Risk and Inequality in Rural Thailand

Chowdhury Rashaad Shabab

Submitted for the degree of Doctor of Philosophy

University of Sussex

August, 2017

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

University of Sussex
Chowdhury Rashaad Shabab
Submitted for the degree of Doctor of Philosophy

Risk and Inequality in Rural Thailand

Summary

The first chapter of this thesis provides an introduction to the issues that will be covered in the remaining chapters, reviews the relevant literature on risk and insurance, and provides an overview of the data on rural Thailand that I will be using.

The second chapter investigates the extent to which households at different parts of the income distribution among these Thai households differ in the extent to which their income streams are affected by droughts. I find that the income streams of relatively rich households are better insured than their poorer counterparts. I am able to empirically link the better insurance possibilities enjoyed by richer households to observable characteristics such as the education level of the head of household, the type of contract the head is likely to be employed in, and the relative youth of the heads of richer households.

The third chapter demonstrates that income inequality among these households is declining, both over the duration of the panel, and over the lifecycles of the heads of these households. I show that this decline cannot be explained by standard lifecycle considerations. Rather, I find that remittances from the adult children of the heads of household account for the entirety of the reduction in income inequality over the lifecycles of the heads of household.

Chapter 4 models the probability with which a household receives remittances (the 'extensive margin') and the share of remittances in household income (the 'intensive margin') as functions of observable household characteristics. Using these models, I construct counterfactual distributions of income which permit me to identify the extent to which the extensive and intensive margins of remittance receipts account for the reduction in inequality that these models are able to explain. Chapter 5 concludes.

Acknowledgements

First, I would like to gratefully acknowledge the guidance and support of my supervisors, Professor Richard Disney and Professor Andy Mckay. It is certainly true that most of what I know about the topics that are studied in this thesis, I have learned from them. But furthering my education on these specific topics is only a small part of the contribution they have made in developing my ability to write this thesis. My previous studies may have had some success in teaching me to think critically, but it has only been under the patient supervision of Richard and Andy that I have learned to think rigorously. Recently, I have had occasion to look back on some of the work I did in the early stages of my candidature. When comparing that work to the work that I submit with these acknowledgements, I cannot help but be struck by the value added by three years of their excellent advice and supervision. It has been my privilege to be their student.

I would like to gratefully acknowledge the contributions of Dr. Julie Litchfied, who was my supervisor in the early stages of this thesis.

The work in this thesis has benefitted immensely from substantive comments from Professor Alan Winters, Professor Michael Lipton, Dr. Sam Marden, Dr. Ingo Bochert, Professor Richard Tol and Dr. Julie Litchfield. I am also grateful to Professor Richard Tol for convening the Economics PhD. programme at Sussex University. The programme has benefitted immensely from his leadership, diligence and thoroughness. I would like to thank Dr. Peter Holmes for helpful comments on an earlier draft of the work contained in this thesis. I also thank the other members of Department of Economics at the University of Sussex for the helpful feedback and advice they have provided over the years. I am grateful to my father, Professor Chowdhury Rafiqul Abrar for (among many other things, which I will mention below) proof-reading the final draft. Any remaining errors are my own.

Among my fellow PhD. students, I would especially like to thank Eva-Maria Egger, who as a fellow migration researcher, has routinely provided me with helpful technical feedback on the substance of this thesis. In addition to this, Eva has not only been an excellent friend to me, but also to my family. The depth of our friendship is best exemplified by the fact that on the night our second child, Niraad, was born it was Eva who looked after our eldest, Radaav. As we are half the world away from familial sources of support, I cannot understate how important it has been to my family and I to be able to depend on someone as kind, capable and generous as Eva.

I am grateful to Ani Rurdra Silwal, whose serene composure and compassionate perspective has often been a source of strength to me. I would also like to thank Edgar Salgado Chavez and Cecilia Poggi who have patiently helped me overcome many difficulties with STATA, and some of the idiosyncrasies of the data that I use. I am also grateful to my other fellow PhD. students who have been a constant source of friendship, empathy and comfort during what has been a very demanding process.

I would like to thank my undergraduate advisor, Professor Paul Johnson, who has over many years provided me with advice and support. It was at his urging that I applied for the Emilie Louise Wells Fellowship from my undergraduate institution, Vassar College. I was the grateful recipient of this fellowship for the second and third years of my candidature. The objective of the fellowship is “To aid in the advancement of knowledge of economics and social activities and to promote social work”. It is open to members of Vassar’s graduating class and to alumni.

I would also like to thank the University of Sussex for the financial support offered by their highly competitive Graduate Teaching Award, of which I was the grateful recipient for 3 years.

Many sacrifices have been made toward the completion of this thesis. Perhaps the most profound of them has been extracted from my sons, Radaav Ninaad Chowdhury and Niraad Raynaav Chowdhury, both of whom were born over the course of my candidature. A lack of funding and the severe working hours demanded by the thesis made it impractical for my family to remain with me for the last year of my candidature. As a result, Radaav has spent most of that year without the company of his father. Niraad, who was born six months ago, has only known the embrace of his father for a handful of weeks over that time. I pledge to do whatever I can to make up for the lost time. For my part, I have missed them profoundly.

I would like to thank my in-laws, Shafiuddin Ahmed and Ila Imam for taking such excellent care of my sons at a time when I could not. Though every moment that I have been away from my sons has been a trial, I have always had the peace of mind that they are benefitting from the capable care and boundless love of their grandparents.

This thesis would not have been possible without the support of my parents, Professor Tasneem Siddiqui and Professor Chowdhury Rafiqul Abrar. They have guided me through the anguish and despair that I was so often plagued by especially (but not exclusively) in the early stages of my candidature. They have made substantial sacrifices to provide the resources that were necessary to fill gaps in my funding at different stages, over the years. They too, have helped raise my sons with love and kindness when the demands of this thesis prevented me from doing so myself. Without their support, this thesis would surely have been impossible.

My sister, Chowdhury Rashaam Raiyan has also been a crucial source of emotional support. Her husband Dr. Ananta Neelim who is a recent Economics PhD. has been a valuable source of advice.

Finally, I come to my wife and partner, Samantha Shuchismita. I am not sure that I can do justice to the numerous ways in which she has contributed to the completion of this thesis. In fact, I am quite sure that I cannot. She has done everything in her power to keep our lives moving forward while still allowing me to remain focused on the PhD. She has given birth to my two sons and for a time, has raised them without my day to day assistance. She has been unwavering in her support of my studies and forgiving beyond measure of all the demands placed on our personal lives by my candidature. Truly, I am in her debt.

I dedicate this thesis to her and to our sons.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This collection of essays studies inequalities and insurance behaviour among village households in rural Thailand using especially high quality data that has been made available by the Townsend Thai Project (Townsend, 2011). In this introductory chapter I present a preview of the research questions that will be addressed in subsequent chapters. I go on to briefly review the relevant strands of the literature on the insurance behaviour of households, with an emphasis on the findings that will be most pertinent to a rural, developing country context. I also introduce the Townsend Thai data which I will use to analyse the economic behaviour of these village households.

In the chapter 2, I investigate the extent to which households at different parts of the income distribution differ from one another in the degree to which their income streams are insured against the materialization of widespread, unanticipated, adverse shocks to income, using the example of droughts. I find that the income streams of relatively rich households are better insured than their poorer counterparts, even though richer households are just as likely as poorer ones to be involved in agriculture. I am able to empirically link the better insurance possibilities enjoyed by richer households to observable characteristics such as the education level of the head of household, the type of contract the head is likely to be employed in, and the relative youth of the heads of richer households.

In the third chapter of this thesis I demonstrate that income inequality among these rural Thai households is declining, both over the duration of the panel, and over the lifecycles of the heads of these households. I show that this decline cannot be explained by either a convergence in the distribution of the earnings of individuals within these households, or by standard lifecycle considerations such as differences in fertility and cohabitation rates of adult children of the heads of household between relatively rich and poor households. Rather, I show that remittances from children are a key channel through which these households insulate their income streams from the retirement-related dip in earnings that usually occurs later on in the lifecycle. I find that remittances from the adult children of the heads of household, who reside outside these villages of origin, account for the entirety of the observed reduction in income inequality over the lifecycles of the heads of household.

In chapter 4, I model the probability with which a household receives remittances (the 'extensive margin' of remittances) and the share of remittances in household income (the 'intensive margin') as functions of observable household characteristics. I identify the extent to which each of these margins varies across the distribution of household income and thereby construct counterfactual income distributions where either the probability of receiving remittances, or the share of remittances in household income is allowed to vary as a function of household income, holding the other fixed. Using these counterfactual distributions, I identify the extent to which the extensive and intensive margins of remittance receipts account for the inequality reducing effect of remittances on the distribution of household income.

I now turn to conducting a brief review of the substantial literature on the economics of the insurance decisions made by households, before introducing the Thai data.

1.1 Economic Paradigms of Insurance

The literature on the degree to which the measured consumption levels of households vary in response to unanticipated changes in measured income is as vast as it is important. It is beyond the scope of this chapter to discuss the numerous, celebrated contributions in this field in a manner which does them justice. Rather, I will briefly touch on those key contributions that help place the remaining chapters of this thesis in their proper context. I start by recalling some of the relevant results that pertain to what has come to be known as the complete markets hypothesis (Arrow 1951, Debreu 1952, Arrow-Debreu 1954), and the permanent income hypothesis (Friedman 1957, Ch 3; Ando and Modigliani 1963), paradigms which Blundell et al. (2008) characterize as being the 'two polar models [that] have highlighted the agenda' on economic inquiry on insurance. Both of these portrayals of the risk environment within which households operate characterise households as not exerting an influence on the degree of risk in their income streams. The possibility that households reduce their exposure to income risk as an insurance strategy is formalized in the 'income smoothing' (Morduch 1994, 1995) literature, which I will also outline.

The Complete Markets Hypothesis

The complete markets paradigm assumes that there are no impediments whatsoever to the ability of households to trade their exposure to risk with one another. Indeed, Arrow and Debreu's (1954) theoretical formulation assumes that households have access to a full set of state-continent claims, which allow them to insure themselves against every possible prospect before uncertainty was resolved. Under these circumstances, if the income streams of any two risk-averse households are less than perfectly correlated, then they can benefit from insuring each other against potential shortfalls in their individual income streams. Indeed, the potential benefits to pooling risk in this way may exist until households have insured one another against all risk that is specific to individual households. If all households are equally risk averse then this benchmark predicts that the consumption level of any individual household varies over time only in response to changes in aggregate output, and not in response to changes in the income level of that individual household. As a result, households will be able to effectively insure one another against idiosyncratic shocks to income which affect individual households, but not against covariate ones which affect entire communities. Arrow and Debreu (1954) show that if markets are complete so that all households can freely trade their exposure to risk in every possible state of the world, then under some stringent assumptions about technologies, information and preferences, the risk pooling equilibrium described above exists and is Pareto efficient.

Clearly, it is not possible for real-world economic agents at any given time to trade on the value of every good, in every possible state of the world, in all future periods. Nonetheless, households and individuals may have access to markets and other institutions which allow them to engage in a considerable degree of welfare enhancing risk pooling. If such institutions are effective, the complete markets benchmark makes distinct predictions about the way in which the consumption levels of individual households should respond to covariate shocks to income as opposed to idiosyncratic ones. Covariate shocks are unpredictable events which affect a whole community, such as drought and flooding. These will cause aggregate output for the community to fall and so cannot be insured against, even in the highly stylized case where markets are complete. As a result, the materialization of such a shock will engender a reduction in the consumption levels of all households in the community (in the case that all households are equally risk averse, this reduction will affect all households proportionately).

Idiosyncratic shocks, which affect individual members of the community can be successfully insured against by pooling risk between households and so, under the complete markets hypothesis, will not engender a change in the level of consumption enjoyed by an affected household. Tests of whether or not households interact with one another within a set of economic and cultural institutions which allow them to approximate this Pareto efficient outcome have been conducted in a wide range of settings.

Cochrane (1991) used data from the U.S. Panel Study of Income Dynamics and found that individual food consumption was not insulated from employment and health status of that individual, leading to a rejection of the complete markets hypothesis. Attanasio and Weber (1995) and Nelson (1994) also reject the complete markets hypothesis in the U.S. using the Consumer Expenditure Survey.

In a developing country context, Townsend (1994) found evidence against the complete markets hypothesis for three Indian villages. Udry (1994) documented that though Nigerian households use informal credit transactions as instruments to pool risk, the degree of insurance offered by this institution was insufficient to fully protect consumption against idiosyncratic income risk, also leading to a rejection of the complete markets hypothesis.

In a recent development, a number of papers have used heterogeneity in risk preferences to argue that earlier tests of the complete markets hypothesis were biased against the null hypothesis of full risk sharing. Accounting for risk preferences using responses to questions about hypothetical jobs in U.S. administrative data, Schulhofer-Wohl (2011) fails to reject the complete markets hypothesis. Mazzocco and Saini (2012) reject perfect risk sharing at the village level using data on India, but fail to do so at the caste level. Chiappori, et al. (2011) measure heterogeneity in risk preferences in Thai villages using a full risk-sharing model. Finding strong evidence for this type of heterogeneity, they fail to reject the complete markets hypothesis, conditioning on the estimated risk-aversion parameters.

This last study is especially salient to this thesis because it is also conducted on data from the Townsend Thai Project (although that study uses the monthly panel which is available for a smaller number of villages than the annual panel which is used in this thesis). Nonetheless, the findings of that paper suggest that covariate shocks to income may be particularly detrimental to the welfare of the rural Thai communities that I study. This is confirmed by Samphantharak and Townsend (2017) who decompose risk in the income streams of the households sampled by the monthly survey into an idiosyncratic and a covariate component and analyse the risk equivalised incomes with each of these types of risk. They find that covariate risks which affect entire communities and are therefore harder to insure against, command a larger risk premium than idiosyncratic risks, as predicted by economic theory. These communities are thus able to protect themselves against idiosyncratic shocks by pooling risk within the community. But the literature has found that such risk pooling does not tend to be effective across communities (Chiappori et al., 2011; Mazzocco and Saini (2012); and Urdy, 1994), so that households in different villages are unlikely to be able to pool their risks with one another. These findings are salient to chapter 2 of this thesis, where I find that the income streams of relatively poor households are particularly susceptible to shortfalls induced by drought, the most widespread of the covariate shocks that these villages report being affected by.

The Permanent Income Hypothesis

In contrast to the complete markets hypothesis, the permanent income hypothesis assumes that only capital markets are perfect. This characterization of the insurance environment in which households operate therefore allows them to borrow and lend freely against their own future income. In the simplest version of the model, the extent to which households discount the future is exactly equal to the interest rate. So long as the preferences of households satisfy diminishing marginal utility, their optimal policy is to set consumption equal to the expected value of lifetime income (inclusive of any assets), averaged over all remaining periods.

