participant mothers. As it can be seen from the table, including village variables, the social capital coefficient declines: this could be due to the fact that *arisan* participation is positively correlated with economic development. Hence, the estimated *arisan* coefficient in the model without proxies for village economic development (SCHOOL PROBIT B) could be capturing also part of the effect of economic development on education.[26]

I briefly describe in more details the results of the richest specifications (SCHOOL PROBIT 1-2-3). Among demographic variables, only age is significantly associated with enrolment: in particular, the probability of being in school decreases with age. Religion is not significantly associated with enrolment, as it is for maternal age and its square. Having more educated parents is strongly positively associated with enrolment in secondary or tertiary education: in particular, children of mothers having secondary or tertiary education are about 20 percentage points more likely to be enrolled, compared to children's of non educated mothers. Parents' literacy has also an important positive association with enrolment.

Among household characteristics, wealth as proxied by availability of electricity is positively related to children's schooling. Female headship is also significantly and positively associated with children's schooling: this could suggest that female headed households, despite being on average poorer than male headed ones, invest a greater amount of resources in children's education. The number of children in the household seems not to influence enrolment, which increases instead with average children's age. This could indicate that older children contribute to household income, combining schooling and working activities, thus making possible for the household to increase the education investment. Among village-level controls, rural residency is negatively and significantly associated with enrolment, which instead is positively linked with the presence of a post office in the village, indicating that enrolment is higher in more developed communities.

The positive association between maternal social capital and children's schooling obtained by the univariate probits is not sufficient to maintain that mothers' participation in *arisans* causes higher children school enrolment. Women who participate in *arisans*, indeed, are likely to differ from women who do not participate for a series of characteristics, as it can be seen from Table 6.1.2, which refers to the sample of mothers of children aged 13-24.

[26] This interpretation is consistent with the results of the model exploring the determinants of *arisan* participation, in which we saw that *arisan* formation is positively associated with economic development (see Section 5).

Table 6.1.2. Average maternal characteristics by *arisan* participation status

			M_Literacy	M_High School_University	M_Worked_Last_Year
1997	Urban	M_Arisan=1	.885	.423	.506
		M_Arisan=0	.764	.217	.387
		Diff	*.121***	*.206***	*.119****
	Rural	M_Arisan=1	.702	.108	.732
		M_Arisan=0	.580	.101	.536
		Diff	*.122****	*.006*	*.196****
2000	Urban	M_Arisan=1	.865	.425	.577
		M_Arisan=0	.787	.187	.538
		Diff	*.078****	*.238****	*.039*
	Rural	M_Arisan=1	.754	.155	.792
		M_Arisan=0	.601	.085	.642
		Diff	*.153****	*.070****	*.150****

Note: *** indicates that the difference between means is significant at the 1% level.

From the table, we see that participant mothers have, on average, a greater probability of being literate, of having secondary or tertiary education, and of having worked during the year previous to the interview, compared to non-participants. While using a set of controls as rich as possible helps in isolating the effect of *arisan* participation from the effect of other *observable* maternal characteristics, there would probably be *unobservable* maternal characteristics, correlated with mothers' *arisan* participation, which also influence children's schooling. This is why, besides including the most comprehensive as possible set of regressors, we use the bivariate model.

6.2 Bivariate probit models of children's enrolment and maternal participation in arisan

As anticipated in the empirical strategy section, I estimate two bivariate probit models without exclusion restrictions, and one bivariate probit including restrictions. The three estimations corresponds to the three univariate probit models which were estimated for mother's participation in *arisan* (Table 5.3), and for children's schooling (Table 6.1.1).

Both equations of the first bivariate probit (BIPROBIT 1) contain the same set of regressors, with the only exception that the schooling equation contains the dependent variable of the equation explaining mothers' participation in *arisans*. In the second bivariate estimation (BIPROBIT 2), I add to the regressors' set of both equations the variable indicating the number of mother's children aged 0-6, and I exclude the regressors indicating the percentage of household having electricity in the village and the dummy saying whether there was an improvement in the village during the year previous to the survey. I exclude these regressors, since both were insignificant in both equations. In the third bivariate probit (BIPROBIT 3), starting from the regressors' set of BIPROBIT 2, I introduce one exclusion restriction for each equation: in particular, I exclude from the schooling equation the variable counting the number of children in the household (so that this equation contains the same regressors of SCHOOL PROBIT 3 in Table 6.1.1), and from the *arisan* equation the

variable indicating child's age (so that this equation contains the same regressors of ARISAN PROBIT 3 in Table 5.3). As explained in Section 3, these exclusion restrictions are not justified theoretically, but only empirically: I exclude from the schooling equation the number of children in the household, since I saw from BIPROBIT 2 that this variable was not relevant in determining schooling, while being strongly significant in explaining mother's participation in *arisans*. Similarly, I exclude children's age from the *arisan* equation since this variable was not significant in that equation, while being strongly significant in the schooling equation. Estimation results are showed in Table 6.2.1.

Table 6.2.1. Bivariate probit for school enrolment and maternal participation in *arisans*
(Marginal Effects Computed at Average Characteristics)

Dependent Variables:
School = 1 if the child is enrolled in secondary or tertiary school, and 0 otherwise.
M_Arisan = 1 if the mother participated in *arisans* in the year previous to the survey, and 0 otherwise.

	BIPROBIT 1		BIPROBIT 2		BIPROBIT 3	
	School	M_Arisan	School	M_Arisan	School	M_Arisan
Individual (Child) Variables						
Female	-.018	.025*	-.018	.024*	-.018	.025*
	(.018)	(.014)	(.018)	(.014)	(.018)	(.014)
Age	-.129***	-.003	-.128***	-.002	-.128***	-
	(.013)	(.002)	(.013)	(.002)	(.012)	
Muslim	-.118**	.229***	-.114**	.225***	-.115**	.224***
	(.052)	(.060)	(.052)	(.059)	(.051)	(.059)
M_Age	.008	.012	.006	.011	.006	.011
	(.012)	(.015)	(.012)	(.015)	(.016)	(.014)
M_Age_sq	-.000	-.000	-.000	-.000	-.000	-.000
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
M_Literacy	.089***	.066*	.090***	.066*	.089***	.065*
	(.028)	(.039)	(.028)	(.039)	(.027)	(.039)
M_Elementary	.032	.032	.029	.033	.029	.033
	(.030)	(.054)	(.030)	(.054)	(.029)	(.054)
M_High	.149***	.147**	.147***	.152**	.145***	.152**
School/University	(.046)	(.059)	(.046)	(.059)	(.046)	(.059)
M_Worked_Last_Year	-.014	.087***	-.016	.084***	-.016	.085***
	(.019)	(.022)	(.019)	(.022)	(.019)	(.022)
M_Arisan	.370***	-	.374***	-	.380***	-
	(.086)		(.084)		(.078)	
F_Literacy	.070**	.107***	.069**	.109***	.067**	.110***
	(.030)	(.035)	(.030)	(.035)	(.030)	(.035)
F_Elementary	-.012	.050	-.010	.049	-.010	.048
	(.035)	(.054)	(.035)	(.054)	(.035)	(.054)
F_High School/University	.109**	.173***	.109**	.173***	.107**	.172***
	(.047)	(.063)	(.047)	(.063)	(.046)	(.062)
F_Worked_Last_Year	.023	-.020	.024	-.024	.023	-.022
	(.027)	(.042)	(.026)	(.041)	(.026)	(.041)
Household Variables						
House	.004	.066*	.002	.061	.002	.061
	(.033)	(.038)	(.032)	(.037)	(.032)	(.037)
H_Electricity	.116***	.013	.124***	.039	.123***	.040
	(.043)	(.065)	(.042)	(.058)	(.041)	(.057)
H_Head_Female	.190**	-.125	.193**	-.126	.193**	-.125
	(.084)	(.116)	(.087)	(.116)	(.087)	(.116)
H_Size	-.000	.020**	-.000	.019**	-.001	.018**
	(.006)	(.008)	(.006)	(.008)	(.004)	(.008)
H_N_Children	-.003	-.039***	-.000	-.034***	-	-.035***

	(.009)	(.011)	(.009)	(.012)		(.012)
H_Age_Children	.013***	.002	.011***	.001	.011***	.001
	(.003)	(.003)	(.004)	(.003)	(.004)	(.003)
M_N_Children_0-6	-	-	-.019	-.019	-.018	-.021
			(.017)	(.026)	(.016)	(.026)
Village Variables	YES, full set	YES, full set	YES, reduced set	YES, reduced set	YES, reduced set	YES, reduced set
N	5,349	5,349	5,349	5,349	5,349	5,349
rho	-.48		-.49		-.50	
P-value of rho	.0042		.0028		.0013	

Notes:
***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively. Robust standard errors clustered at district level are reported in parentheses. Province fixed effects and year dummy included.

Considering BIPROBIT 1, we see that mother's participation in *arisan* positively influences enrolment: in particular, the effect is about .37, thus greater than that obtained from SCHOOL PROBIT 1 (.068). Hence, the univariate model's social capital coefficient is downward biased with respect to the bivariate model's one: the unobservables of the *arisan* equation are negatively correlated with the unobservables of the schooling equation. This is confirmed by the statistically significant negative estimate of the *rho* coefficient from the bivariate probit model. A possible interpretation of these results is the following: among the unobservables of the univariate model there could be household vulnerability to shocks. In more vulnerable households, once controlling for household wealth and village characteristics, mothers could increase their participation in *arisans*, as a risk-coping device. At the same time, they invest less in children's education, since this investment is perceived by them as highly risky. Conversely, less vulnerable households are characterized by a smaller need to insure against shocks, compared to more vulnerable ones. At the same time, less vulnerable households can invest a higher amount in education, compared to more vulnerable ones. Hence, these results confirm the role of *arisans* as insurance devices. These estimates are also consistent with the role of *arisans* as savings mechanisms: households having more difficulties in accumulating savings will utilize *arisan* more, ceteris paribus, compared to households which accumulates savings more easily. In this case, *arisan* represents a way to afford durable goods' expenditures, among which education can be considered. At the same time, mother's in households having less savings, will, ceteris paribus, invest less in children's education.

The size of mothers' past social network could also be among the unobservables: mothers' having a smaller network, in fact, would participate more in arisan, in order to increase their network. At the same time, they would invest less in children's education, because they would perceive the investment as more risky, compared to better connected mothers.

BIPROBIT 2 includes among regressors the number of children aged 0-6, while excluding two village variables which were not significant in the previous specification (the percentage of households having electricity in the village and the occurrence of improvements during the year previous to the survey). The

estimated *arisan* coefficient and marginal effect slightly increase. This estimate, thus, confirms the role of *arisan* as a determinant of the investment in children's education.

Results concerning the other covariates confirm the conclusions of the univariate analysis: parents' education is a significant determinant of children's schooling and this is particularly true for mother's education: the impact effect is .15, indicating that having a mother with secondary or tertiary education increases the probability the child is in school by 15 percentage points, compared to having a mother with no education. Wealth as proxied by the availability of electricity in the household is also significantly positively associated with schooling, as it is female headship. As previously pointed out, female headed households, though being on average poorer than male headed ones, may invest more in children's education, because women tend to spend a greater share of household income for children's education, if compared to men.

BIPROBIT 3 considers one exclusion restriction for each equation: in particular, age is excluded from the *arisan* equation and the number of children is excluded from the schooling equation. The presence of exclusion restrictions further improves precision (the p-value of the estimated *rho* coefficient declines from .0028 to .0013). Furthermore, results of BIPROBIT 1 and 2 are confirmed.

As a further robustness check, I estimate the bivariate probit models restricting the sample to the Javanese provinces, thus dropping the Sumatra ones. I do this because the Javanese provinces are the one with the strongest tradition of community participation, and of *arisan* in particular.[27] Results are reported in Table 6.2.2.

[27] Geertz (1962).

Table 6.2.2. Bivariate Probit for School Enrolment and Maternal Participation in *Arisan*
(Marginal Effects Computed at Average Characteristics, Javanese Provinces Only)

Dependent Variables:
School = 1 if the child is enrolled in secondary or tertiary school, and 0 otherwise.
M_Arisan = 1 if the mother participated in *arisan* during the year previous to the survey, and 0 otherwise.

	BIPROBIT 1, JAVA		BIPROBIT 2, JAVA		BIPROBIT 3, JAVA	
	School	M_Arisan	School	M_Arisan	School	M_Arisan
Individua (Child)l Variables						
Female	-.047**	.035**	-.048***	.037**	-.048***	.037**
	(.018)	(.016)	(.018)	(.016)	(.018)	(.016)
Age	-.121***	-.002	-.122***	-.002	-.124***	-
	(.015)	(.002)	(.015)	(.002)	(.015)	
Muslim	-.096	.154*	-.094	.151*	-.094	.151*
	(.066)	(.080)	(.065)	(.081)	(.066)	(.081)
M_Age	.004	.008	.002	.009	.003	.008
	(.013)	(.015)	(.012)	(.015)	(.012)	(.015)
M_Age_sq	-.000	-.000	-.000	-.000	-.000	-.000
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
M_Literacy	.086***	.067	.088***	.067	.089***	.067
	(.029)	(.041)	(.028)	(.041)	(.029)	(.041)
M_Elementary	.029	.031	.028	.032	.029	.032
	(.029)	(.057)	(.029)	(.057)	(.029)	(.057)
M_High	.145***	.137**	.144***	.139**	.146***	.139**
School/University	(.048)	(.062)	(.048)	(.062)	(.049)	(.062)
M_Worked_Last_Year	-.027	.085***	-.027	.082***	-.027	.083***
	(.017)	(.023)	(.018)	(.023)	(.013)	(.023)
M_Arisan	.415***	-	.412***	-	.406***	-
	(.076)		(.074)		(.077)	
F_Literacy	.044	.122***	.042	.129***	.043	.129***
	(.028)	(.036)	(.028)	(.035)	(.030)	(.035)
F_Elementary	.003	.056	.003	.055	.004	.055
	(.033)	(.055)	(.034)	(.055)	(.034)	(.055)
F_High School/University	.140***	.168***	.142**	.168**	.145***	.168**
	(.052)	(.069)	(.052)	(.069)	(.051)	(.069)
F_Worked_Last_Year	-.000	-.051	.024	-.055	.000	-.054
	(.027)	(.044)	(.026)	(.044)	(.027)	(.044)
Household Variables						
House	.001	.147***	.003	.142***	.004	.142***
	(.035)	(.045)	(.035)	(.045)	(.035)	(.045)
H_Electricity	.089*	-.061	.079	.003	.080***	.004
	(.052)	(.082)	(.053)	(.070)	(.053)	(.070)
H_Head_Female	.244**	-.212*	.243**	-.209*	.242**	-.208*
	(.101)	(.120)	(.103)	(.120)	(.103)	(.120)

Table 6.2.2. (Continued)

(Marginal Effects Computed at Average Characteristics, Javanese Provinces Only)

Dependent Variables:
School = 1 if the child is enrolled in secondary or tertiary school, and 0 otherwise.
M_Arisan = 1 if the mother participated in *arisans* in the year previous to the survey, 0 otherwise.

	BIPROBIT 1, JAVA		BIPROBIT 2, JAVA		BIPROBIT 3, JAVA	
	School	M_Arisan	School	M_Arisan	School	M_Arisan
Household Variables						
H_Size	.001	.016*	.000	.016*	.002	.015*
	(.007)	(.008)	(.007)	(.008)	(.005)	(.008)
H_N_Children	.001	-.041***	.003	-.041***	-	-.040***
	(.009)	(.013)	(.009)	(.014)		(.014)
H_Age_Children	.014***	.000	.012***	.002	.012***	.001
	(.003)	(.003)	(.004)	(.003)	(.004)	(.004)
M_N_Children_0-6	-	-	-.019	-.016	-.017	-.014
			(.018)	(.025)	(.017)	(.025)
Village Variables	YES, full set	YES, full set	YES, reduced set	YES, reduced set	YES, reduced set	YES, reduced set
N	4,147	4,147	4,147	4,147	4,147	4,147
rho	-.55		-.54		-.53	
P-value of rho	.0007		.0005		.0011	

Notes:
***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively. Robust standard errors clustered at district level are reported in parentheses. Province fixed effects and year dummy included.

Results regarding the restricted Javanese sample are consistent with findings from the non-restricted sample. The estimated coefficient of the variable of interest slightly increases and the marginal effects is now .40, indicating that maternal participation in *arisan* determines a 40 percentage points increase in the probability a child is enrolled in secondary or tertiary education.

7. Conclusions

The univariate probit analysis show a positive statistically significant association between children's enrolment in secondary and tertiary education and maternal participation in *arisans*. This result is confirmed by the bivariate probit estimations, which also highlight a negative significant correlation between the unobservables of the schooling equation and the unobservables of the *arisan* equation. On the basis of the analysis so far, however, this negative correlation has not a clear-cut interpretation. For example, the negative correlation is consistent with the presence of households' vulnerability to economic shocks among the unobservables. The size of mother's past social network could also be among the omitted variables.

The absence of a clear explanation of the results is a limitation of the analysis, besides the problem of potential poor identification, which could stem from the absence of proper exclusion restrictions in the bivariate probit model.

Future research will aim at giving a more appropriate interpretation of the results, and at understanding the channels through which mothers' *arisan* participation influences children's schooling. From the analysis of *arisan* determinants we see that a certain level of wealth may be necessary for participation. Hence, the benefits of *arisan* would be highest for intermediate level income households. However, this seems to hold mainly for men's participation, while for women and especially for mothers wealth is a less important determinant of participation. This suggests that *arisans* could influence household behavior not only through a monetary channel, but also via information exchange and the acquisition by mothers of a greater awareness of the importance of education, and of a higher weight in the household decision making.

In order to understand the channels through which mothers' *arisan* participation influences children's schooling, I will look for more information about the actual functioning of *arisans*. In particular, it is important to understand which is the prevalent rule utilised for the 'pot' allocation. If the pot were allocated mainly randomly, in fact, the insurance role of *arisans* would be limited. Another step of future analysis consists in estimating the models in sub-samples identified on the basis of wealth: a higher *arisan* effect for the poorer sample could give support to the insurance role of *arisan.*

Bibliography

Anderson, S., Baland, J. (2002). 'The economics of ROSCAs and intrahousehold resource allocation', *The Quarterly Journal of Economics* 117(3), 963-995.

Ardener, S. (1964). 'The comparative study of rotating credit associations', *The Journal of the Royal Anthropological Institute* 94(2), 201-229.

Ardener, S., Burman, S. (1995). Money-Go-Rounds. The Importance of Rotating Savings and Credit Associations for Women, Ardener and Burman (eds), Berg Oxford.

Bandiera, O., Rasul., I. (2006). 'Social Networks and Technology Adoption in Northern Mozambique', *The Economic Journal* 116(Oct): 869-902.

Bertrand, M., Luttmer, E.F.P, Mullainathan, S. (2000). 'Network effects and welfare cultures', *The Quarterly Journal of Economics* 115(3), 1019-1055.

Coleman, J.S. (1988). 'Social Capital in the Creation of Human Capital', *The American Journal of Sociology* 94 (Supplement), S95-S120.

Durlauf, S.N. (2002). 'On the empirics of social capital', *The Economic Journal* 112(483), F459-F479.

Fafchamps, M., Minten, B. (2002). 'Returns to social network capital among traders', *Oxford Economic Papers* 54(2), 173-206.

Fernandez, R., E. Castilla, Moore, P. (2000). 'Social Capital at Work: Employment at a Phone Center', *The American Journal of Sociology* 105, 1288-1356

Frankenberg, E., Karoly, L. (1995). 'The 1993 Indonesian Family Life Survey: Overview and Field Report', RAND. DRU-1195/1-NICHD/AID.

Frankenberg, E., Thomas, D. (2000). 'The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2', RAND. DRU-2238/1-NIA/NICHD.

Geertz, C. (1962). 'The Rotating Credit Association: a "Middle Rung" in Development', *Economic Development and Cultural Change* 10(3), 241-63.

Glaeser, E.L., Laibson, D., Sacerdote, B. (2002). 'An Economic Approach to Social Capital', *The Economic Journal* 112(Nov.), F437-F458.

Greene, W. H. (2003). Econometric Analysis, London: Prentice Hall.

Granovetter (1985). 'Economic Action and Social Structure: The Problem of Embeddedness', *The American Journal of Sociology* 91(3), 481-510.

Hagan, J. and B. McCarthy, (1995). 'Getting Into Street Crime: The Structure and Process of Criminal Embeddedness', *Social Science Research*, 24, 63-95.

Helliwell, J.F., Putnam, R.D. (2007). 'Education and Social Capital', *Eastern Economic Journal* 33(1), 1-19.

Isham, J. (2002). 'The Effect of Social Capital on Fertilizer Adoption: Evidence from Rural Tanzania', *Journal of African Economies* 11(1), 39-60.

Knack, S., Keefer, P. (1997). 'Does Social Capital Have an Economic Payoff? A Cross-Country Investigation', *Quarterly Journal of Economics* 112(4), 1251-1288.

Levenson Alec R., Besley, T. (1996). 'Rosca Participation in Taiwan', *Journal of Development Economics* 51(Jan), 45-68.

Manski, C.F. (1993). 'Identification of endogenous social effects: the reflection problem', *The Review of Economic Studies* 60(3), 531-542.

Miguel, E., Gertler, P., and Levine, D. I. (2005). 'Does Social Capital Promote Industrialization? Evidence From a Rapid Industrializer', *The Review of Economics and Statistics* 87(4), 754-762.

Monfardini, C., Radice, R. (2008). 'Testing Exogeneity in the Bivariate Probit Model: A Monte Carlo Study', *Oxford Bullettin of Economics and Statistics* 70(2), 271-282.

Pitt, M., Khandker, S. (1998). 'The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter?', *Journal of Political Economy* 106 (5), 958-96.

Putnam, R. (2000). 'Bowling alone': The Collapse and Revival of American Community, New York, Simon and Schrster.

Putnam, R., with Leonardi, R., Nanetti, R. (1993). Making Democracy Work: Civic Traditions in Modern Italy, Princeton: Princeton University Press.

Strauss, J., Beegle, K. (1996). 'Intrahousehold allocations: a review of theories, empirical evidence and policy issues', MSU International Development Working Paper No. 62, Michigan State University.

Strauss, J., Beegle, K., Sikoki, B., Dwiyanto, A., Herawati, Y. Witoelar, F. (2004). 'The Third Wave of the Indonesia Family Life Survey (IFLS3): Overview and Field Report', WR-144/1-NIA/NICHD.

Thomas, D. (1990). 'Intrahousehold Resource Allocation: An Inferential Approach', *Journal of Human Resources* 25(4), 635-664.

Varadharajan, S. (2003). 'Explaining Participation in Rotating Savings Associations (ROSCAs): Evidence from Indonesia', Department of Economics, Cornell University.

Wilde, J. (2000). 'Identification of multiple equation probit models with endogenous dummy regressors', *Economics Letters* 69, 309-312.

Appendix

Variable Definitions

Children

Age. Age of individual (years).

M/F_Elementary. Dummy=1 if mother's/father's highest education level is elementary.

M/F_High School_University. Dummy=1 if mother/father has secondary or tertiary education.

M/F_Literacy. Dummy=1 if mother/father can read in Indonesian or another language.

M/F_Worked_Last_Year. Dummy=1 if mother/father worked during the year previous to the survey.

M_N_Child_0-6: Number of mother's children aged 0-6.

Muslim. Dummy=1 if the individual is Muslim, and zero otherwise.

School. Dummy=1 if the child is enrolled in secondary or tertiary education.

Sex. Dummy=1 if the individual is a male.

Adults

Elementary. Dummy=1 if mother's/father's highest education level is elementary.

High School_University. Dummy=1 if mother/father has secondary or tertiary education.

Literacy. Dummy=1 if mother/father can read in Indonesian or another language.

Sex. Dummy=1 if the individual is a male.

Worked_Last_Year. Dummy=1 if mother/father worked during the year previous to the survey.

Household

H_Electricity. Dummy=1 if the household utilize electricity.

H_Head_F. Dummy=1 if the household head is a female.

House. Dummy=1 if the households owns the house it occupies.

N_H_Child. Number of children in the household.

Age_H_Child. Average age of children in the household

Community

Credit_Formal. Dummy=1 if at least one formal credit institution is present in the community.

Post. Dummy=1 if a post office is present in the community.

Rural. Dummy=1 if the household lives in a rural area.

Improvement/Disaster. Dummy=1 if during the year previous to the survey there was an improvement/a disaster in the community, which benefited/damaged at least 30% of the population (e.g. road construction, construction of a health facility, a school/fire, flood, earthquake).

V_Electricity. Percentage of households using electricity in the community.

Chapter 3. Patterns of Internal Migration in Indonesia: Evidence From A Longitudinal Survey

Valeria Groppo
Università degli Studi di Milano

1. Introduction

This chapter analyses the characteristics of *internal* migrants in Indonesia, comparing them to those of non-migrants. The analysis entails the estimation of a model of individuals' migration decisions, as a function of a set of individual- and household-level characteristics. The study sheds light on the role of different factors in shaping internal mobility decisions in Indonesia. The main contribution of the analysis is the distinction among different forms of internal mobility. In particular, I distinguish between individual and family moves.

The analysis is performed using the Indonesia Family Life Survey (IFLS) dataset. This dataset entails a section on migration, in which respondents provide their complete migration histories, by reporting all their movements across villages which lasted at least six months, since age 12. I use the second and third waves of dataset (IFLS2 and IFLS3), corresponding to years 1997 and 2000. Taking advantage of the longitudinal nature of the dataset, I estimate individuals' migration decisions between 1997 and 2000 as a function of their individual- and household-level characteristics in 1997, before the considered migration episodes. In particular, I estimate a Multinomial Probit model, in which the individuals choose among the following three alternatives: do not migrate, migrate alone, and migrate with the family. The estimates are performed on women and males sample separately.

The analysis shows that migrants select differently according to gender and migration type. Some dimensions of selection are common to both genders, differing only by migration types (for instance, marital status and the value of household business landholdings). Other dimensions of selection, instead, have a differential impact both by migration type and gender.

This chapter also highlights that the main motivations of individual moves, both for males and females, are own work and education. Hence, this study underlines that there might be a lack of educational infrastructures and job opportunities in migrants' sending areas. Hence, from a policy perspective, the study shows the importance of interventions which increase job opportunities and educational services in regions where they are scarcer.

Our results also confirm the strong link between migration and insurance market imperfections, which is one of the key insights of the New Economics of Labour Migration.

The paper is structured as follows. Section 2 outlines the main theoretical and empirical studies on rural-urban migration. Section 3 describes the main characteristics of internal migration in Indonesia. In Section 3, the characteristics of the Indonesia Family Life Survey dataset are illustrated. Sections 5 and 6 provide the statistical and econometric analyses of individual and family migration determinants, while Section 7 concludes.