In this framework permanent and transitory shocks to income will have very different effects on lifetime income and thus current consumption (Hall, 1978; Hall and Mishkin, 1982). A permanent shock to income will have a substantial effect on remaining lifetime income and therefore on the level of consumption, as households consume an amount that is in proportion to their expected remaining lifetime income. Since a transitory shock has a very small impact on lifetime income, consumption should respond negligibly to such a shock. Thus the textbook permanent income hypothesis predicts that consumption is not insured against permanent shocks, and almost fully insured against transient ones.

Campbell and Deaton (1989) observed that macro-data on consumption in the U.S. for the post-war period was characterised by too much insurance against permanent shocks to income to be consistent with the standard permanent income theory. Deaton (1992), and Attanasio and Pavoni (2011) have also confirmed the 'excess smoothness' of consumption to permanent income shocks in micro data from the U.S. and U.K. respectively.

In contrast, where the theory predicts that transient shocks should have little impact on consumption, Hall and Mishkin (1982) have found that consumption exhibits 'excess sensitivity,' to these short term fluctuations. Deaton (1991) shows theoretically, that even if all shocks to income are temporary, liquidity constraints can cause the optimal consumption stream to track changes in income.

Some of the most insightful implications of the permanent income hypothesis become apparent when one considers the way that the incomes of individuals evolve over the lifecycle[1]. On average, earnings decline precipitously when individuals reach retirement age, but since this decline is anticipated, it should not engender a corresponding decline in an individual's level of consumption. Chapters 3 and 4 of this thesis establish the role of remittances from the adult children of the ageing heads of these households in insulating their income streams from such lifecycle related declines in earnings.

Insuring Consumption by Smoothing Income

When markets for risky assets and credit are both incomplete, households can neither trade their exposure to risk with one another, nor insure themselves by borrowing against their own future incomes. Under these circumstances households may prefer income prospects with lower variances, even at the expense of potentially substantial reductions in the mean. Morduch (1994 and 1995) used the term 'income smoothing' to describe this type of insurance strategy, in contrast to the 'consumption smoothing' that is consistent with the complete markets or permanent income paradigms.

'Income smoothing' as an insurance strategy has typically been studied among the poor and vulnerable, who may be more likely to be credit constrained. Much of the early evidence was based on data from rural India that was gathered by the International Crops Research Institute for the Sami-Arid Tropics (ICRISAT) between 1975 and 1984. Morduch (1990) (cited in Morduch 1995) showed that asset-poor households, whose

[1] These insights relate macroeconomic variables such as the savings rate and measures of inequality to the ageing of a population. The link between lifecycle theory and the evolution of inequality will be discussed at length in chapter 3, and so to avoid repetition I do not discuss them here. Since this thesis does not speak directly to the literature on aggregate savings, I also refrain from discussing that literature here.

consumption is most vulnerable to income shocks devote a greater proportion of their land to safer, but lower yielding, traditional varieties of crops than richer households. Rosenzweig and Binswanger (1993) demonstrated that households surveyed by ICRISAT who were in lower wealth quartiles used production techniques that were less susceptible to rainfall variation, even though on average these techniques were less productive. Kochar (1999) found that households surveyed by ICRISAT protected their consumption levels by diverting labour from farm employment to off-farm employment when faced with a crop shock, thereby reducing income variability.

These existing narratives of income smoothing assume that credit is the main dimension along which the insurance strategies available to poorer households are constrained, relative to their better-off peers. Chapter 2 of this thesis complements this narrative by observing that the choice of income generating activities available to richer households may offer greater insurance possibilities than those available to poorer households. The chapter documents that certain observable household characteristics are associated with better insured income streams. These characteristics are more common among relatively rich households. The chapter goes on to present evidence that the income streams of richer households are indeed better insured against covariate shocks than their poorer counterparts in this panel of village households.

The findings of the third chapter of this thesis are also consistent with this insurance paradigm. Households in societies with well-functioning capital markets are likely to accumulate savings during the productive phase of the lifecycle, and run down those savings in retirement; that is, they use savings to buffer their consumption against an anticipated decline in income. In contrast, I find that ageing households in rural Thailand instead depend on transfers from younger generations to smooth out the slump in household income that might otherwise have occurred. Thus household income is itself protected, perhaps because these households cannot depend on formal credit markets to protect consumption from this anticipated decline in income. The importance of migration as an instrument to protect the household against low productivity has been documented by previous studies in the Thai context. Paulson (2000) using data from the Thai Socio Economic Survey, demonstrates that migration patterns in Thailand are consistent with an insurance motive, that is, migrants tend to select destinations where incomes are negatively correlated with incomes in their communities of origin. Also using the Socio Economic Survey data, Yang (2004) demonstrates that there is a great deal of variation in output between Thai districts, but relatively little variation in household incomes. Yang finds that remittance transfers from high output areas to low output ones, supplement household income in low output areas, thereby accounting for the relatively even distribution of household income. One of the findings reported by Townsend (2013) uses the monthly series of the Townsend Thai Data to show that in certain areas, households headed by the elderly rely on remittances to supplement their income. Thus the fact that households in rural Thailand use remittances to smooth out fluctuations in productivity is already well established in the literature. Chapter 3 of this thesis will demonstrate that differentials in the receipt of these remittances across the distribution of income are sufficient to drive a downward trend in the dynamics of income inequality among these households.

1.2 The Townsend Thai Data

All the empirical work in this thesis is conducted on a panel of households that are sampled annually by the Townsend Thai Project (Townsend, 2011). The project also publishes a set of data for a different group of households representing a different (and smaller) set of villages that are observed on a monthly basis, which I do not use. I have been able to access these data because Townsend Thai project has made them freely available for download from the “Harvard Dataverse” website[2].

The project started interviews in 1997, when 2,880 rural households comprised the baseline survey. A total of 4 provinces (*Changwats*) – 2 in the semi-arid Northeast and 2 in the more developed central region near Bangkok – were chosen for survey. These particular provinces were chosen because of the availability of pre-existing data, that the project intended to use for comparison. Within each of the provinces, 12 sub-provinces (*Tambons*) were chosen at random. Within each sub-province, four villages were selected, also at random. Thus in total, 192 villages were surveyed, covering 2,880 households.

In 1998 one third of the original sample was chosen for resurvey. All four original provinces were covered, but only 4 out of the 12 sub-provinces in each were selected randomly for resurvey. Accordingly, from 1998 the sample size shrinks to 960 households in each year. Additional areas were then added to the survey in 2004 (when the sample size rises to 1,080) and 2005 (when the sample size is 1,320), but due to the outbreak of conflict in some of the newly added areas, the sample size drops to 1,200 in 2006 which is sustained until 2011.

I will use a subset of this data set in the analysis here. To ensure that I am comparing like with like, I restrict my attention to the 64 villages that are sampled in all fifteen years for which data are available, that is I work with the panel of 960 households. Within this subsample, the year on year attrition rate reaches a maximum value of 6%, but is usually closer to 3%. These households are replaced with households from the same village to keep the number of observations in each cross section at 960.

[2] Https://dataverse.harvard.edu

Depending on the particular research question being addressed, I will either use this unbalanced panel of 960 households per year, or a subset of 609 households which constitute the balanced panel. The difference between the balanced and unbalanced panel is accounted for by 288 households that are subject to attrition between 1997 and 2011; 51 households that report missing values for key variables of interest – namely income or consumption – in at least one year; and 10 observations which I drop by hand for appearing spurious. The households that comprise the last of these groups are listed in appendix 1.

In chapter 2, I will identify differences in the extent to which changes in income engender changes in consumption among these Thai households. Because the goal of this exercise is to identify household-level changes in variables over time, differences in the composition of the sample from one period to the next may pose a considerable threat to identification. Accordingly, this analysis is conducted on the balanced panel of households.

The methodology I employ in chapter 3 is to follow cohorts of households, grouped by the decade of birth of the head of household, over the 15 years of the panel. Since cohort analysis averages over many households headed by people of similar age, the cost incurred due to comparing a slightly different set of households from one wave of the survey to the next is minimal. The benefits in terms of increased numbers of observations are substantial. For these reasons, my preferred specifications in chapter 3 are based on the unbalanced panel. Nonetheless, I check that the core insights of the paper are robust to using only the balanced panel.

Chapter 4 disaggregates the inequality reducing effect of remittances (that will be documented in chapter 3) into extensive and intensive margins. In that context too, the threats to identification posed by using an unbalanced panel are small, relative to the potential benefits of increased sample size.

An attractive aspect of the annual series of the Townsend Thai data is that detailed household level information is complemented by information from interviews with a key informant from each village (typically the village headman) about a range of issues. These interviews provide information on the number of households in each village that are affected by droughts in each year. In chapter 2 I use this variable as an exogenous, covariate shock to household income. These key-informant interviews are also the source of the information on the mass migration histories of individual villages which I use as an instrument for current remittance receipts from the migrant children of the heads of these household in chapter 4.

Empirical applications of the Townsend Thai data

The annual series of the Townsend Thai data have been used extensively, in a wide variety of applications[3]. Perhaps the most celebrated contributions evaluate the effects of the "Million Baht Village Fund" – a large scale government initiative to provide microcredit to over 77,000 villages that was rolled out in 2001. Kaboski and Townsend (2011 and 2012) exploit the quasi-random nature of the per-capita effect of this policy (villages received the same amount of money regardless of their population) and concluded that the widespread provision of microcredit had a positive effect on consumption levels, but an ambiguous effect on investment. They present evidence of positive spill-over effects from entrepreneurs who take up microcredit to the wage growth of those who they

[3] As have the monthly data, but since I do not use the monthly series in any of the analysis in this thesis, I do not review the numerous contributions that were made on the basis of that data, except to point out those applications which are directly related to the research agenda of this thesis.

employ, but also note that the intervention did not appear to create more and better businesses, as proponents of microcredit might have hoped.

Poggi (2015) uses the quasi-random nature of the Million Baht Fund to identify the effect of credit on internal migration from these rural communities, and concludes that a greater provision of credit in a community decreases the likelihood of that community producing internal migrants. In chapter 3 of this thesis, which studies the effect of remittances on inequality among these rural households, I check to see that the results are not driven by this, or any other village-specific process, checking that the results are robust to a fully interacted set of village and time fixed effects.

Kaboski and Townsend (2005) also study microcredit institutions, but this study is conducted on the first wave of the data, from 1997, before the Million Baht Fund was announced. They find that the presence of microfinance institutions can promote asset growth, consumption smoothing and occupational mobility, so long as these institutions provide a certain set of services, namely savings services, emergency services and training and advice.

The annual series of the Townsend Thai data has also been used extensively to study the effects of financial liberalization and financial deepening in transitioning economies. Alem and Townsend (2014) study the effect of different financial institutions on household level outcomes. They find that the presence of a government development bank results in substantially smoother consumption and investment streams, thereby embedding an insurance function in their lending activities. Gine and Townsend (2004) show that increases in financial intermediation not only benefit entrepreneurs, by relieving their credit constraints and allowing them to invest and expand their businesses; but also those who come to be employed by these larger and more productive businesses. Paulson and Townsend (2004) found that financial constraints in these villages were associated with reduced entrepreneurial activity, particularly in the relatively poor North-eastern region.

The data have also shed light on the geographic concentration of enterprises in developing countries. Felkner and Townsend (2011) show that high levels of firm concentration in an area predict subsequent growth near that area. Furthermore, they report that the physical geography of an area is correlated with the level of enterprise activity, leading some areas to be prone to an agglomeration of enterprises while others are left behind.

Thus the Townsend Thai data provide a detailed account of the economic lives of these rural Thai households. These data have been extensively used in highly regarded empirical work that has substantively enhanced our understanding of such communities. Nonetheless, there remain aspects of the economic behaviour of these village households that have not yet been studied.

While the extent to which the consumption streams of these households are insured against income shocks has been studied thoroughly using these data (for example by, Kaboski and Townsend (2011, 2012) and Alem and Townsend (2014), the extent to which the income streams themselves perform an insurance function has not. Empirical studies of this type of insurance behaviour have typically been conducted using data with a narrower temporal range and geographic coverage. Dercon (2006) used data 15 from Ethiopia villages that spanned 5 years. Rosenzweig and Binswanger (1993), Morduch (1990, 1995) and Kochar (1999), are able to exploit variation across 3 villages in rural India over 9 years from 1975 – 1984. In comparison, the Townsend Thai data allows me to investigate insurance behaviour across 64 different villages over 15 years from 1997 to 2011. Not only do the Thai data cover a wider range of villages over a longer time span, but these villages are at a different stage of economic development from the Indian villages studied by Rosenzweig and Binswanger (1993), Morduch (1994, 1995) and Kochar (1999). In the late 1970s and early 1980s the Indian economy was still poor and insular. This contrasts sharply with the economy of Thailand in the late 1990s and the first decade of the 2000s. At this time, the Thai economy was already integrated into the global economy and by the end of the survey period, in 2011, had already achieved upper-middle income status (World Bank, 2016). This thesis therefore, contributes to the income smoothing literature not only by using data that has wider temporal and geographic coverage, but also by studying this form of insurance in an economy that is at a relatively advanced stage of development, over a much more recent time period.

The annual series of the Townsend data have been used to study the effect of relaxing credit constraints on migration behaviour, but they have not been used to study the extent to which remittance receipts are linked to the lifecycles of the heads of these households. Pawasutipaisit and Townsend (2010) and Townsend (2016) use the monthly series to study the lifecycle dynamics of remittance receipts in a different set of Thai villages (which are fewer in number than those covered by the annual data series). This set of studies observe that wealth inequality is declining in these communities but that this "decline is largely due to savings rather than incoming gifts and remittances" (Pawasutipaisit and Townsend, 2010: 56). The authors note that these findings are "not robust to annual data" (page 57), which is the series I study here. The research question of chapter 3 of this thesis is thus distinct from this body of empirical work in two ways: first, where these studies have sought to understand wealth inequality, I am interested in income inequality; and second I document and explain declining inequality in the villages covered by the annual series of the Townsend Thai project whereas those papers study the monthly series. Furthermore, the work in this thesis is novel among the broader literature on the effect of remittances on income inequality field (which will be discussed in more detail in chapter 3) in that it explicitly studies these effects in the context of lifecycle theory. Because I observe these households over a much longer period than preceding studies in this, I am able to credibly discuss lifecycle effects, where previous studies were constrained to either cross sectional effects or relatively short panels.

This thesis will fill these particular gaps in our understanding of this set of rural households, so as to further our understanding of insurance and inequalities in rural communities more generally.

Chapter 2

Income Smoothing Among Thai Households

Introduction

When markets for credit and insurance are incomplete, households have an incentive to insure their consumption by obtaining their income from less volatile sources. In an influential pair of papers, Morduch (1994, 1995) used the term ‘income smoothing’ to describe this phenomenon and contrasted it with the ‘consumption smoothing’ that we consider more commonly. The literature on income smoothing has focused on documenting the use of this type of insurance among the poor and vulnerable (Morduch, 1994 and 1995; Rosenzweig and Binswanger, 1993; and Kochar, 1999), under the maintained hypothesis that credit is the only dimension along which the insurance decisions of risk averse households are constrained. Hence, poorer households, with less access to credit markets are more likely to need to insure themselves using income smoothing. In this paper, I allow for the possibility that poorer households may also be constrained in their ability to secure low-risk income. Then, on the one hand, richer households are less likely to be liquidity constrained and therefore less likely to insure themselves through smooth income; while on the other, they may have privileged access to low-risk income streams which makes them more likely to insure themselves using smooth income. It is then an empirical question as to whether or not richer households are more likely to depend on low-risk income to satisfy their insurance needs than their poorer counterparts. In this chapter, I present evidence from rural Thailand that the income streams of relatively well-off households are better insured against covariate shocks than their worse-off counterparts.

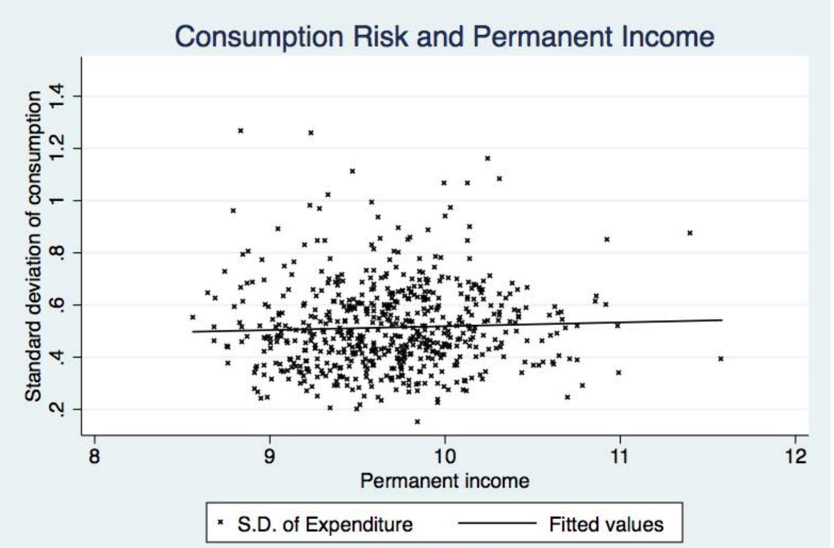


Figure 2.1

Figures 2.1 to 2.3 provide some descriptive statistics that motivate this paper. In the first figure, the y-axis measures the sample standard deviation in the log of each individual household's real[1], equivalised[2] level of consumption over the fifteen years of the panel, $\hat{\sigma}_c$. This is a crude measure of the extent to which the amount consumed by individual households varies over time, or equivalently the extent to which the consumption streams of these households exhibit a lack of insurance. The variable on the x-axis is intended to convey a sense of the level of well-being enjoyed by each of these households, relative to one another. Temporary fluctuations in income will cause any ordering based on the level of income observed in a particular period to be an unreliable measure of underlying

[1] All figures are inflated to 2011 Thai Baht using Consumer Price Index data from the Bank of Thailand website
[2] Per-adult equivalence is calculated using the 'old OECD' scale where the first adult receives a weight of 1, all additional adults receive a weight of 0.7 and each child is weighted by 0.5.

relative wellbeing. Averaging over observed incomes for the duration of the panel partially addresses this concern, although in the permanent income view (Friedman 1957, Ch3; Ando and Modigliani 1963) averaging over consumption may be an even more reliable signal of relative wellbeing. In this view, the amount consumed by a household in each period will be equal to the amount of income that the household expects to generate over the remainder of their lives, net of any assets they may have, divided by the number of remaining periods for which they expect to survive. Thus the comparison of averages over time for each household may provide a misleading impression of relative wellbeing if these households are at different stages of the lifecycle. This is because households headed by retirees who, for example may consume out of savings rather than income, will have systematically higher wellbeing than their incomes indicate. Thus the average of a household's observed level of (equivalised, real) consumption is likely to provide a more reliable signal of that household's relative wellbeing than the average of observed income. For these reasons, on the x-axis, I measure the average over time of each individual household's log real, equivalised consumption. Figure 1 plots the measure of consumption variability against this measure of permanent income for the 609 households that comprise the balanced panel[3] in the annual series of the Townsend Thai data. The figure also plots a line of best fit through these points. I defer a discussion of some of the subtler points around the use of this measure of relative wellbeing until section 2.4.4, later in this chapter.

Figure 2.1 suggests (and a *t*-test reported in the first column of table 2.1 confirms) that the level of insurance, as measured by the variability in household consumption, does not differ significantly across the spectrum of households ranked by permanent income.

[3] The specifics of how I identify the balanced panel were discussed on page 10

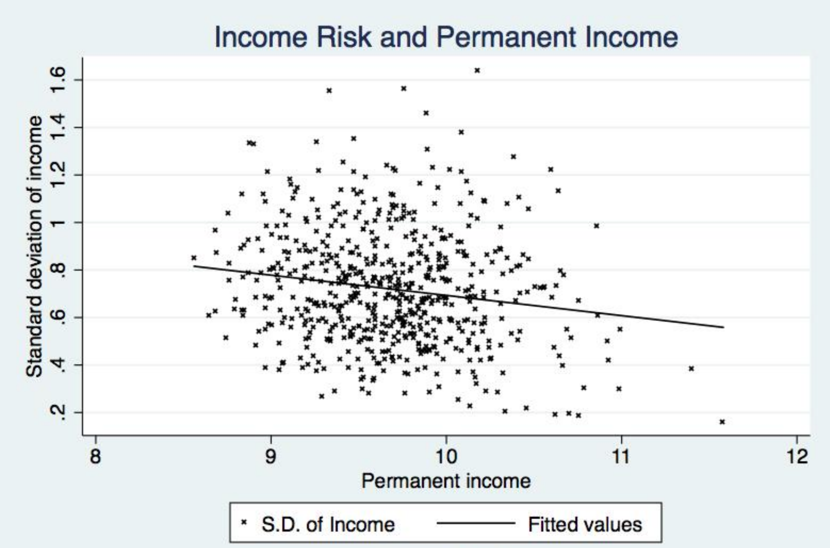


Figure 2.2

Bearing in mind Morduch's (1995: 104) advice that 'One cannot simply look at the smoothness of consumption and know which type of smoothing is at work,' figure 2.2 plots the estimated standard deviation of log income, $\hat{\sigma}_y$, against mean consumption for the same 609 households. Here, there is a clear downward trend, the coefficient for which is reported in the second column of table 2.1. Thus relatively well-off households appear to have smoother income streams than their worse-off counterparts.

Table 2.1: Income Smoothing and Consumption Smoothing

	(1)	(2)	(3)
Dependent variable	*SD(c)*	*SD(y)*	*SD(c)/SD(y)*
Mean Consumption	0.0147 (1.02)	-0.0852*** (-4.06)	0.206*** (6.21)
Constant	0.372*** (2.65)	1.545*** (7.60)	-1.201*** (-3.73)
N	609	609	609

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Recalling that the standard deviation of consumption is similar across households and that 'the two types [of insurance] can act as substitutes for one another' (Morduch, 1994: 104), I then compute the ratio, $\hat{\sigma}_c/\hat{\sigma}_y$ to gauge the extent to which consumption is smoothed, relative to income. Figure 2.3 plots estimates of this consumption smoothing measure against mean household consumption. The proportion of income variation that is allowed by households to pass uninsured into consumption variation increases with the level of mean consumption. Poorer households smooth consumption more, relative to income, whereas better-off households seem to achieve the same degree of insurance in their consumption stream (from figure 2.1) by relying more heavily on low variance income streams. Appendix 2 demonstrates that these patterns are robust to a more restrictive definition of consumption that includes only food, alcohol, tobacco and gasoline expenditures, and an alternative measure of variability, namely the coefficient of variation.

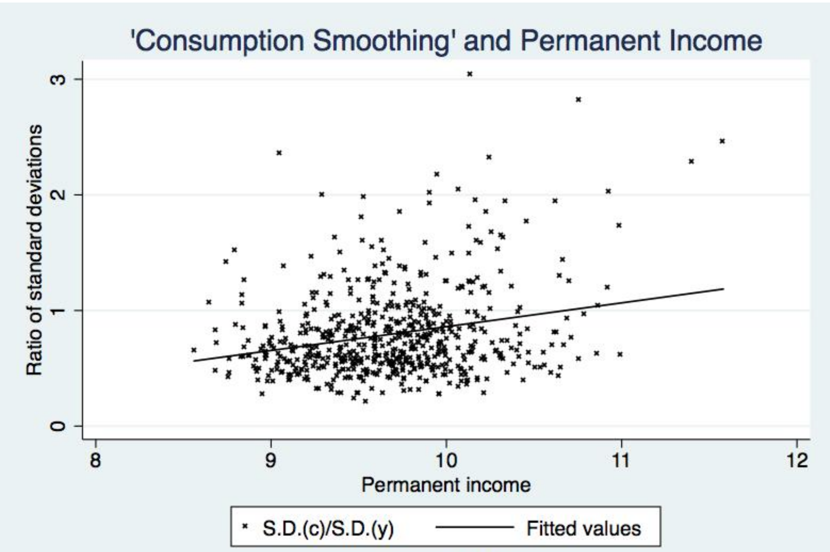


Figure 2.3

These patterns of insurance are not compatible with theories of income smoothing that assume different access to credit markets. Such models predict that the relatively poor, who have less access to credit markets will use more income smoothing for their insurance needs. Rather, these descriptive statistics suggest that in rural Thailand, it is the relatively rich who rely more heavily on low-risk income streams to satisfy their insurance needs than the relatively poor, possibly because they have access to income streams that offer greater insurance possibilities. In other words, households may be constrained in the set of income generating activities that they have access to. Dercon (2002) and Dercon and Krishnan (1996) document how constraints to occupation choice which bind for relatively poor households in Ethiopia and Tanzania cause their income streams to exhibit higher risk for a given return than their relatively rich counterparts. This chapter draws similar conclusions from rural Thailand.

To pursue these issues further and to understand what forces are at work, I combine household-level income data with information on external shocks measured at the village level. Covariate shocks will be defined as shocks which are common to villages and have a demonstrable adverse effect on the income levels of households within those villages. Other studies of the Townsend Thai data have demonstrated that covariate risks are especially detrimental to the welfare of households in rural Thailand. Chiappori et al. (2014) show that household consumption exhibits a lack of insurance against aggregate shocks and Samphantharak and Townsend (2017) find that aggregate risk commands a greater risk premium than idiosyncratic risk, to compensate for this lack of insurance possibilities. This paper will show that such shocks have a disproportionate adverse effect on the income levels of certain groups of households, namely those that are headed by people with low levels of education, those that are employed in jobs that do not pay a monthly wage, those that are headed by the elderly, those that are headed by women, those that are headed by people with multiple jobs and those that own a business. The chapter will also document that a number of these characteristics are more commonly observed among the relatively poor households in the sample. Furthermore, this chapter also finds that there is an important regional component to this phenomenon – drought has a particularly strong adverse effect on village incomes in the relatively poor Northeastern region, but no statistical effect on incomes in the relatively rich central region. As a result of these forces, opportunities to insure their income streams against covariate shocks disproportionately benefit richer households, and I demonstrate that the income streams of these households are indeed better insured against covariate shocks. These forces explain the patterns of insurance observed in figures 2.1, 2.2 and 2.3.

The rest of the paper is organized as follows. In section 2.1 I review the income smoothing literature, noting how this type of insurance behaviour has typically been studied among the relatively poor and vulnerable, in contrast to these preliminary findings. In section 2.2, I discuss the possibility that households may differ in the expected returns they must sacrifice in order to access risk-free income, within the context of Morduch's (1994) theoretical model of income smoothing. Section 2.3 describes the data from the Townsend Thai Project (Townsend, 2011). Section 2.4 uses the Thai data to examine if the income streams of relatively well-off households are indeed better insured against covariate shocks than their worse-off counterparts. I also identify observable characteristics of better-off households that are linked in the data to their ability to reduce the exposure of their income streams to these covariate shocks. Section 2.5 concludes.

2.1 Literature Review

The standard models of insurance such as the text book versions of the permanent income hypothesis (Friedman 1957, Ch3; Ando and Modigliani 1963) and the complete markets hypothesis (Arrow 1951, Debreu 1952, Arrow and Debreu, 1954) assume perfect capital markets. Such a characterization of markets implies that even risk-averse households will make production decisions that maximize the mean of net income, and use borrowing and lending to insure consumption against the risk in the resulting income stream. Thus different forms of savings can play a crucial role in household insurance decisions. In the Thai context however, Chiappori et al. (2014), Chiappori et al. (2014), Townsend (2013) and Paulson (2000) have established that consumption smoothing is achieved by a somewhat different mechanism. They find that consumption smoothing in this context is achieved by households pooling their exposure to risk with one another, so that when a household experiences an adverse income shock, a transfer from another household that has not experienced the same adverse income shock is used to smooth out the effect of the shock on consumption. As such, savings are not the primary channel through which consumption smoothing is achieved. Savings, however, are important to the welfare of these households because Townsend (2013) has found that they vary systematically over the lifecycles of these households and also because differentials in savings rates have important implications for the evolution of wealth inequality among these households. These issues are discussed in detail in chapter 3 which deals explicitly with the evolution of inequality over the lifecycles of these households.

Morduch (1994 and 1995) notes that when capital markets are imperfect so that households are unable to insure consumption against shocks to income, they have an incentive to seek income-earning activities which themselves have lower variances.

The use of income smoothing as an insurance strategy has far reaching implications for a wide range of important research areas in economics. As Morduch (1995) observes, attempts to estimate underlying income processes that neglect this form of insurance will underestimate the true extent of rick in the underlying income process and the degree of insurance used by households. In the Morduch model, such low variance is only available at the expense of lower mean income. If a large proportion of households in the economy use income smoothing[4], economy-wide mean income will almost surely be lower than it would be under perfect capital markets, where each of these households would choose the income earning activity with the highest mean available to them. Thus income smoothing may be a source of substantial inefficiencies on the production side of such economies (Morduch, 1995).

Rosenzweig and Binswanger (1993, page 58) observe that when income smoothing is common among households at the lower end of the wealth distribution, not only "Average incomes are […] lower, but income inequality is exacerbated". This is because income smoothing among households which are already at the lower portion of the distribution of income implies that these households earn lower mean incomes than they would have, if they were not trying to reduce their exposure to income risk.

Within households and over time, such inequality may be exacerbated because the choice of low-risk but low-return income streams may also affect the ability of households to save and accumulate assets. Carter and Barrett (2006) present a framework where there is a threshold value of assets, below which the rate of return per unit of assets decreases. This non-convexity in the payoff to assets creates a poverty trap – households with small asset holdings receive low returns on average and are therefore unable to save enough to accumulate sufficient assets to access high returns. Thus the inability to smooth consumption in response to shocks to income and the consequent need to smooth income itself can be a cause of persistent poverty and poverty traps (Dercon, 2006).