2. Literature on the determinants of rural-urban migration in developing countries

This section reviews the main theoretical and empirical studies on the determinants of migration decisions. I mainly follow the framework provided by Taylor and Martin (2001), and, given that the object of this chapter are primarily internal migrants in Indonesia, I mainly include studies dealing with internal rather than with international mobility.

2.1 Theories of rural-urban migration

The majority of the theoretical studies of internal migration focus on rural-urban moves.[28] The first strands of literature examining internal migration decisions focus on the role of sectoral or spatial income differentials.

The classical model of Lewis (1954) assumes that a capitalist and a non-capitalist sector exist, which have then been associated with urban and rural areas. The capitalist sector hires labour and sells output for a profit, and, as it expands, it draws labour from the non-capitalist sector. If the two sectors are located in different areas, the growth of the capitalist sector determines migration of workers across areas. The main assumption of this model is that labour is available in unlimited supply to the capitalist sector, at a fixed real wage. The unlimited surplus of labour in the non-capitalist sector implies that, in this sector, the marginal product of labour is zero and labour can be transferred without loss of agricultural output. Key testable assumptions of the model are that rural out-migration is not associated with a decrease in agricultural output or with an increase in urban or rural wages. The neoclassical two-sector model of Ranis and Fei (1961) extends the model of Lewis (1954), by assuming that through migration the marginal products of labour (and thus wages) are equalized across areas. The key testable prediction of this model is the convergence of wages across sectors. The above considered models assume full employment in the urban sector and focus on inter-sector wage differences in explaining rural-urban migration. These models have been criticised mainly because they are not consistent with empirical evidence of continuing rural-urban migration even in the presence of high urban unemployment and of persistent differences in the returns to homogeneous labour across sectors.

The models of Todaro (1969) and Harris and Todaro (1970), by incorporating *labour market imperfections*, provide an explanation of why migration continues even in the presence of high urban unemployment. Also these models emphasize the role of income differentials in explaining rural-urban

[28] However, there are exceptions. For instance, the theoretical model considered by Rosenzweig and Stark (1989), explicitly refers also to rural-to-rural migration. In my analysis about Indonesian migration, I will consider rural-to-urban, urban-to-urban, and rural-to-rural migration.

moves, but depart from the neoclassical models because they do not assume full employment in the urban sector. In the Todaro models, the urban sector is indeed characterised by unemployment, which implies that migrants' earnings depend not only on the level of urban wages, but also on the probability to find a job. Indeed, migration decisions depend on the differences between rural and urban *expected* rather than *actual* earnings, where expected earnings are defined as the *present value* of urban real income streams over the worker time horizon.[29] Expected earnings are determined by the actual rural-urban wage differentials and by the probability for migrants to find a job at the urban destination. Given these quantities, rural-urban migration continues as long as rural earnings are lower than expected urban ones. In particular, the rural-urban gap in actual earnings may be so high that, even in the face of high urban unemployment (and thus lower probability for migrants to find a job), urban expected income is higher than the rural one and thus migration is economically rational. This model is consistent with migration happening even in the face of low probability to find a job in the initial period in urban areas, as long as the present value of future income streams compensate for the initial losses. Therefore, Todaro (1969) concludes that any effort devoted to decreasing urban unemployment might be worthless if it is not accompanied by a concomitant effort in reducing the gap in living standards between urban and rural locations.

The models above adopt a macro perspective, and do not explain how migrants self-select, i.e. why certain individuals migrate and others do not, given a certain level of wage differentials. The human capital models of migration adopt instead a micro perspective, focusing on the roles of human capital in determining migrants' selection. According to the seminal contribution of Sjaastad (1962), spatial earnings differentials have certainly a role in determining migration decisions. However, there are cases in which earnings differentials are not necessarily associated with migration from the lowest to the highest earnings locations, at least within certain age groups. Hence, earnings differentials are only part of the story.[30] To clarify which other factors influence migration decisions, the author treats migration as a form of investment, entailing private and social costs and returns. Therefore, another important factor influencing migration is *age*: migration is especially undertaken by young individuals, given that they have a longer time horizon over which to amortize investment costs and to reap its benefits. However, the author observes a sharp decline in migration rates at very young ages, which is not completely explained by migration costs and length of amortization period. In this respect, the study underlines the role of *skills*: it may be possible that, upon migration, migrants are required to change occupation. This means that migration only happens if the complementary human capital investment is undertaken besides the migration one. It is clear that only the youngest (among the young) would find it profitable to invest in further human capital accumulation. Hence the author concludes that the migration investment cannot be treated in isolation, and any analysis of migration decisions must take into account complementary investments that the individual may concomitantly decide to (or be required to) undertake, most importantly investments in human capital. The

[29] In other words, the decision to migrate is taken on the basis of a "permanent income calculation" Todaro (1980).
[30] Differently from Todaro (1969) and Harris and Todaro (1970), Sjaastad (1962) refers to *current*, not to *expected* earnings differentials.

importance of human capital investments besides the migration one can also explain why in certain instances migration does not take place in face of spatial income differences.

The models outlined above focus on the individual as the decision-making unit, and are unable to explain a range of common phenomena in LDCs, such as the presence of geographically extended families and the consistent flow of remittances from higher earnings migrants to lower earnings families at origin. The New Economics of Labour Migration (NELM) explains these facts, modelling migration decisions as the result of complex interactions between households as a whole an their individual members, and focusing on the role of insurance and capital market imperfections.[31] In particular, departing from Todaro's assumption of a *risk neutral individual* as decision making entity, Stark and Levhari (1982) underline the role of migration as a strategy undertaken by *risk averse households* to spread and thus mitigate risk. According to this study, households which lack risk diversification opportunities and access to insurance markets in rural areas are able to reduce total income risk by sending members in urban locations characterised by risk profiles that are imperfectly correlated with the rural ones.[32] This study, which maintains the so called 'portfolio diversification approach', underlines that increasing gains from longer migration distance ('*diversification gains'*) come at increasing costs from monitoring and enforcement ('*contractual costs'*). However, the family ties which characterise household's members migration loosen this trade-off, thus making the migration investment enforceable (Lucas and Stark, 1985; Rosenzweig, 1988). Rosenzweig and Stark (1989), starting from the observation that in India rural to rural women's moves predominate relatively to rural to urban men's moves, focus on the role of women's marriage migration as a means allowing households to diversify income risk. The authors highlight that in rural developing settings this type of migration improves households' welfare even in the absence of inter-village wage differentials, and in the presence of identical village-specific risk distributions, as long as the timing of events are not perfectly synchronized.[33] This 'risk-theoretic approach to marriage migration' predicts that individuals of the same household will go to different destinations and that wealthier households (which are those with better insurance) will invest less in marriage migration, choosing shorter distances locations. As underlined by the authors, these predictions contrast with the ones of the standard marriage models of Becker (1973, 1974). These models, indeed, predict that individuals from the same household will go to the same destination, and that individuals with relatively more rare traits (e.g. wealthier individuals) will invest more in search, by choosing longer distances locations. Hence, the distance-wealth relationship represents a relatively strong test of the two approaches to marriage.[34]

Katz and Stark (1986), assuming a *risk averse individual* as decision making unit, focus on the role of *capital market imperfections* and *relative wealth* in determining rural-urban migration. The authors underline

[31] See Stark and Bloom (1985).
[32] The role of migration as a risk diversification strategy equally holds when migration is viewed as an individual decision. In the face of a highly risky rural environment, a risk-averse individual would find it optimal to accept a temporary increase in risk by migrating, as long as the expected risk in future periods in urban areas is lower than the one prevailing in rural areas Stark and Levhari (1982).
[33] Rosenzweig and Stark (1989), p. 909.
[34] Rosenzweig and Stark (1989), p. 911.

that in rural developing settings, characterised by limited information, high transaction costs, and marked barganing power asymmetries, poor individuals may be unable to access capital markets, and thus to realize the full potential values of their wealth. In these contexts, migration may provide an alternative means to fully reap these gains. They also provide a second explanation of migration, focusing on the fact that individuals' utility depends also on the status provided by wealth. With this assumption, migration represents a way to increase individual's utility, through the individual's advancing in his or her position in the wealth ranking relative to other individuals.

The study of Hoddinott (1994) includes both the individual's and the household's role in determining migration and remittances behavior, generalizing the Todaro and NELM approaches. This study shows that both household-level and individual-level characteristics influence the migration decision. In particular, the study considers the role of household *land* in determining migration decisions, and highlights that from a theoretical point of view households' landholdings have an ambiguous effect on the probability of migration by household members. Indeed, higher landholdings may require more labour, thus discouraging migration. However, they may also be associated with higher wealth, which favours migration by allowing households to pay the costs of their members' moves. In the context of Kenya, the author finds that the second implication is verified, i.e. that higher household landholdings increase migration.

The above theories are not mutually exclusive. Instead, they are complementary, in that they underline different aspects of the migration decision process. Therefore, in my empirical anaylsis I will take into account the main hints of the different theoretical approaches above outlined, highlighting whether and to what extent they are verified in the case of internal migration in Indonesia. Before passing to the Indonesian case, however, the following section reviews the main empirical studies on the determinants of rural-urban migration.

2.2 Empirical evidence on the determinants of rural-urban migration

Several econometric studies provide supportive evidence of the Todaro models (see, for instance, Perloff *et al.*, 1998).

In the NELM framework, Rosenzweig and Stark (1989) test the hypothesis that households facing exogenously higher agricultural profit variability invest more in marriage migration, as a means to reduce risk. Using data on 6 Indian villages observed over 10 years, the authors find that households with higher levels of profit variability have a higher number of migrants, and send them to relatively more distant locations.[35]

The study of Mendola (2008), based on the NELM theoretical framework, focuses on the role of migration in households' technology adoption decisions. Using cross-sectional data on rural Bangladesh, the

[35] In particular, Rosenzweig and Stark (1989) consider a model having as dependent variable a measure of migration (the number of migrants or mean marriage distance), and as independent variable of interest the variance of agricultural profits over 10 years (1975-84). Since profit variance is endogenous, it is instrumented using village-levels means and variances of rainfall in the considered period, as well as interaction terms between rainfall statistics and head's dry and irrigated landholdings at inheritance.

study dinstinguishes different types of migration (temporary and permanent domestic, and international), and highlights that the patterns of households' selection in migration depend on the type of migration. For instance, households with international migrants are on average wealthier and larger than households without migrants, while households with temporary migrants are relatively worse off.

Kuhn (2002) analyzes the different determinants of individual and family migration in rural Bangladesh. The author underlines the main differences between family and individual migration: (i) family migration includes more women and children, (ii) it also requires the acquisition of a permanent dwelling instead of living with relatives or in a hostel, (iii) it separates a migrants' parents from a daugther-in-law, a primary source of elderly support, (iv) it weakens rural-urban ties, and (v) it transfers not only the production of a family, but also its consumption. The study concludes that family migration is more likely than individual migration among men from households that lack land and the support of other relatives.

In my analysis, I consider both individual-level and household-level determinants of migration, thus considering the implications of the NELM theoretical framework. As Kuhn (2002), I distinguish between family and individual migration. My study differs from the one of Kuhn (2002), mainly for the following reasons: while the above author only includes married male adults in his sample of analysis, I also consider females and unmarried individuals, pointing out the differences between the selection patterns of males and females. Moreover, the above study only includes rural-urban moves, while I also consider urban-urban moves. I also analyze the role of individuals'and households' income in influencing migration decisions, which is not considered by Kuhn (2002). Kuhn (2002) uses an unbalanced panel dataset, in which person-year observations are pooled and migration in year $t+1$ is modelled as a function of contemporaneous year $t+1$ characteristics. I instead use a balanced panel, estimating migration choices in period $t+1$ as a function of individual and household characteristics in period t (see Sections 5 and 6). Finally, the above study uses a Multinomial Logit model, which requires that the hypothesis of Independence of Irrelevant Alternatives (IIA) is verified. The analysis of my study is instead performed using a Multinomial Probit, which does not require the IIA property.

3. The context: migration in Indonesia

This section describes the main patterns of migration in Indonesia. Given that the focus of the chapter is on internal migrants, here the characteristics of internal migration are primarily considered.

Indonesia is the largest country in South East Asia, and the fourth most popolous country in the world. Its population steadily grew in the last fifty years, from about 92 millions in 1960, to 150 millions in 1980, reaching 239 millions in 2010.[36] The country is constituted by about 17,000 islands, characterised by enourmous diversity along many dimensions, such as economic development, population density, ethnic composition and culture.[37] Historically, economic development and population density were higher in the central island of Java, followed by the islands of Sumatra, Sulawesi, and Kalimantan.

[36] Source: World Bank Development Indicators.
[37] About 300 ethnic groups and six officially recognized religions exist in the countries (MacAndrews, 1978).

Migration in Indonesia is not a recent phenomenon. Since 1905, indeed, Indonesia had a public *transmigration program*, which transferred people from Inner to Outer regions of the country. According to MacAndrews (1978), two main phases of this program can be distinguished. The first pre-independence colonial program moved about 190,000 persons between 1905 and 1941, with Sumatra as main destination. In the period 1950-1974, the Indonesian National Program moved about 787,000 persons, with the Outer Islands becoming increasingly important as transmigrants destinations. The change in destination coincided with the shift in the primary aims of the program. In the period 1950-1970, indeed, the transmigration program was mainly aimed at reducing population pressure in Java, providing labour in sparsely populated provinces, and increasing national security (MacAndrews, 1978). During the 1970s, the government stressed more the role of transmigration as a means to spur regional development, rather than to limit population pressure in Java (Tirtosudarmo, 2009). This change of objectives concided with a period of increased public investments in infrastructure, communication and education, which also encouraged people mobility. Despite its ambitious objectives, the program was closed in 2000 due to lack of funds which was mainly linked to the 1997-98 Asian financial crisis. Two types of transmigrants can be distinguished (MacAndrews, 1978): "general transmigrants", which are fully supported by the government from their departure to their arrival and settlement at destination, and "spontaneous transmigrants", who move on their own expenses, but participate in government projects once at destination. Besides these two main groups, there was a large flow of spontaneous migrants who settle either in empty land or in land they rent from other occupants. As for the program evaluation, it is clear that the program had only a minimal impact in diminishing population pressure in Java. In the period 1905-1977, indeed, the program moved about 990,000 individuals from Java, where the population increased by 35 million people in the same period (MacAndrews, 1978). As for regional development objectives, the program did design regional development plans, but their implementation, as the implementation of the whole transmigration program, often suffered from lack of funds, authority, and coordination among the responsible agencies. MacAndrews (1978) also underlined the program's weaknesses in increasing food production, given the low fertility and productivity of Outer Islands' land compared to the land in Java. Notwithstanding these weaknesses, the program succeeded in increasing transmigrants' living standards, by providing income and facilities to them. Moreover, the program has the merit of having indirectly stimulated the increase of voluntary migration in a society which was highly immobile until the 1960s.

From the 1970s, migration within and from Indonesia steadily increased, driven by multiple forces, which are mainly the following (ILO, 2004):

- spatial mismatches between the areas where job opportunities were arising and those where workers were living. For instance, investments in manufacturing activities were concentrated in Java, especially from the second half of the 1980s, while extraction activities were mainly concentrated in the Outer Islands. In this regard, it is important to underline that inequality had a relatively smaller role as a determinant of internal migration in Indonesia, where access to land and education are more equal compared to other settings, such as China (IOM, 2008);

- marked improvements in education levels, which spurred young people to leave rural areas in search for alternatives to agricultural employment;
- development of infrastructures, public transportation, and increasing ownership of motorvehicles;
- increased commercialization and mechanization of agriculture, with labour inputs progressively replaced by capital inputs;
- presence of local and regional conflicts, which spurred peoples' to move in seach for a securer living areas.

In the early 1990s, the majority of internal migrants moved from Outer Islands to Inner Indonesia. This flow was partially reverted with the financial crisis that hit the country in 1997-1998, when emigration out of Java increased and immigration to Java declined. Ethnicity represents another element which influenced people mobility in Indonesia, where different ethnic groups have indeed traditionally different propensities to move.

In the last two decades, the average distance over which migrants are moving increased, with the number of inter-provincial movers drastically growing in this period. One of the most important features of Indonesian internal migration is the fact that it is mainly non-permanent in nature. As such, it is more difficult to be recorded by standard surveys. As underlined by ILO (2004), temporary migration has several advantages over permanent migration. For instance, it is more flexible and it allows individuals to divide their time among different activities (in the village and in the urban areas), thus diversifying risk. Moreover, given the higher costs of living in urban than in rural areas, working in urban settings temporarily while maintaining the family in rural areas, increases households' purchasing power in those areas. In addition, rural settings are often perceived more appropriate as places where to grow children, given that urban areas are frequently characterised by relevant environmental and security problems. Temporary migration also allows individuals to maintain social ties with the village of origin, which could be a source of financial and psychological support. In this regard, it has to be emphasized that social networks are crucial in determining the patterns of internal migration in Indonesia: most temporary migrants indeed, make their first movement with other experienced migrants or join friends or relatives at destination (ILO, 2004). Another important feature of recent migration flows in Indonesia is the increase of migration by women, with females' movements outnumbering males' ones in some areas. Both unskilled and skilled women move: women with lower education generally work in the informal sector, mainly as domestic servants, while women with secondary education work in the formal sector, mostly in the factories of Jakarta's metropolitan region (ILO, 2004).

4. Data

4.1 IFLS sampling scheme

IFLS is a longitudinal survey which was first fielded in 1993 (IFLS1). Subsequent waves were administered in 1997 (IFLS2), 1998 (IFLS2+), 2000 (IFLS3), and 2007 (IFLS4).[38]

In this paper, I only use IFLS2 and IFLS3, since these are the only two waves for which it is possible to have comparable values of real monetary variables, deflated both spatially and temporally. Indeed, spatial and temporal deflators constructed in the same way (with base Jakarta December 2000) are only available for IFLS2 and IFLS3. For IFLS1 and IFLS2, there is a general CPI by province, that converts nominal values to 1986 values. This index, however, is only for urban areas and it is only temporal. Therefore, with the available IFLS data, it is not possible to convert 1993 to 2000 values. However, in the overview that follows, I will describe also IFLS1, since it provides the basis for all the successive waves.

IFLS1 was administered in 13 out of 27 Indonesian provinces and it is representative of 83% of the 1993 Indonesian population.[39] Frankenberg and Karoli (1995) provide details on the IFLS1 sampling scheme: the survey stratified on provinces and rural-urban locations, then randomly sampled within these strata. Within each of the 13 provinces, Enumeration Areas (EAs) were randomly chosen from a sample frame used in another socioeconomic survey administered in Indonesia, the 1993 SUSENAS.[40] Urban EAs and EAs in smaller provinces were oversampled to ease rural-urban and Javanese-non Javanese comparisons. Within EAs, households were randomly selected on the basis of the 1993 SUSENAS listing. In IFLS1, interviews were conducted with 7,224 households, defined '*panel* households'. The total number of individuals living in these households in 1993 was 33,081. Basic information was recorded for each household member. However, to reduce survey costs, a subsample of household members was defined to be eligible for *detailed* interviews. This subsample is constituted by the following individuals, defined '*panel respondents*':

- the household head and his/her spouse
- up to two of their children aged 0 to 14, who were randomly selected
- an individual aged 50 or older and his/her spouse, randomly selected from remaining members
- for a randomly selected 25% of households, an individual aged 15 to 49 and his/her spouse randomly selected from the remaining members.

Eligible panel respondents were 22,588 in IFLS1. Of these, 97.5% completed the interviews, and 2.5% refused (Thomas *et al.*, 2010).

For panel respondents the dataset includes both household-level data, such as location, demographic composition, expenditure and assets, and individual-level data, including educational, marital, employment and migration histories. For all IFLS waves , community-level data were also collected, by interviewing the

[38] IFLS2+ was fielded in 1998, to evaluate the immediate impact of the Asian economic crisis that hit Indonesia since January 1998. It only included 25% of the IFLS household sample, in 7 out of the 13 IFLS provinces.

[39] The Indonesian provinces sampled in IFLS1 were the following: four provinces on Sumatra (North Sumatra, West Sumatra, South Sumatra and Lampung), the five Javanese provinces (DKI Jakarta, West Java, Central Java, DI Yogyakarta, and East Java), and four provinces covering the remaining major islands groups (Bali, West Nusa Tenggara, South Kalimantan, and South Sulawesi).

[40] The 1993 SUSENAS sample includes 60,000 households.

head of the community about, among others, available transportation means, communication, sanitation infrastructure, credit opportunities, and economic changes.

4.2 IFLS household-level attrition rate

IFLS2 had the goal of re-contacting all 7,224 IFLS1 households, defined '*panel households*' (Frankenberg and Thomas, 2000). If an entire panel household had moved in the periods between different waves, it was tracked, as long as it still resided in one of the 13 IFLS1 provinces.[41] The rule of tracking only those households who moved to another IFLS province could be a source of bias, if the number of households who moved to a non-IFLS province is high. However, in the period between IFLS1 and IFLS2, less than 1% of IFLS1 households moved to a non-IFLS province or to another country, while 1% moved to an unknown location.[42] Hence, we can exclude that the tracking rules biased the selection of the IFLS2 sample, with respect to the original IFLS1 sample design. In 1997, 94.4% of households were re-contacted and 93.3% of IFLS1 households were re-interviewed, in the sense that at least one person of the IFLS1 household was re-interviewed. This re-contact rate is very high compared to the one of other surveys.[43] As it can be seen from table 1 below, of the IFLS1 households that were re-interviewed in 1997, 90.9% had not moved at all, and another 3% had moved within village. A total of 6% were interviewed outside their 1993 village.

Table 1. Location of IFLS1 origin households re-interviewed in IFLS2

	N	%
Did not move	6,125	90.8
Moved of whom	617	9.2
Within village	203	3.0
Within sub-district (*kecamantan*)	99	1.5
Within district (*kabupaten*)	120	1.8
Within province	122	1.8
Another IFLS province	73	1.1
Total IFLS1 households re-interviewed in IFLS2	6,742	100.0

Source: IFLS2.

[41] In 2000, also individuals who moved to the province of Riau (which is not one of the 13 original IFLS1 provinces) were tracked. A small number of households were also followed in Southeast Sulawesi and Central and East Kalimantan, because their locations were near the borders of IFLS provinces and thus within cost-effective reach (Strauss *et al.*, 2004).

[42] Source: ILFS2.

[43] A similar re-contact rate is also obtained by the Kagera Health and Development Survey (KHDS), used by Beegle *et al.* (2011). As underline by these authors, the attrition rate of the KHDS dataset is very low compared to that of other panel surveys, summarized in Alderman *et al.* (2001) and ranging from 17.5% to 1.5%.

IFLS3 had the goal to re-contact all IFLS1 original households having living members the last time they were contacted, plus IFLS2 and IFLS2+ split-off households. A total of 10,574 households were contacted in 2000, of which 7,928 were IFLS3 target households and 2,646 were new split-off households. Thus, 95% of IFLS3 target households were actually re-contacted. A total of 10,345 households were actually re-interviewed, with the difference between re-contacted and re-interviewed being due to the presence of households whose all members died between the survey rounds or to households that merged with others. The following table reports the household-level re-contact rates among IFLS1-2-3.

Table 2. Household-level re-contact rate between IFLS1 and IFLS2[a]

	IFLS1	IFLS2 target HHs contacted	Recontact rate (%)	IFLS3 target HHs[b]	IFLS3 target HHs contacted	Recontact rate (%)[c]
IFLS1 HHs	7,224	6,821	94.4	7,138	6,800	95.3
IFLS2 split-off HHs	-	877	-	865	819	94.7
IFLS2+ split-off HHs	-		-	344	309	89.8
IFLS3 target HHs	-		-	8,347	7,928	95.0
IFLS3 split-off HHs	-		-	-	2,646	-
Total HHs contacted	7,224	7,698	-	-	10,574	-

Source: Strauss et al. (2004)
Notes:
a. The number of re-contacted households includes those households whose members all died and households that recombined into other households since the last survey.
b. IFLS3 target households are IFLS1 households, IFLS2 split-off households, and IFLS2+ split-off households known to have at least some members living in the last survey.
c. Re-contact rates are out of IFLS3 target households.

As underlined by Strauss et al. (2004), because of movers, the geographical distribution of households in 2000 changed since 1997 and 1993. As indicated in the following table, of the IFLS2 households that were re-interviewed in 2000, 91.6% had not moved since 1993, and another 3.4% had moved within the village. Only 5% of tracked households were found outside the origin 1997 village. By contrast, over half of split-off households are in a different village from the 1993 one, 16.5% moved to a different district and about 13% moved to a different province. Besides showing the importance of tracking, the tables below indicate that "there has been a considerable amount of moving by split-off households, though not by non-split-offs".[44] The same pattern emerges by comparing IFLS1 and IFLS3 locations of IFLS1 origin and split-off households. It is important to underline that episodes of migration across villages of the entire IFLS1 origin household as a unit are very rare. This implies that when we refer to 'family migration' in the next sections, we focus on moves of nuclear families from origin households whose other members, in most cases, do not move.

[44] Strauss et al. (2004), p. 7.

Table 3. Households re-interviewed in IFLS3: Relocations since IFLS2

	All HHs interviewed		IFLS1 origin HHs		Split-off HHs	
	N.	%	N.	%	N.	%
Did not move	7,634	73.2	6,192	91.6	1,442	39.2
Moved within village	683	6.5	228	3.4	455	12.4
Moved within district	360	3.4	85	1.3	275	7.5
Moved within sub-district	483	4.6	75	1.1	408	11.1
Moved within province	708	6.8	101	1.5	607	16.5
Moved to another IFLS province	496	4.8	65	1.0	431	11.7
Moved to non-IFLS province	71	0.7	12	0.2	59	1.6
Total	10,435	100.0	6,758	100.0	3,677	100.0

Source: Strauss et al. (2004).

Table 4. Households re-interviewed in IFLS3: Relocations since IFLS1

	All HHs interviewed		IFLS1 origin HHs		Split-off HHs	
	N.	%	N.	%	N.	%
Did not move	6,098	58.4	5,573	82.5	470	12.8
Moved within village	1,278	12.2	474	7.0	859	23.4
Moved within district	601	5.8	204	3.0	397	10.8
Moved within sub-district	693	6.6	174	2.6	519	14.1
Moved within province	1,001	9.6	194	2.9	807	21.9
Moved to another IFLS province	690	6.6	126	1.9	564	15.3
Moved to non-IFLS province	74	0.7	13	0.2	61	1.7
Total	10,435	100.0	6,758	100.0	3,677	100.0

Source: Strauss et al. (2004).

4.3 IFLS individual-level attrition rate

IFLS2 had the goal to re-interview the 22,588 IFLS1 panel respondents, as well as those IFLS1 household members who were at least 26 years old in 1993. These two categories of individuals were defined '*target respondents*' (Frankenberg and Thomas, 2000). Individuals who left their 1993 household by 1997 were tracked, if: (i) they were target respondents and (ii) available information indicated that they were residing in one of the IFLS1 provinces. The new households in which the individuals were found were defined '*split-off households*'.[45] As it can be seen in the next table, in IFLS2 there were 23,049 target respondents, thus eligible to be tracked. Of these, 91% were interviewed, 1% refused, and 7.5% were not found.[46]

[45] Once a household was found, the rules for interviewing its members differ between origin and split-off households. In origin households, all members including new entrants since 1993 were eligible for interview, while in split-off households only target respondents, their spouses and their biological children were to be interviewed.

[46] As a comparison, in the Kagera Health and Development Survey (KHDS), which is another tracking survey characterized by low attrition, 82% of the baseline respondents were re-interviewed.

Table 5. IFLS completion rates: individual respondents

	IFLS1 (1993)		IFLS2 (1997)		IFLS3 (2000)	
	N	%	N	%	N	%
1. Eligible for survey			33,081		39601[a]	
2. Died between the waves			854		790	
3. Eligible for survey / alive at survey date (row 1-row 2)	33081[b]		32,227		38,811	
4. Assessed (% = row 04 / row 03)			26,948	83.6	32,586	84.0
5. Eligible to be tracked and contacted, of whom	22588[c]		23,049		32,189	
6. Interview conducted	22,019	97.5	21,073	91.4	29,440	91.5
7. Refused	569	2.5	244	1.1	261	0.8
8. No interview conducted			1,732	7.5	2,488	7.7
9. New entrants in this wave	-		5,404		6,104	
10. Total sample interviewed in this wave (row 4+row 9)	22,019		32,352		38,690	
11. Total potential sample for the next wave (row 3+row 9)	33,081		37,631		44,915	

Source: Thomas et al. (2010), table 1, p. 5.
Notes:
a. There are 1,970 new entrants in 2000 from a special 1998 survey of a sub-sample of respondents.
b. Respondents included in 1993 baseline.
c. Respondents selected for individual interview in 1993.

IFLS3 was fielded in 2000 with the aim of re-contacting both IFLS2 respondents and individuals who were eligible in IFLS2, but who were not interviewed in 1997. Therefore, a respondent who had not been found in a wave could be recovered in a subsequent wave (Thomas et al., 2010). The total potential sample of IFLS3 included almost 40,000 individuals. As for household, the tracking rule implied that only individuals who moved within IFLS provinces were eligible to tracking. The definition of target respondents was expanded in IFLS3, including not only IFLS1 panel respondents and IFLS1 household members who were at least 26 years old in 1993, but also other individuals, such as those born since 1993 in IFLS1 households.[47] As it can be seen from the previous table, in IFLS3 91.5% of target respondents were interviewed, 0.8% refused, and 7.7% were lost to follow-up. Hence, IFLS3 completion rates are very similar to those of IFLS2. The lowest re-interview rates were found for individuals aged 15-19 in 1993. This was due to the high mobility rates in this age group, but also to IFLS tracking rules, which did not consider individuals in this age range in the past, and only a random subsample of them in 2000 (Strauss *et al.*, 2004). A total of about 16,000 individuals were added since 1993, due to the presence of spouses, children, and other relatives who moved into existing households. There were 17,990 individuals who had individual interviews in IFLS1, 2 and 3. This corresponds to about 82% of IFLS1 panel respondents.

IFLS4 used the same re-contact protocols of IFLS3, and registered 12% of individuals lost to follow-up.

[47] In particular, the category of 'target respondents' as defined in IFLS3 included: IFLS1 'panel respondents', IFLS1 household members who were at least 26 years old in 1993, individuals born since 1993 in origin IFLS1 households, individuals born after 1988 in origin IFLS1 households, IFLS1 household members born between 1968 and 1988 if interviewed in 1997, a 20% random sample of IFLS1 household members born between 1968 and 1988 if not interviewed in 1997. The reason of the expansion of the target categories was to follow children and young adults born in panel households, thus keeping the sample representative of the 1993 population in the 13 considered provinces.

4.2 Migration data in IFLS

IFLS provides both household-level and individual-level information on migration. As anticipated in the previous section, the survey includes a variable indicating whether a household moved between the survey rounds. This variable also indicates whether the move was within village, sub-district, district, province, or across provinces.

Individual-level information on migration is provided in two different sources in IFLS: first, the household roster indicates, for each household member that was listed in a previous round, whether he or she is in the household, or not. If not, the month and year of departure are recorded, together with the reason of the move, the current location (whether in the same village, sub-district, district, province or in another province), and other information such as education level and total monthly income. This source is only available for individuals who changed household between two waves (either entering another existing household, or forming a new split-off household), and only indicates one episode of migration (the one of departure from the household).

The second source of individual-level mobility information is the migration history module, which is answered by target respondents, as defined in the previous paragraph. Individuals are asked to state whether they moved crossing village boundaries, staying at least 6 months since age 12. For each move, individuals reported the year, the destination, the distance (km), the reason, and who moved together with the migrant at the time of the move.

The sample used in this paper mainly comes from the migration history module, given that it was answered by target individuals, and not by proxy respondent. Hence, the information it gathers should be more accurate than the one of the roster. In the sample of analysis, I define migrants those individuals who reported at least one move in the period between IFLS2 and IFLS3 (1997-2000).

I take into account two relevant migration characteristics, which are the reason of migration and whether the individual moved alone or with his or her family. If the individual reported more than one move in the period 1997-2000, I only consider the first move that he or she did. For instance, if the individual moved, say in 1998 alone and then again in 2000 with the family, I only consider the move he or she did alone and classify the individual as individual migrant.

In 2000, migrants were interviewed either in the same households they were in 1997 (in this case they were already back by the time of the 2000 interview), or in a new household (in this case they were still at destination by 2000 and they were tracked). As above mentioned, in IFLS when an individual was tracked, the household in which he or she was found was given a new identification code and classified as a new split-off household. This implies that split-off households are not necessarily occupied by an individual who moved with the spouse for marriage reasons. Differently, we can find split-off households which are occupied by only one individual, if he or she moved alone (for any reason) and was tracked at destination.[48]

[48] IFLS3 includes a variable saying whether the household in which the individual was interviewed in 2000 was a split-off household or not. However, this variable only says that between 1997 and 2000 the individuals changed household. We do not know the year of the split-off, so we cannot associate it to the migration episodes we observe for individuals, for which we also have the year in which the move happened.

The sample includes both individuals who migrated for the first time in their lives in the considered period, and individuals who migrated for the second or third time. We can call the individuals in the first category *new migrants* and those in the second *past migrants*. In the econometric analysis, I will estimate both models which do not control for past migration (i.e. migration before 1997), and models which control for it.

Table 6 below reports the reasons of individuals' moves for the same sample used in the statistical and econometric analyses, with the only difference that this table also includes marriage migration, which is instead excluded in the empirical analysis that follows. I excluded individuals who moved for marriage, pregnancy or divorce, because moves related to these events are likely to respond mainly to non-economic factors. This is especially true for marriage moves, which, in Indonesia, are highly influenced by traditional laws and customs (Buttenheim and Nobels, 2009).

Table 6. Reasons of migration (first move done in the period 1997-2000)

	Men		Women	
	N.	%	N.	%
Work (self)	352	40.9	205	20.9
Work (others)	19	2.2	51	5.2
Education (self)	99	11.5	99	10.1
Education (others)	1	0.1	4	0.4
Marriage/Pregnancy/Divorce	89	10.3	141	14.3
Sickness/Death/Shocks	26	3.0	26	2.6
Migration with family	56	6.5	139	14.1
To be close to family	69	8.0	115	11.7
Other	149	17.3	203	20.7
Total	860	100.0	983	100.0

Source: IFLS2-3.
Note: the category "Other" aggregates the following original IFLS categories: like the destination, new housing opportunities, other reasons.

As it can be seen, the most common reason for both males' and females' migration are own work, and other reasons.[49] The third most common migration reason is own education for males and marriage/pregnancy/divorce for females. The category "Migration with family", which is associated with 6.5% and 14.1% of males' and females' moves respectively, does not allow to understand the specific reason of the move. Indeed, an individual could have migrated with his or her family because of work, or for the education of children, or because of a new housing opportunity. Therefore, the IFLS classification does not allow to disentangle properly migrants by the reason of their moves. However, the data entail the questions: "Did you move with other householders?" and "Who moved together with you at the time of the move?". On the basis of these questions, I distinguish individuals who moved together with their (nuclear) family, from those who did not. In particular, I define 'family migrants' individuals who moved with their spouse and

[49] I specified the category "Other" by aggregating the following original IFLS categories: like the destination, new housing opportunities, other reasons.

children (in case they have), or with their parents while 'individual migrants' are those migrants who did not move with other householders, or who moved with householders other than the spouse o parents.

As previously explained, in the analysis that follows, I do not consider those individuals who moved for marriage/pregnancy/divorce reasons. Hence, the category 'individual migrants' includes individuals who moved alone for any of the above categories, excluded "Marriage/Pregnancy/Divorce".[50] Instead, family migrants include individuals who moved with spouse and/or children and/or parents for any reason excluded "Marriage/Pregnancy/Divorce".

The category "To be close to family" mainly regards return movers, who previously migrated for work, education or other reasons.

The following table shows the reasons of migration, distinguishing between individual and family migration, and excluding the marriage/pregnancy/divorce category. We can see that, for both men and women, the primary reasons for moving alone are own work and education, followed by the willingness to be close to family and other reasons. Family migration is mainly due to "Other" reasons, followed by own work for men and "Migration with family" for women.

Table 7. Reasons of migration by type of migration (first move done in the period 1997-2000)

	Individual Migration				Family Migration			
	Men		Women		Men		Women	
	N.	%	N.	%	N.	%	N.	%
Work (self)	262	57.7	162	33.8	90	28.4	43	11.8
Work (others)	2	0.4	15	3.1	17	5.4	36	9.9
Education (self)	94	20.7	99	20.7	5	1.6	0	0.0
Education (others)	1	0.2	2	0.4	0	0.0	2	0.6
Sickness/Death/Shocks	5	1.1	12	2.5	21	6.6	14	3.9
Migration with family	13	2.9	39	8.1	43	13.6	100	27.5
To be close to family	41	9.0	81	16.9	28	8.8	34	9.4
Other	36	7.9	69	14.4	113	35.6	134	36.9
Total	454	100.0	479	100.0	317	100.0	363	100.0

Source: IFLS2-3.
Note: the category "Other" aggregates the following original IFLS categories:
like the destination, new housing opportunities, other reasons.

[50] There are also some individuals who were classified as 'individual migrants' while stating "Migration with family" as reason for migration. These individuals are those who reported having moved with family members other than the spouse, or the children, or the parents (for instance, individuals who moved with one brother).

5. Descriptive statistics

5.1 Methodology

In the analysis that follows, I compare migrants and non-migrants on the basis of their pre-migration characteristics, distinguishing by individual and family migrants. Among pre-migration characteristics, I consider both individual-level variables (such as age, education, sector of occupation and income) and household-level variables (households' landholdings).

In this statistical analysis, as in the econometric one of the next section, I restrict the sample to individuals who are in the age range from 15 to 75 years.[51]

The methodology of the statistical analysis is mainly based on Fernández-Huertas Moraga (2011), who analyses the selection pattern of migration from Mexico to the United States. The above study employs the data originated from the *Encuesta Nacional de Empleo Trimestral* (ENET), which records the wage (and other characteristics) of (future) emigrants in the quarter before they leave. In that paper, 'emigrants' are individuals who are in Mexico at quarter *t* and are reported to have left for the United States when the interviewer returns to their same household at quarter *t+1*. On the other hand, 'non-migrants' are those who are in Mexico in both quarters. Comparing these two categories, Fernández-Huertas Moraga (2011) evaluates the direction of selection of Mexican emigrants on the basis of income and education, and concludes that they are negatively selected. Following this study, I take advantage of the fact that IFLS allows to observe migrants before they move. More precisely, I consider the individuals interviewed both in IFLS2 and in IFLS3, and see who among them moved between the two waves. I define those who moved 'migrants' and those who did not 'non-migrants'. I then compare the characteristics of (prospective) migrants and non-migrants at the baseline (1997), before the migration episodes happening from 1997 to 2000. Hence, in my analysis I explore the determinants of recent migrations. In this way, I also avoid using data referring to very distant moments in time, which could be affected by recall bias.

Compared to the ENET data, IFLS has the drawback that the time span between *t* and *t+1* is not a quarter, as in the ENET case, but it ranges between a few months and three years. Therefore, if the migrant moved, say, in 1999, his or her wage can be significantly different from the wage of the same person that I observe in 1997. This disadvantage of IFLS with respect to ENET data implies that comparing migrants and non-migrants on the basis of their pre-migration characteristics to understand the patterns of selection may be a less 'robust' procedure in the case of IFLS than in the case of ENET. However, IFLS has also a relevant advantage with respect to ENET. Indeed, ENET does not record as 'migrants' individuals whose complete household emigrated. In other words, in ENET an individual is classified as migrant only if at least one of the members of his or her household remains in Mexico. Differently, IFLS also records as migrants individuals who migrated with the whole family, because the survey entails the tracking of moving

[51] Other studies on migration in developing countries restricted the age range of sample individuals between 15 and 65 (e.g. Beegle et al. 2011, Fernandez-Huertas Moraga, 2011). However, I extended the sample age range to include individuals from 15 to 75 years old, mainly for two reasons. First, having a larger sample allows a more precise estimate of the effects, and second individuals in Indonesia work until the elderly age. Thus they presumably also move (for work or other reasons) in the same age range.

households. Moreover, as we saw in the previous section, this feature of IFLS allows to distinguish between family and individual migration, and this distinction will be key to our analysis.

5.2 Location, age and marital status

Table 8 below shows that, on average, migrants come predominantly from urban areas, irrespective of gender and migration type. If we look at the type of migration, we can see that a higher fraction of individual migrants come from rural areas, compared to family migrants.[52] These figures may indicate a 'wealth effect', by which individuals living in rural areas are poorer than those living in urban ones, and thus they have less resources to pay the costs associated to family migration. Compared to individuals in rural settings, individuals in urban settings may also have greater access to capital which provide resources to bear migration costs, and greater networks, which could reduce migration costs. These figures may also be related to the fact that the skills of individuals living in urban areas are more "transferable" (to other urban settings) than those of individuals living in rural areas Partha and Seck (2009).

Table 8. Location, age and marital status

	Non-migrants	Individual migrants	Family migrants	Non-migrants	Individual migrants	Family migrants
N obs	13,973	935	681			
% on total	89.6	6.0	4.4			
% male	45.3	48.6	46.6			
	Men			Women		
Live in rural	0.53	0.45	0.34	0.52	0.39	0.37
	(0.49)	(0.49)	(0.48)	(0.49)	(0.48)	(0.48)
Age	38.70	23.38	34.46	37.81	24.98	30.99
	(15.52)	(10.45)	(12.64)	(14.99)	(13.28)	(10.55)
Unmarried	0.24	0.76	0.19	0.17	0.68	0.11
	(0.42)	(0.42)	(0.39)	(0.38)	(0.47)	(0.32)

Source: IFLS2-3.

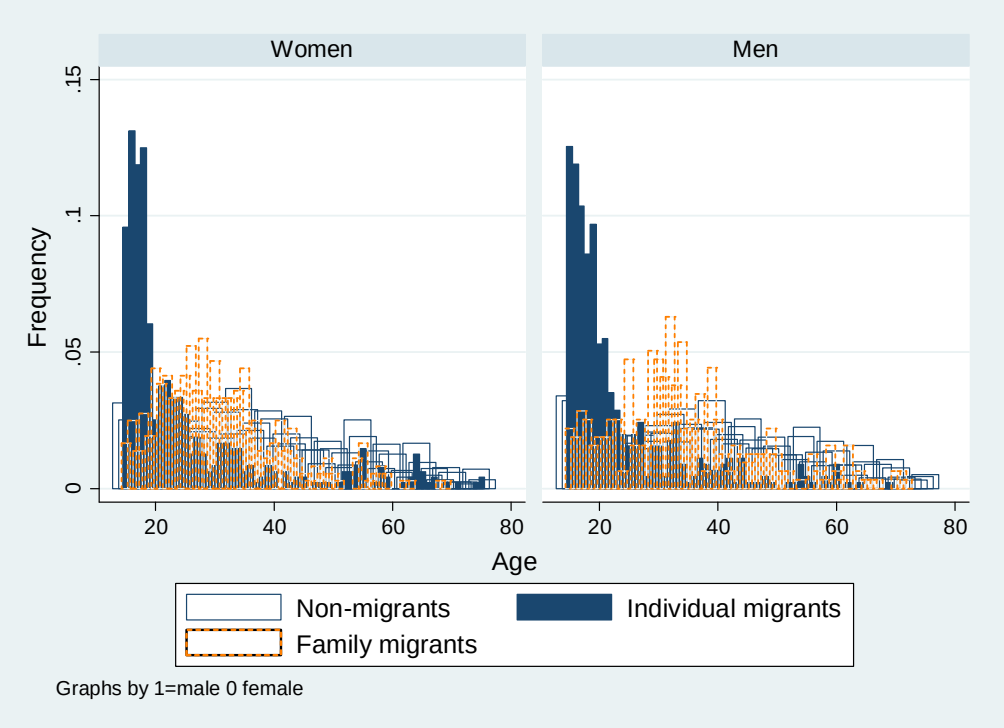
Both among men and women, migrants are, on average, more likely to be unmarried than non-migrants. However, this is mainly driven by the marital status of individual migrants, while the majority of family migrants are married. The presence of unmarried individuals in the category migration with family is due to the presence in this category of youths who migrate with their parents.

On average, migrants are younger than non-migrants and, for both sexes, individual migrants are significantly younger than family migrants. By focusing only on individual migrants, hence, one is likely to consider only a very specific group of individuals, i.e. very young individuals, at about age 20. However, as the table above shows, also older individuals can migrate, most likely with their families. The table hence shows once again the importance of considering both individual and family migrants, distinguishing the two categories.

[52] The difference is statistically significant at the 5% for males and 1% for females.

Migrants' selection with respect to age can be shown also graphically. Following Fernández-Huertas Moraga (2011), I obtain histograms of age for non-migrants, individual and family migrants, by gender. Figure 1 below shows that the three categories are characterized by different age profiles. In particular, irrespective of gender the probability of individual migration is highest at about age 20 and that of family migration is highest around age 30. After age 40, individuals are more likely to stay, regardless of gender.

Figure 1. Age distribution of migrants and non-migrants (1997)



Source: IFLS2-3.

5.3. Education

Table 9 below indicates that migrants in Indonesia are positively selected in terms of education: both among women and men, indeed, migrants are significantly more educated than non-migrants, and the type of migration that is associated with the highest education level is individual migration. Moreover, there is a statistically significant difference between the education level of male and female non-migrants, with females being less educated than males.

Table 9. Individuals' education, by migration type and sex (1997)

	Non-migrants	Individual migrants	Family migrants	Non-migrants	Individual migrants	Family migrants
	Men			Women		
Years of education	6.97	9.38	8.98	5.65	9.27	8.09
	(4.59)	(3.98)	(4.95)	(4.56)	(4.40)	(4.77)
N.	6,334	454	317	7,639	480	363

Source: IFLS2-3.

5.4 Employment and income

Table 10 shows that, both among men and women, individual migrants are less likely to work than non-migrants and family migrants.[53] This is related to the fact that among those who move alone there are also individuals who move for education, who are unlikely to be working. The table also shows that non-migrants are significantly more likely to be working in the primary sector than migrants, both among men and women. Moreover, male individual migrants are associated with a significantly higher probability of working in the primary sector than male family migrants. This may be due to the fact that a household engaged in primary activities often owns farm business assets that deter it from moving as a whole.

Table 10 also shows the differences in mean real incomes between the different categories of individuals. Total income includes monthly earnings from primary and secondary job (in case the individual has one) and non-labour income. It is measured as a fraction of provincial average, to avoid time trend effects (Fernández-Huertas Moraga, 2011), and in logarithms. The table shows that, with the exception of women in rural areas, the migration type which is associated with the highest income is family migration. This is likely to be related to the fact that family migration requires higher costs than individual migration, and thus individuals earning higher incomes are more likely to be able to afford family moves.

From the table below, we see that individual migrants are more likely to be unemployed at the baseline, with respect to non-migrants and family migrants.[54]

Table 10. Individuals' employment, by migration type and sex (1997)

	Non-migrants	Individual migrants	Family migrants	Non-migrants	Individual migrants	Family migrants
	Men			Women		
Primary	0.30	0.09	0.15	0.11	0.03	0.04
	(0.45)	(0.29)	(0.36)	(0.31)	(0.16)	(0.19)
Manufacturing	0.11	0.09	0.14	0.09	0.07	0.07
	(0.31)	(0.29)	(0.34)	(0.28)	(0.26)	(0.25)
Construction	0.07	0.05	0.06	0.00	0.00	0.01
	(0.25)	(0.21)	(0.24)	(0.06)	(0.06	(0.10)
Services	0.31	0.19	0.46	0.22	0.17	0.22
	(0.46)	(0.39)	(0.49)	(0.42)	(0.38)	(0.42)
Not working	0.21	0.56	0.19	0.58	0.72	0.66
	(0.41)	(0.49)	(0.39)	(0.49)	(0.45)	(0.47)
N.	6,334	454	317	7,639	480	363

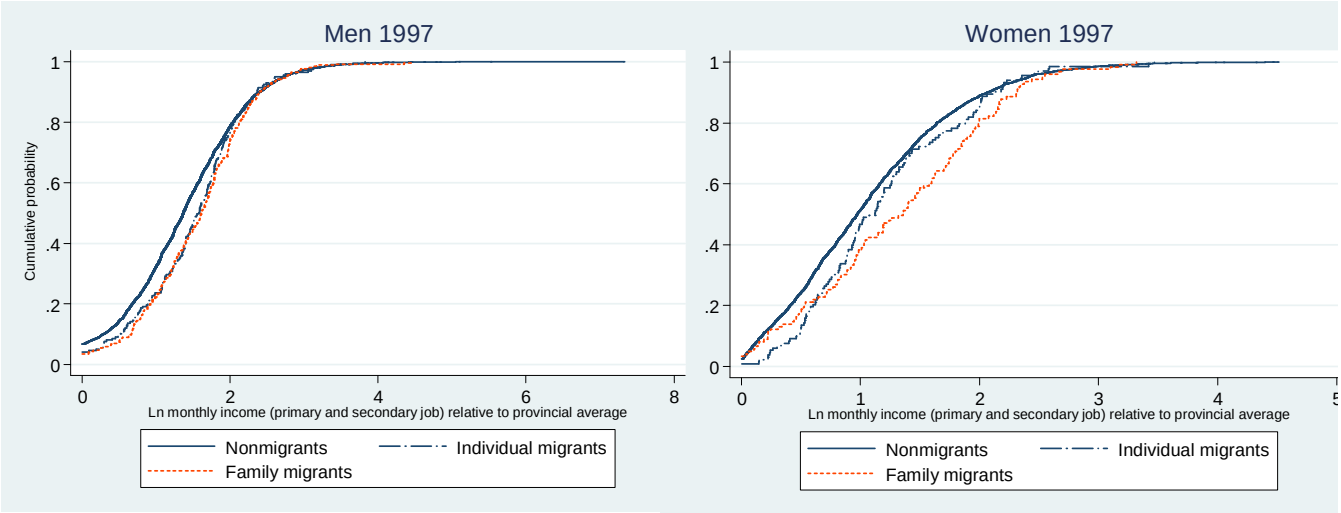
Source: IFLS2-3.

[53] The difference in the working probability between individual migrants and non-migrants is statistically significant for both men and women at the 1% level, while the difference in the working probability between family migrants and non-migrants is statistically significant for men at the 1% level and for women at the 5% level.

[54] The differences in employment probability between individual migrants and non-migrants are statistically significant at the 1% level. [54] The differences in employment probability between individual migrants and family migrants are statistically significant at the 1% level for males and at the 10% level for females.

We use a graphical analysis to visualize the different selection patterns of the various categories of migrants, according to income. The graphs below plot the cumulative density functions of individuals' income, distinguishing by gender and type of migration. This analysis only considers *working individuals*.

Figure 2. Individuals' income distribution, by migration type and sex (1997)



Source: IFLS2-3.
Note: monthly own income is the logarithm of monthly earnings from primary and secondary job (in case the individual has one). Income refers to profits if the individual is self-employed and to wages if he/she is a private or government worker. In this graph, income is measured as a fraction of provincial average, to avoid time trend effects (Fernández-Huertas Moraga, 2011).

Figure 2 above shows that on average migrants earn more than non-migrants, both among men and women. Following Fernández-Huertas Moraga (2011), I performed a Kolmogorov-Smirnov (K-S) test, to verify whether the earnings distributions of migrants and non-migrants indeed differ. If we do not distinguish by migration type, the null hypothesis of equality of income distribution between migrants and non-migrants is rejected at the 0.1% significance level, for both men and women. Disentangling the different selection patterns by migration type, the K-S rejects equality of income distributions between non-migrants and individual migrants, at the 0.1% and 5% for males and females, respectively. Instead, for both sexes, equality of income distributions between family migrants and non-migrants is still rejected at the 1% significance level. Hence, family and individual migrants are positively selected in terms of own income, both among men and women. Equality of income distributions between individual and family migrants is only rejected for females, at the 5% significance level.

5.5 Household land ownership

Table 11 below shows average per capita household land values for migrants and non-migrants at the baseline, differentiating by gender and migration type.[55] As the income variable, land values are measured in logarithms and in real terms, using spatial and temporal deflators (with base Jakarta in December 2000). I consider landholdings only, since it is the asset category that most realistically can be considered exogenous, while other assets, such as house or other buildings may be influenced by the household's past migration experience.[56]

Table 11. Households' land ownership by migration type and sex (1997)

	Non-migrants	Individual migrants	Family migrants	Non-migrants	Individual migrants	Family migrants
		Men			Women	
P.c. Business land	5.07	4.41	2.37	4.59	3.98	2.39
	(7.03)	(6.77)	(5.35)	(6.85)	(6.70)	(5.31)
P.c. Non-business land	4.40	4.57	3.48	4.21	4.27	3.75
	(6.72)	(6.79)	(6.26)	(6.62)	(6.78)	(6.44)
N.	6,334	454	317	7,639	480	363

Source: IFLS2-3.
Note:
Variables are measured as the logarithm of land values. Land values are measured in per capita real terms (base Jakarta 2000). Given the presence of households not possessing land (for which land values have been set to zero), before taking logs land variables (x) have been replaced by (x+1).

Considering business land for men, family migrants own a significantly lower level of business land than individual migrants or non-migrants. This pattern suggests that in the presence of a farm business, individual migration rather than family migration is chosen. In other words, when the household owns a farm business, work on the land is required and thus individual migration is preferred. It is optimal for the household to send individual male members to different locations, thus spreading the risk associated with household's activities (Rosenzweig, 1988). For women, we again observe that the category which is associated with the lowest business value is that of family migrants.

From the above analysis, we can conclude that different categories of migrants are selected along many dimensions in different ways, depending on gender. We also saw that migrants are positively selected in terms of education, with the differences in education between migrants and non-migrants being particularly relevant in the case of individual migrants.