Much of the existing empirical evidence on income smoothing, guided by Morduch's (1994) theoretical model, which assumes that access to credit markets is the only dimension along which household insurance is constrained, has searched for this type of

[4] The argument assumes that the resultant income streams are independently distributed across households.

insurance strategy among the relatively poor and vulnerable, who are more likely to be excluded from credit markets. Morduch (1990), using household-level data on Indian villages gathered by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) between 1975 and 1984, finds that asset-poor households, whose consumption is most vulnerable to income shocks, devote a greater proportion of their land to safer, but lower yielding, traditional varieties of crops than richer households.

Also using the ICRISAT data on rural India, Rosenzweig and Binswanger (1993) find that Indian households in lower wealth quartiles make systematically different decisions with respect to production inputs than households in higher quartiles. By adjusting the ratios of farm inputs such as irrigated and unirrigated land, draught animals, milk animals, other animals, farm implements, modern machinery, liquid capital and consumption assets; households achieve profit functions which differ in their responsiveness to weather shocks. Poorer households in areas with high rainfall variability employ asset portfolios whose returns are more insulated against this source of risk than their better-off counterparts. But as the authors note, this safety comes at a cost.

> “A one-standard-deviation increase in the onset date coefficient of variation lowers average profits by 264 rupees or 4.5% [...] while for farmers with holdings below the 25th percentile, average profits are lowered by 555 rupees. This cost of risk reduction represents 35% of average profits for the lowest quartile of farmers.”

Bliss and Stern (1982, cited in Morduch 1994) find that Indian farmers could greatly increase expected profits by increasing the application of fertilizer, which sells at a price that is less than one-third its marginal product. They conclude that farmers abstain from using this productive input with a view to minimizing risk – in the event that the harvest fails, farmers reduce their losses by not having borne the cost of fertilizer, a clear example of households reducing income variability at the expense of its mean.

The literature has shed some light on the mechanisms households use to smooth their income. Rosenzweig and Binswanger (1993) show that among their sample of Indian farmers, households adjust the ratios of farm inputs, to insulate their profits against weather variability.

Also using the ICRISAT data from India, Kochar (1999) finds that households protect consumption by diverting labour from farm employment to off-farm employment when faced with a crop shock. Kochar presents evidence that this mechanism is used to protect consumption against crop shocks, but does not compare the extent to which income itself is insulated across households with differing levels of welfare, as I do here.

In section 2.4, I check to see if different types of labour contract account for differences in income variability. One such contract is what Morduch (1995: 100) calls 'Perhaps the most extreme case of income smoothing', in which Bardhan (1983) (cited in Morduch, 1995) analyses 'tied labour,' where the labourer accepts a very low wage to avoid bearing any risk. More recently, Gutierrez (2014) finds that increased prevalence of salaried jobs in Mexico plays an important role in insuring workers against shocks to productivity, such as illness. In this paper, I find that contract types that are more common among better-off households, are also better insured against certain types of shock.

In the present paper, I add a number of potential income smoothing mechanisms to those considered in previous papers. The literature on income smoothing has focused almost entirely on agricultural households. In the Thai data, a greater variety of households are represented, so that the characteristics which enable households to smooth income are also potentially more varied. In Section 2.4.5 of this paper I attempt to uncover the characteristics of better-off households in rural Thailand that are associated with their ability to protect income against covariate shocks. I check if different characteristics of the household including the level of education, the type of wage contract for their primary occupation, sex, birth cohort, whether or not they have multiple occupations and the employment rate among household members, can be statistically linked to the extent to which household income is insulated against covariate shocks. I find that the level of education of the household head, the type of wage contract in the household head's primary occupation, whether or not household is dependent on agriculture, and the cohort of birth of the household head are systematically related to the ability of households to insure their income streams against a shock.

These results confirm and extend the findings of Dercon and Krishnan (1996) who explicitly studied constraints to occupation choice and the implications of such constraints for the type of risk faced by the poor and vulnerable in rural Ethiopia and Tanzania. They found that poorer households engaged in activities which had lower entry costs such as collecting firewood and dung-cakes, making charcoal and working as day labourers. Entry into high return activities such as cattle rearing and shopkeeping were restricted to richer households, even though the poor stated a desire to enter these activities. In the Thai context, I find that households across the distribution of income are equally likely to be business owners, in contrast to these results from Ethiopia and Tanzania. Dercon and Krishanan (1996) also find that a lack of education restricts the ability of relatively poor households to gain low risk salaried employment. I confirm that relatively poor Thai households are similarly constrained.

2.2 A Theoretical Model of Income Smoothing

In this section I explore the implications of heterogeneity in the cost at which risk-free income is available to households within the context of Morduch's (1994) theoretical model of income smoothing. I find that when households are required to sacrifice a smaller proportion of mean risky income to secure risk-free income, they use a greater degree of income smoothing than they otherwise would. If I assume that relatively well-off households have a greater choice of income streams that offer more insurance possibilities (such as access to jobs which pay a monthly wage), they may depend more heavily on income smoothing to satisfy their insurance needs than their worse-off counterparts. It then becomes an empirical issue, whether relatively rich households are more or less likely to use income smoothing than poorer counterparts.

I now recall the key elements of Morduch's (1994) theoretical model of income smoothing. The model describes an agricultural household that is set in two periods. The household may choose between an absolutely safe activity that yields a return s and a risky activity that yields a return of h when in state $1H$ and a negative return l in state $1L$. Morduch assumes that each state occurs with probability 0.5 and that $(h+l)/2 \geq s > 0$ so that on average the risky activity is more remunerative than the riskless one. All uncertainty is resolved in the first period and income in the second period is known with certainty to be Y_2. The household chooses $\theta \in [0,1]$, the proportion of its portfolio that is dedicated to the safe income earning option. Thus when state $1H$ is realized, income $Y_{1H} = \theta s + (1-\theta) h$; when state $1L$ is realized, $Y1L = \theta s + (1-\theta)\, l$. Given state $1L$ in the first period, the household's utility maximization problem may be written as:

$$Max\ U(C_{1L}) + U(C_{2L})$$

$$\text{s.t. } C_{2L} \leq (Y_{1L} - C_{1L})(1+r) + Y_2 \qquad (2.1)$$

When there are no credit constraints, consumption will be optimized where the marginal utilities are equal so that $U'(C_{1L}) = (1+r)\ U'(C_{2L})$. Let U_L denote lifetime utility in the low-income state and U_H denote lifetime utility in the high-income state, and b^*_L denote the optimum level of borrowing in this case. As Morduch notes, this result depends on there being no limit to borrowing.

Supposing that the borrowing constraint takes the form that a household can only borrow a fraction α of Y_2 in period 1, so that the constraint binds whenever state 1_L is realized and $b^*_L < \alpha Y_2$. Let the welfare loss associated with the binding liquidity constraint be denoted by δ, so that

$$\delta = U_L - [U(Y_{1L} + \alpha Y_2) + U(Y_2 - \alpha Y_2[1+r])] > 0, \qquad (2.2)$$

recalling that U_L is lifetime utility when there is no borrowing constraint, the first term in the square brackets is first period utility in state $1L$, where the household has exhausted all borrowing opportunities and the second term is the utility in the second period, where the household consumes all of Y_2 that was not borrowed to finance first-period consumption. Assuming, as Morduch does, that the borrowing constraint does not bind in state $1H$, the household chooses the proportion β to maximize lifetime utility:

$$\mathrm{Max}_\beta\ 0.5\ U_H + 0.5(U_L - \delta) \qquad (2.3)$$

The first order condition for this problem is:

$$0.5(s-h)U'_H + 0.5(s-l)U'_L - 0.5\ (s-l)\ \delta = 0 \qquad (2.4)$$

Rearranging and simplifying yields:

$$U_H' / [U_L' - \delta'] = (l-s)/(s-h) \qquad (2.5)$$

As Morduch notes, the borrowing constraint is captured explicitly in this equation by the δ' expression in the denominator of the left hand side. An increase in y_1, decreases the utility lost due to the liquidity constraint, so that δ' is negative. Its presence therefore causes the denominator on the left hand side to be larger than it would otherwise be. Therefore, the numerator must also be larger, for equality with the right hand side to hold. Thus marginal lifetime utility in the good state of the world is higher than it would otherwise be, so lifetime utility must be lower. This happens only if more resources are allocated to the less risky asset with payoff $s<h$. Thus $d\theta/d\alpha <0$, that is, when the borrowing constraint is tightened, a household will dedicate more of its resources to safe activities.

This yields the result that when constraints leave households unable to smooth consumption in the face of income fluctuations, households respond by choosing lower risk, and potentially lower return income streams.

Given the plausible assumption that relatively well-off households are less likely to be liquidity constrained, this model of income smoothing does not account for the pattern observed in figures 2.1 – 2.3 where better-off households use more income smoothing. This may be because better-off households have access to income streams that offer greater insurance possibilities than their worse-off counterparts. This can be formalized in Morduch’s model, either by assuming that better off households have access to a higher s than others, or by assuming they have higher l. The effects of these two, alternative explanations are discussed in turn.

The effect of heterogeneity in s on the optimum level of income smoothing can be easily understood with reference to equation 2.5. When s is high, so that household incomes streams offer insurance at a relatively low cost, the numerator on the right hand side of equation 2.5 is a relatively large, negative number and the denominator is a relatively small, negative number. The right hand side then, is positive and relatively large. Holding liquidity constraints constant, households that can access risk-free income streams at a relatively low cost, will have high marginal utility in the good state of the world. Thus total utility must be relatively low, which can only happen in the good state of the world if the household generates a relatively large proportion of their income from the riskless stream.

Alternatively, if we assume that better-off households have higher *l*, the distance *h*-*l* decreases. Thus under this assumption, the income streams of relatively well off households are inherently better insured than their worse off counterparts. A relatively high *l* causes the right hand side of equation 2.5 to be a relatively small, negative number so that marginal utility in the high state of the world is relatively low (holding liquidity constraints constant), which occurs when utility in the high state is relatively high. Thus under this assumption, these households use less income smoothing, because their income streams are inherently smoother than their worse off counterparts, so they require less total insurance.

Under either theoretical explanation, the degree to which the income streams of households are insured against risk at different welfare levels is subject to two conflicting forces: on the one hand, poorer households are more likely to be liquidity constrained, decreasing their ability to consumption smooth; but on the other, their income streams may be less suited to providing insurance, increasing their need to consumption smooth.

This observation – that households may differ not only in the extent to which they are liquidity constrained but also in the degree to which their income streams are suited to providing insurance – potentially reconciles the theory of income smoothing with the insurance patterns observed in figures 2.1-2.3. These figures suggest that relatively well-off households appear to rely more heavily on low-risk income streams for their insurance needs than on consumption smoothing. But this explanation hinges on relatively well-off households differing from their poorer counterparts in ways that enable them to more easily access low-risk income. In so far as these differences are observable, the theory presents a number of testable hypotheses. First, does the degree to which household income is insured differ across the income distribution? Second, are differences in the observable characteristics of households empirically linked to these differences in insurance behaviour? In the following sections of this paper I test these hypotheses using household and village level data gathered by the Townsend Thai Project (Townsend, 2011). The next section of the paper describes the data.

2.3 The Data

The analysis in this chapter will be done on data gathered by the Townsend Thai Project from the 64 villages that are interviewed in every year from 1997 to 2011, i.e. the panel of 960 households. Within this subsample, the year on year attrition rate reaches a maximum value of 6%, but is usually closer to 3%. These households are replaced with households from the same village to keep the number of households in each cross section at 960. Because I am interested in the dynamics of income and consumption, I focus on the balanced panel of 609 households so as not to conflate these dynamics with the entry and exit of households to and from the panel.

Household composition accounts for a significant proportion of the variation in the household income and consumption figures. For this reason, the variables presented here are all per adult-equivalents, using the 'old OECD scale' (OECD, 1982). Using this scale, the head of household receives a weight of 1, additional adults (16 and over) receive weights of 0.7, while children (under 16) are weighted by 0.5.

I have adjusted for inflation using Bank of Thailand figures for headline consumer price inflation. All figures are inflated to 2011 Baht.

The key outcome variables discussed in this paper are income and consumption, which are defined as follows. Net income is the difference between a household's gross income and agricultural and business expenses over the last 12 months. In the data, the interviewer notes their perception of the accuracy of this figure, and if inaccurate, they write what they believe to be a more accurate figure, after consulting the respondent. Where appropriate, I have used the latter (presumably more accurate) figure. These numbers are then revalued to allow for inflation and equivalised. The summary statistics for the resulting income data are reported in the first row of table 2.2.

The consumption variable that I use is constructed from two distinct parts of the questionnaire: annual consumption items and monthly consumption items. The annual consumption items include spending on household and vehicle repairs, education, clothing and eating outside the home. Monthly consumption items include various food items, gasoline, alcohol and tobacco. Information on these is elicited on the basis of a typical month in the last year. To make these figures comparable to expenditure on annual consumption goods and income, I multiply these values by 12. These figures are then inflated to 2011 Baht and equivalised as described above, and the resulting summary statistics are presented in the second row of table 2.2. For some of the analysis, I use the value of household assets as an explanatory variable. These include the amount of money spent over the last year on the purchase of a television, refrigerator, motorcycle, boat, bicycle, and the like. The questions used to solicit these values are consistent over the duration of the survey with one exception. From 1998 to 2011 there is a question specifically on household assets purchased over the last year. In 1997 (the baseline year), however, the question is on the stock of all assets owned by the household, with a sub-question on when each asset was purchased. For the sake of comparability, in the 1997 data I only keep the value of assets purchased in the last year. Following Alem and Townsend (2014) I impose an annual depreciation rate of 10% on all household assets. I also account for inflation in asset values for new purchases over the duration of the panel using Bank of Thailand data. However, some assets are very old, and for historical asset purchases (those bought before 1997) I use an annual inflation rate of 3%, the sample average.

Table 2.2: Summary Statistics of Household Data				
Variable	*Mean*	*Std. Dev.*	*Min*	*Max*
Income	49,510	71,087	175	1,956,726
Consumption	21,128	22,817	466	667,738
Assets	83,494	160,960	0	2,870,177
Birth Year	1948	13	1903	1988
Sex	0.300	0.458	0	1
Second Occupation	0.640	0.480	0	1
Involved in Agriculture	0.904	0.295	0	1
Number of household members in employment	2.645	1.247	0	10
Contract Type				
Business Owner	0.639	0.480	0	1
Unpaid Family Worker	0.037	0.189	0	1
Daily Wage	0.140	0.348	0	1
Monthly Wage	0.031	0.173	0	1
Piece Rates	0.015	0.123	0	1
Government Work	0.021	0.144	0	1
Other	0.003	0.054	0	1
Maximum Schooling:				
None	0.109	0.312	0	1
Kindergarten	0.048	0.214	0	1
Primary	0.762	0.426	0	1
Secondary	0.053	0.224	0	1
Vocational	0.011	0.104	0	1
University	0.017	0.128	0	1
Other	0.000	0.010	0	1

* *31.5147 Thai Baht = 1 U.S. Dollar on 31st Dec. 2011 (source: exchangerates.org.uk)*

These data are combined with the stock of business assets and agricultural assets to compute the stock of total assets the household owns. The summary statistics for these asset holdings are also presented in table 2.2. I note here that whenever assets are used as a conditioning variable in a regression with income and consumption, I use the lagged value of assets to address the obvious endogeneity problem.