As regards income, we saw different selection patterns not only by gender, but also by migration type.

Concerning land, we saw that family migration is associated with lower levels of household's business land ownership, but only in rural areas. Moreover, only for women in rural areas, individual migrants own significantly more non-business land than non-migrants. Hence, in the econometric analysis that follows I consider both business and non-business land values.

[55] Given the presence of households not possessing land (for which land values are set to zero), before taking logs relative land variables (x) have been replaced by (x+1).
[56] See Stacey (2011).

6. Econometric analysis

6.1 Methodology

In the econometric analysis, migration between 1997 and 2000 is analyzed as a function of individuals' and household characteristics in 1997. This methodology resembles the one of Beegle et al. (2011). These authors use two survey rounds of a tracking panel survey in Tanzania and model individuals' migration decisions between two survey waves as a function of individual- and household-level characteristics at the baseline year.[57]

Differently from the above study, I distinguish between individual and family migration, estimating a Multinomial Probit model, where the individuals' choice set entails three alternatives: non-migration, individual migration and family migration. Indexing with *i* the individual, *j* his or her household and *k*=1, 2, 3 the chosen alternative, the estimated model is the following:

$$M_{ijk,97\text{-}00} = \beta_{k0} + \beta_{k1} X_{ijk,97} + \beta_{k2} H_{jk,97} + \varepsilon_{ijk,97} \quad (1)$$

where $M_{ijk,97\text{-}00}$ is a dummy which equals one if the individual chose alternative *k* between 1997 and 2000, and $X_{ijk,97}$ and $H_{jk,97}$ are vectors of individual- and household-level characteristics in 1997. The sample is the same as that considered in the previous statistical analysis: it includes the individuals who were interviewed both in 1997 and 2000, and that reported their full migration history in the 2000 interview (see Section 4). I considered the 15-75 years age range, and performed the estimates on men and women samples separately.

On the basis of the literature on the determinants of migration and of the previous descriptive analysis, the baseline specification (table 12 below, column 1) includes both individual- and household-level characteristics. In particular, the individual-level variables are the following: a dummy which equals one if the individual is unmarried, age and its square, and the deviation in years of schooling from the peers' provincial average.[58] Household-level variables include as a dummy equal to one if the individual is the household head or his or her spouse, a dummy indicating whether the individual is child of the head, and the number of his or her biological children in the age ranges 0-5, 6-10, 11-15, 16-20, and over 20 years.

In column 2, we add controls for households' landholdings, distinguishing between business and non-business land.

[57] The analysis of Beegle et al. (2011) is based on the Kagera Health and Development Survey (KHDS), which consisted initially of about 900 households interviewed up to four times from 1991 to 1994. The KHDS 1991-1994 serves as the baseline survey for the study. In 2004, another round of the KHDS was administered, aiming at interviewing all individuals who were household members in any round of the survey in 1991-94. 93% of the baseline households were re-contacted, tracking those who moved in the period 1991/1994-2004. The authors model individuals' migration between 1991-94 and 2004 as a function of individual and household characteristics in 1991-94.

[58] Following (Beegle et al. 2011), I include "years of schooling completed relative to peers" instead of "years of schooling completed", for two reasons: first, the sample also includes individuals younger than 18 at the baseline, who have not necessarily completed their education. Second, considering years of schooling relative to peers avoids that the education variable captures part of the age effect. The variable is constructed as the deviation of the individuals' completed years of schooling from the average of his/her peers. Age-specific peers are considered for individuals younger than 18, while all other individuals are considered for adults. Differently from (Beegle et al. 2011), I calculate the averages within province, because of the considerable heterogeneity across Indonesian provinces.

In column 3, I also consider individuals' working conditions at the baseline, as well as their income. In particular, I include the logarithm of individuals' real total monthly income, and dummies for his or her employment sector (manufacturing, construction, services, not working, with the baseline category being primary sector employment).

Finally, I also consider further household-level controls, i.e. dummies indicating the quintiles of households' per capita real income at the baseline year 1997.

All models also include controls for location, which are a rural/urban dummy, province dummies, and the interactions between the rural and the province dummies. Standard errors are corrected for heteroskedasticity and clustered at the household-level. Results are displayed in the form of *average marginal effects* on the probability of being in each migration category, with the base category in the estimation being that of non-migrants.

6.2 Results

Table 12 concerns the probability of migrating alone for males.

Table 12. Marginal effects on the probability of *individual* migration for men

	Outcome: individual migration (men)			
	(1)	(2)	(3)	(4)
Unmarried	0.073***	0.074***	0.070***	0.068***
	(0.02)	(0.02)	(0.02)	(0.02)
Age	-0.003***	-0.003***	-0.003***	-0.003***
	(0.00)	(0.00)	(0.00)	(0.00)
HH Head/Spouse	0.009	0.010	0.010	0.009
	(0.01)	(0.01)	(0.01)	(0.01)
Head's child	-0.038***	-0.038***	-0.038***	-0.037***
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 0-5	-0.019**	-0.019**	-0.019**	-0.019**
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 6-10	-0.018**	-0.018**	-0.019**	-0.018**
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 11-15	-0.009	-0.009	-0.010	-0.010
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 16-20	-0.017*	-0.018*	-0.018*	-0.017*
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 21+	-0.018*	-0.018*	-0.020*	-0.019*
	(0.01)	(0.01)	(0.01)	(0.01)
Education	0.003***	0.003***	0.003***	0.003***
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Business land		0.000	0.000	0.000
		(0.00)	(0.00)	(0.00)
P.c. Non-business land		0.001	0.001	0.001
		(0.00)	(0.00)	(0.00)
Manufacturing			0.015	0.017
			(0.01)	(0.01)
Construction			0.019	0.022
			(0.02)	(0.02)
Services			0.016*	0.018*
			(0.01)	(0.01)
Not working			0.040***	0.042***
			(0.02)	(0.02)
Monthly own income			0.001	0.001
			(0.00)	(0.00)
2nd quintile HH income				-0.009
				(0.01)
3rd quintile HH income				-0.023**
				(0.01)
4th quintile HH income				-0.019**
				(0.01)
5th quintile HH income				-0.011
				(0.01)
N	7,105	7,105	7,105	7,105

Notes:
Robust standard errors in parentheses, clustered at the HH level. *** 1%, ** 5%, * 10% significance level.
All models also include a constant term, rural and province dummies and their interactions.

Column 1 includes basic demographic individual- and household-level characteristics. Being unmarried raises the probability of individual migration by about 7 percentage points, on average and *ceteris paribus*. The role of marital status in determining males' individual migration is confirmed when controls for household land ownership, individuals' working status, and household wealth are added. Consistently with previous descriptive statistics, the table also shows that individual male migrants are significantly younger than non-migrants. As for household composition, we see that being the household head or his/her spouse does not significantly influence the probability of males' individual migration, which is instead significantly decreased when the individual is the head's child. In particular, a higher number of children in the age ranges 0-5, 6-10, 16-20 and 21+ significantly reduces the probability of men's individual migration. More precisely, having an additional child in the above specified age ranges decreases males' individual migration probability by 1.7 to 2.0 percentage points, on average and *ceteris paribus*. The negative effect of the number of children on the probability of men's individual migration is maintained when further covariates are added.

Households' land holdings, added in column 2, are not statistically significant determinants of males' individual migration probability.

In column 3, I add controls for whether individuals working status, together with his own real monthly income. Being unemployed or working in the services sector in the baseline year 1997 increases males' probability of moving alone by 4 and 1.6 percentage points, respectively. This effect holds when controls for household income and expenditure are included, in column 4. The effect of working status suggests that one of the primary motivations for moving alone for males is to look for a job. The result is also related to the fact that, within individual movers, there are also those who moves for education, which are likely not to be working. The estimated coefficient for income is not statistically significant, indicating that once demographic, household structure, employment sector and land controls are included, own income does not influence the probability of individual migration for males.[59]

The model in column 4 includes quintiles for total household-level per capita monthly income in 1997. The results indicate that belonging to households whose income is in the third and fourth quintiles of the distribution significantly reduces the probability of men migrating alone. This result is consistent with Banerjee and Newman (1998), whose model predicts that individuals belonging to the intermediate part of household income distribution are less likely to migrate. This is due to the fact that individuals at the intermediate levels of income are more likely to count on informal insurance mechanisms at the village-level, which reduce their incentives to move. These mechanisms are instead less likely to be used by the poorest and the richest segments of the population. Indeed, the richest can to rely on more formal coping devices, while the extremely poor are unlikely to have access to informal village-level groups.

The next table reports the marginal effects on the probability of men's family migration.

[59] It could be that even individuals who do not earn high income are given the resources needed to pay migration costs by other members of the family.

Table 13. Marginal effects on the probability of *family* migration for men

	Outcome: family migration (men)			
	(1)	(2)	(3)	(4)
Unmarried	-0.050***	-0.051***	-0.053***	-0.053***
	(0.01)	(0.01)	(0.01)	(0.01)
Age	-0.001**	-0.001**	-0.001**	-0.001**
	(0.00)	(0.00)	(0.00)	(0.00)
HH Head/Spouse	-0.011	-0.015	-0.014	-0.015
	(0.01)	(0.01)	(0.01)	(0.01)
Head's child	0.004	0.005	0.006	0.005
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 0-5	0.017***	0.017***	0.017***	0.018***
	(0.00)	(0.00)	(0.00)	(0.00)
N. children 6-10	-0.000	-0.000	-0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
N. children 11-15	-0.007*	-0.007	-0.007*	-0.006
	(0.00)	(0.00)	(0.00)	(0.00)
N. children 16-20	0.001	0.001	0.001	0.001
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 21+	-0.001	-0.001	-0.001	-0.001
	(0.01)	(0.01)	(0.01)	(0.01)
Education	0.002***	0.002***	0.002***	0.002***
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Business land		-0.002***	-0.002***	-0.002***
		(0.00)	(0.00)	(0.00)
P.c. Non-business land		-0.001*	-0.001	-0.001*
		(0.00)	(0.00)	(0.00)
Manufacturing			0.015	0.014
			(0.01)	(0.01)
Construction			0.008	0.007
			(0.01)	(0.01)
Services			0.022***	0.021***
			(0.01)	(0.01)
Not working			0.003	0.001
			(0.01)	(0.01)
Monthly own income			-0.001	-0.001
			(0.00)	(0.00)
2nd quintile HH income				0.000
				(0.01)
3rd quintile HH income				0.007
				(0.01)
4th quintile HH income				0.001
				(0.01)
5th quintile HH income				0.012
				(0.01)
N	7,105	7,105	7,105	7,105

Notes:
Robust standard errors in parentheses, clustered at the HH level. *** 1%, ** 5%, * 10% significance level.
All models also include a constant term, rural and province dummies and their interactions.

From column 1, we see that being unmarried decreases the probability of men's family migration by about 5 percentage points, on average and *ceteris paribus*. As for individual men's migration, age increases significantly reduce the probability of men's family migration. Being the household head or his or her child does not significantly impact the probability of family migration for men.

As for household composition, we see that having one more child in the age range 0-5 increases the probability of family migration for males by 1.7 percentage points, on average and *ceteris paribus*. The effect is stable, even if slightly increased, when controls for household wealth are included, in column 4.

Education exerts a statistically significant and positive influence on the probability of family migration for men. In particular, having one year of education more than the average peers in the province increases the probability of men's family migration by 0.2 percentage points, on average and *ceteris paribus*.

Column 2 adds controls for households' land holdings: we see that a higher value of land (business or non-business) significantly decreases the probability of men's family migration. This result is consistent with the descriptive statistics of the previous section, which show a negative association between household business land value and family migration.

Regarding the labour market variables of column 3, being employed in the services sector at the baseline year increases the probability of men's family moves by 2.2 percentage points, on average and *ceteris paribus*. Differently from the case of individual men's migration, in the case of men's family migration, being unemployed at the baseline does not have a significant impact on the probability of moving.

From column 4, we can also see that moving with the family is not significantly associated with household income. This is an important difference with respect to the case of individual men's migration.

The following table reports the results for women's individual migration.

Table 14. Marginal effects on the probability of *individual* migration for women

	Outcome: individual migration (women)			
	(1)	(2)	(3)	(4)
Unmarried	0.051***	0.051***	0.052***	0.050***
	(0.01)	(0.01)	(0.01)	(0.01)
Age	-0.003***	-0.003***	-0.003***	-0.003***
	(0.00)	(0.00)	(0.00)	(0.00)
HH Head/Spouse	-0.013	-0.013	-0.013	-0.015*
	(0.01)	(0.01)	(0.01)	(0.01)
Head's child	-0.025***	-0.025***	-0.025***	-0.022***
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 0-5	-0.017**	-0.017**	-0.017***	-0.017***
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 6-10	-0.017**	-0.017**	-0.017***	-0.017***
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 11-15	-0.002	-0.002	-0.002	-0.001
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 16-20	-0.019**	-0.019**	-0.019**	-0.018**
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 21+	0.001	0.001	0.001	0.003
	(0.01)	(0.01)	(0.01)	(0.01)
Education	0.003***	0.003***	0.003***	0.003***
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Business land		-0.000	-0.000	-0.000
		(0.00)	(0.00)	(0.00)
P.c. Non-business land		0.000	0.000	0.000
		(0.00)	(0.00)	(0.00)
Manufacturing			0.007	0.009
			(0.02)	(0.02)
Construction			0.017	0.019
			(0.04)	(0.05)
Services			0.022*	0.024*
			(0.02)	(0.02)
Not working			0.021*	0.021*
			(0.01)	(0.01)
Monthly own income			0.000	0.001
			(0.00)	(0.00)
2nd quintile HH income				-0.020***
				(0.01)
3rd quintile HH income				-0.027***
				(0.01)
4th quintile HH income				-0.024***
				(0.01)
5th quintile HH income				-0.019**
				(0.01)
N	8,482	8,482	8,482	8,482

Notes:
Robust standard errors in parentheses, clustered at the HH level. *** 1%, ** 5%, * 10% significance level.
All models also include a constant term, rural and province dummies and their interactions.

Column 1 in table 14 indicates that being unmarried increases women's probability of moving alone by 5 percentage points, on average and *ceteris paribus*. As for the case of men's individual and family moves, age increases significantly reduce the probability of females' individual migration. The effects of marital status and age are maintained, once land, working status and household wealth values are controlled for. As for household composition, we see that the number of children in the age ranges 0-5, 6-10 and 16-20 significantly decreases women's probability of moving alone. Differently from the case of men's individual migration, the number of children over age 20 does not have a statistically significant impact on the probability of migrating alone for women.

Regarding education, we see that having one year of education more than the peers in the province increases the probability of women's individual migration by 0.3 percentage points, on average and *ceteris paribus*.

As for males, also for females land variables, included in column 2, are not statistically significant determinants of individual migration.

Own income is not a statistically significant determinant of individual migration for females, as we saw for males. Moreover, also for females being unemployed or being employed in the services sector at the baseline year is significantly and positively associated with the probability of individual migration.

The others controls for household wealth, included in column 4, show that women whose household belongs to the second to the fifth quintile of the distribution of household income are significantly less likely to migrate alone, compared to those whose household is in the lowest quintile. This indicates that women who move alone belong to the poorest segment of the distribution of household *current* income. This result suggests that poorest women move in order to find better opportunities, when income sources lack in their origin village.

Table 15 shows the results for females' family migration.

Table 15. Marginal effects on the probability of *family* migration for women

	Outcome: family migration (women)			
	(1)	(2)	(3)	(4)
Unmarried	-0.047***	-0.048***	-0.048***	-0.048***
	(0.01)	(0.01)	(0.01)	(0.01)
Age	-0.002***	-0.002***	-0.002***	-0.002***
	(0.00)	(0.00)	(0.00)	(0.00)
HH Head/Spouse	0.006	0.004	0.005	0.004
	(0.01)	(0.01)	(0.01)	(0.01)
Head's child	0.012	0.013	0.012	0.013
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 0-5	0.013***	0.012***	0.012***	0.012***
	(0.00)	(0.00)	(0.00)	(0.00)
N. children 6-10	0.002	0.002	0.002	0.002
	(0.00)	(0.00)	(0.00)	(0.00)
N. children 11-15	-0.006	-0.006	-0.006	-0.005
	(0.00)	(0.00)	(0.00)	(0.00)
N. children 16-20	0.004	0.004	0.004	0.004
	(0.00)	(0.00)	(0.00)	(0.00)
N. children 21+	-0.004	-0.003	-0.004	-0.004
	(0.01)	(0.01)	(0.01)	(0.01)
Education	0.002***	0.002***	0.002***	0.002***
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Business land		-0.001***	-0.001***	-0.001***
		(0.00)	(0.00)	(0.00)
P.c. Non-business land		-0.000	-0.000	-0.000
		(0.00)	(0.00)	(0.00)
Manufacturing			0.012	0.011
			(0.01)	(0.01)
Construction			0.068*	0.059*
			(0.05)	(0.05)
Services			0.018*	0.016
			(0.01)	(0.01)
Not working			0.017*	0.017*
			(0.01)	(0.01)
Monthly own income			-0.000	-0.001
			(0.00)	(0.00)
2nd quintile HH income				0.000
				(0.01)
3rd quintile HH income				0.006
				(0.01)
4th quintile HH income				0.000
				(0.01)
5th quintile HH income				0.016*
				(0.01)
N	8,482	8,482	8,482	8,482

Notes:
Robust standard errors in parentheses, clustered at the HH level. *** 1%, ** 5%, * 10% significance level.
All models also include a constant term, rural and province dummies and their interactions.

From column 1, we see that being unmarried decreases the probability of family migration for women by 4.7 percentage points, on average and *ceteris paribus*. The influence of marital status is maintained when other controls are added. As for the cases seen above, age has a negative and statistically significant impact on the probability of women's family migration. As for men, the probability of women's family migration increases with the number of children aged 0-5.

Education is also significantly related to women's family migration decisions. In particular, having one year more of education compared to the peers in the province increases a woman's probability of migrating with her family by 0.2 percentage points, on average and *ceteris paribus*.

In columns 2, I add the land controls. We see that the probability of women's family migration significantly decreases with the value of business land. Non-business land, instead, does not significantly influence the probability of women's family migration. From column 3, we see that the women who are employed in the construction or service sectors, or unemployed at the baseline are significantly more likely to move with the family, compared to women who are working in the primary sector at the baseline. However, only the effect of the construction sector remains statistically significant once household wealth controls are added, in column 4. This results could be related to the fact that in the cases in which women work in the households' farm land at the baseline, they are less likely to follow their spouses at destination, were they would have low probabilities of finding a job.

As for men, also for women households' income does not represent a significant determinant of family migration.

In the next table, I only report the results of models 4 above for males and females, by migration type.

Table 16. Determinants of *individual* and *family* migration, by gender

	Individual migration		Family migration	
	Men	Women	Men	Women
Unmarried	0.068***	0.050***	-0.053***	-0.048***
	(0.02)	(0.01)	(0.01)	(0.01)
Age	-0.003***	-0.003***	-0.001**	-0.002***
	(0.00)	(0.00)	(0.00)	(0.00)
HH Head/Spouse	0.009	-0.015*	-0.015	0.004
	(0.01)	(0.01)	(0.01)	(0.01)
Head's child	-0.037***	-0.022***	0.005	0.013
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 0-5	-0.019**	-0.017***	0.018***	0.012***
	(0.01)	(0.01)	(0.00)	(0.00)
N. children 6-10	-0.018**	-0.017***	0.000	0.002
	(0.01)	(0.01)	(0.00)	(0.00)
N. children 11-15	-0.010	-0.001	-0.006	-0.005
	(0.01)	(0.01)	(0.00)	(0.00)
N. children 16-20	-0.017*	-0.018**	0.001	0.004
	(0.01)	(0.01)	(0.01)	(0.00)
N. children 21+	-0.019*	0.003	-0.001	-0.004
	(0.01)	(0.01)	(0.01)	(0.01)
Education	0.003***	0.003***	0.002***	0.002***
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Business land	0.000	-0.000	-0.002***	-0.001***
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Non-business land	0.001	0.000	-0.001*	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Manufacturing	0.017	0.009	0.014	0.011
	(0.01)	(0.02)	(0.01)	(0.01)
Construction	0.022	0.019	0.007	0.059*
	(0.02)	(0.05)	(0.01)	(0.05)
Services	0.018*	0.024*	0.021***	0.016
	(0.01)	(0.02)	(0.01)	(0.01)
Not working	0.042***	0.021*	0.001	0.017*
	(0.02)	(0.01)	(0.01)	(0.01)
Monthly own income	0.001	0.001	-0.001	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)
2nd quintile HH income	-0.009	-0.020***	0.000	0.000
	(0.01)	(0.01)	(0.01)	(0.01)
3rd quintile HH income	-0.023**	-0.027***	0.007	0.006
	(0.01)	(0.01)	(0.01)	(0.01)
4th quintile HH income	-0.019**	-0.024***	0.001	0.000
	(0.01)	(0.01)	(0.01)	(0.01)
5th quintile HH income	-0.011	-0.019**	0.012	0.016*
	(0.01)	(0.01)	(0.01)	(0.01)
N	7,105	8,482	7,105	8,482

Notes:
Robust standard errors in parentheses, clustered at the HH level. *** 1%, ** 5%, * 10% significance level.
All models also include a constant term, rural and province dummies and their interactions.

From the above table, we see that there are some determinants which play an analogous role across genders and migration types. For instance, increases in age significantly reduce mobility, either alone or with the family, and regardless of gender. Being more educated, instead, increases the probability of both individual and family moves for both genders.

Other determinants are, instead, typical of migration type, regardless of gender. For instance, being married or being the head's child reduces the probability of individual migration for both males and females. Similarly, for both sexes, having a higher number of children in the age ranges 0-5 significantly reduces the probability of individual migration, while increasing that of family migration. Also the role of business land as a determinant of migration is common for men and women, varying by migration type. Indeed, the results indicate that for both men and women higher values of per capita households' land holdings does not significantly influence the probability of individual migration, while significantly decreasing that of family migration. Moreover, both for men and women the probability of moving alone increases for individuals who are unemployed or working in the services sector at the baseline. Services employment might also include informal and low-skill jobs, such as small-scale retail trade activities. Thus, the results indicate that both among men and women, individuals undertaking lower earnings activities at the baseline year 1997 are more likely to move alone in the period 1997-2000, on average and *ceteris paribus*. Thus, the results suggest that one of the main motivation for individual moves is the search for better job opportunities.

The results also show that for both men and women there is a negative and statistically significant association between the probability of individual migration and current household income.

There are instead some selection patterns that are specific to both gender and migration type. For instance, the number of children in the age older than 20 years significantly and negatively influences only males' individual moves, without significantly impact females individual migration or family migration of either men or women.

As previously stated, the sample of analysis includes both individuals who migrated for the first time in the period 1997-2000, and individuals who migrate for the second, third, or n^{th} time. Hence, I also estimate models controlling for past migration experience, including a dummy which is equal to one if the individual migrated (either alone or with his/her family) also before 1997. However, it is only an exploratory analysis, as past migration is endogenous. From table A.1 in the Appendix, we can see that the variable indicating past migration experience is positive and significant for both genders and migration types. It is also higher for family than individual migration. Other coefficient estimates do not vary in a relevant way from previous estimates not controlling for past migration.

7. Conclusions

Internal migration represents a key phenomenon in developing countries, which has been shown by past literature to have important implications, both for migrants themselves, and for their origin households. Therefore, it is important to understand the determinants of individuals' moves. In particular, understanding the determinants of internal migration helps to point out which market imperfections exist in origin areas, and may thus be a reasonable starting point for the design of policy interventions in developing settings.

In this analysis, I focused on one important aspect of internal migration, which has not been considered by the literature so far. In particular, I analyzed migrants' selection pattern, distinguishing between individuals who moved alone and individuals who move with their families.

To do this, I used the Indonesia Family Life Survey dataset. Indeed, Indonesia represents an ideal setting to analyze internal migrants' characteristics, given that the country recently had relevant increases in mobility within its national borders. I took advantage of the longitudinal nature of IFLS, by using two survey waves (1997 and 2000). Following recent empirical studies on migration (Beegle *et al.*, 2011; Fernández-Huertas Moraga, 2011), I analyzed individuals' migration choices between period *t* (year 1997), and period *t + 1* (year 2000), as a function of individual- and household-level variables at *t* (1997), i.e. before the considered migration episodes. This methodology increases the robustness of our results, by reducing time-constant individual-level heterogeneity.

In order to distinguish the different roles of migration determinants according to migration type, I estimated a Multinomial Probit of individuals' migration decisions, where the individuals' choices' set entailed three alternatives: not moving, moving alone, moving with the family. I performed the estimates on men's and women's sample separately.

Our results indicate that it is possible to identify some migration determinants which are common across migration types and genders, while for other dimensions of selection there are significant differences both by gender and migration type. Age and education, for instance, are significantly related to all migration types, regardless of gender, with higher education and lower age increasing individuals' mobility.

Among the factors that are typical of the migration type, and invariant across genders, we can identify marital status, with unmarried individuals being significantly more likely to move alone and less likely to move with the family. The number of children in the age ranges 0-5 and 6-10 significantly and positively influences family migration, while influencing negatively individual migration. The value of business land negatively influences family migration, both for men and women. There are instead other factors that are not only specific to migration types, but also to gender. For instance, the working status influences negatively individual migration and has a positive influence on family migration, only for males.

The analysis also highlighted that the main motivations of individual moves, both for males and females, are own work and education. Hence, this study underlined that there might be a lack of educational infrastructures and job opportunities in migrants' sending areas. Therefore, from a policy perspective, the analysis shows the importance of interventions which increase job opportunities and educational services in regions where they are scarcer.

Moreover, for individual men's migration, we found evidence of lower mobility at intermediate levels of the household income distribution. As shown by Banerjee and Newman (1998), this might be related to the fact that individuals whose household has an intermediate income have a lower incentive to move, compared to individuals in households at the extremes of the income distribution. Indeed, intermediate levels of income at the village-level are likely to give access to a set of village-based informal insurance mechanisms, which would not be available in the city. Therefore, the poorest and the richest would be more likely to move, given that they are less affected by the absence of informal insurance mechanisms in urban areas. This effect might create poverty traps, for those individuals at intermediate household income levels who decide not to move, even in the face of economically advantageous job opportunities at destination.

This effect underlines the importance of the presence of insurance mechanisms, both at origin and at destination. Hence, the results confirm the strong link between migration and insurance market imperfections, which is one of the key insights of the New Economics of Labour Migration.

The above analysis, moreover, confirms that internal migration is a complex phenomenon, which should be analyzed along different dimensions. In this chapter, in particular, I distinguished the determinants of individual from that of family migration. In the next chapter, I will distinguish migration types according to whether migrants are currently away or back at origin.

Bibliography

Alderman, H., Behrman, J. R., Kohler, H.-P., Maluccio, H. A., and Watkins, S. C. (2001). Attrition in Longitudial Household Survey Data. *Demographic Research*, 5(4), 79-124.