The remaining rows of table 2.2 present summary statistics on other variables of interest. The mean year of birth of a household head is 1948, while the oldest was born in 1903 and the most recent in 1988. Since most of the analysis is performed on a balanced panel of households, it is important to note that over the duration of the panel heads of households are, on average getting older. Women head 30% of the households and 64% of household heads have more than one occupation. By far the most common type of employment is 'business owner', followed by those who are in a contract that pays them a daily wage. 90% of households are involved in some kind of agricultural activity. The most common education level of household heads is to the primary level, accounting for 76% of the sample while the next most common is 'none'. The mean number of employed

Table 2.3: Observable Characteristics of Households with Below and Above Median Permanent Income

Level of permanent income:	*Below median*	*Above median*
Education level of head (%)		
None	16.68	4.59
Kindergarden	0	0.11
Primary	80.90	81.03
Secondary	1.97	9.06
Vocational	0.30	1.78
University	0.12	3.40
Other	0.00	0.02
Primary contract type of head (%)		
Government work	0.24	4.03
Other monthly wages	1.15	4.92
Daily wages	17.86	10.50
Piece rates	1.32	1.80
Business owner	59.84	68.02
Other	0.16	0.35
Decade of birth of head (%)		
1930s	26.05	14.45
1940s	23.51	18.51
1950s	22.65	37.22
1960s	13.49	21.62
Other characteristics		
More than one breadwinner (%)	62.68	64.74
Head holds multiple jobs (%)	58.60	71.27
Headed by women (%)	36.14	24.54
Involvement in agriculture (%)	90.30	91.20
Lagged log of assets	9.73	11.10

people in a household is 2.6, and there is substantial variation in this number. The maximum number of employed people in one household is 10, but households who have 5 for fewer employed members account more than 99% of the data. The median number of employed people per household is 2, accounting for almost 42% of the data.

Table 2.3 presents some descriptive statistics which assess whether or not these observable characteristics differ across the distribution of income. To construct this table, I use mean log of equivalised real consumption of households over the duration of the panel as a proxy for permanent income (as was done in figures 2.1 - 2.3), and divide households into two groups based on whether or not their mean consumption is above or below the sample median. The table shows that households which have permanent income higher than the sample median differ in many observable characteristics from those which do not. Such households are more likely to be in government work or other jobs that pay steady monthly wages, as opposed to varying daily wages or unpaid family work. They are likely to be better educated than their worse-off counterparts. They are less likely to be dependent on a sole breadwinner. The head of household is more likely to have more than one source of income. On average, they are headed by younger people. Unsurprisingly, they have larger asset holdings than their worse off counterparts. In the analysis that follows I will test if any of these differences in observable characteristics can be statistically linked to the ability of relatively well off households to insure their income against shocks.

Potential heterogeneity in risk preferences could in principle pose a threat to the central thesis of this chapter – the proposition that richer households have privileged access to better insured income streams. If poorer households happened to have a greater risk tolerance, they may select relatively risky income streams even if they were afforded the same insurance possibilities as their richer counterparts. This narrative however, is not consistent with the descriptive statistics in table 2.3. Households headed by the elderly tend to be more risk averse than their younger counterparts. The fact that households headed by people who were born in the 1930s and 1940s are much more common among the relatively poor, whereas those headed by people born in the 1950s and 1960s are more common among the relatively rich, suggests that if anything poorer households that are on average headed by older people, are likely to be more risk averse than the richer ones. In section 3.5.2 of the next chapter, I will also present evidence that poorer households are more likely to exhibit higher fertility, another demographic factor that is generally associated with high risk aversion. These characteristics suggest that poorer households are unlikely to be more risk averse than their richer counterparts. Furthermore, using the monthly series of the Townsend Thai Data, Chiappori et al. (2014) estimate household risk preferences in rural Thailand (albeit on a distinct, but similar set of villages to those covered by the annual survey which I use here) using a full risk-sharing model and they find that though there is substantial evidence of heterogeneity in risk preferences, this heterogeneity is unrelated to household wealth, or indeed any other characteristic. Chiappori et al. (2013) use an alternative methodology, which exploits variation in the different income portfolios chosen by different households, to identify risk preferences, also in the monthly series of the Townsend Thai data. Their results are qualitatively similar to those of the preceding paper, in that they also find no evidence that risk preferences are correlated with demographic variables or household wealth. Taken together, these factors lead me to conclude that heterogeneity in risk preferences is unlikely to pose a credible threat to the key contributions of this paper.

2.3.1 Measures of Exogenous 'Shocks' to Household Income

To identify differences in the degree to which the income streams of different households are insured, I require variables which are likely to exogenously affect household productivity. The literature has met this requirement by using observations of involuntary job loss (Cochrane 1991), adverse weather (Rosenzweig and Binswanger, 1993; Dercon 2004; Kochar 1999), or adverse health outcomes (Cochrane, 1991; Gutierrez, 2014). Unfortunately, not all of these shocks are identified in the Townsend Thai Project. I now turn to identifying exogenous variation in household income from those potential shocks which are measured in the data.

The Townsend Thai Project does measure information on involuntary job loss. However, because of the agricultural and entrepreneurial nature of households in the Thai sample, involuntary job loss in the form of layoffs and factory closures is very rare: in the twelve years for which this information is available, there were only two reported layoffs, and four jobs lost due to factory closures.

The Townsend Thai data also collects information on the occurrence of adverse weather and communicable disease, though these are measured at the village level, rather than the household level. These data are collected from interviews with a key informant, typically the village headman. They were not gathered in the first year of the survey (1997), but are available for all subsequent years (1998-2011)[5].

Data are available on the number of households in each village affected by crop disease, drought, flood, polluted drinking water, polluted irrigation water, soil erosion, storm, HIV/AIDS, cholera, hemorrhagic fever, malaria, other disease, smallpox and typhoid. Table 2.4 presents some summary statistics of these variables.

[5] Nonetheless, information on household characteristics in 1997 remains useful because some of the covariates are potentially endogenous to adverse weather and health outcomes, which leads me to use their lagged values as instruments in the empirical work which follows.

My goal is to identify exogenous shocks to household income. Because the data on these potential shocks are measured at the village level, the best I can do is to identify 'covariate shocks' to household income. For a village level occurrence to serve as a useful 'covariate shock', it must satisfy two properties:

1. When the shock materializes, it must be widespread enough to be justifiably considered 'covariate'
2. It must exhibit enough variation over time within each village as to be considered a 'shock'.

The first column of table 2.4 reports that drought is the most wide-spread among these potential shocks, affecting on average 2,082 households a year. Crop disease and flood are also reasonably wide-spread, affecting on average 1,194 and 452 households, respectively. These are unconditional averages. To understand if a potential shock satisfies point 1 above, I calculate the average proportion of village households affected by each potential shock, given the occurrence of that shock in that village in that year.

These are presented in the second column of table 2.4. These figures show that the average outbreak of crop disease, drought, flood, polluted drinking water, polluted irrigation water or soil erosion, affect a substantial proportion of households. Among these the average flood is the least widespread, still affecting 28% of households in a village. Therefore, I consider these potential shocks, as being sufficiently 'covariate' to use in my empirical analysis.

Table 2.4: Summary Statistics of Village-level Data					
Shock	*Average number of households affected per year*	*Minimum households affected in one year*	*Maximum households affected in one year*	*Mean proportion of villages affected, given occurrence of shock*	*Measure of variation over time within villages*
Crop disease	1194.07	238	2,149	0.351	0.465
Drought	2081.5	492	6,396	0.519	0.492
Flood	451.93	90	1,345	0.281	0.345
Polluted drinking water	54.86	0	256	0.590	0.039
Polluted irrigation water	87.21	0	398	0.510	0.061
Soil erosion	199.71	0	1,255	0.295	0.104
Storm	103.07	24	180	0.053	0.327
HIV/AIDS	31.71	17	54	0.018	0.295
Cholera	16.79	0	110	0.054	0.087
Haemorrhagic Fever	98.79	23	407	0.028	0.430
Malaria	21.64	0	99	0.037	0.122
Other disease	61.57	0	218	0.052	0.277
Smallpox	0.43	0	3	0.007	0.013
Typhoid	3.93	0	14	0.019	0.075

The rest of column 2 of table 2.4 illustrates that storms and the average outbreaks of HIV/AIDS, cholera, haemorrhagic fever, malaria, other disease, smallpox and typhoid, affect a relatively small proportion of village households. Cholera is the most widespread among these potential shocks, yet the average outbreak affects only 5.4% of households in a village. This prevents me from justifiably using any of these occurrences as ‘covariate’ shocks.

The third and fourth columns of the table present the minimum and maximum number of households affected by each potential shock in any year. While these numbers are illustrative summary statistics, they are not sufficient to determine the extent of variation over time in village exposure to the remaining potential shocks. For example, if some villages were persistently affected by soil erosion, but others were never affected, then soil erosion would be more appropriately classified as a time invariant village characteristic, rather than an exogenous shock.

To measure the magnitude of variation within villages over time exhibited by each of these potential shocks, for each of the 64 villages, I construct a dummy variable that takes on a value of 1 if a shock occurs in a given year and 0 otherwise. I measure the variability in each potential shock over time, within each village by computing the standard deviation of this dummy variable for each of the 64 villages. Finally, I compute the mean of these standard deviations, over the 64 villages, and the resulting number is presented in the 6th column of table 2.4. The widespread shocks identified above, are readily classified into two groups: the materialization of crop disease, drought and flood exhibit a relatively large amount of variation over time within the average village; whereas the materialization of polluted drinking water, polluted irrigation water and soil erosion do not.

Guided by these characteristics of the data, I retain drought, crop disease and flood as potential covariate shocks to income.

2.4 Estimating the Effect of Shocks on Income

My goal is to test for differences in the degree to which household income is insured against risk. This requires that I identify sources of variability in the income dynamics of different households. To this end, I assume that the dynamics of household income are driven by:

1. A permanent component, which may be a function of 'fixed' household characteristics, such as education and year of birth of the head of household.
2. Village–level characteristics, such as soil fertility and distance to nearest city.
3. A transitory, household-specific component.
4. Village-level shocks, such as the occurrence of a drought.

Thus the observed income of household *i* in village *v* at time *t*, which I denote by y_{ivt} is composed of the following:

$$y_{ivt} \equiv \bar{y}_i + \bar{y}_v + y_{it} + y_{vt} + \varepsilon_{ivt}, \tag{2.6}$$

where $\bar{y}_i$, $\bar{y}_v$, y_{it}, and y_{vt} denote the components described by 1, 2, 3 and 4 above, respectively. Finally, ε_{ivt} is a mean zero error term.

The average level of household i's income, $\bar{y}_i$, is determined by that household's 'innate,' characteristics, such as the year of birth, sex and the level of education of the head of household.

Holding these characteristics constant, the average level of household income may vary systematically with the village in which a household is resident. Some villages may be more fertile than others, increasing returns to agricultural labour; some may be better connected to urban centres, increasing returns to labour employed in services. Such heterogeneity between villages (if it exists) will be a component of permanent household income for all households in that village, and is denoted by $\bar{y}_v$[6]. Section 2.4.1 will model permanent component of household income.

The transitory, household-specific component of income, y_{it}, will include shocks to household income such as involuntary unemployment (Cochrane, 1991) and waves of illness (Cochrane, 1991 and Gutierrez, 2014). Observed y_{it} will also include endogenous household responses to these unanticipated shortfalls to income, such as increases in the labour brought to market by other household members (Mincer, 1962). Furthermore, in practice it is difficult to separate y_{it} from the error term, ε_{ivt}, because of measurement error[7] and unobserved heterogeneity[8].

For these reasons, and given the available data, I focus the analysis on village level transient income, y_{vt}. In section 2.4.3, I extend the model of household income using information from the village level shocks, identified in the section 2.4.2, using lags of potentially endogenous variables as instruments in a two stage least squares (2SLS) estimation strategy. Having adequately identified the effect of a covariate shock on the income of the mean household, in section 2.4.4 I test if the income streams of relatively well-off households are better insured against the covariate shock than their poorer counterparts. All the results for different models of household income are presented in table 2.5.

---

[6] It is of course possible that over time, households respond to heterogeneous payoffs by migrating to villages that offer better employment opportunities, driving down wages at these destinations, until such differences no longer exist.

[7] If measurement error is independently and identically normally distributed, across households and over time, and 'shocks' are to be used as explanatory variables, these errors decrease the efficiency of estimates. If, however they are systematically related to any variable of interest (for example, if richer households are more likely to underreport their income) then estimated coefficients may be biased and inconsistent.

[8] Insofar as such heterogeneity is time invariant, one could take the first difference of panel data during analysis to remove the effect of this potential bias. But if the effects of heterogeneity varied over time, for example if poorer households were more likely to receive 'gifts' during lean times than their wealthier counterparts, y_{it} and ε_{ivt} would be correlated.

Table 2.5: Models of Household Income

	(1)	(2)	(3)	(4)	(5)	(6)
Model	*'Permanent' household income*	*Baseline model of the effect of drought on income*	*The effect of drought after accounting for endogeneity*	*Permeant income moderates the effect of drought*	*Households with above median permanent income*	*Households with below median permanent income*
Estimation Strategy	*OLS*	*OLS*	*2SLS*	*2SLS*	*2SLS*	*2SLS*
Drought		-0.000777**	-0.00101***	-0.00118***	-0.000470	-0.00168***
		(-2.45)	(-3.12)	(-4.26)	(-1.17)	(-4.06)
Drought × permanent income				0.00260***		
				(4.11)		
Head has more than one job		0.0751**	0.0822	0.0637	-0.0941	0.111
		(2.44)	(1.26)	(1.12)	(-0.90)	(1.33)
Employment rate within household		0.367***	0.396***	0.327***	0.522***	0.290**
		(4.86)	(3.63)	(3.74)	(3.37)	(2.35)
Less than primary education	-0.181***	-0.160**	-0.150**	-0.0720	-0.0959	-0.101*
	(-2.75)	(-2.53)	(-2.42)	(-1.44)	(-0.64)	(-1.83)
More than primary education	0.400***	0.402***	0.401***	0.147**	0.304***	-0.00502
	(4.88)	(4.88)	(4.91)	(2.48)	(3.26)	(-0.05)
Sex of head	0.0138	0.0482	0.0563	0.0241	-0.0399	0.0656
	(0.38)	(1.32)	(1.63)	(0.84)	(-0.70)	(1.45)
Lagged log assets	0.149***	0.145***	0.144***	0.0306***	0.187***	0.0387**
	(12.71)	(12.19)	(11.54)	(2.82)	(8.31)	(2.46)
Monthly wages	0.572***	0.587***	0.607***	0.399***	0.487***	0.293*
	(6.27)	(6.36)	(6.47)	(7.66)	(5.36)	(1.78)
Household size	-0.0626***	-0.0470***	-0.0413***	-0.0248**	-0.0253	-0.0414**
	(-6.26)	(-4.10)	(-3.30)	(-2.24)	(-1.41)	(-2.56)
Involvement in agriculture	0.0433	-0.0612	-0.0583	-0.0597	-0.108	0.0366
	(0.77)	(-1.06)	(-0.92)	(-1.08)	(-1.10)	(0.55)
Time	0.0725***	0.0691***	0.0753***	0.0750***	0.0730***	0.0770***
	(19.77)	(17.68)	(18.70)	(19.18)	(15.09)	(14.89)
Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Permanent income				0.848***		
				(15.07)		
Constant	9.299***	9.135***	7.888***	9.392***	7.854***	8.963***
	(30.96)	(33.10)	(28.81)	(41.25)	(26.83)	(35.09)
N	8230	8229	7653	7653	3943	3710

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.4.1 Modelling the 'Permanent' Component of Household Income

To model the 'permanent' component of household income I estimate:

$$y_{iv} = \alpha_i + \boldsymbol{\beta}'\boldsymbol{X_i} + \gamma V + \tau T + e_{iv} \qquad (2.7)$$

where y_{iv} is the average over time of household income from 1997 to 2011. $\boldsymbol{X_i}$ is a vector of innate household characteristics and V is a matrix of village dummies, T is a time trend and e is a mean zero error term. The parameters that will be estimated are α, $\boldsymbol{\beta}$, γ and τ. The results of this regression are presented in the first column of table 2.5[9].