Banerjee, A., and Newman, A. (1998). Information, the Dual Economy, and Development. *Review of Economic Studies*, 65(4), 631-654.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4), 813-846.

Becker, G. S. (1974). A Theory of Marriage: Part II. *Journal of Political Economy*, 82(2), S11-S26.

Beegle, K., Weerdt, J. D., and Dercon, S. (2011). Migration and Economic Mobility in Tanzania: Evidence From a Tracking Survey. *Review of Economics and Statistics*, 93(3), 1010-1033.

Buttenheim, A. M., and Nobles, J. (2009). Ethnic Diversity, Traditional Norms, and Marriage Behaviour in Indonesia. *Population Studies*, 63(3), 277-294.

Frankenberg, E., and Karoli, L. A. (1995). *The 1993 Indonesia Family Life Survey: Overview and Field Report*: RAND.

Frankenberg, E., and Thomas, D. (2000). *The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2*: Rand.

Harris, J. R., and Todaro, M. P. (1970). Migration, Unemployment and Development: A Two-Sector Analysis. *The American Economic Review*, 60(1), 126-142.

Hoddinott, J. (1994). A Model of Migration and Remittances Applied to Western Kenya. *Oxford Economic Papers, New Series*, 46(3), 459-476.

Katz, E., and Stark, O. (1986). Labor Migration and Risk in Less Developed Countries. *Journal of Labor Economics*, 4(1), 134-149.

Kuhn, R. (2002). The Logic of Letting Go: Family and Individual Migration from Rural Bangladesh, *Institute of Behavioral Science Working Paper*. Boulder: University of Colorado at Boulder, Population Aging Center.

Lewis, A. W. (1954). Economic Development with Unlimited Supply of Labour. *Manchester School of Economic and Social Studies*, 22.

Lucas, R. E. B., and Stark, O. (1985). Motivations to Remit: Evidence from Botswana. *The Journal of Political Economy*, 93(5), 901-918.

Mendola, M. (2008). Migration and Technological Change in Rural Households: Complements or Substitutes? *Journal of Development Economics*, 85(1-2), 150-175.

Mincer, J. (1978). Family Migration Decisions. *The Journal of Political Economy*, 86(5), 749-773.

Fernández-Huertas Moraga, J. (2011). New Evidence on Emigrant Selection. *Review of Economics and Statistics*, 93(1), 72-96.

Partha, D., and Seck, P. (2009). *Internal Migration, Selection Bias and Human Development: Evidence from Indonesia and Mexico*: United Nations Development Programme (UNDP).

Perloff, J. M., Lynch, L., and Gabbard, S. M. (1998). Migration of Seasonal Agricultural Workers. *American Journal of Agricultural Economics*, 80(1), 154-164.

Ranis, G., and Fei, J. C. H. (1961). A Theory of Economic Development. *The American Economic Review*, 51(4), 533-565.

Rosenzweig, M. R. (1988). Risk, Private Information, and the Family. *The American Economic Review*, 78(2), 245-250.

Rosenzweig, M. R., and Stark, O. (1989). Consumption Smoothing, Migration and Marriage: Evidence from Rural India. *The Journal of Political Economy*, 97(4), 905-926.

Sjaastad, L. A. (1962). The Costs and Returns of Human Migration. *Journal of Political Economy*, 70(5), 80-93.

Stacey, D. (2011). Tenure Insecurity, Adverse Selection, and Liquidity in Rural Land Markets, *Queen's Economics Department Working Paper*. Kingston, Ontario, Canada: Queen's University.

Stark, O. (1991). *The Migration of Labor*. Cambridge, MA: Blackwell.

Stark, O., and Bloom, D. E. (1985). The New Economics of Labor Migration. *The American Economic Review*, 75(2), 173-178.

Stark, O., and Levhari, D. (1982). On Migration and Risk in LDCs. *Economic Development and Cultural Change*, 31(1), 191-196.

Strauss, J., Beegle, K., Sikoki, B., Dwiyanto, A., Herawati, Y., and Witoelar, F. (2004). *The Third Wave of the Indonesia Family Life Survey (IFLS3): Overview and Field Report*: Rand.

Taylor, J. E., and Martin, P. (2001). Human Capital: Migration and Rural Population Change. In B. L. Garner and G. C. Rausser (Eds.), *Handbook for Agricultural Economics*. New York: Elsevier Science.

Thomas, D., Witoelar, F., Frankenberg, E., Sikoki, B., Strauss, J., Sumantri, C., et al. (2010). Cutting the Costs of Attrition: Results from the Indonesia Family Life Survey. *Journal of Development Economics, In Press, Corrected Proof, Available online.*

Todaro, M. P. (1969). A Model of Labor Migration and Urban Unemployment in Less Developed Countries. *The American Economic Review*, 59(1), 138-148.

Todaro, M. P. (1980). Internal Migration in Developing Countries: A Survey. In R. A. Easterlin (Ed.), *Population and Economic Change in Developing Countries* (pp. 361-402): University Of Chicago Press.

Appendix

Table A.1 Determinants of *individual* and *family* migration, by gender
(controlling for past migration)

	Individual migration		Family migration	
	Men	Women	Men	Women
Past migrant	0.037***	0.045***	0.030***	0.045***
	(0.01)	(0.01)	(0.01)	(0.01)
Unmarried	0.072***	0.056***	-0.049***	0.056***
	(0.02)	(0.01)	(0.01)	(0.01)
Age	-0.003***	-0.003***	-0.001***	-0.003***
	(0.00)	(0.00)	(0.00)	(0.00)
HH Head/Spouse	0.009	-0.015*	-0.011	-0.015*
	(0.01)	(0.01)	(0.01)	(0.01)
Head's child	-0.029***	-0.010	0.013***	-0.010
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 0-5	-0.019**	-0.018***	0.018**	-0.018***
	(0.01)	(0.01)	(0.00)	(0.01)
N. children 6-10	-0.017**	-0.015**	0.000**	-0.015**
	(0.01)	(0.01)	(0.00)	(0.01)
N. children 11-15	-0.009	0.001	-0.006	0.001
	(0.01)	(0.01)	(0.00)	(0.01)
N. children 16-20	-0.016	-0.017**	0.002	-0.017**
	(0.01)	(0.01)	(0.01)	(0.01)
N. children 21+	-0.018*	0.002	-0.002*	0.002
	(0.01)	(0.01)	(0.01)	(0.01)
Education	0.002***	0.002***	0.001***	0.002***
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Business land	0.000	0.000	-0.002	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Non-business land	0.001	0.000	-0.001	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Manufacturing	0.014	0.007	0.014	0.007
	(0.01)	(0.02)	(0.01)	(0.02)
Construction	0.020	0.018	0.006	0.018
	(0.02)	(0.04)	(0.01)	(0.04)
Services	0.014	0.020	0.020	0.020
	(0.01)	(0.02)	(0.01)	(0.02)
Not working	0.036**	0.018	-0.001**	0.018
	(0.02)	(0.01)	(0.01)	(0.01)
Monthly own income	0.001	0.000	-0.002	0.000
	(0.00)	(0.00)	(0.00)	(0.00)

Table A.1 (Continued).

	Individual migration		Family migration	
	Men	Women	Men	Women
2nd quintile HH income	-0.005	-0.019***	0.001	-0.019***
	(0.01)	(0.01)	(0.01)	(0.01)
3rd quintile HH income	-0.021**	-0.026***	0.008**	-0.026***
	(0.01)	(0.01)	(0.01)	(0.01)
4th quintile HH income	-0.017*	-0.023***	0.000*	-0.023***
	(0.01)	(0.01)	(0.01)	(0.01)
5th quintile HH income	-0.010	-0.019**	0.010	-0.019**
	(0.01)	(0.01)	(0.01)	(0.01)
N	7,105	8,482	7,105	8,482

Notes:
Robust standard errors in parentheses, clustered at the HH level. *** 1%, ** 5%, * 10% significance level.
All models also include a constant term, rural and province dummies and their interactions.

Chapter 4. Internal Migration and Consumption in Indonesia

Valeria Groppo
Università degli Studi di Milano

Mariapia Mendola
Università degli Studi di Milano Bicocca

1. Introduction

Understanding what leads households in poor contexts to smooth consumption and improve living standards is a fundamental issue within development economics and policy. This question has assumed greater relevance in light of work on the importance of (ex-ante and ex-post) family strategies to manage uncertainty, diversify income sources and alleviate liquidity constraints (Morduch, 1995; Dercon, 2002). In particular, increased attention has been given to the role of labor mobility for household insurance and welfare considerations (Rosenzweig and Stark, 1989; Stark, 1991; Yang and Choi, 2007).

This paper aims at assessing the impact of internal migration on consumption growth at origin in Indonesia. The importance of domestic migration in this country - especially from rural to urban areas - is substantial since almost half of the population experienced internal mobility at least once (Deb and Seck 2009). Using the detailed Indonesia Family Life panel data Survey (IFLS) we focus on migrant-sending households and explore to what extent current and return domestic migration improve the living conditions of those who stay behind.

Several recent studies have investigated the role of migration in shaping household behavior in communities at origin, with a particular concentration on the effects of international migration on the one hand, and (international) remittances on the other (Adams, 2005; Yang, 2008; Gibson et al. 2011). Yet, while domestic (temporary or permanent) mobility is also an important livelihood strategy in several countries, the short and long-term impact of these migration flows has received less attention in recent research, particularly in Asia (main exceptions are de Brauw and Harigaya 2007 for Vietnam and Du et al 2005 for China). This is due in part to the sheer magnitude of international migration and financial flows worldwide, as well as to the shortage of detailed data on internal (seasonal and/or informal) flows (deBraw *et al.*, 2011, Mendola, 2008). On the other hand, given the different features of the internal migration process with respect to international mobility in developing countries (e.g. less positive self-selection with respect to skills or income, lower remittance rates *via* mostly informal channels), a better understanding of the welfare effects of internal labor mobility on people left behind is fundamental from the point of view of national poverty reduction strategies. Moreover, the focus on the welfare impact of remittances only, i.e. in isolation from migration, has been largely questioned as inappropriate given the complexity of the relationships involved in the migration process (see McKenzie and Sasin, 2007). Indeed, migration entails (monetary and non-monetary) costs for households and communities at origin (e.g. forgone income and other household inputs

that migrants would have generated) so that empirical analysis is needed in order to shed light on ambiguous *a priori* predictions.

Thus, in this study we measure the short-term impact of having at least one migrant household member in the period 1997-2000 on Indonesian origin households' consumption path. The comprehensive panel data survey as well as the high response and tracking rates allow us to distinguish between current and return migrants, as to uncover channels through which migration welfare effects may operate (McKenzie and Sasin, 2007).

A significant challenge related to migration research based on non-experimental data, though, is controlling for unobserved heterogeneity which is likely to affect both the migration decision and outcomes of people left behind. To do so we use longitudinal data, which are relatively scant in developing contexts, and a difference-in-difference (DID) estimator. More in detail, our empirical strategy considers households' per capita consumption growth between 1997 and 2000 as a function of the migration behavior of their members during the considered time span, controlling for a large set of time-varying household characteristics and location fixed effects.

A DID specification removes all sources of time-constant unobserved heterogeneity. Yet, if unobserved characteristics shaping both the migration process and the living conditions of remaining household members vary in time and space (e.g. ambition), DID estimates will be still biased and inconsistent. Thus, we also use an instrumental variable (IV) method as to take time-varying unobserved heterogeneity into account. A similar approach has been used by Beegle *et al.* (2008) to measure the migration gains for those who move internally in Tanzania. As we said, though, our interest is not on migrants themselves but on the migration impact on household members left behind - impact that critically shapes the overall effect of migration and thereby the design of appropriate migration and welfare policies.

Our results show that households with current internal migrants have either a small or no significant impact on consumption growth compared to household without migrants in the considered period. Differently, origin households having return migrants had significantly lower growth of per capita consumption, compared to non-migrant households. In particular, our IV results indicate that households having return migrants had 60% lower per capita consumption growth than households not having return migrants in the considered period.

The paper is structured as follows: Section 2 provides a summary of the literature on the economic effects of migration at origin. Section 3 describes the characteristics of the dataset, and especially of the consumption and migration modules. Section 4 provides descriptive statistics on the evolution of consumption between 1997 and 2000, by households' migration status. Section 5 entails the econometric analysis on the effect of migration on consumption growth. Section 6 describes the characteristics of migrants and non-migrants in 2000, aiming at interpreting the results, while Section 7 concludes.

2. Literature review

This section reviews the main theoretical and empirical contributions on the micro-economic effects of migration. Given the focus of the paper, we review in particular the literature on *internal* migration and its economic effects on people left behind.

2.1 Theoretical studies on the determinants and effects of migration

The classical and neoclassical models of Lewis (1954) and Ranis and Fei (1961), as well as the Todaro models focus on the *labour market effects* of migration. The Lewis model assumes the existence of a capitalist and a non-capitalist sector, which is associated with urban and rural areas. The capitalist sector hires labour and sells output for a profit, and, as it expands, it draws labour from the other sector. The main assumption of the model is that rural labour is available in unlimited supply to the urban sector, at a fixed real wage. This implies that in the rural sector the marginal product of labour is zero and individuals' move from rural areas do not decrease agricultural production, nor increase rural wages. The only potential effect of migration is an increase in the average product of those who remain in rural areas. These predictions have been invalidated by evidence that, even in areas generally characterized by surplus of labour, the marginal product of labour may be positive in seasonal peaks (Gregory, 1986).[60]

The two-sector model of Ranis and Fei (1961) extends the model of Lewis (1954), by assuming that migration equalizes the marginal products of labour (and thus wages) across areas. However, these models are not consistent with evidence of persistent wage differentials across sectors, even in the presence of migration.

The above considered models assume full employment in the urban sector and focus on inter-sector wage differences in explaining rural-urban migration. They have been criticised mainly because they are not consistent with empirical evidence of continuing rural-urban migration even in the presence of high urban unemployment and of persistent differences in the returns to homogeneous labour across sectors.

The models of Todaro (1969) and Harris and Todaro (1970) incorporate *labour market imperfections*, and provide an explanation of why migration continues even in the presence of high urban unemployment. Also these models emphasize the role of income differentials in explaining rural-urban moves, but depart from the neoclassical models because they do not assume full employment in the urban sector. In the Todaro models, the urban sector is indeed characterised by unemployment, which implies that migrants' earnings depend not only on the level of urban wages, but also on the probability to find a job. Indeed, migration decisions depend on the differences between rural and urban *expected* rather than *actual* earnings, where expected earnings are defined as the *present value* of urban real income streams over the worker time horizon. Expected earnings are determined by the actual rural-urban wage differentials and by the probability for migrants to find a job at the urban destination. Given these quantities, rural-urban migration continues as long as rural earnings are lower than expected urban ones. In particular, the rural-urban gap in actual

[60] The Lewis' model is also subjected to the more general critique of treating the rural sector as a "black box", without analyzing the potential consequences that labour moves may have on it (Taylor and Martin, 2001).

earnings may be so high that, even in the face of high urban unemployment (and thus lower probability for migrants to find a job), urban expected income is higher than the rural one and thus migration is economically rational. This model is consistent with migration happening even in the face of low probability to find a job in the initial period in urban areas, as long as the present value of future income streams compensate for the initial losses. Moreover, in these models employment creation in urban areas pushes a higher number of individuals to leave their rural dwellings, and may eventually increase rather than decrease urban unemployment, especially if the elasticity of migration with respect to urban jobs is high. Therefore, Todaro (1969) concludes that any effort devoted to decreasing urban unemployment might be worthless if it is not accompanied by a concomitant effort in reducing the gap in living standards between urban and rural locations.

The Lewis's and Todaro's models do not analyze the consequences of the migrants' *capital stock* move, on welfare at origin. Differently, Grubel and Scott (1966) and Berry and Soligo (1969) examine the role of the capital moves associated with migration, in determining welfare consequences for origin countries.[61] The first of these two studies focuses on the role of emigrants' human capital in determining welfare at origin. By encompassing the role of remittances, the study concludes that emigration of highly skilled individuals causes only marginal welfare losses at origin, which are primarily concentrated in the short run and are mainly due to the training costs associated with migrants' replacement. Hence, according to Grubel and Scott (1966), in the long run, human capital migration is likely to bring sizeable benefits to emigrants' origin countries, and these benefits are mainly related to the fact that the emigrants' productivity increases due to migration spread to emigrants' origin countries, through the diffusion of technological knowledge. Berry and Soligo (1969) challenge this conclusion, in a model which refers to emigrants' capital stock in general, not only to their human capital. In particular, assuming constant return to scale, perfect markets, and absence of remittances, the authors show that emigration decreases the income of the non-emigrant population, except for the case in which migrants own a relatively large proportion of the capital stock, and they leave at least part of their capital at origin. In this framework, as underlined by Taylor and Martin (2001), human capital emigration determines a welfare loss for origin countries.

The model of Lucas (1987) departs from the classical models by assuming that surplus of labour might not be present, and thus migration might increase domestic wages and reduce crop production at origin in the short run. Moreover, this study encompasses the New Economics of Labour Migration (NELM) assumptions of imperfect insurance and capital markets, allowing for the possibility that migrants' earnings might ease capital accumulation at origin and thus enhance productivity in the long run.

[61] Grubel and Scott (1966) and Berry and Soligo (1969) refer to international migration, but their conclusions are relevant also for internal migration.

2.2 Empirical evidence on the economic effects of migration at origin

This section describes key concerns arising when estimating the effects of migration and/or remittances, and reviews the main solutions which have been used in the literature. Moreover, we review the most significant contributions providing evidence on the economic effects of migration at origin, focusing on the analyses of consumption and well-being at origin.

2.2.1 Methodological concerns in the estimation of the impact of migration

The first key concern when estimating the effect of migration is to correctly formulate the research question. In this respect, McKenzie and Sasin (2007) underline the importance of posing the research question about the impact of migration in a broad sense, rather than restricting the analysis to the impact of remittances. Indeed, analyzing the impact of remittances only could be misleading, because migration has a series of implications on origin and destination areas, which are not all channeled through remittances.

For instance, as highlighted by Taylor and Mora (2006), migration alters prices and incomes constraints faced by households: it indeed implies a loss of family labour, which increases the opportunity cost of time for household members; migration also eases households' integration with distant markets, lowering the transaction costs and altering the prices faced by households. Because of migration, hence, households may substitute purchased goods for home-produced goods and shift from more to less time-intensive home produced goods. Migration also alters the information and human capital constraints faced by households: the change in households' cultural standards and the reorganization of households' finances implied by migration may indeed vary their preferences and create demand for new goods and services, with an increase of non-local goods in households' consumption set. Finally, through its risk-reducing impact, migration can also alter households' production technologies (Mendola, 2008).

From the above, it is clear that limiting the analysis to the impact of remittances could provide partial or misleading results.[62] Therefore, in the econometric analysis that follows we will focus on the broader impact of migration on origin households' consumption.

The analysis of migration effects is complicated by the fact that households' decisions about migration and other investments are made simultaneously. Moreover, elements which determine one decision could also affect the others. For instance, households engaging in migration, may decide to cut other forms of investments in order to finance migration costs. In this framework, the main problems which arise estimating the impact of migration are reverse causality, selection bias, and omitted variables.

Reverse causality is present when the outcome of interest influences migration and not vice-versa. For instance, if a household reduces its expenditure in some items to finance migration costs, one might find a negative relationship between migration and expenditure and conclude that migration has a negative impact on the latter.

[62] This is particularly true for the studies analyzing how remittances are spent. This approach, indeed, ignores the fact that money is fungible (i.e. income from different sources are generally pooled) and thus is likely to provide only a partial picture of the true impact of remittances on households' spending patterns.

Selection bias arises because migrants are not a random sample of the population, but are instead selected according to specific dimensions, some of which are observable (e.g. education level), while others are unobservable (e.g. ability, entrepreneurial spirit). Failure to control for these selection criteria may bias the estimated effect of migration on the outcome of interest. If, for instance, wealthier households are more likely to send migrants, any significant effect of migration on a certain outcome would be biased, if estimated without controlling for pre-migration household wealth.

Omitted variables problems arise because there could be factors not observed by the researcher, which influence both migration decisions and the outcome of interest. Not considering this, may produce biased estimates of the impact of migration on the outcome of interest. For instance, a positive economic shock in the area where a household lives (e.g. an increase in rice prices for a rice producer household) may increase both its expenditure, and its number of migrants, given the increased possibility to finance migration costs. Any estimated positive effect of migration on household consumption would be upward biased, if obtained without taking into account the effect of the shock.

Different econometric strategies can be used to solve the above problems, with various degrees of success and different costs. Randomized experiments represent the methodology which provides the best results, in terms of unbiasedness of the estimated causal effect of migration. They imply the distinction of a control and a treatment group, both randomly selected among the population. The treatment group is given the right to migrate, which is instead denied to the control group. The impact of migration is then evaluated by confronting the outcomes for control and treatment groups. Examples of studies which adopt this methodology are Gibson et al. (2011), and McKenzie *et al.* (2010). However, these types of experiments are quite rare and non-experimental econometric methods are more common in the literature.

Among non-experimental studies, the most common techniques involve the reconstruction of counterfactuals, and the use of instrumental variables (McKenzie and Sasin, 2007).

The counterfactual can be obtained in different ways, among which we find the 'manual' reconstruction or the propensity score matching technique. In the case of manual reconstruction, we focus on one type of households, say households having migrants, and we ask which characteristics would these household have had in case they did not had migrants. For instance, to obtain the counterfactual household income, we would deduct remittances from households' income and impute a hypothetical income had the households not had migrants. The imputation is problematic, because the decision of which function to use to predict earnings is not clear-cut. The propensity score matching technique implies the comparison of a migrant household, with a non-migrant household with the same propensity of having migrants. This method entails the estimation of a migration decision regression. Propensity score matching can lead to a considerable improvement over OLS. However, the selection of 'similar' households to compare is made on the basis of observable characteristics only. Therefore, this method does not correct for possible biases which can arise due to unobservable variables.

The use of panel data addresses addresses the problem of time-invariant unobservables, while the use of instrumental variables also corrects for the presence of time-variant unobservables.

As highlighted by McKenzie and Sasin (2007), the choice of which instrumental variables (IVs) to use for the migration variable is determined both by data availability and by the outcome of interest. The outcomes of interest can be classified in two main groups: outcomes which regard origin countries/households (e.g. origin households' total or education expenditure), and outcomes related to the destination countries (e.g. migrants' employment and income at destination). We follow McKenzie and Sasin (2007) in reporting examples of instruments that are more appropriate in the two cases above outlined:

1. *IVs for outcomes at origin:*
    a. Historical or current migration networks at the community-level at origin are used as instrument for migration, on the basis that networks have been shown to decrease migration costs, thus directly easing migration. Network variables are good instruments for analyses of migration impacts on origin countries, while they are likely to perform poorly when analysing outcomes at destination. This is because migration networks directly influence also migrants' situation at destination (for instance, they have a direct impact on migrants' probability of finding a job at destination). Network variables have been used, among others, by Hildebrandt and McKenzie (2005) and Mansuri (2006), when analyzing the impact of migration on children's health in Mexico and Pakistan, respectively. In particular, Hildebrandt and McKenzie (2005) used historical (1924) US migration rates for the states in which households are located, and underlined that historical migration rates, reflecting the early development pattern of railroads in Mexico, do not influence children's health seventy years later, a part from their effect through current migration.[63] Mansuri (2006) used current village-level migration networks, measured as the percentage of households in the village having a current migrant. The author acknowledged that village-level networks are correlated with village-level unobservable variables which could have an independent impact on children's health. To solve this problem, she constructed an instrument varying within villages, by interacting the network variable with dummies indicating households' land ownership group. The underlying assumption is that access to migrant networks vary within landholding groups, with landholding being an exogenous characteristics of households.[64] Migration network variables have also been used as instruments for migration by Acosta (2011), Beaudouin (2006), Mendola (2008).
    b. Cultural factors or norms that impact migration have also been used in the literature. For instance, Mansuri (2006) used the number of adult males in the household as an instrument for migration, when analyzing the impact of migration on children's health in rural Pakistan. The author underlined that, in the considered context, the presence of at least an adult male in the household is generally required, and that households with a single adult male are unlikely to undertake migration. To maintain the exclusion restriction, she showed that, conditional on a set

[63] Hildebrandt and McKenzie (2005), p. 267.
[64] The author highlighted that in rural Pakistan land can be considered exogenous to households, since it is mainly inherited. Thus, it is unlikely that more entrepreneurial households both acquire more land and invest more in their children (Mansuri, 2006, p. 9).

of household characteristics, the number of adult males in the household does not influence the outcome of interest, except for its effect through migration. Among cultural factors and norms, also age and the relationship to the household head have been used as instruments. For instance, Beegle et al. (2011) used individuals' age rank in the household among those aged 5-15 interacted with a dummy indicating that the individual is in the age range 5-15 as an instrument for individuals' migration decisions in rural Tanzania. The same authors also used indicators for being the household head or her/his spouse, and for being the son of the head. The use of these instruments derives from the fact that, by cultural norms, in rural Tanzania (as in other developing settings) the relationship of an individual with the household head influences his/her probability of moving.[65]

c. Economic shocks at destination have been used for instance by Yang and Martinez (2005) and Yang (2008) to analyse the impact of remittances on poverty and household investments, respectively. In particular, these authors exploited the fact that the 1997 Asian financial crisis determined a wide variation in the depreciation rates of the Philippino peso across different currencies at destination. Hence, remittances received by households at origin were subject to divergent exogenous exchange rate shocks, according to the source country.

2. *IVs for outcomes at destination:*

a. Distance: for instance, McKenzie et al. (2010) use distance from the New Zealand consulate in Tonga to instrument migration, when analysing the impacts of moving on migrants in New Zealand. Distance variables are not appropriate when analysing the impact of migration at origin, because distance is likely to have a direct effect on a number of origin countries/households' characteristics.

b. Natural shocks at origin: Munshi (2003) use rainfalls in Mexico to instrument migration, when analysing outcomes abroad. As distance, natural shocks at origin are not an appropriate instrument when analysing an outcome at origin, since they are likely to have a direct effect on the outcome of interest, besides their effect on migration.

In the econometric analysis that follows we rely on the use of panel data and instrumental variables, in order to address the endogeneity of migration. Given that we are evaluating the impact of migration on origin households' consumption, we focus on those instruments which are more appropriate in the analysis of outcomes at origin, using in particular migration networks and cultural factors variables.

[65] In particular, in many developing countries (adult) children of the household head are generally less likely to move from their origin communities, relative to other household members with similar characteristics. Moreover, as explained by Beegle et al. (2011), in rural Tanzania local norms on marriage are patrilocal, which implies that a girl is expected to move to her husband's community after marriage, while the husband is expected to stay where his father was based. Hence, sons of the head are likely to be significantly less mobile than household members with similar characteristics, but having a different relationship with the household head.