Equivalised real income is on average, 16.5%[10] lower in households that are headed by people who have not completed a primary education, than in households headed by people who completed only primary school. Those who have attained an educational qualification higher than the primary level (for example, secondary, vocational or a university degree) on average receive equivalised real income that is 49% more than those who have completed only a primary education. Households headed by people whose primary occupation pays a monthly wage (including government employees) on average earn 77.1% more real income per adult equivalent than those which are headed by people who do not receive monthly wages. On average, an increase in household size of one person, is associated with a fall in real equivalised income of 6.1%. Real incomes are growing at the rate of 7% per annum over the sample period. There are no statistically significant differences in income levels between households that are headed by men and women. Nor is there a statistically significant difference between the level of income enjoyed by households which are and are not involved in agriculture.

The 63 village fixed effects are jointly highly significant. A test of the null hypothesis that the coefficients of all the village fixed effects are jointly equal to zero yields an F(63,

[9] Standard errors are clustered at the village level, since the variables include a mix of observations at the household and village level.

[10] exp[-0.181]-1 = -0.165

1851) value of 16.99[11]. Therefore, I strongly reject the null hypothesis, in favour of the alternative that the level of average household income varies systematically from village to village. For the sake of brevity, the coefficients of the 63 village fixed effects are not reported here. Nonetheless, they are economically as well as statistically significant – even after controlling for 'fixed' household characteristics (as I have done here), average real income per adult equivalent in the wealthiest village is 73.3% higher than the sample mean, whereas that in the most deprived village is 38.2% lower than the sample mean.

The remaining columns of table 2.5 will present models of household income which identify the effect of a covariate shock on the income streams of different households. Before I can construct these models, I must first identify a covariate shock to household income.

2.4.2 Identifying a Covariate Shock to Transient Income

Section 2.3 concluded that crop disease, drought and flood be retained as potential shocks to village level income. I now identify if any of these potential shocks are associated with significantly low transient household income.

Equation 2.7 modelled the permanent components of household income. I now work with the residuals from that model, which contain information on transient income at the household level, y_{it}; transient income at the village level, y_{vt}; and an error term, ε_{ivt}. I assume that village level shocks are mutually uncorrelated with the error term, i.e. that:

$$\text{Cov}\,(y_{vt}, \varepsilon_{ivt}) = 0, \qquad (2.8)$$

In table 2.6 I test the null hypothesis that mean transient income (the residual from equation 2.7) is equal to zero when villages report instances crop disease, drought or flood

[11] $\frac{(RSS_R - RSS_U) \div q}{RSS_U \div (n-k)} = \frac{(5282.25-4469.24)\div 63}{4469.24 \div (8230-79)} = 16.99$

against the alternative that it is not. I find that drought is associated with significantly lower levels of transient income. For crop disease and flood, I fail to reject the null hypothesis that mean transient income is the same whether or not villages experience these events.

Table 2.6: Effect of Potential Shocks on Transient Income

Potential shock	*Difference in residual (given shock – without shock)*	*t-statistic*
Crop disease	0.027	1.554
Drought	-0.046	-2.700
Flood	0.000	0.012

Given these findings I conclude that of the three potential 'covariate shocks' shortlisted in section 2.3, drought is the only one which is associated with significantly lower levels of transient income. Dercon (2004) demonstrated that drought was a very persistent covariate shock to household income in rural Ethiopia. Rosenzweig and Binswanger (1993) found that rainfall variability in a semi-arid setting (which captures the effect of drought) was strongly associated with household insurance behaviour. In accordance with this literature, and guided by the properties of the data at hand, I now use drought as a covariate shock to village level income.

2.4.3 Extending the Model of Income to Include Transient Components

I can now construct a model of household income, including its time varying components, with the goal of identifying differences in the extent to which the income streams of different households are insured against covariate shocks. The number of household members in employment and the number of jobs held by the head of household were not included as explanatory variables in the model of 'permanent' household income presented above, because the literature has documented evidence that households adjust these variables over time (Mincer, 1962; and Kochar, 1999 are examples that will be discussed below). Now, I introduce these variables into my model of household income, and also include my primary variable of interest, the proportion of households in the village a household is resident in which are affected by a drought in a given year, d_{vt}. Thus I estimate the equation:

$$y_{ivt} = \alpha_i + \boldsymbol{\beta}'\boldsymbol{X_{it}} + \delta d_{vt} + \gamma V + \tau T + e_{ivt}\,, \tag{2.9}$$

where δ is a parameter to be estimated and all other variables are as they have been defined before. The second column of table 2.5 presents the results of this regression. First, I note that introducing these transient components of household income has no significant effect on the parameter estimates of the 'permanent' component. The coefficient on the variable for the proportion of households in each village affected by drought has the correct sign, and a reasonable magnitude: at the sample mean, a ten percentage point increase in the proportion of households affected by drought decreases the real, equivalised income of the mean household in that village by 0.7%. Households which are headed by people who have more than one job enjoy on average 7.8% more income per adult equivalent than households that are not. A ten percentage point increase in the employment rate within the household is associated with an increase in income per adult equivalent of 3.7%.

Importantly, households are not passive recipients of an income such as drought, as alluded to above. There is a large literature on 'the added worker effect', going back at least as far as Mincer (1962), who documents that the labour supply of married women in the U.S. is greater when their husbands experience spells of unemployment. Households also respond to temporarily low returns to labour in one market by selling their labour in another, as documented by Kochar (1999), who finds that when farm earnings in rural India are hit by an adverse shock, households divert labour to off-farm activities. So while the time varying component of household income may be affected by exogenous shocks, this effect is potentially confounded in the data by endogenous household responses to the shocks. Failing to adequately account for this endogeneity may cause OLS estimates to be biased and inconsistent. In particular, if the head of household takes on a non-agricultural job in the event of drought, the OLS estimates of the effect of drought on income would be biased towards zero. A similar argument applies to the number of household members in employment.

To address this problem, I use lags of the endogenous variables as instruments in a 2SLS estimation strategy. While these lags are likely to be highly correlated with the contemporaneous values, they are not likely to vary in response to unpredictable shocks to transient income that have not yet materialized. Indeed, I find that the first and second lags of both instruments are highly relevant. When I perform two, separate regressions with the employment rate within the household as the dependent variable, and its first and second lags as independent variables, the regressions yield R-squared values of 0.498 and 0.330, respectively. Similar regressions performed on the variable for whether or not the head of household has multiple jobs yield R-squared values of 0.352 and 0.238. A Sargan test of the null hypothesis that the instruments are jointly uncorrelated with the error term fails to reject the null for both the employment rate within the household (p-value: 0.558) and for the whether the head holds multiple jobs (p-value: 0.279). Accordingly, I proceed with 2SLS, using two lags of the potentially endogenous variables to instrument for their current values.

In the first stage I use OLS to estimate the following equations:

$$\boldsymbol{x^1}_{it} = \boldsymbol{c} + \boldsymbol{bx^1}_{it-l} + \boldsymbol{e}_{\boldsymbol{it}} \qquad l \in \{1,2\} \qquad (2.10)$$

where l is either the first or second lag and x^1 denotes the subset of household characteristics that are likely to vary in response to shocks. These are the number of household members in employment and the number of jobs held by the head of household. I then compute the predicted values from this estimation:

$$\widehat{\boldsymbol{x}}'_{it} = \hat{\boldsymbol{c}} + \widehat{\boldsymbol{b}}\boldsymbol{x}^{\boldsymbol{1}}_{it-l} \tag{2.11}$$

and use them in place of the original variables for the number of household members in employment and the number of jobs held by the head of household in equation 2.9, which thus becomes:

$$y_{ivt} = \alpha_i + \boldsymbol{\beta}'\boldsymbol{X}_{\boldsymbol{it}} + \theta'\widehat{\boldsymbol{x}_{\boldsymbol{it}}} + \delta d_{vt} + \gamma V + \tau T + e_{ivt} \tag{2.12}$$

where the vector $\boldsymbol{X}_{it}$ no longer includes the potentially endogenous variables and θ is a parameter to be estimated.

The third column of table 2.5 presents these 2SLS results. The estimated coefficient of drought on income increases in magnitude, so that a 10 percentage point increase in the number of households affected by drought decreases income per adult equivalent by 1%. This increase in the estimated coefficient when compared that of the second column suggests that heads of household may indeed take on additional jobs or that other household members may bring their labour to market to reduce shortfalls to income brought about by a drought.

2.4.4 Are the Incomes of the Better off more Insured against Drought?

The preceding section estimated the effect of the prevalence of drought on the mean household. I now extend that model to answer the key research question of this paper, namely to test for heterogeneity in the extent to which household income streams are insured against shocks, across the distribution of income.

I adopt a reduced form approach to answering this question with the data at hand by interacting drought with a measure of relative wellbeing and testing whether or not this interaction term is statistically significant. This frames the problem as a moderated relationship (Jaccard and Turrisi, 2003; Aiken and West, 1991), where I test if the effect of drought on income is moderated by relative well-being[12].

To implement this approach, I need to identify a suitable variable that summarizes the relative wellbeing of each household over the 15-year duration of the panel. One candidate is the household's permanent income, which was modelled in equation 2.7. The use of income as an outcome variable obviously rules out this choice. However, under the textbook permanent income hypothesis (Friedman 1957), consumption in each period is equal to the expected value of a household's permanent income. Therefore, the value of consumption at any one point provides a measure, not only of household welfare at that particular point in time, but also of how well-off that household expects to be in the future. As was argued in the introduction to this chapter, (equivalised) consumption is also likely to provide a cleaner measure of relative well-being for households which are at different stages of the lifecycle. It is true that in any one period, a household may misjudge their future prospects, or be temporarily liquidity constrained. There may also be issues of measurement error. Together, these issues imply that consumption in any one period is at best a noisy indicator of permanent income. The Townsend Thai data, contain a balanced panel of 609 households for which I observe consumption data in all 15 periods. Averaging over the observed values of consumption is likely to cause errors in individual periods to cancel each other out, yielding a cleaner signal of relative wellbeing than a single observation. These benefits of employing this measure of relative wellbeing

[12] Quantile regressions are an attractive alternative estimation strategy. However, they are not used as the primary strategy here because existing statistical software packages do not have established routines for two-step quantile estimation which is required to properly to address endogenous regressors. Nonetheless I present results of quantile regressions using lags of endogenous variables in appendix 3.

must be weighed against some potential costs. Households may exhibit heterogeneity in their rates of time preference. Such differences can cause the consumption streams of households to evolve differently even if their income streams are held constant. Thus early in the lifecycle, a relatively patient household may have a low level of consumption compared to a relatively impatient one, even if they both expected to enjoy the same level of lifetime income, and thus relative wellbeing. Similar issues arise if we are to make allowances for heterogeneous risk preferences and risky income streams. Furthermore, while using consumption instead of income and averaging over multiple periods can help address issues arising due to transient shocks, they cannot address the possible presence of permanent shocks. The presence of permanent shocks to household income would imply that household permanent income, and hence relative well-being was changing over the duration of the panel, raising fundamental issues with a time-invariant ranking of households by relative well being. Thus using the average of log real, equivalised consumption as a proxy for household permanent income implicitly makes a number of very strong assumptions about the preferences of households and the type of risk in their income streams. Nonetheless in my opinion, these disadvantages are more than balanced by the advantages listed above and so I proceed with the use of this proxy for household permanent income.

Figure 2.4 plots the values of average of the log of (real, equivalised) consumption and income against one another. There is a strong, positive association between the two. This is confirmed by running a regression with mean, log household consumption over the duration of the panel as the dependent variable, and mean, log household income (and a constant) as the independent variable, which yields a correlation coefficient of 0.588, and a t-statistic of 34.5 (distributed with 607 degrees of freedom). Thus there is evidence of a strong, positive association between income and consumption. The coefficient is significantly different from one (so that the measured propensity to consume out of permanent income is less than unity) most likely because the consumption variable does not include purchases of household assets and durable goods. The R-squared of this regression is also reasonably high, with mean income explaining over 66% of the variation in mean consumption. Supported by this highly significant coefficient and this large R-squared value, I proceed by assuming that the mean of log consumption over the 15-year duration of the panel is an acceptable proxy for the level of household permanent income, and thus the level of wellbeing a household enjoys, the concerns raised earlier notwithstanding. For each of these households, I compute the average level of log consumption, $\bar{c}_i$.