2.2.2 Empirical studies on the impact of migration on consumption at origin

The seminal paper of Rosenzweig and Stark (1989) provides evidence of women's marriage migration being a strategy used by households to smooth consumption over time. In particular, this study shows that both the number of married women and the distance between the origin household and that of the marital partner significantly reduce the variability of the origin household's food consumption. Using a 10-years time series longitudinal dataset on households in 6 Indian villages, the study estimates both the impact of migration on household consumption variability and the impact of profit variability on migration decisions within households. In estimating the impact of migration on consumption variability, the authors do not address the endogeneity of the migration variables, and therefore the estimated effect of migration on consumption cannot be directly given a causal interpretation. However, when estimating the impact of profit variability on the number of migrants in the household, the authors instrument profit variability with rainfalls variance interacted with households' landholdings, finding that it is those households having (exogenously) greater profit variability that invest more in migration. Since higher profit variability implies a greater need of consumption smoothing, the study provides support to the hypothesis that migration is used to smooth consumption, and thus to a causal interpretation of the negative relationship between consumption variability and marriage migration in India.

Rosenzweig and Stark (1989) looked at marriage *cum* migration as this is the main form of geographical mobility in India, whilst (rural-to-urban) economic migration is a relatively small component of total migration in that context. This is not the case in several developing countries, though, and there has been increasing empirical evidence on the economic impact of migration and remittances on household well-being at origin. Studies focusing on international people and remittance flows include Adams (2005, 1998, 1991), Yang and Choi (2007), Yang (2008) and Gibson et al. (2011) among others. The latter study in particular considers out-migration from the Kingdom of Tonga and provides experimental estimates of the impact of international migration on household members left behind. The study considers a set of different outcomes, such as household (current) income, household size and composition, durable assets and financial assets and health. It finds that emigration leads to a significant reduction in household labour per capita income, which is only partially offset by the increase in remittances. Emigration also has a negative impact on assets' ownership by household members who remained at origin. Without experimental data instead, Yang (2008) makes use of exogenous shocks to the income and wealth of Philippine migrants worldwide due to the 1997 Asian financial crisis to show that, in contrast with previous studies, exogenously-determined changes in remittances have negligible effects on household consumption but large effects on various types of household investments. Households experiencing more remittances, though, raise their investment-related disbursements in several areas, in particular in educational expenditures. Nevertheless, as argued above, the focus on remittances only may be misleading plus international and internal migration are far from being similar in their own patterns and characteristics- hence also their effects are likely to be different.

In a recent work Taylor and Mora (2006) evaluate whether migration influenced households' expenditure shares, distinguishing between international and domestic migration. Using cross-section household-level

data on Mexican households, and instrumenting household-level migration through the number of family members at foreign or internal migration destinations 12 years prior to the survey, the study finds that households with international migrants have a significant larger marginal budget share on investment and consumer durables than otherwise similar non-migrant households. Differently, households with internal migrants have a marginal budget share for investment and consumer durables that is lower than that of non-migrant households. On the other hand, by using panel data from Vietnam, deBraw and Harigaya (2007) show that migration increases annual household expenditures and decreases the poverty headcount in the country, concluding that migration played an important role in the improvement of living standards of people left behind.

Another paper which is relevant for our study is the one by Witoelar (2005), which answers the question whether extended families in Indonesia pool their incomes to smooth their consumption. An extended household is defined as the set of households originating from the same original household, through the detachment of some of the original members.[66] The author performs two main analyses: first, he performs a test of perfect income pooling at the extended household-level. This estimation implies evaluating the impact of households' own income on their consumption, once the resources of the extended family are controlled for, through the inclusion of extended-family fixed effects. If households perfectly pool their incomes, a household's own income should not have any effect, once the extended family's resources are controlled for. Consistently with previous literature testing risk-sharing, the results of this study show that changes in households' own income do affect households' own consumption among different units in an extended household. This means that housheolds do not perfectly pool their incomes. However, the magnitudes of the estimated income coefficients are small, suggesting that households may pool their incomes to some degree. This study also entails the regression of other households' incomes on a household's own consumption (within an extended family): overall, incomes of other households do not seem to influence own household consumption. The author also estimates a reduced form regression of household consumption on all exogenous characteristics of other households, finding that other households' variables do have some effect, if only small, on a household's own consumption. This study is relevant for our analysis because extended households can originate from the migration process. However, the study does not distinguish between extended households whose components all reside in close locations, from those extended households having units in distant locations (extended and spatially diversified housholds). It could be that the degree of income pooling is higher among extended households that are also spatially diversified.[67] In our analysis on the impact of migration on consumption at origin, our sample households are both extended and spatially diversified. Therefore, our analysis, even if performed through a different methodology not entailing a direct test of income pooling as the one of Witoelar (2005), can shed some light on the presence of income pooling between extended and spatially diversified households in Indonesia. Indeed, if the presence of a current

[66] An extended household is thus formed by an original household, plus one or more split-off households.

[67] The degree of income pooling could be particularly higher for extended households formed through temporary migration. The intention to return, indeed provides temporary migrants with an incentive to remit more, in order to maintain better linkages with their origin households.

migrant is found to have a significant impact on the consumption of household members at origin, this could signal a transfer of resources from the split-off household established by the migrant to his/her origin household.

3. Data

The objective of this chapter is to evaluate whether the fact of having migrants in the period 1997-2000 had an impact on origin households' consumption growth in the same period. In order to do this, we used data on consumption and on the presence of migrants for origin households in the period 1997-2000. Both types of data are provided by the Indonesia Family Life Survey, of which we use the second and third waves (IFLS2 and IFLS3). The main characteristics of the dataset were explained in the previous chapter. This section explains how the relevant variables on household-level migration and consumption were obtained, starting from the original dataset.

3.1 Consumption data

IFLS contains data on monthly household consumption, expressed in real terms. For the waves IFLS2 and IFLS3, temporal and spatial deflators are available, which transform the nominal values in real values with base Jakarta in December 2000.[68]

IFLS provides the following aggregates for expenditure:

- *Food*: it is recorded with reference to the week prior to the interview and it includes both the value of purchased food items, and the value of the food produced by the household or received as a gift. The data also entail a variable recording the value of food given to other parties outside the household during the week preceding the interview.
- *Non-food*: as for the food category, both market purchases and own-production or gifts are included in the non-food consumption aggregate, which entails:
    - Frequently purchased goods and services: they are recorded with reference to the past month and include electricity/water/phone, personal toiletries, household items (e.g. cleaning supplies), domestic services, recreation and entertainment, transportation, sweepstakes, Rotating Savings and Credit Associations (ROSCAs), the values of non-food items give to others outside the household on a regular basis.
    - Less frequently purchased goods and services: they are recorded with reference to the past year, and then converted to monthly figures by dividing by 12. They include clothing, furniture, medical costs, ceremonies, taxes and other.
- *Education*: schooling expenditure were asked with reference to the past year, then converted to monthly figures. They are divided in expenditures for children in the household, and outside the household.

[68] Witoelar (2009) provides further details on the construction of IFLS temporal and spatial deflators.

- *Housing*: households reported the value of their monthly rent, or, in the cases in which they owned their house, they reported how much money they would have paid if they were renting the house.

The objective of our analysis is to evaluate the impact of having a migrant on total household consumption at origin. Therefore, to obtain our dependent variable, we sum all the above items of expenditure, included expenditures for ROSCAs and ceremonies. We only exclude transfers to members out of the household, given that they are likely to be sent to current migrant members. Therefore, if we include transfers out, any increase of household per capita consumption may be related to migrants' consumption, not to the consumption of those individuals who remain at origin. The measure of consumption we use is expressed in monthly real terms, per capita.

3.2 Household-level data on the presence of migrants in the period 1997-2000

To obtain data on the presence of migrants in the household during the period 1997-2000, we started from the IFLS3 migration module, in which target individuals provided all their moves crossing village boundaries and lasting at least 6 months, from age 12 up to year 2000.[69] Using the information on the year of migration, we constructed a dummy variable indicating whether the individual moved in 1997-2000. We define *migrants* those individuals who moved in the period 1997-2000 for at least 6 months.

Out of 22,024 individuals, 3,321 (15.1%) moved at least once in 1997-2000 for periods of at least 6 months. It is important to underline that here we are considering only those individuals who were interviewed both in 1997 and in 2000. We do not consider individuals who entered the survey as new members in 2000, because for them we would not have the information on his or her location in 1997. This information is crucial to distinguish between current and return migrants. Current and return migrants can indeed be identified by looking at migrants' households' identification code and location in 1997 and 2000. Some of the individuals in our sample were present also in 1993: for them, we take into account also their 1993 location, when defining their return status. In particular, current and return migrant are defined as follows. A *current migrant* is an individual who, in year 2000 resided (at least) in a different village with respect to his or her location in 1997. A *return migrant* is instead an individual who moved between 1997 and 2000, returning by 2000 either to the 1997 or the 1993 location (if he or she was also interviewed in 1993). Therefore, according to our definitions, those individuals who are classified as non-migrants are individuals who did not move in the period 1997-2000.

Once obtained the information of the migration status of individuals between 1997 and 2000, we summed by 1997 origin households the number of migrants. We define *migrant households* those households having at least one migrant, current and/or return, in the period 1997-2000. A *return (current) migrant household* is

[69] Therefore, all individuals who moved for less than 6 months are classified as non-migrants. This could induce a bias in our results, since our control group (formed by non-migrants) includes individuals who might have experienced migration for less than 6 months. Therefore, a positive estimated effect of migration on expenditure from our dataset could represent a downward biased estimate of the overall impact of migration, if also migration for less than 6 months has a positive impact on origin household expenditure.

a household having at least one return (current) migrant. Out of 6,770 origin households, 2,055 (30.4%) had at least one migrant (either current or return) in the period 1997-2000.

In this chapter, we aim at isolating the effect of having a migrant member for origin households. In order to have a clear evaluation of this effect, we have to be sure that the origin households did not move in the considered period. Indeed, if a household both moved and had migrants, the outcome of interest could be influenced both by the move of the household and by the presence of migrants, and we would not be able to precisely evaluate the impact of having a migrant. Therefore, we dropped those origin households who moved between survey waves (1508 households, 22.3 % of origin households). The survey also gathers information about the reason of each move. First, we will evaluate the impact of migration considering migrants who moved for any reason. Then, we will distinguish between households having 'tied' movers only, and households having at least one 'non-tied' mover. Tied movers are defined as those individuals who moved following others (e.g. a child who migrated following her parents), while non-tied movers are those who moved for own work or education.

We also lost some information due to missing data on the other variables used in the analysis. Therefore, our final sample includes 5,270 households, of which 1,388 (26.3%) had at least one migrant (either current or return) in the period 1997-2000.

The following table summarizes the number and proportion of migrants at the individual- and household-level.

Table 1. Individual-level and household-level migration status

Individuals				Their origin households	
	N	%		N	%
Non-migrants	15,106	89.0	Non-migrants	3,882	73.7
Migrants, of which	1,860	11.0	Migrants, of which	1,388	26.3
Current	1,134	6.7	At least one current	919	17.4
Return	726	4.3	Return only	469	8.9
Total	16,966	100.0		5,270	100.0

Source: IFLS2-3.

Note: Information about individuals' migration status is obtained from the migration history module. Household-level information is obtained by aggregating over origin households information about the individuals.

4. Descriptive statistics

Before passing to the econometric analysis, we provide statistics on migrant and non-migrant households, describing among others their composition and wealth. This analysis, by shedding light on how migrant households are 'selected' in Indonesia, complements the one of the previous chapter, which focused on the individual-level migration estimates. Here we are analyzing the characteristics of migrant and non-migrant households at the baseline year 1997, thus before the considered migration episodes, which, as above mentioned, refer to the period 1997-2000.

As the next table shows, households having migrants have a significantly higher number of members than non-migrant households, and the difference is higher between non-migrant and current migrant households. These figures are counter-intuitive, since one would expect that households having members away have a lower number of members at origin. However, the table also shows that households having current migrants have a significantly higher number of children in the age 0-5 than non-migrant households, while having a lower number of individuals older than 11 years. Taken together, the figures in table 2 suggest that households having current migrants have a significantly higher size than non-migrant households, because they entail a greater number of children in the age range 0 to 5 years, compared to non-migrant households. This is confirmed by the fact that the average age of household members is significantly lower in households having current migrants than in non-migrant households.

Table 2 also shows that in migrant households (either current or return) the head is significantly older than in non-migrant households. This is likely to be related to the fact that more often are the youngest who move.

Table 2. Household structure, by migration status

	HH migration status				
	Non-migrant	Return	Current		
	(1)	(2)	(3)	Diff (1)-(2)	Diff (1)-(3)
Household size[a]	4.29	4.87	5.38	-0.57***	-1.08***
	(1.91)	(2.06)	(2.07)	(0.09)	(0.07)
N. children 0-5	0.50	0.49	0.43	0.01	0.07**
	(0.71)	(0.72)	(0.67)	(0.03)	(0.03)
N. children 6-10	0.51	0.46	0.47	0.05	0.04
	(0.70)	(0.69)	(0.68)	(0.01)	(0.03)
N. children 11-15	0.52	0.67	0.81	-0.15***	-0.29***
	(0.71)	(0.79)	(0.82)	(0.04)	(0.03)
N. children 16-20	0.37	0.60	0.79	-0.24***	-0.43***
	(0.65)	(0.76)	(0.82)	(0.03)	(0.03)
N. males 21+	1.07	1.21	1.33	-0.15***	-0.26***
	(0.61)	(0.79)	(0.82)	(0.03)	(0.03)
N. females 21+	1.25	1.35	1.47	-0.10***	-0.22***
	(0.61)	(0.68)	(0.78)	(0.03)	(0.02)
Average age in HH	32.25	30.20	29.61	2.05***	2.64***
	(14.87)	(10.24)	(9.47)	(0.71)	(0.51)
Head's age	48.23	49.83	50.45	-1.60**	-2.22***
	(13.99)	(12.00)	(11.37)	(0.67)	(0.49)
Head is unmarried	0.01	0.01	0.01	0.01	0.01
	(0.11)	(0.08)	(0.09)	(0.00)	(0.00)
N.	3,882	469	919		

Source: IFLS2-3.

Note:

a. Household size is measured excluding current migrants (i.e. only individuals physically present in the household are considered). The same holds for the other variables of household composition included in the table.

The next table informs on the direction of migrant households' selection, in terms of education, employment sector of the household head, and number of workers in the household at the baseline. Both in households having return and in households having current migrants the head is significantly less likely to having no education non-migrant households. Moreover, the heads of households having current migrants are significantly more likely of having senior secondary (or university) education than the heads of non-migrant households. These figures confirm that, as we saw in the previous chapter, migration involves a positive selection process in terms of education.

As for the head's employment, table 3 indicates that the household head is more likely to be employed in the primary sector in non-migrant than in migrant households (either current or return). Moreover, the heads of households having current migrants are significantly more likely to be employed in the services sector than the head of non-migrant households. These figures are consistent with the positive selection in terms of education we just described. From the table, we can also notice that non-migrant households are significantly

less likely to have a non-working head than return or current migrant households. This might be related to the fact that, as we saw in the previous table, the head of a household having return migrants is significantly older than the head of a non-migrant household. Moreover it could signal the fact that the head who was not working at the baseline migrated in the period 1997-2000 to find a job.

Finally, from table 3 we can see that households having migrants (either current or return) have a higher number of private sector workers than non-migrant households in 1997. Moreover, households having current migrants have a significantly higher number of government workers than non-migrant households. These figures confirm the positive selection of migrant households in terms of education.

Table 3. Household head's education and employment, by household migration status

	Household's migration status				
	Non-migrant	Return	Current		
	(1)	(2)	(3)	Diff (1)-(2)	Diff (1)-(3)
Head's education					
None	0.23	0.19	0.15	0.04**	0.07***
	(0.42)	(0.39)	(0.36)	(0.02)	(0.02)
Primary/Jr. Secondary	0.62	0.64	0.62	-0.02	-0.01
	(0.48)	(0.48)	(0.48)	(0.02)	(0.02)
Senior secondary	0.12	0.14	0.15	-0.02	-0.03***
	(0.32)	(0.34)	(0.36)	(0.02)	(0.01)
University	0.03	0.03	0.06	0.00	-0.03***
	(0.18)	(0.18)	(0.25)	(0.01)	(0.01)
Head's employment					
Primary	0.38	0.34	0.33	0.04*	0.05***
	(0.49)	(0.47)	(0.47)	(0.02)	(0.02)
Manufacturing	0.10	0.08	0.08	0.02*	0.02*
	(0.30)	(0.27)	(0.28)	(0.01)	(0.01)
Construction	0.06	0.08	0.06	-0.02*	0.00
	(0.24)	(0.27)	(0.23)	(0.01)	(0.01)
Services	0.30	0.31	0.35	-0.01	-0.04**
	(0.46)	(0.46)	(0.48)	(0.02)	(0.02)
Not working	0.15	0.19	0.18	-0.04**	-0.02*
	(0.36)	(0.39)	(0.38)	(0.02)	(0.01)
N. workers in household					
Private sector	0.58	0.69	0.71	-0.11***	-0.14***
	(0.81)	(0.86)	(1.01)	(0.04)	(0.03)
Government	0.12	0.13	0.19	-0.01	-0.08***
	(0.38)	(0.39)	(0.49)	(0.02)	(0.01)
N.	3,882	469	919		

Source: IFLS2-3

In the next table, we explore the selection of households in the migration process, on the basis of their total per capita real monthly income and expenditure. We can see that households having current migrants have a significantly lower probability of belonging to the first quintile of the distribution of total per capita

household income compared to non-migrant households. Moreover, households whose income belong to the highest quintile of the distribution have a significantly higher probability of having current migrants compared to non-migrant households. Hence, current migrant households are positively selected in terms of total per capita household income. The same can be said looking at households per capita expenditure. For return migrant households, we find a different pattern: indeed, non-migrant households have a significantly higher probability of belonging to the fourth and fifth quintiles of the distribution of per capita income, compared to households only having return migrants. Hence, return migrant households are negatively selected in terms of monthly per capita household income at the baseline. Return migrant households appear to be negatively selected also in terms of expenditure: indeed households having return migrants are more likely to belong to the second lowest quintile of the distribution of per capita household expenditure at the baseline, compared to non-migrant households.

Table 4. Households' per capita income and expenditure, by migration status

	Household's migration status				
	Non-migrant	Return	Current		
	(1)	(2)	(3)	Diff (1)-(2)	Diff (1)-(3)
Quintiles of total per capita income					
1st	0.22	0.21	0.19	0.02	0.03**
	(0.42)	(0.40)	(0.39)	(0.02)	(0.02)
2nd	0.22	0.23	0.20	-0.00	0.01
	(0.41)	(0.42)	(0.40)	(0.02)	(0.01)
3rd	0.21	0.21	0.18	-0.00	0.02*
	(0.40)	(0.41)	(0.39)	(0.02)	(0.01)
4th	0.19	0.16	0.19	0.03*	-0.01
	(0.39)	(0.37)	(0.39)	(0.02)	(0.01)
5th	0.16	0.19	0.22	0.02*	-0.06***
	(0.37)	(0.39)	(0.41)	(0.02)	(0.01)
Quintiles of total per capita expenditure					
1st	0.22	0.19	0.19	0.02	0.03**
	(0.41)	(0.39)	(0.39)	(0.02)	(0.02)
2nd	0.21	0.25	0.21	-0.04**	0.00
	(0.41)	(0.43)	(0.40)	(0.02)	(0.01)
3rd	0.21	0.19	0.19	0.00	0.01
	(0.41)	(0.39)	(0.39)	(0.02)	(0.01)
4th	0.19	0.17	0.21	0.02	-0.01
	(0.39)	(0.39)	(0.41)	(0.02)	(0.01)
5th	0.16	0.18	0.19	-0.02	-0.03**
	(0.37)	(0.38)	(0.39)	(0.02)	(0.02)
N.	3,882	469	919		

Source: IFLS2-3.

The difference in the selection patterns according to income and expenditure between return and current migrant households may be related to the fact that current migrants may have undertaken a more permanent migration, which is likely to imply higher costs than return migration. Hence, households which are positively selected in terms of wealth are more likely to afford to send permanent migrants, and also to provide support to the migrants in case of need. Moreover, more permanent forms of migration are also typical of migrants with higher human capital, who are more likely to be found in richer households. Indeed, those migrants who have a higher level of education are more likely to find a stable job at destination. These individuals have a lower probability of returning at origin, both because they are likely to having definitively quitted previous economic activities at origin, and because they have a lower risk of job loss.

We also consider the role played by assets in households' migration behavior. In particular, we consider the role of business and non-business land, and that of jewelries.[70] From table 5 below, we can see that non-migrant households had a significantly lower value of non-business land and jewelries at the baseline, compared to current migrant households. This, again, can be related to the fact that households having current migrants had to finance a more stable form of migration and households having a higher value of assets are more likely to sustain these costs.

Table 5. Households' assets in 1997, by migration status in 1997-2000

	Household's migration status				
	Non-migrant	Return	Current		
	(1)	(2)	(3)	Diff (1)-(2)	Diff (1)-(3)
P.c. Business land	5.25	4.97	5.53	0.27	-0.28
	(7.11)	(7.02)	(7.12)	(0.34)	(0.26)
P.c. Non-business land	4.36	3.97	5.00	0.38	-0.65***
	(6.69)	(6.51)	(6.91)	(0.32)	(0.25)
P.c. Jewelries	6.54	6.40	6.85	0.14	-0.31*
	(5.94)	(6.01)	(5.90)	(0.29)	(0.22)
N. observations	3,882	469	919		

Source: IFLS2-3.
Note: assets are measured in real (base Jakarta 2000) per capita terms, and in logs.

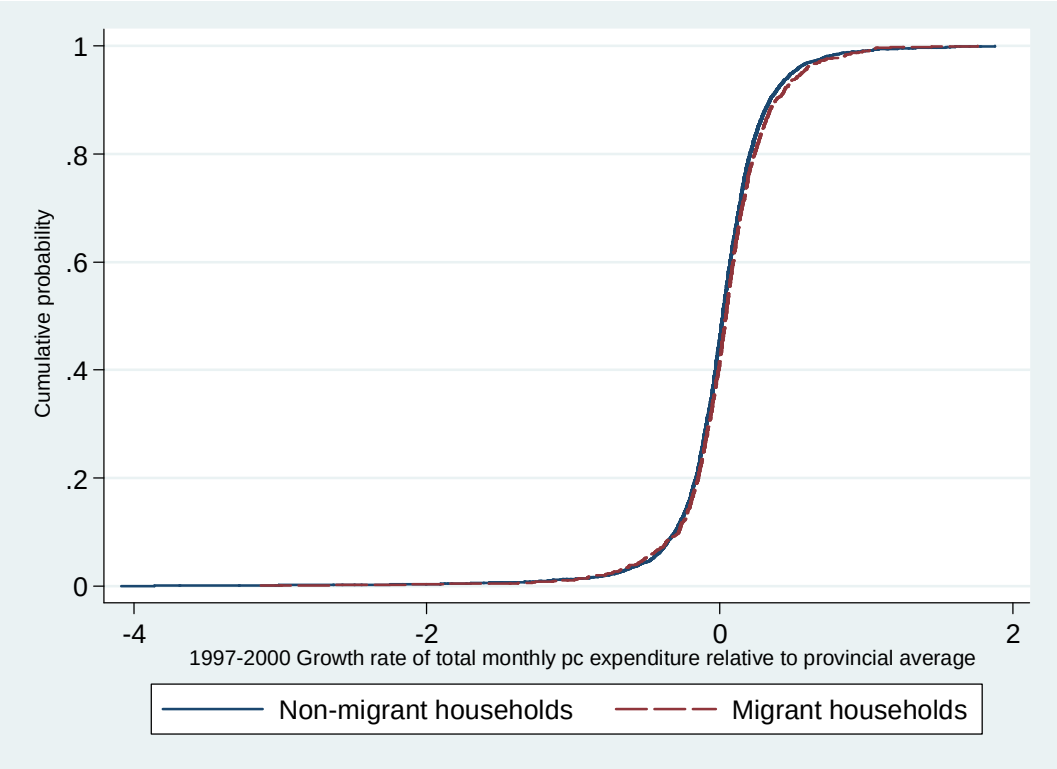
After having described the main differences between migrant and non-migrant households at the baseline, we focus on our outcome of interest, which is households per capita real expenditure growth in the period 1997-2000. In particular, we show the cumulative distribution function of the growth rate of per capita total monthly household expenditure in 1997-2000 relative to provincial average, by households' migration status.[71] We see that migrant households had a higher per capita expenditure growth than non-migrant

[70] The acquisition of gold as an investment and savings device is very important in Indonesia. Indeed, in our sample of analysis 56.2% of households possessed some wealth in form of jewelries at the baseline year 1997. Moreover, the sale of jewelries represented a relevant consumption smoothing mechanism during the economic crisis which hit the country in 1997-1998 (Frankenberg *et al.*, 2003).

[71] In the econometric analysis of the next section, the dependent variable (growth rate of expenditure) is not calculated using expenditure as a fraction of provincial average, but using simply expenditure. Province fixed effects are included

households. A Kolmogorov-Smirnov (K-S) test of equality of the expenditure distribution functions between migrant and non-migrant households rejects the null hypothesis of equality of distributions at the 1% significance level.

Figure 1. Cumulative distribution function of the 1997-2000 p.c. household expenditure growth rate (migrant and non migrant households)

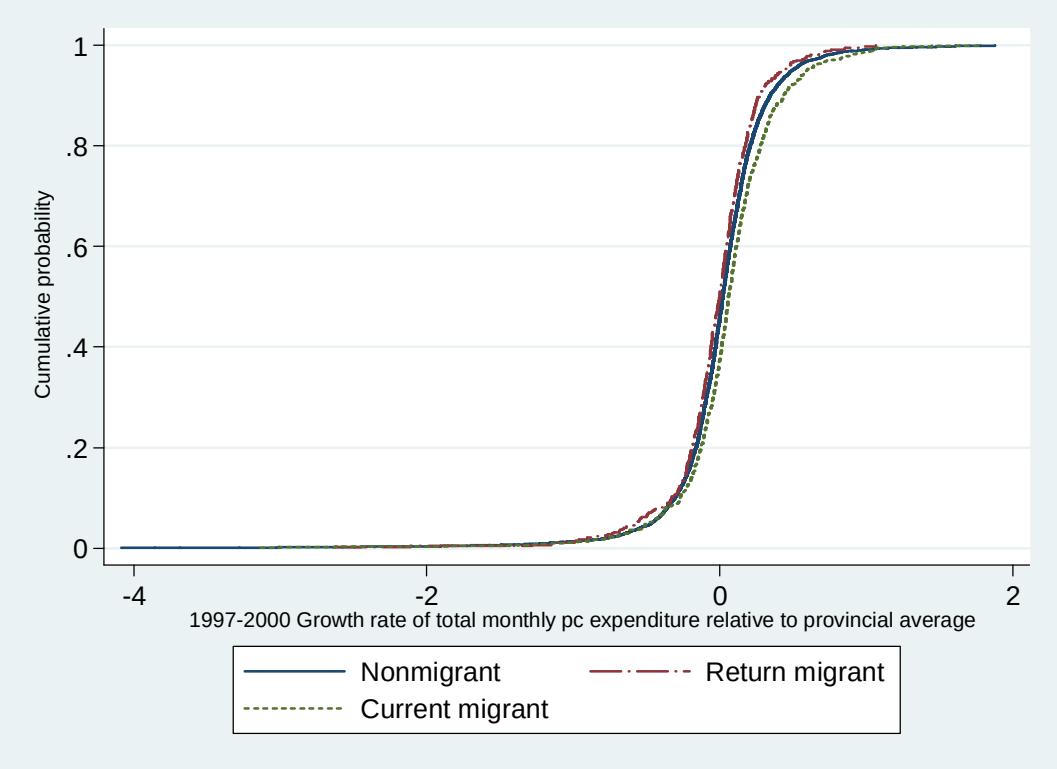


Source: IFLS2-3.

When we take into account the distinction between current and return migrants (Figure 2 below), we difference in expenditure growth between migrant and non-migrant households is due to the higher per capita expenditure growth for current migrant households with respect to non-migrant households, while return migrant households' expenditure growth was significantly below that of non-migrant households. A K-S test of equality of growth rate distributions is rejected at the 10% significance level when we compare non-migrant and return migrant households (with return migrant households having lower expenditure than non-migrant households) while it is rejected at the 0.1% when the comparison is between non-migrant and current migrant households (with current migrant households having higher expenditure growth than non-migrant households). We take into account the distinction between return and current migrant households also in the econometric analysis of the next section.

in the right hand side. In this descriptive analysis, however, we show graphs which consider the expenditure relative to provincial average, to control for province effects and thus make visible the relationships of interest.

Figure 2. Cumulative distribution function of the 1997-2000 p.c. household consumption growth rate (current/return migrant households)



Source: IFLS2-3.

5. Econometric analysis

In this section, we perform an econometric analysis of the impact of having migrants on origin households' total monthly per capita expenditure in the period 1997-2000. The previous statistics highlighted that current migrant households are positively selected in terms of various factors, such as income and head's education. Therefore, the higher consumption growth we observe in the considered period for current migrant households with respect to non-migrant and return migrant households may be due to households' characteristics other than migration status. To understand whether there is indeed a causal effect of migration on expenditure growth, we use a difference-in-differences estimator, coupled with an instrumental variable method. In particular, we estimate a regression having as dependent variable the difference in natural logarithm of households' total monthly per capita expenditure between 1997 and 2000, and as independent variable of interest the households' migration status in the same period. We include on the right hand side a set of household-level controls, measured at the baseline year 1997, before the considered migration episodes.

The use of controls measured at the baseline year is an element of similarity between our analysis and that of Beegle *et al.* (2011). However, that study answered to the question whether migration improved the living conditions of migrants themselves on their own households, not of their *origin* households. In the analysis of Beegle *et al.* (2011), the dependent variable is in fact the difference in expenditure between origin and destination households. Moreover, their main focus is on the individuals. In particular, individuals were

interviewed in two periods, reporting their migration behavior and the expenditure of the households they belonged to in these two periods. If individuals moved between survey rounds, they were tracked and interviewed at destination. This analysis estimated the consumption growth of the household to which individuals belonged to, as a function of individuals' migration between the two waves, and a series of individuals' characteristics at the baseline period. In this type of analysis, individuals' initial and final consumption may refer to the same household or to different households. They will refer to the same households, for individuals who did not move. Differently, for individuals who moved, and are tracked and interviewed at destination, initial and final consumption generally refer to different households: individuals' initial consumption refers to their origin households in the first period, while individuals' final consumption refers to the new (split-off) households in which they live and are found in the second period.[72] Beegle *et al.* (2011) also performed their analysis at the household-level, but their main focus does not change: they evaluated the consumption change before and after migration, with final consumption referring to the destination household. Differently, our analysis focuses on *origin households* to which migrants belong and answers to the question whether the fact of having had migrants improved their living conditions. Also the recent experimental estimates of Gibson *et al.* (2011) analyze the impact of migration on origin households. However, as mentioned in the literature review, this study refers to international rather than internal migration. Moreover, it does not explicitly consider the distinction between return and current migration, which is crucial in our analysis. The study of de Brauw and Harigaya (2007) focused on the impact of migration households' consumption levels, considering internal migration in Vietnam. However, also this study does not consider the distinction between current and return migration, which is instead considered in our paper.

More precisely, we estimate the following regressions:

$$\Delta \ln C_{j97\text{-}00} = \beta_0 + \beta_1 M_j + \beta_3 H_{j97} + \varepsilon_{j97} \quad (1)$$

where *j* indicates migrants' origin households (*j*=1, …, 5,270). $\Delta \ln C_{j97\text{-}00}$ measures the difference in total log monthly per capita origin household expenditure between 1997 and 2000, M_j is a dummy variable which equals one if the household had at least one migrant (either current or return) in the considered period, and H_{j97} is a set of household-level controls referred to household *j* and measured in 1997.[73] As above said, we distinguish between current and return migration, by estimating the following model:

[72] Initial and final consumption may refer to the same households also in cases in which individuals moved, if they moved with all other household members (i.e. if the household moved as a unit). However, this case is of relatively lower importance with respect to the one considered above. Indeed, Beegle *et al.* (2011) underline that the second period survey round included more than 2,700 households, starting from 912 households at the baseline period. This means that 1,788 new (split-off) households formed as a result of individuals' moves. There could be cases in which, even if individuals do not move, their initial and final consumption refer to different households: this could happen for individuals who married and formed a new household in the same village. Also this case is of relatively lower importance with respect to the one of split-off households who also moved (at least) in another village.

[73] Differently, Beegle *et al.* (2011) estimate the following regression: $\Delta \ln C_{it+1,t} = \beta_0 + \beta_1 M_i + \beta_3 X_{it} + \delta_{ih} + \varepsilon_{it}$, in which *i* indicates the individual and the dependent variable measures the growth rate of per capita consumption in the *household*

$$\Delta \ln C_{j97\text{-}00} = \beta_0 + \beta_{1a} \text{return} M_j + \beta_{1b} \text{current} M_j + \beta_2 H_{j97} + \varepsilon_{j97} \quad (2)$$

where $\Delta \ln C_{j97\text{-}00}$ and H_{j97} are defined as in equation (1), while $\text{return}M_j$ and $\text{current}M_j$ are two dummy variables which are equal to one if household *j* has at least one return or one current migrant, respectively.

We first obtain OLS estimates, and then perform IV estimates, to address the endogeneity of the migration variables. We estimate the above models first considering all migrants together, and then including in the sample only households having at least one non-tied migrants, i.e. one migrant who moved for own purposes and not to follow others.

5.1 OLS estimates of the impact of migration on origin households

In this section, we obtain OLS estimates of equations (1) and (2) above, including the following 1997 household-level controls: a set of controls for household structure (the average age of household members, household head's age and marital status, household size, the number of children aged 0-5, 6-10, 11-15 and 16-20, and the number of males and females over 20 in the household). We also include a series of controls for household wealth at the baseline. The importance of including these controls derives from the fact that the period 1997-2000 was a period of serious economic crisis in Indonesia, and the crisis mainly hit richer households (Strauss *et al.*, 2004). We include the following controls for wealth:

- Three dummy variables indicating the household head's education level, corresponding to primary or junior secondary, senior secondary, and university (the excluded category is no education).
- Four dummy variables indicating the household's quintile of per capita expenditure in 1997 (the excluded category refers to households in the lowest quintile).
- Four dummy variables indicating the household's quintile of per capita income in 1997 (the excluded category includes households in the lowest quintile).[74]
- Four variables measuring households' landholdings in 1997 (the log of the per capita values of business and non-business land and their squares).
- A variable measuring the value of households' assets held in form of jewelries. We include this variable because available evidence of the impact of the Indonesian economic crisis showed that the sale of gold represented an important means to smooth consumption (Frankenberg et. al, 2003).

Our estimates also include the following controls for 1997 labour supply in the household:

- Four dummy variables indicating whether the household head was working in manufacturing, construction, services or was not working in 1997 (the excluded category is agricultural or mining work).

---

in which *i* is residing in the two periods. X_{it} and δ_{ih} are individual-level controls and initial household fixed effects, respectively (Beegle *et al.* 2011, p. 1017). As a robustness check, the authors also estimated a household-level variant of this equation, in which all individual-level controls in X_{it} are replaced by household-level averages. Also in this model, the focus of their analysis is consumption growth between origin and destination households.

[74] The inclusion of dummies for expenditure controls for *permanent income*, while the income dummies controls for *current income*.

- Two variables indicating the number of market workers and government workers in the household in 1997. These variables were included because it is likely that the crisis hit more seriously market workers than government workers.

All estimates also include a dummy controlling for rural or urban location, province dummies, and the interaction terms between rural location and province dummies. Standard errors are corrected for heteroskedasticity and clustered at the district-level.

Table 5. OLS estimates of the impact of migration on origin household consumption growth 1997-2000

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
HH has a migrant	0.051**		0.055**	
	(0.02)		(0.02)	
HH has a current migrant		0.131***		0.143***
		(0.03)		(0.03)
HH has a return migrant		-0.096***		-0.091***
		(0.03)		(0.03)
HH size	0.034	0.035	0.033	0.034
	(0.06)	(0.06)	(0.06)	(0.06)
N. children 0-5	-0.078	-0.078	-0.074	-0.074
	(0.07)	(0.07)	(0.07)	(0.07)
N. children 6-10	-0.093	-0.094	-0.092	-0.093
	(0.07)	(0.07)	(0.07)	(0.07)
N. chidlren 11-15	-0.016	-0.019	-0.014	-0.018
	(0.06)	(0.06)	(0.07)	(0.07)
N. children 16-20	-0.056	-0.060	-0.058	-0.061
	(0.06)	(0.06)	(0.07)	(0.07)
N. males 21+	-0.032	-0.035	-0.033	-0.033
	(0.05)	(0.05)	(0.06)	(0.06)
N. females 21+	-0.023	-0.026	-0.022	-0.025
	(0.06)	(0.06)	(0.06)	(0.06)
Average age in HH	0.001	0.000	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Head's age	0.001	0.001	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Head unmarried	0.034	0.037	0.034	0.038
	(0.11)	(0.11)	(0.12)	(0.12)
Head's education				
Elementary/Jr. secondary	0.024	0.026	0.029	0.030
	(0.03)	(0.03)	(0.03)	(0.03)
Senior secondary education	0.207***	0.209***	0.211***	0.211***
	(0.04)	(0.04)	(0.04)	(0.04)
University	0.366***	0.364***	0.385***	0.381***
	(0.13)	(0.13)	(0.13)	(0.13)
Head's sector				
Manufacturing	-0.067*	-0.070*	-0.077*	-0.078*
	(0.04)	(0.04)	(0.04)	(0.04)
Construction	-0.076**	-0.073**	-0.075**	-0.071**
	(0.03)	(0.03)	(0.04)	(0.04)
Services	0.015	0.014	0.018	0.017
	(0.03)	(0.03)	(0.03)	(0.03)
Not working	0.002	0.002	0.005	0.007
	(0.03)	(0.04)	(0.04)	(0.04)

Table 5. (Continued)

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
N. workers in HH				
Martket	-0.041***	-0.040***	-0.041***	-0.041***
	(0.01)	(0.01)	(0.02)	(0.01)
Government	0.024	0.022	0.025	0.021
	(0.03)	(0.03)	(0.03)	(0.03)
Quintile of HH income				
2nd	0.065*	0.068*	0.062*	0.064*
	(0.04)	(0.04)	(0.04)	(0.04)
3rd	0.050	0.051	0.060	0.062
	(0.04)	(0.04)	(0.04)	(0.04)
4th	0.159***	0.159***	0.162***	0.162***
	(0.04)	(0.04)	(0.04)	(0.04)
5th	0.316***	0.315***	0.311***	0.312***
	(0.05)	(0.05)	(0.05)	(0.05)
Quintile of HH expenditure				
2nd	-0.414***	-0.411***	-0.414***	-0.412***
	(0.03)	(0.03)	(0.04)	(0.04)
3rd	-0.600***	-0.599***	-0.600***	-0.600***
	(0.04)	(0.04)	(0.04)	(0.04)
4th	-0.865***	-0.865***	-0.864***	-0.865***
	(0.04)	(0.04)	(0.04)	(0.04)
5th	-1.396***	-1.396***	-1.397***	-1.398***
	(0.06)	(0.06)	(0.06)	(0.07)
P.c. Business land	-0.017	-0.018	-0.021	-0.022*
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Non-business land	0.008	0.007	0.012	0.011
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Business land squared	0.001	0.001*	0.002*	0.002*
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Non-business land squared	-0.000	-0.000	-0.001	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Jewelries	0.011***	0.010***	0.011***	0.010***
	(0.00)	(0.00)	(0.00)	(0.00)
N	5,270	5,270	5,027	5,027
R-squared	0.222	0.225	0.219	0.222

Notes:
Robust standard errors in parentheses, clustered at the district-level.
*** 1%, ** 5%, * 10% significance level.
The dependent variable is the difference in log per capita household total monthly consumption between 1997 and 2000, excluding transfers to non-household members and to members who are not in the household. Consumption is measured in real terms, with base Jakarta 2000. Estimates include a constant term. Province and rural dummies and their interactions are also included.

From the model in column 1, we see that the dummy indicating that the household had a migrant in 1997-2000 is positive and statistically significant. In column 2, we distinguish between return and current migrants: the current migration dummy is statistically significant and positive, while the return migration dummy is statistically significant and negative. This result may be related to the fact that return migrants might go back at origin because they were not successful in finding a job, or because they experienced a negative economic shock at destination, such as illness. However, our results may also be related to the fact that return migration might be very different from current migration, in terms, for instance, of migrants' employment. Current migrants, as previously mentioned, might be engaged in a more permanent migration, and are thus likely to having more stable and better paid jobs than return migrants. Therefore, they can transfer a higher amount of resources at origin. Return migrants, differently, might be *seasonal* migrants who move to other settings on a more temporary basis, and are employed in low-skill jobs at destination. This implies that they are less likely to provide a significant positive contribution to their origin households' consumption, while they are away or when they are back. Hence, our results suggest that the type of migration in terms of duration and skill-intensity of employment at destination are crucial in determining the impact of migration on origin households' consumption growth.

As for the other factors considered in the estimation, we can see that the household structure has not a significant impact on households' per capita consumption growth, once other factors are considered, such as household's head's education. Households' whose head had senior secondary (university) education in 1997 had 20 percent (37 percent) higher per capita consumption growth compared to households where the head had no education. Also the labour market variables included in the model display the expected impacts: households whose head was employed in the construction or manufacturing sector had significantly lower per capita consumption growth than households whose head was employed in agriculture in 1997. Moreover, a higher number of private sector workers also decreased households' consumption growth. These results are consistent with the fact that the economic crisis hit more seriously those workers employed in construction or other private sectors, compared to government officials.

The estimated coefficients on the 1997 expenditure quintiles dummies are all negative and statistically significant, indicating that households having the highest expenditure in 1997 had the sharpest decline in expenditure in the considered time period.[75] The coefficients on the income quintiles dummies, however, are positive and statistically significant. These results suggest that, while richest households experienced the highest cut in per capita expenditure during the crisis, those households having higher incomes within a certain expenditure category were more able to increase their consumption.

As for the role of assets in households' consumption growth, our estimates show a significant and positive impact of business land holdings on household per capita consumption growth (even though only the squared term in column 2 is statistically significant). As expected on the basis of Frankenberg *et al.* (2003), the estimated coefficient on the variable measuring the value of households' jewelries is positive and

[75] This result is consistent with existing studies on the impact of the 1997-1998 economic crisis in Indonesia. See, for instance, Frankenberg *et al.* (2003) and Strauss *et al.* (2004).

statistically significant, confirming the role of jewelries' sales as a way for households to smooth their consumption.

Regarding location variables, among the province dummies (not shown), some have a significant impact on households consumption growth, while the rural dummy is not statistically significant, indicating that, once household-level and province-level characteristics are considered, the fact of living in a rural or urban area was not significantly related to expenditure growth.

Considering the differences between 'tied' and 'non-tied' migrants, we can see the coefficient on the dummies indicating the presence of a migrant and of a current migrant are higher in the case of non-tied migration (columns 3 and 4), than in the case of all migrants. In other words, the positive association between migration (or current migration) and household consumption growth is stronger for households having at least one non-tied migrants, than for households having tied migrants only. This could be related to the fact that tied-migrants are less likely to send transfers back to their origin households, compared to non-tied migrants, who may have moved also to support their origin household through their own work. The magnitude and significance level of the other estimated coefficients do not vary in a relevant way when the migrant sample is restricted to households only having at least one non-tied migrant.

5.2 2SLS estimates of the impact of migration on origin households

In this section, we perform instrumental variables estimation of the impact of having a migrant for origin households. We first consider the impact of having a migrant, and then separate the impact of having a current or return migrant. Given that the majority of migrants in our dataset are individual migrants, we use as instruments variables that, in the previous chapter, have been shown to predict individual migration. Our instruments for the presence of *migrants* or *current migrants* in the period 1997-2000 in the household are the following:

1. A dummy variable which equals one if in 1997 the household had at least one unmarried member aged 15-20 years.
2. A dummy variable which equals one if in 1997 the head has at least one child and in the household there was at least one unmarried individual aged 12-22 years.
3. A dummy variable which is equal to one if in 1997 the head had at least one child and in the household there was at least one individual aged 12-22, not working and having at least 7 years of education.
4. The predicted values of a probit model of migration including the above instruments: in particular, for both current migrant and migrant three probit models where estimated, including either (i) instrument 1. above, or (ii) instrument 2. above, or (iii) instruments 2. and 3. above. By obtaining predicted probabilities of households' migration status, hence, we generated other three instruments, besides instruments 1., 2., and 3. above. The methodology of using probit predicted values as instruments is suggested by Wooldridge (2002), who underlines that this procedure generally leads to more precise

estimates of the endogenous coefficient of interest, even if at the cost of stronger identifying assumptions.[76]

The use of the above instruments is justified by the fact that the majority of migrants in our sample are individual migrants, and we saw in the previous chapter that individual migrants are on average of age 20, and are more likely to be unmarried. Using this instruments, the identifying assumption is that the presence of at least one member with the characteristics above described in a 1997 household influences the probability that a household member will move between 1997 and 2000, without having a direct impact on household per capita expenditure growth in the same time span. This assumption is likely to be verified, given that our second stage estimates control for a set of household composition variables, such as the number of children and adults in the household. Therefore, the direct impact of household structure on consumption should be captured by these controls.

As instruments for *return* migration, we considered the following variables:

5. A dummy variable which equals one if in 1993 the household had at least one unmarried individual aged 12-21;
6. A dummy variable which equals one if in 1993 the head had at least one child and in the household there was at least one unmarried individual aged 12-22 years.

Hence, our instruments for households having a return migrant in the period 1997-2000 refer to households' composition in 1993. Therefore, they capture the fact that households may have had a migrant who left between 1993 and 1997. We calculated that, for individuals in our sample who did more than one move in the course of their lives, the average duration of migration was 4.7 years. Therefore, an individual who left in 1993 and will return, is likely to return between 1997 and 2000.

We considered different sets of instruments, including or excluding the predicted values from probit. We only report results which include the predicted values from probit as instruments, since they allowed a better identification. As for the OLS case, we estimate the models for two separate samples, the full sample and the one excluding households which only had tied migrants.

[76] Wooldridge (2002), pp. 621-625.

Table 6. 2SLS estimates of the impact of migration on origin household consumption growth 1997-2000

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
HH has a migrant	-0.103		-0.076	
	(0.12)		(0.12)	
HH has a current migrant		0.267		0.323
		(0.26)		(0.26)
HH has a return migrant		-0.619**		-0.616**
		(0.31)		(0.31)
HH size	0.040	0.044	0.037	0.041
	(0.06)	(0.06)	(0.06)	(0.06)
N. children 0-5	-0.091	-0.090	-0.082	-0.081
	(0.07)	(0.07)	(0.07)	(0.07)
N. children 6-10	-0.103	-0.108	-0.098	-0.104
	(0.07)	(0.07)	(0.07)	(0.07)
N. chidlren 11-15	-0.014	-0.029	-0.011	-0.028
	(0.06)	(0.06)	(0.07)	(0.07)
N. children 16-20	-0.047	-0.062	-0.050	-0.060
	(0.06)	(0.06)	(0.06)	(0.07)
N. males 21+	-0.031	-0.040	-0.031	-0.032
	(0.05)	(0.06)	(0.06)	(0.06)
N. females 21+	-0.020	-0.033	-0.019	-0.033
	(0.05)	(0.05)	(0.06)	(0.06)
Average age in HH	-0.000	-0.001	0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Head's age	0.002	0.002	0.001	0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Head unmarried	0.033	0.061	0.035	0.064
	(0.11)	(0.12)	(0.12)	(0.12)
Head's education				
Elementary/Jr. secondary	0.031	0.036	0.034	0.037
	(0.03)	(0.03)	(0.03)	(0.03)
Senior secondary education	0.219***	0.224***	0.220***	0.222***
	(0.04)	(0.04)	(0.04)	(0.04)
University	0.382***	0.374***	0.400***	0.387***
	(0.13)	(0.13)	(0.13)	(0.14)
Head's sector				
Manufacturing	-0.069*	-0.080*	-0.079**	-0.084**
	(0.04)	(0.04)	(0.04)	(0.04)
Construction	-0.070**	-0.058	-0.068*	-0.055
	(0.03)	(0.04)	(0.04)	(0.04)
Services	0.019	0.016	0.021	0.019
	(0.03)	(0.03)	(0.03)	(0.03)
Not working	0.011	0.010	0.011	0.019
	(0.03)	(0.04)	(0.04)	(0.04)

Table 6. (Continued)

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
N. workers in HH				
Market	-0.041***	-0.038***	-0.041***	-0.039***
	(0.01)	(0.01)	(0.01)	(0.01)
Government	0.026	0.017	0.026	0.012
	(0.03)	(0.03)	(0.03)	(0.03)
Quintile of HH income				
2nd	0.063*	0.073**	0.061*	0.070*
	(0.04)	(0.04)	(0.04)	(0.04)
3rd	0.042	0.049	0.054	0.063
	(0.04)	(0.04)	(0.04)	(0.04)
4th	0.151***	0.150***	0.156***	0.155***
	(0.04)	(0.04)	(0.04)	(0.05)
5th	0.317***	0.308***	0.313***	0.311***
	(0.05)	(0.05)	(0.05)	(0.05)
Quintile of HH expenditure				
2nd	-0.407***	-0.395***	-0.408***	-0.400***
	(0.03)	(0.03)	(0.03)	(0.04)
3rd	-0.593***	-0.586***	-0.594***	-0.590***
	(0.04)	(0.04)	(0.04)	(0.04)
4th	-0.858***	-0.854***	-0.858***	-0.860***
	(0.04)	(0.04)	(0.04)	(0.04)
5th	-1.386***	-1.385***	-1.388***	-1.390***
	(0.06)	(0.06)	(0.06)	(0.07)
P.c. Business land	-0.016	-0.022*	-0.020	-0.024*
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Non-business land	0.009	0.006	0.012	0.010
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Business land squared	0.001	0.002*	0.002*	0.002**
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Non-business land squared	-0.000	-0.000	-0.001	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Jewelries	0.010***	0.010***	0.010***	0.010***
	(0.00)	(0.00)	(0.00)	(0.00)
N	5,270	5,270	5,027	5,027

Table 6. (Continued)

	All migrants		At least one 'non-tied' migrant	
F-excluded instr. (migrant)[a]	41.45		42.87	
F-excluded instr. (current)[a]		25.09		29.02
F-excluded instr. (return)[b]		17.07		16.12
Angrist-Pischke F-excluded instr. (current)		13.00		16.76
Angrist-Pischke F-excluded instr. (return)		12.41		10.59
Hansen J P-value	0.42	0.77	0.48	0.88

Notes:
Robust standard errors in parentheses, clustered at the district-level.
*** 1%, ** 5%, * 10% significance level.
The dependent variable is the difference in log per capita household total monthly consumption between 1997 and 2000, excluding transfers to non-household members and to members who are not in the household. Consumption is measured in real terms, with base Jakarta 2000. Estimates include a constant term. Province and rural dummies and their interactions are also included.
[a.] To instrument the dummies indicating whether the household had a migrant in 1997-2000 and a current migrant in the same period, the predicted values from probit estimations described at point 4. in Section 5.2 above were used as excluded instruments.
[b.] To instrument the dummy indicating whether the household had a return migrant in 1997-2000, the variables indicated at points 5. and 6. in Section 5.2 above were used as excluded instruments.

The results show that the instruments are valid: in all cases, the F-test of excluded instruments is above the threshold level of 10, below which the instruments are considered weak (Stock *et al*., 2002). The Hansen test of overidentifying restrictions in all cases fails to reject the null hypothesis of validity of the instruments. The variable indicating the presence of a migrant in the household is negative and statistically insignificant, both in the full sample and in the sample which excludes households only having tied migrants. When we distinguish between the presence of return and current migrants in the households, we see that the presence of current migrants in the households does not have a statistically significant impact on origin households' consumption growth

Regarding current migration, we see that both in the full sample and in the sample excluding households which only had tied migrants the estimated coefficient of the variable indicating the presence of at least one return migrant in the household is statistically significant and negative. In particular, households having a return migrant in the period 1997-2000 experienced 60% lower growth of per capita expenditure, compared to households not having return migrants. The estimated coefficients of the other variables are analogous for magnitude and significance levels to the OLS estimates of the previous table.

The negative effect of return migration may be related to the fact that return migrants are likely to be individuals with limited ability to contribute to the process of household income creation. Indeed, they may be individuals whose migration investment failed and who had to return to their origin households (for instance, because of a negative economic shock, such as job loss or illness). We will explore the characteristics of return migrants in section 6, in order to understand the channels through which return migration may impact origin households' consumption growth.

The above results are obtained considering the expenditure variables excluding transfers to individuals outside of the household (either household members not at home, or non-household members). We also

estimated a model considering the expenditure including transfers out. The results, reported in Appendix tables A.1 and A.2, show that, when considering expenditure including transfers out, having a current migrant significantly increases per capita expenditure growth in 1997-2000, in the case of households which have at least one non-tied migrant. These results suggest that origin households might being 'financing' the migration investment of *current* migrants, by spending resources for them, either to pay their education expenses, if they moved for education purposes, or to support them during the job search period, in case migrants moved to look for a job. We also include in Appendix table A.3 OLS estimates of the impact of having current and return migrants on households' transfers' consumption growth in 1997-2000. As expected, results show that current migration is statistically significantly positively associated with households' transfers' growth, while return migration is negatively significantly related to households' transfers' growth.