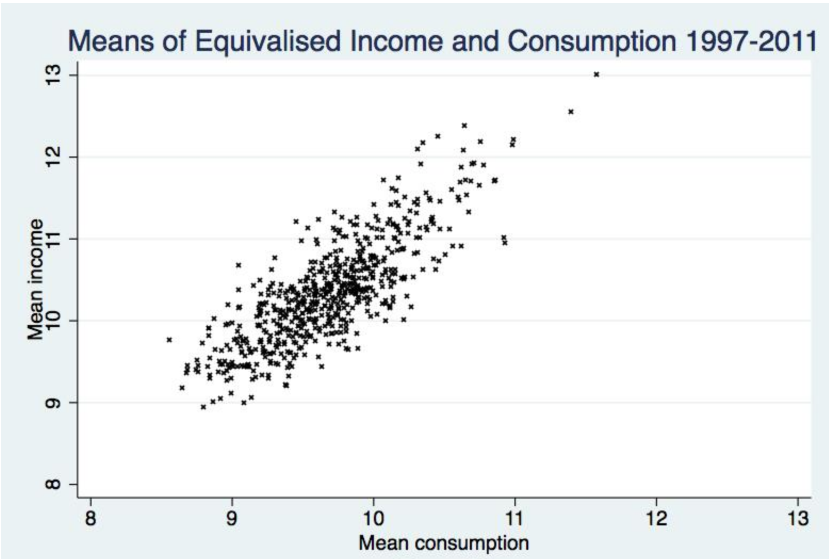


Figure 2.4

If I were to simply use $\bar{c}_i$ in an interaction with d_{vt}, the resulting coefficient on drought would have to be interpreted as being conditioned on a household consuming 1 Thai Baht per year, on average. Following the advice of Jaccard and Turrisi (2003), I subtract the mean value of $\bar{c}_i$ over the entire sample from each household's estimated permanent income. That is, I define the quantity,

$$\bar{c}_{\mu i} = \bar{c}_i - \mu, \tag{2.13}$$

where μ is the average $\bar{c}_i$ observed in the sample. Thus $\bar{c}_{\mu i}$ continues to increase monotonically with permanent income, and therefore remains a valid measure of household i's wellbeing, relative to the sample average. Using $\bar{c}_{\mu i}$ as a proxy for permanent income, I estimate:

$$y_{ivt} = \alpha_i + \boldsymbol{\beta}'\boldsymbol{X_{it}} + \theta'\widehat{\boldsymbol{x_{it}}} + \delta d_{vt} + \chi(d_{vt} \times \bar{c}_{\mu i}) + \gamma V + \tau T + e_{ivt} \quad (2.14)$$

where $\bar{c}_{\mu i}$ is also an element of the vector of household characteristics, $\boldsymbol{X_i}$. Again, I account for the endogeneity of the number of jobs held by the household head and the employment rate within the household by instrumenting for these variables using their first and second lags, as discussed in section 2.4.3. Village fixed effects and a time trend also remain present.

Here, the key parameter of interest is the coefficient of the interaction term, χ. If the income streams of households with higher levels of permanent income are indeed better insured against covariate shocks, χ will be positive and significant.

The fourth column of table 2.5 presents the results. Indeed, χ is positive and significant, so that the adverse effect of drought on income is moderated by high levels of permanent income. The estimated effect of drought at the sample mean is unchanged, while that of mean consumption is large and significant, as expected.

Thus I strongly reject the null hypothesis that the effect of a drought is the same for households with different levels of permanent income, in favour of the alternative that the income streams of richer households are better insured against this shock. Indeed, the magnitude of the coefficient is such that at the sample mean, a one standard deviation increase in mean consumption (0.459 log points) nullifies the impact of drought on income, while a one standard deviation decrease from the sample mean, doubles the effect of drought on income.

These results are corroborated when I split the sample into households with permanent income above or below the sample median and re-estimate equation 2.12. The resulting estimates are presented in column five and six of table 2.5, respectively. Column five shows that there is no statistically significant effect of drought on income for households whose permanent income is above the sample median. Column six, however, confirms that there is a strong negative effect of drought for households with permanent income below the sample median. Indeed, the coefficient in this sub-sample is almost double that estimated at the mean of the entire sample.

Thus I find evidence that relatively well-off households in rural Thailand enjoy income streams that are better insured against this covariate shock. The next section of this paper attempts to identify the characteristics of better-off households that enable them to better insulate their income streams against drought.

2.4.5 The Effect of Drought by Households' Observable Characteristics

The preceding section found evidence that was consistent with the hypothesis that the income streams of better-off households in rural Thailand are better insulated against a covariate shock. Table 2.3 illustrated that households which have levels of permanent income higher than the median differ in many observable characteristics from those which do not. Such households are more likely to be in government work or other jobs that pay steady monthly wages, as opposed to varying daily wages or unpaid family work. They are likely to be better educated than their worse-off counterparts. They are less likely to be dependent on a sole breadwinner. The head of household is more likely to have more than one source of income. On average, they are headed by younger people, who may be better able to adapt to changes in the agricultural markets. They have larger asset holdings

Table 2.7(a): Effect of drought on income by household characteristic							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sub-group	*Whole sample*	*Monthly wage earners*	*More than primary education*	*Born in 1930s*	*Born in 1940s*	*Born in 1950s*	*Born in 1960s*
Head has more than one job	0.0822 (1.26)	-0.470*** (-2.77)	0.346 (0.96)	0.0734 (0.51)	0.283* (1.75)	0.158 (0.97)	-0.0400 (-0.19)
Employment rate in household	0.396*** (3.63)	0.364 (1.37)	0.863** (2.36)	0.715*** (2.91)	0.249 (0.88)	0.278* (1.86)	0.362 (1.43)
Drought	-0.00101*** (-3.12)	-0.000345 (-0.28)	-0.000925 (-0.96)	-0.00193*** (-3.29)	-0.000885 (-1.27)	-0.000537 (-1.28)	-0.000575 (-0.89)
Time	0.0753*** (18.70)	0.0372*** (3.52)	0.0267** (2.31)	0.0690*** (8.46)	0.0650*** (8.65)	0.0833*** (14.40)	0.0787*** (10.68)
Less than primary education	-0.150** (-2.42)	0 (.)	0 (.)	-0.124 (-0.69)	-0.106 (-0.75)	-0.356** (-2.35)	-0.464*** (-2.59)
More than primary education	0.401*** (4.91)	0.682*** (3.19)	0 (.)	0.116 (0.46)	0.183 (0.62)	0.580*** (5.30)	0.339 (1.47)
Cohort of birth	Yes	Yes	Yes	No	No	No	No
Sex of head	0.0563 (1.63)	0.0230 (0.22)	0.377*** (2.69)	0.121 (1.20)	0.200 (1.55)	-0.0646 (-1.11)	0.0974 (0.80)
Lagged log assets	0.144*** (11.54)	0.238*** (4.47)	0.251*** (3.86)	0.120*** (4.31)	0.141*** (3.40)	0.173*** (7.53)	0.136*** (4.05)
Monthly wages	0.607*** (6.47)	0 (.)	0.0750 (0.95)	0.778*** (3.62)	0.433*** (3.42)	0.657*** (5.29)	0.116 (0.87)
Household size	-0.0413*** (-3.30)	-0.0521 (-1.46)	-0.0597 (-1.44)	0.00667 (0.32)	-0.0892*** (-4.32)	-0.0472** (-2.26)	-0.0908** (-2.38)
Involved in agriculture	-0.0583 (-0.92)	0.135 (0.91)	-0.548** (-2.47)	-0.203 (-1.42)	-0.0929 (-0.81)	-0.123 (-0.86)	0.0946 (0.59)
Village fixed effects	Yes	Yes	Yes	No	No	No	No
_cons	7.888*** (28.81)	8.499*** (15.04)	7.800*** (9.76)	8.333*** (24.13)	8.279*** (18.06)	8.153*** (26.26)	8.377*** (21.03)
N	7653	409	664	1516	1598	2368	1378

t statistics in parentheses

* p<0.10 ** p<0.05 *** p<0.01

Table 2.7(b): Effect of drought on income by household characteristic

	(1)	(2)	(3)	(4)	(5)	(6)
Sub-group	*Whole sample*	*Business owners*	*Female head*	*Head has more than one job*	*More than one bread winner*	*Not in agriculture*
Head has more than one job	0.0822 (1.26)	0.129 (1.58)	-0.134 (-0.97)	0 (.)	0.0686 (0.96)	-0.303 (-0.66)
Employment rate in household	0.396*** (3.63)	0.327*** (2.61)	0.470*** (2.60)	0.278** (2.20)	0.385*** (3.01)	0.985*** (2.65)
Drought	-0.00101*** (-3.12)	-0.00109*** (-2.79)	-0.000943* (-1.78)	-0.00105*** (-2.81)	-0.00102*** (-2.83)	0.000983 (0.84)
Time	0.0753*** (18.70)	0.0778*** (17.48)	0.0781*** (12.23)	0.0786*** (19.18)	0.0739*** (18.06)	0.0479*** (2.82)
Less than primary education	-0.150** (-2.42)	-0.206** (-2.24)	-0.305*** (-2.75)	-0.191** (-2.12)	-0.125** (-2.03)	-0.167 (-1.15)
More than primary education	0.401*** (4.91)	0.277*** (3.29)	0.453** (2.57)	0.346*** (4.36)	0.362*** (4.36)	0.656*** (2.63)
Cohort of birth	Yes	Yes	Yes	Yes	Yes	Yes
Sex of head	0.0563 (1.63)	0.0426 (0.91)	0 (.)	0.0223 (0.48)	0.0408 (1.16)	0.229 (1.35)
Lagged log assets	0.144*** (11.54)	0.152*** (9.08)	0.133*** (5.35)	0.168*** (10.79)	0.155*** (11.56)	0.0986** (2.50)
Monthly wages	0.607*** (6.47)	0 (.)	0.455*** (3.39)	0.472*** (5.87)	0.616*** (6.64)	0.700*** (2.95)
Household size	-0.0413*** (-3.30)	-0.0630*** (-4.39)	-0.0403* (-1.85)	-0.0659*** (-4.53)	-0.0469*** (-3.39)	0.0756* (1.83)
Involved in agriculture	-0.0583 (-0.92)	-0.293*** (-2.62)	-0.0187 (-0.22)	-0.168* (-1.86)	-0.107 (-1.40)	0 (.)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
_cons	7.888*** (28.81)	8.312*** (20.78)	7.907*** (21.27)	8.067*** (21.28)	7.818*** (23.96)	6.921*** (7.75)
N	7653	4896	2305	5034	6713	708

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01

which may be leveraged to engage in production techniques that insulate their revenue streams against covariate shocks.

To understand if any of these observable characteristics can be related to the ability of households to insure their income streams against drought, I re-estimate equation 2.12 on groups of households which exhibit each of these characteristics. The results are presented in tables 2.7(a) and 2.7(b). To ease comparison, the first column of each of these tables reproduces the results of estimating equation 2.12 on the full sample.

In the second column, I restrict attention to those households who are in government work or other jobs that pay monthly wages. The impact of drought on the income streams of these households is statistically indistinguishable from zero. Tellingly, in this subgroup there are insufficient observations of household heads with less than a primary education to identify a parameter for the corresponding variable. This supports the hypothesis that the income streams of better-off households are better insulated against shocks because their educational qualifications enable them to access government or other jobs which pay steady, monthly wages.

The results in column 3 complement this finding, by restricting attention to households that have completed an educational qualification greater than the primary level, that is one of secondary, university or vocational degrees. The effect of drought on the income streams of these households is statistically indistinguishable from zero. Relatively well-off households are much more likely to have these high levels of education, as illustrated by table 2.3. An important caveat here, however is that these households only account for 8% of the sample.

The remaining columns of table 2.7(a) restrict the sample to households which are headed by people born in the 1930s, 1940s, 1950s and 1960s[13] respectively. I find that drought

[13] Households headed by people born before 1930 account for only 7.8% of the sample while those born after 1970 account for only 3.4% of the sample

has a strong, negative effect only on the income streams of households headed by the eldest cohort. For households headed by cohorts born after 1940, the estimated effect of drought is smaller than for the cohort born in the 1930s, and never significantly different from zero. The data thus suggest that households headed by people born after 1940 are better able to adapt to the occurrence of drought, though I cannot be sure why this is. Nonetheless, the heads of relatively well-off households tend to be younger than those of poorer households (from table 2.3), so the age of the household head may help explain why the income streams of richer households are better insured against covariate shocks than their poorer counterparts.

Table 2.7(b) presents results from regressions which restrict attention to business owners, female headed households, households where the head has more than one job and households where there is more than one breadwinner. The degree to which drought affects the income streams of all these subgroups is similar to the result for the whole sample, and thus I am unable to relate these characteristics to differences in the extent to which the income streams relatively well-off households are insured against drought.

In the sixth column, I focus on households which report that they are not involved in any kind of agricultural activity. For obvious reasons, I find that the income streams of this subgroup are also well insured against crop disease. This finding, however, may not be linked to the ability of better-off households to access low-risk income because households above and below the median level of permanent income are almost equally likely to be involved in agriculture. Again, it is important to note that these households account for less than 10% of the sample.

The villages sampled by the Townsend Thai Project are drawn from two regions in Thailand. Some of the villages are from the semi-arid Northeastern region while the others villages from in the more developed central region. Other studies of the Townsend

Thai Data have found important differences between these regions with regard to the types of risks faced by households (Samphantharak and Townsend, 2017), the ways in which households utilize savings (Townsend 2013) and the factors which drive household wealth accumulation (Pawasutipaisit and Townsend, 2011). In light of this literature, Table 2.8 allows for drought to affect village households differently in each of these areas. The resulting estimates from the relatively rich central region are presented in the second column of the table, while those from the relatively poor Northeastern region are are in the third column. Again, for comparison I have included the full sample results in the first column. In the relatively affluent central region, the presence of drought does not exert a statistically significant effect on household income. In contrast, the presence of a drought exerts a strong, negative effect on household income in the relatively poor Northeastern region. Thus there is some evidence that geographic factors may constrain the ability of relatively poor households to generate low risk income streams.

This section has demonstrated ability of better-off households in rural Thailand to insure their income streams against drought is empirically linked to a number of observable characteristics of these households. Better-off households are more likely to hold an educational qualification above the primary level, and to be in jobs which pay relatively stable monthly wages, thereby helping insure their income streams against covariate shocks. Intuitively, households that are in no way dependent on agriculture also have income streams that are better insured against drought, though this characteristic is not more common among relatively well-off households. Better-off households are more likely have younger heads, who may be better able to adapt to changing conditions in agricultural markets. Finally, better off households are more likely to reside in the developed central region which may offer lower risk income generating possibilities than the less developed Northeastern region.

Table 2.8: Effect of Drought on Income by Region

	(1)	(2)	(3)
Sub-group	*Whole sample*	*Central region*	*Northeast region*
Head has more than one job	0.0822	-0.0719	0.167**
	(1.26)	(-0.77)	(2.12)
Employment rate in household	0.396***	0.756***	0.0179
	(3.63)	(5.50)	(0.15)
Drought	-0.00101***	-0.000170	-0.00197***
	(-3.12)	(-0.48)	(-4.84)
Time	0.0753***	0.0702***	0.0798***
	(18.70)	(14.81)	(14.76)
Less than primary education	-0.150**	-0.242***	-0.0419
	(-2.42)	(-2.77)	(-0.51)
More than primary education	0.401***	0.341***	0.414***
	(4.91)	(3.01)	(3.56)
Cohort of birth	Yes	Yes	Yes
Sex of head	0.0563	0.0944*	0.0445
	(1.63)	(1.65)	(1.09)
Lagged log assets	0.144***	0.161***	0.160***
	(11.54)	(9.77)	(8.53)
Monthly wages	0.607***	0.623***	0.616***
	(6.47)	(4.55)	(5.90)
Household size	-0.0413***	0.00396	-0.0940***
	(-3.30)	(0.25)	(-6.85)
Involved in agriculture	-0.0583	-0.0600	0.0801
	(-0.92)	(-0.81)	(0.63)
Village fixed effects	Yes	Yes	Yes
_cons	7.888***	7.605***	7.869***
	(28.81)	(17.74)	(28.51)
N	7653	3460	4239

t statistics in parentheses

* p<0.10 ** p<0.05 *** p<0.01

2.5 Discussion and Conclusions

This paper started by observing that the use of low-risk income streams as a means of insurance has typically been studied among the poor and vulnerable. It speculated that this may be because the literature has been guided largely by Morduch's theoretical contribution which assumed that credit was the only dimension along which household insurance decisions were constrained, implying that the poor, who are more likely to be credit constrained, are also unable to smooth consumption in response to shocks to income. This standard model of income smoothing was at odds with the patterns of insurance observed in figures 2.1-2.3, which suggest that richer households in rural Thailand use low-risk income streams to satisfy a greater proportion of their insurance needs than poorer households. In section 2.2, the paper showed analytically that if poorer households were also constrained in their ability to secure low-risk income, then whether or not they used a greater degree of consumption smoothing to satisfy their insurance needs than richer households, became an empirical question.