6. Migrants' and non-migrants' characteristics in year 2000

This section provides descriptive evidence of the characteristics of migrants and non-migrants in year 2000, that is when return migrants were back at origin. We distinguish between current migrants, who are tracked and interviewed at destination, and return migrants, who are interviewed at origin after their return. The information on migrants and non-migrants in 2000 are only available for a subset of the individuals who refer to the origin households we considered in the above analysis, i.e., only for a subset of the individuals considered in table 1 above. However, we can provide information for 13,934 non-migrants (92% of the total number of non-migrants in the sample), 646 return migrants (89% of the total number of return migrants in the sample), and 1071 current migrants (94% of the total number of current migrants in the sample).

The next table shows individuals' age, marital status and rural residency. As expected, current migrants are, on average, younger, less likely to be married and more likely to reside in urban areas with respect to return migrants.

Table 7. Migrants' and non-migrants' age, marital status, and residency in 2000

	Individuals' migration status 1997-2000					
	Non-migrant	Return	Current			
Individuals' characteristics in 2000	(1)	(2)	(3)	Diff (1)-(2)	Diff (1)-(3)	Diff (2)-(3)
Age	36.32	24.66	21.87	11.65***	14.43***	2.79***
	(16.96)	(10.12)	(9.24)	(0.44)	(0.52)	(0.48)
Unmarried	0.29	0.63	0.72	-0.31***	-0.43***	-0.12***
	(0.45)	(0.49)	(0.45)	(0.01)	(0.01)	(0.02)
Rural	0.52	0.54	0.38	-0.22	0.13***	0.15***
	(0.49)	(0.49)	(0.48)	(0.02)	(0.02)	(0.02)
N.	13,934	646	1,071			

Source: IFLS2-3.

Table 8 shows migrants' education and employment characteristics. We can see that return migrants are, on average, significantly more educated than non-migrants. However, return migrants are significantly less educated than current migrants. Moreover, return migrants are significantly more likely to be unemployed in 2000, compared to non-migrants. Considering income, table 8 shows that return migrants are significantly more likely than non-migrants or current migrants to be in the lowest income quintile in 2000. These figures, thus, suggest that return migrants may have a limited ability to contribute to total household income. This explains our results on the negative impact of return migration on household expenditure growth.

Table 8. Migrants' and non-migrants' education, employment and income in 2000

	Individuals' migration status 1997-2000					
Individuals' characteristics in 2000	Non-migrant (1)	Return (2)	Current (3)	Diff (1)-(2)	Diff (1)-(3)	Diff (2)-(3)
Education						
None	0.15	0.04	0.02	0.11***	0.13***	0.01*
	(0.36)	(0.19)	(0.15)	(0.01)	(0.01)	(0.00)
Elementary/Jr. Secondary	0.58	0.52	0.42	0.06***	0.16***	0.10***
	(0.49)	(0.50)	(0.49)	(0.02)	(0.02)	(0.02)
Senior secondary	0.21	0.35	0.38	-0.13***	-0.17***	-0.03
	(0.41)	(0.48)	(0.48)	(0.02)	(0.01)	(0.02)
University	0.06	0.10	0.19	-0.04***	-0.13***	-0.09***
	(0.49)	(0.30)	(0.39)	(0.00)	(0.01)	(0.02)
Employment						
Primary	0.22	0.15	0.08	0.07***	0.14***	0.07***
	(0.41)	(0.35)	(0.27)	(0.02)	(0.01)	(0.02)
Manufacturing	0.08	0.11	0.12	-0.02**	-0.04***	-0.01
	(0.28)	(0.31)	(0.37)	(0.01)	(0.01)	(0.02)
Construction	0.03	0.05	0.03	-0.02**	0.00	0.02**
	(0.17)	(0.21)	(0.16)	(0.01)	(0.01)	(0.01)
Services	0.28	0.25	0.34	0.04**	-0.06***	-0.09***
	(0.45)	(0.43)	(0.47)	(0.02)	(0.02)	(0.02)
Not working	0.38	0.45	0.43	-0.07***	-0.05***	0.02
	(0.48)	(0.49)	(0.49)	(0.02)	(0.01)	(0.02)
Quintile of monthly income						
1st	0.34	0.41	0.38	-0.07***	-0.03**	0.03*
	(0.48)	(0.49)	(0.49)	(0.02)	(0.02)	(0.02)
2nd	0.05	0.05	0.04	0.01	0.02**	0.01
	(0.22)	(0.20)	(0.18)	(0.01)	(0.01)	(0.00)
3rd	0.22	0.17	0.16	0.05***	0.06***	0.01
	(0.41)	(0.37)	(0.32)	(0.01)	(0.01)	(0.02)
4th	0.20	0.24	0.22	-0.03**	-0.02**	0.01
	(0.40)	(0.42)	(0.42)	(0.02)	(0.01)	(0.02)
5th	0.18	0.14	0.21	0.05***	-0.02**	-0.07***
	(0.38)	(0.34)	(0.40)	(0.01)	(0.01)	(0.02)
N.	13,934	646	1,071			

Source: IFLS2-3.

7. Conclusions

Using a detailed panel data survey from Indonesia, this paper explores the short-run impact of internal migration on per capita consumption of household members left behind. Within-country labor mobility, on a temporary or permanent basis, is a common and widespread family strategy with implications for people left behind in terms of forgone (monetary and non-monetary) inputs of migrants, diversification of income sources and potential cash (remittances) inflows.

While assessing the impact of migration, we tackle both endogeneity and migrants self-selection concerns by using a DID approach coupled with the IV method. While doing so, we are able to distinguish between current vs return migration, as to disentangle the heterogeneity of the migration process and the different channels through which it affects the remaining household members.

We find that origin households having internal return migrants in the period 1997-2000 have significantly lower per capita real consumption growth, compared to households not having migrants. In particular, our estimates indicate that return migrant households have 60% per capita consumption growth lower than households not having return migrants, on average and *ceteris paribus*. Differently, the presence of current migrants has a non-significant impact on origin households' members consumption growth.

These results are likely to be related to the fact that returning migrants have a lower income potential in 2000, when compared to non-migrants and current migrants. Data for the year 2000, indeed, show that returning migrants have significantly lower employment probabilities than non-migrants or current migrants. This may be related to the fact that return migrants may be those individuals whose migration investment failed, i.e. those who, for instance, lost their jobs. Alternatively, returning migrants may also be seasonal migrants, whose return, have a negative impact on origin households' consumption, due to the fact that the size of the household increase when they are back, without an increase in household income sufficient to compensate for the reduction in consumption due to the household size increase.

These results have been obtained by considering origin households' consumption which excludes transfers given to individuals outside the households. When these transfers are instead included in our dependent variable, we find that having at least a current non-tied migrant significantly increases origin households' consumption growth in the considered period. Hence, our results suggest that origin households increase their transfers when having current migrants, i.e. there is evidence of origin households 'financing', at least, in part the migration investments of members outside the household. This is also confirmed by the fact that, when estimating a model of transfers out as a function of households' migration status, we find a statistically significant and positive association between transfers out and presence of current migrants.

Bibliography

Acosta, P. (2011). School Attendance, Child Labour, and Remittances from International Migration in El Salvador. *Journal of Development Studies*, 47(6), 913-936.

Adams, R. H. (2005). Remittances, Household Expenditure and Investment in Guatemala, World Bank Policy Research Working Paper No. 3532.

Adams, R. H. (1998). Remittances, Investment and Rural Asset Accumulation in Pakistan, *Economic Development and Cultural Change* 47, 155-73.

Adams, R. H. (1991). The Economic uses and Impact of International Remittances in Rural Egypt, *Economic Development and Cultural Change* 39, 695-722.

Beaudouin, P. (2006). Economic Impact of Migration on a Rural Area in Bangladesh, *mimeo*: Centre d' Économie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne.

Beegle, K., Weerdt, J. D., and Dercon, S. (2011). Migration and Economic Mobility in Tanzania: Evidence From a Tracking Survey. *Review of Economics and Statistics*, 93(3), 1010-1033.

Berry, R. A., and Soligo, R. (1969). Some Welfare Aspects of International Migration. *The Journal of Political Economy*, 77(5), 778-794.

Deb, P. and Seck, P. (2009). Internal Migration, Selection Bias and Human Development: Evidence from Indonesia and Mexico. Human Development Research Paper 2009/31.

de Brauw, A. and Mueller, V. and Woldehanna, T. (2011). Insurance motives to remit: Evidence from a matched sample of Ethiopian internal migrants, IFPRI discussion papers 1090, International Food Policy Research Institute (IFPRI).

de Brauw, A. and Harigaya, T. (2007). Seasonal migration and improving living standards in Vietnam. *American Journal of Agricultural Economics*, 89(2), 430–447.

Dercon, S. (2002). Income Risk, Coping Strategies and Safety Nets', *World Bank Research Observer*, vol. 17, 141-166.

Du, Yang; Park, Albert and Wang, Sangui (2005). Migration and Rural Poverty in China " *Journal of Comparative Economics*, 33(4), 688-709.

Frankenberg, E., Smith, J. P., and Thomas, D. (2003). Economic Shocks, Wealth and Welfare. *The Journal of Human Resources*, 38(2), 280-321.

Gibson, J., McKenzie, D., and Stillman, S. (2011). The Impacts of International Migration on Remaining Household Members: Omnibus Results from a Migration Lottery Program. *The Review of Economics and Statistics*, 93(4), 1297-1318.

Gregory, P. (1986). *The Myth of Market Failure: Employment and the Labour Market in Mexico*. Baltimore: The John Hopkins University Press.

Grubel, H. B., and Scott, A. D. (1966). The International Flow of Human Capital. *The American Economic Review*, 56(1/2), 268-274.

Harris, J. R., and Todaro, M. P. (1970). Migration, Unemployment and Development: A Two-Sector Analysis. *The American Economic Review*, 60(1), 126-142.

Hildebrandt, N., and McKenzie, D. (2005). The Effects of Migration on Child Health in Mexico. *Economia*, 6(1), 257-289.

Lewis, A. W. (1954). Economic Development with Unlimited Supply of Labour. *Manchester School of Economic and Social Studies*, 22.

Lucas, R. E. B. (1987). Emigration to South Africa's Mines. *The American Economic Review*, 77(3), 313-330.

Mansuri, G. (2006). Migration, Sex Bias, and Child Growth in Rural Pakistan, *Policy Research Working Paper Series 3946*: The World Bank.

McKenzie, D., Gibson, J., and Stillman, S. (2010). How Important is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration. *Journal of the European Economic Association*, 8(4), 913-945.

McKenzie, D., and Sasin, M. J. (2007). Migration, Remittances, and Human Capital: Conceptual and Empirical Challenges, *World Bank Policy Research Working Paper*: The World Bank.

Mendola, M. (2008). Migration and Technological Change in Rural Households: Complements or Substitutes? *Journal of Development Economics*, 85(1-2), 150-175.

Morduch, J. (1995). Income Smoothing and Consumption Smoothing, *Journal of Economic Perspectives*, Vol.9 (Summer), 103–14.

Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *The Quarterly Journal of Economics*, 118(2), 549-599.

Ranis, G., and Fei, J. C. H. (1961). A Theory of Economic Development. *The American Economic Review*, 51(4), 533-565.

Rosenzweig, M. R., and Stark, O. (1989). Consumption Smoothing, Migration and Marriage: Evidence from Rural India. *The Journal of Political Economy*, 97(4), 905-926.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics*, 20(4), 518-529.

Strauss, J., Beegle, K., Dwiyanto, A., Herawati, Y., Pattinasarany, D., Satriawan, E., et al. (2004). *Indonesian Living Standards Before and After the Crisis: Evidence from the Indonesia Family Life Survey*. RAND Corporation, Santa Monica, CA.

Taylor, J. E., and Martin, P. (2001). Human Capital: Migration and Rural Population Change. In B. L. Garner and G. C. Rausser (Eds.), *Handbook for Agricultural Economics*. New York: Elsevier Science.

Taylor, J. E., and Mora, J. (2006). Does Migration Reshape Expenditures in Rural Households? Evidence from Mexico, *World Bank Policy Research Working Paper 3842*. Washington, DC: The World Bank.

Todaro, M. P. (1969). A Model of Labor Migration and Urban Unemployment in Less Developed Countries. *The American Economic Review*, 59(1), 138-148.

Witoelar, F. (2005). Inter-household Allocations within Extended Family: Evidence from the Indonesia Family Life Survey, *Economic Growth Center Discussion Paper No. 912*: Yale University.

Witoelar, F. (2009). Note on the Construction of the IFLS Consumption Expenditure Aggregates, *mimeo*. Washington, DC: The World Bank.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, Massachusetts: MIT Press.

Yang, D. (2008). International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks. *The Economic Journal*, 118(528), 591-630.

Yang, D., and Martinez, C. (2005). Remittances and Poverty in Migrants' Home Areas: Evidence from the Philippines, *Background paper for World Bank's International Migration and Development Research Program*: The World Bank.

Appendix

Table A1. OLS estimates of the impact of migration on origin household consumption growth 1997-2000
(expenditure including transfers out)

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
HH has a migrant	0.061***		0.066***	
	(0.02)		(0.02)	
HH has a current migrant		0.146***		0.159***
		(0.03)		(0.03)
HH has a return migrant		-0.095***		-0.089***
		(0.03)		(0.03)
HH size	0.043	0.044	0.043	0.044
	(0.06)	(0.06)	(0.06)	(0.06)
N. children 0-5	-0.094	-0.095	-0.092	-0.091
	(0.07)	(0.07)	(0.07)	(0.07)
N. children 6-10	-0.107	-0.109	-0.107	-0.108
	(0.07)	(0.07)	(0.07)	(0.07)
N. chidlren 11-15	-0.021	-0.025	-0.020	-0.024
	(0.06)	(0.06)	(0.07)	(0.07)
N. children 16-20	-0.065	-0.069	-0.068	-0.071
	(0.06)	(0.06)	(0.07)	(0.07)
N. males 21+	-0.039	-0.041	-0.040	-0.041
	(0.05)	(0.05)	(0.06)	(0.06)
N. females 21+	-0.028	-0.032	-0.028	-0.032
	(0.06)	(0.06)	(0.06)	(0.06)
Average age in HH	0.000	-0.000	0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Head's age	0.001	0.001	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Head unmarried	0.015	0.018	0.015	0.018
	(0.11)	(0.11)	(0.11)	(0.11)
Head's education				
Elementary/Jr. secondary	0.020	0.021	0.024	0.025
	(0.03)	(0.03)	(0.03)	(0.03)
Senior secondary	0.192***	0.194***	0.194***	0.195***
	(0.04)	(0.04)	(0.05)	(0.04)
University	0.297***	0.294***	0.312***	0.308***
	(0.11)	(0.11)	(0.11)	(0.11)

Table A1. (Continued)

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
Head's sector				
Manufacturing	-0.061	-0.064	-0.070*	-0.071*
	(0.04)	(0.04)	(0.04)	(0.04)
Construction	-0.074**	-0.071**	-0.073**	-0.069*
	(0.03)	(0.03)	(0.04)	(0.04)
Services	0.013	0.012	0.014	0.014
	(0.03)	(0.03)	(0.03)	(0.03)
Not working	-0.006	-0.006	-0.004	-0.002
	(0.04)	(0.04)	(0.04)	(0.04)
N. workers in HH				
Market	-0.047***	-0.047***	-0.048***	-0.048***
	(0.01)	(0.01)	(0.01)	(0.01)
Government	0.045	0.043	0.047*	0.043
	(0.03)	(0.03)	(0.03)	(0.03)
Quintile of HH income				
2nd	0.064*	0.066*	0.060	0.062*
	(0.04)	(0.04)	(0.04)	(0.04)
3rd	0.037	0.039	0.047	0.049
	(0.04)	(0.04)	(0.04)	(0.04)
4th	0.154***	0.153***	0.156***	0.156***
	(0.04)	(0.04)	(0.04)	(0.04)
5th	0.299***	0.297***	0.294***	0.295***
	(0.05)	(0.05)	(0.05)	(0.05)
Quintile of HH expenditure				
2nd	-0.397***	-0.394***	-0.395***	-0.394***
	(0.03)	(0.03)	(0.03)	(0.03)
3rd	-0.584***	-0.583***	-0.583***	-0.584***
	(0.04)	(0.04)	(0.04)	(0.04)
4th	-0.843***	-0.843***	-0.841***	-0.843***
	(0.04)	(0.04)	(0.04)	(0.04)
5th	-1.366***	-1.366***	-1.365***	-1.366***
	(0.06)	(0.06)	(0.06)	(0.06)
P.c. Business land	-0.017	-0.019	-0.021	-0.022*
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Non-business land	0.005	0.004	0.009	0.008
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Business land squared	0.001	0.001*	0.002*	0.002*
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Non-business land squared	-0.000	-0.000	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Jewelries	0.010***	0.010***	0.010***	0.010***
	(0.00)	(0.00)	(0.00)	(0.00)
N	5,027	5,027	5,027	5,027
R-squared	0.22	0.23	0.22	0.23

Table A2. 2SLS estimates of the impact of migration on origin household consumption growth 1997-2000

(expenditure including transfers out)

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
HH has a migrant	-0.128		-0.105	
	(0.12)		(0.12)	
HH has a current migrant		0.532		0.568*
		(0.37)		(0.33)
HH has a return migrant		-1.026**		-0.985**
		(0.50)		(0.42)
HH size	0.051	0.057	0.049	0.053
	(0.06)	(0.06)	(0.06)	(0.06)
N. children 0-5	-0.110	-0.108	-0.102	-0.100
	(0.07)	(0.07)	(0.07)	(0.07)
N. children 6-10	-0.119*	-0.127*	-0.115*	-0.123*
	(0.07)	(0.07)	(0.07)	(0.07)
N. chidlren 11-15	-0.019	-0.045	-0.015	-0.045
	(0.06)	(0.07)	(0.07)	(0.07)
N. children 16-20	-0.054	-0.081	-0.056	-0.075
	(0.06)	(0.07)	(0.06)	(0.07)
N. males 21+	-0.037	-0.053	-0.038	-0.040
	(0.05)	(0.06)	(0.06)	(0.06)
N. females 21+	-0.025	-0.049	-0.024	-0.047
	(0.06)	(0.06)	(0.06)	(0.06)
Average age in HH	-0.001	-0.001	-0.000	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Head's age	0.001	0.002	0.001	0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Head unmarried	0.014	0.058	0.016	0.062
	(0.11)	(0.12)	(0.11)	(0.12)
Head's education				
Elementary/Jr. secondary	0.027	0.036	0.030	0.035
	(0.03)	(0.03)	(0.03)	(0.03)
Senior secondary	0.206***	0.214***	0.207***	0.207***
	(0.04)	(0.04)	(0.05)	(0.04)
University	0.317***	0.299**	0.332***	0.308**
	(0.11)	(0.12)	(0.12)	(0.12)

Table A2. (Continued)
(expenditure including transfers out)

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
Head's sector				
Manufacturing	-0.063	-0.082*	-0.072*	-0.081*
	(0.04)	(0.05)	(0.04)	(0.05)
Construction	-0.066*	-0.047	-0.064*	-0.043
	(0.03)	(0.04)	(0.04)	(0.04)
Services	0.018	0.012	0.019	0.015
	(0.03)	(0.03)	(0.03)	(0.03)
Not working	0.005	0.003	0.003	0.017
	(0.03)	(0.04)	(0.04)	(0.04)
N. workers in HH				
Market	-0.048***	-0.042***	-0.048***	-0.044***
	(0.01)	(0.01)	(0.01)	(0.01)
Government	0.047*	0.032	0.049*	0.025
	(0.03)	(0.03)	(0.03)	(0.03)
Quintile of HH income				
2nd	0.060*	0.079**	0.059	0.073*
	(0.04)	(0.04)	(0.04)	(0.04)
3rd	0.028	0.042	0.039	0.054
	(0.04)	(0.04)	(0.04)	(0.04)
4th	0.143***	0.143***	0.148***	0.147***
	(0.04)	(0.05)	(0.04)	(0.05)
5th	0.299***	0.284***	0.296***	0.293***
	(0.05)	(0.05)	(0.05)	(0.05)
Quintile of HH expenditure				
2nd	-0.387***	-0.369***	-0.388***	-0.375***
	(0.03)	(0.03)	(0.03)	(0.03)
3rd	-0.575***	-0.565***	-0.575***	-0.571***
	(0.04)	(0.04)	(0.04)	(0.04)
4th	-0.834***	-0.829***	-0.833***	-0.838***
	(0.04)	(0.04)	(0.04)	(0.04)
5th	-1.353***	-1.354***	-1.353***	-1.358***
	(0.06)	(0.06)	(0.06)	(0.06)
P.c. Business land	-0.016	-0.027*	-0.021	-0.027*
	(0.01)	(0.02)	(0.01)	(0.01)
P.c. Non-business land	0.006	-0.001	0.009	0.004
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Business land squared	0.001	0.002**	0.002*	0.002**
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Non-business land squared	-0.000	0.000	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)
P.c. Jewelries	0.010***	0.009***	0.010***	0.009***
	(0.00)	(0.00)	(0.00)	(0.00)
N	5,270	5,270	5,027	5,027

Table A2. (Continued)
(expenditure including transfers out)

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
F-excluded instr. (migrant)[a]	68.51		69.41	
F-excluded instr. (current)		38.52		43.05
F-excluded instr. (return) [b]		25.75		25.57
Angrist-Pischke F-excluded instr. (current)		17.17		27.83
Angrist-Pischke F-excluded instr. (return)		11.75		12.68
Hansen J P-value	0.16	0.56	0.15	0.69

Notes:
Robust standard errors in parentheses, clustered at the district-level.
*** 1%, ** 5%, * 10% significance level.
The dependent variable is the difference in log per capita household total monthly consumption between 1997 and 2000, including transfers to non-household members and to members who are not in the household. Consumption is measured in real terms, with base Jakarta 2000. Estimates include a constant term. Province and rural dummies and their interactions are also included.
[a.] To instrument the dummies indicating whether the household had a migrant in 1997-2000 and a current migrant in the same period, the predicted values from probit estimations described at point 4. (iii) in Section 5.2 above were used as excluded instruments.
[b.] To instrument the dummy indicating whether the household had a return migrant in 1997-2000, the variables indicated at points 5. and 6. in Section 5.2 above were used as excluded instruments.

Table A.3 Transfers as a function of households' migration status

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
HH has a migrant	0.737***		0.816***	
	(0.25)		(0.26)	
HH has a current migrant		0.789***		0.905***
		(0.28)		(0.30)
HH has a return migrant		0.312		0.281
		(0.29)		(0.30)
HH size	0.309	0.313	0.452	0.454
	(0.36)	(0.36)	(0.40)	(0.40)
N. children 0-5	-0.619	-0.622	-0.750	-0.751
	(0.44)	(0.44)	(0.48)	(0.48)
N. children 6-10	-0.380	-0.389	-0.541	-0.549
	(0.42)	(0.42)	(0.45)	(0.45)
N. chidlren 11-15	-0.343	-0.351	-0.484	-0.494
	(0.40)	(0.40)	(0.43)	(0.44)
N. children 16-20	-0.242	-0.252	-0.399	-0.407
	(0.41)	(0.41)	(0.45)	(0.45)
N. males 21+	-0.223	-0.229	-0.365	-0.366
	(0.38)	(0.38)	(0.41)	(0.41)
N. females 21+	-0.519	-0.527	-0.619	-0.627
	(0.39)	(0.39)	(0.42)	(0.42)
Average age in HH	-0.012	-0.012	-0.014	-0.015
	(0.01)	(0.01)	(0.02)	(0.02)
Head's age	0.005	0.005	0.003	0.003
	(0.01)	(0.01)	(0.01)	(0.01)
Head unmarried	0.188	0.164	0.232	0.210
	(1.00)	(1.00)	(1.01)	(1.01)
Head's education				
Elementary/Jr. secondary	0.083	0.085	0.140	0.142
	(0.26)	(0.26)	(0.27)	(0.28)
Senior secondary education	0.357	0.365	0.313	0.320
	(0.41)	(0.41)	(0.41)	(0.41)
University	0.641	0.633	0.417	0.410
	(0.61)	(0.61)	(0.60)	(0.60)
Head's sector				
Manufacturing	0.267	0.266	0.222	0.221
	(0.36)	(0.36)	(0.35)	(0.35)
Construction	0.245	0.255	0.209	0.224
	(0.39)	(0.38)	(0.39)	(0.39)
Services	0.447	0.447	0.355	0.355
	(0.28)	(0.28)	(0.28)	(0.28)
Not working	0.001	0.003	-0.005	-0.001
	(0.32)	(0.32)	(0.34)	(0.34)

Table A.3 (Continued)

	All migrants		At least one 'non-tied' migrant	
	(1)	(2)	(3)	(4)
N. workers in HH				
Market	-0.122	-0.121	-0.156	-0.156
	(0.13)	(0.13)	(0.13)	(0.13)
Government	-0.184	-0.188	-0.139	-0.148
	(0.28)	(0.28)	(0.28)	(0.28)
Quintile of HH income				
2nd	-0.083	-0.082	-0.093	-0.093
	(0.27)	(0.28)	(0.28)	(0.28)
3rd	-0.621**	-0.619**	-0.610**	-0.606**
	(0.27)	(0.27)	(0.29)	(0.29)
4th	0.527*	0.522*	0.568*	0.562*
	(0.31)	(0.31)	(0.33)	(0.33)
5th	-0.006	0.000	0.116	0.126
	(0.39)	(0.39)	(0.41)	(0.41)
Quintile of HH expenditure				
2nd	-0.351	-0.345	-0.312	-0.307
	(0.28)	(0.28)	(0.29)	(0.29)
3rd	-0.693**	-0.697**	-0.737**	-0.744**
	(0.31)	(0.31)	(0.32)	(0.32)
4th	-1.051***	-1.059***	-0.984***	-0.995***
	(0.34)	(0.34)	(0.35)	(0.35)
5th	-1.504***	-1.505***	-1.553***	-1.553***
	(0.36)	(0.36)	(0.38)	(0.38)
P.c. Business land	0.126	0.123	0.100	0.098
	(0.11)	(0.11)	(0.11)	(0.11)
P.c. Non-business land	-0.170	-0.171	-0.189	-0.192
	(0.13)	(0.13)	(0.14)	(0.14)
P.c. Business land squared	-0.007	-0.007	-0.005	-0.005
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Non-business land squared	0.011	0.011	0.012	0.013
	(0.01)	(0.01)	(0.01)	(0.01)
P.c. Jewelries	0.026	0.026	0.024	0.024
	(0.02)	(0.02)	(0.02)	(0.02)
N	5,263	5,263	5,020	5,020
R-squared	0.060	0.060	0.062	0.062

Notes:
Robust standard errors in parentheses, clustered at the district-level.
*** 1%, ** 5%, * 10% significance level.
The dependent variable is the difference in log of total household transfers out between 1997 and 2000.Transfers are measured in real terms, with base Jakarta 2000. Estimates include a constant term. Province and rural dummies and their interactions are also included.