Section 2.4 identified drought as a covariate shock to income from among the village level information collected by the Townsend Thai Project. The paper then used the mean level of household consumption over the duration of the panel as a proxy for permanent income, and found that the effect of this covariate shock was moderated by the level of household permanent income.

Section 2.4.5 attempted to discern the characteristics of better-off households which enable them to insure their income streams against covariate shocks. It found that access to government jobs and other jobs which paid monthly wages; high levels of education; and relatively young household heads were associated income streams that were insured against drought.

The evidence supports the hypothesis that the income streams of relatively well-off households in rural Thailand are indeed better insured against covariate shocks than their poorer counterparts. This is a novel contribution within the income smoothing literature, which has usually focused on identifying this type of insurance among the relatively poor and vulnerable, in exclusively agrarian settings. These findings suggest that the poor may be at a disadvantage relative to their richer counterparts not only in their ability to secure credit, but also in their ability to access low-risk income opportunities.

Where the existing literature on income smoothing has warned about underestimating the true extent of risk in the income streams of the poor and vulnerable (Morduch, 1995) the analysis in this chapter suggests we run the same risk if we fail to account for the insurance role played by low-risk income at the top end of the distribution, as well. In rapidly industrializing parts of the world, particular attention should be paid in evaluating the impact of increasing access to jobs that pay monthly wages. These jobs have the potential to contribute to household welfare not only by increasing average earnings, but also by serving a potentially crucial insurance function for households. If we acknowledge that there is an element of insurance inherent in regular, monthly wages, and then estimates of insurance that relay exclusively on consumption smoothing may underestimate the total amount of insurance used by households in a wider range of settings than previously thought.

Chapter 3

Inequality and Remittances in Rural Thailand: A Lifecycle Perspective

Introduction

This paper will study the dynamics of income inequality in a panel of rural Thai households, from a lifecycle perspective, using especially high quality income data made available by the Townsend Thai Project (Townsend, 2011). It finds that income inequality between these households is decreasing over time, even within groups of households headed by people from the same cohort of birth. It fails to find evidence to support the hypotheses that declining inequality is driven either by a convergence in the distribution of individual incomes, or by differences in the dynamics of household composition. Rather, it presents evidence that differences in the receipts of remittances from adult children of the heads of these households account for the entirety of the observed convergence in the distribution of income in the sample of village households.

The paper finds that the reason remittances from children generate falling income inequality over the lifecycle of the heads of household derives from two key characteristics of the distribution of remittances between households: first, remittances from children constitute a larger proportion of the incomes of relatively poor households than relatively rich ones; and second, they become an important component of household income only later on in the lifecycle of the head of household. Thus, as households in the panel (or indeed, any fixed membership group such as a cohort) age over time, households within the group become increasingly likely to receive this inequality reducing transfer, explaining why inequality declines over time. The paper demonstrates that these findings are not driven by differences in the propensity to receive remittances between villages, and that they are robust to a variety of different measures of inequality.

The proportion of remittances in the incomes of poorer households is greater than that of richer households, in part because poorer households have a larger number of children who reside outside the village of origin and remit back to their parent's households, but also because the average annual amount remitted by each child from a relatively poor household is a greater proportion of household income than that remitted by their richer peers.

The paper contributes to the literature which studies the link between remittances and income inequality, a path of inquiry started (at the University of Sussex) by Michael Lipton (Lipton, 1980). The perspective of this paper, however is slightly different from what has since been done in this literature: rather than ask whether remittances increase or decrease income inequality, this paper documents a pattern of decreasing inequality in net household income (including remittances from children) over the lifecycle of the sample heads of households and demonstrates that remittances account for the entirety of the observed reduction in inequality. To the best of my knowledge, it is the first paper to study the effect of remittances on income inequality from a lifecycle perspective.

This paper also contributes to the literature that studies the role of intergenerational, intra-family transfers on the applicability of the lifecycle theory, particularly in the developing country context. It does so by documenting the importance of remittances from the children of the heads of households over the lifecycles of the parents. The previous literature (Willis, 1979; Kotlikoff and Spivak, 1981; and Deaton, 1989; among others) found that the cohabitation of adult children with their parents helped insure[1] the household against the dip in lifecycle earnings associated with the age-related decline in the productivity of the household head. This paper confirms that remittances from adult children who live outside the family home also serve this purpose. This literature suggests that the extent of insurance provided by transfers between family members may be important because the household can overcome many of the informational and commitment issues that can impede market-based insurance solutions. This paper provides evidence that supports this hypothesis, as the quality of insurance offered by remittances from children is sufficient to reverse the increase in income inequality that is typically observed over the course of the lifecycle (for example, by Deaton and Paxson, 1994; Blundell, Pistaferri and Preston, 2008; and Jappelli and Pistaferri, 2010; among many others). This is an interesting counterpoint to studies that have found that transfers from older generations to younger ones typically tend to perpetuate inequality (Becker and Tomes, 1979 and Piketty, 2013; among others) whereas the current paper documents a transfer in the reverse direction, that is from younger generations to older ones, which reduces income inequality.

Where this paper studies income inequality, earlier work on the Townsend Thai Data have found that wealth inequality is declining rural Thailand. Pawasutipaisit and Townsend (2010) and Townsend (2013) present evidence that in the monthly series of the Townsend Data (which is drawn from a different and smaller set of villages from the annual data that I use here) differences in savings rates and differences in returns on assets between richer and poorer households account for the reduction in wealth inequality that they document. The authors explicitly state that incoming remittances do not account for the

[1] To the extent that the decline in productivity and longevity can be perfectly predicted, this is a relatively broad use of the word 'insurance'. Nonetheless, it is not difficult to imagine circumstances in rural Thailand which could cause there to be substantial uncertainty around each of these variables so that the term 'insurance' may indeed be appropriate.

reduction in wealth inequality, though they note that that finding is 'not robust to the annual data'. In the present paper, I use the annual series of the data and document declining income inequality. Furthermore, I find that this decline is indeed explained by remittances originating from the migrant children of the head of household.

The remainder of the paper is organized as follows. Section 3.1 reviews the related literature, while section 3.2 introduces the data. Section 3.3 establishes that inequality in household income is decreasing, not only in the balanced panel of households, but also within year of birth cohorts of the head of household. It also confirms that this decrease is driven neither by convergence in the distribution of individual incomes, nor by the dynamics of household composition, as previous literature might lead one to suspect. Section 3.4 establishes that differences in the receipt of remittances from the children of the heads of these village households explain the entirety of the observed reduction in income inequality within decade-of-birth cohorts of the heads of household. The section also demonstrates that this result is robust to a range of different measures of inequality and is not driven by differences in income or remittance dynamics between villages. Section 3.5 studies the characteristics of the distribution of remittances across households and over time which explain their redistributive effect, while Section 3.6 concludes.

3.1 Literature Review

The Dynamics of Inequality and the Permanent Income Hypothesis

It is a robust prediction of the permanent income (Friedman 1957, Ch3) and lifecycle (Ando and Modigliani 1963) hypotheses, that income inequality will be increasing in any fixed-membership group. The theory behind this result is simple: suppose that innovations to individual incomes consist of a permanent component (typically, modelled as a random walk[2]) and a transitory component. Then to the extent that permanent shocks are not correlated between individuals, the distribution of incomes within any group of individuals will diverge. Indeed, since permanent shocks affect not just contemporaneous but permanent income, under the permanent income hypothesis, consumption inequality too, will increase over time.

Deaton and Paxson (1994) demonstrate that these (and other) predictions of the permanent income hypothesis hold in repeated cross sectional data in countries as diverse as Taiwan, the United States and the United Kingdom. Recent papers have reported similar findings from Australia (Chatterjee et al., 2015), Germany (Bonke et al., 2015), Italy (Rosati, 2003; and Jappelli and Pistaferri, 2010) and Japan (Yamada, 2009).

Developing Countries May Have Different Inequality Dynamics

It is no coincidence that the studies cited above rely almost exclusively on data from rich, industrializing countries. Income in poorer countries tends to consist of a larger share of small-holder agriculture. Two aspects of uncertainty in the income stream in communities which are heavily dependent on agriculture, such as rural Thailand (where 91% of households in the balanced panel receive at least some part of their income from agriculture), are particularly salient to a discussion of the evolution of income inequality: their covariate[3] nature and lack of persistence.

The literature has documented a number of instances where agricultural shocks have been demonstrated to include a strong covariate component (Rosenzweig and Wolpin, 1993;

---

[2] A random walk is a stochastic process where the value of a random variable in a particular period, say ε_t, is equal to its value in the immediately preceding period plus a mean-zero innovation term. That is, $\varepsilon_t = \varepsilon_{t-1} + \vartheta_t$. The innovation term, ϑ_t is typically assumed to be independently and identically distributed over time. Most economic applications also assume that it is normally distributed.

[3] Shocks to agricultural productivity such as droughts, floods and pestilence affect whole villages or areas at a time rather than individual households. As a result, these may change the level of village income but not its cross-sectional variance.

Udry 1994; Morduch, 1994; Townsend, 1994 and Dercon 2006, among others). But even the highly stylized case where all shocks to household income are covariate, only explains why income inequality within cohorts does not increase – it cannot account for the reduction in inequality that this paper documents.

Deaton (1989, 1991) observes that in agricultural contexts, where income risk is very much driven by weather, innovations to income will be predominantly temporary, rather than persistent, lending some plausibility to models where income is a mean-reverting process. If we are prepared to make the extreme assumption that all shocks to household income are transient (so that there is no permanent component), a positive shock to a household in one period will on average be offset by a negative shock in a future period, leaving the distribution of income unchanged. But again, this extreme assumption can only explain why income inequality does not increase; it cannot account for the convergence in the distribution of income over time that this paper will document.

If the observed convergence in household income inequality cannot be generated by placing reasonable restrictions on the exogenous stochastic processes that determine household income[4], then they must be linked to some margin of adjustment within the household. Added worker effects (Mincer 1962) and the substitution of household labour from farm to non-farm activities in the presence of an agricultural shock (Kochar, 1999) are well known channels through which households may smooth out temporary fluctuations in income, which have been studied elsewhere in this thesis. These forces however, do not appear to account for the persistent decline in inequality that is documented here (as will be demonstrated later).

It is well known that one of the crucial differences between the nature of households in developed and developing countries is the increased likelihood of observing multiple

[4] In section 3.3, I use the limited data available on the earnings of individual members of the household to look for direct evidence of such convergence in the distribution of individual earnings, and fail to find any.

generations of adult members within the same household in the latter. Within the context of the lifecycle model, the literature has understood this type of household structure to internalize an insurance function that would otherwise require hump-shaped lifecycle saving: parents invest in their children when parental productivity is high, and children support their parents later on in the lifecycle when parents' productivity declines (Deaton, 1989; Cai et al., 2006; Banerjee et al., 2010; Oliveira, 2016). If the children of relatively poor households were more likely to stay on and cohabitate with their parents after entering adulthood, these households would have a larger number of potential breadwinners, possibly explaining the convergence in the distribution of household income noted above. However, this paper finds that differences between richer and poorer households in the rates of cohabitation with adult children of the head of household do not vary in ways which explain the observed reduction in household income inequality.

There is a related macroeconomic literature which studies the effect of fertility and various transfers from older generations to younger ones of human, financial and physical assets (Lam, 1986; Kremer and Chen, 2002; Mare, 2011 and Piketty 2013). In general, this literature finds that such transfers tend to increase inequality, both directly and through general equilibrium effects. As such, these forces are unlikely to explain the convergence in the distribution of income that I observe among this sample of Thai households.

Remittances and Inequality

Cohabitation with younger generations is only one strategy that households can use to insure themselves against low productivity later in the lifecycle. Children may attempt to uphold their end of this intergenerational bargain by sending remittances to their parents, even when they no longer cohabit with them. Remittances from the children of the head of household prove to be particularly important in rural Thailand, as the average of the proportion of household income accounted for by this particular transfer[5] is one quarter (even more, if we restrict attention to that part of the lifecycle where heads are likely to have children of working age, as will be documented in section 3.4)[6].

The relationship between income inequality and the receipt of remittances has been a rich area of economic research, with mixed results. Lipton (1980) reasoned that migration from rural areas was likely to increase rural income inequality, because the available evidence at the time suggested that remittance flows were likely to disproportionately benefit households that were better-off to begin with. Stark et al. (1986) on the other hand, found that Gini coefficients in two Mexican villages calculated with the inclusion of remittance flows were lower than those calculated without them. They hypothesized that the diffusion of information on migration possibilities and early migration outcomes across households in migrant-sending regions reversed the initial increase in income inequality documented by Lipton (1980). Adams (1989) noted that simply excluding remittances from income data does not adequately describe the counterfactual of 'no migration', as people who migrate would presumably have been working in their home communities, had they not migrated. By comparing observed income with predicted household income if migrants had stayed, he finds that remittances increase income inequality in three Egyptian villages. McKenzie and Rapoport (2007) note that migration may impact inequality through a host of other channels such as multiplier effects on goods and services produced in the migrant sending communities and other general equilibrium effects of remittance flows. Attempting to account for these effects, they find that the

[5] *A priori*, we may expect that government assistance and retirement compensation also play an important role in supporting income later in the lifecycle. In the balanced panel, only 6.44% of households receive the former, and only 4.67% receive the latter, so that their contribution to the income of the average household is very small.

[6] The proportion of the average household's income, by contrast is 14.3%, and also increases when we restrict attention to later in the lifecycle.

overall effect of migration among their sample of Mexican villages is to reduce inequality, so long as communities have sufficiently high levels of past migration.

In some respects, the goals of this paper are quite modest, relative to the state of the art in this literature. I do not attempt to construct a counterfactual distribution of income for a whole community or group of communities. Rather, I observe a reduction in income inequality in a specific set of households that does not appear to be easily explained by the standard lifecycle considerations outlined above. This leads me to consider remittance flows as an intra-family, intergenerational transfer that insures members against lifecycle-related declines in productivity. The 'quality' of the insurance provided by this transfer is sufficient to explain the entirety of the observed reduction in inequality. As mentioned in the introduction, to the best of my knowledge, this is an innovative attempt to study the effect of remittances on income inequality within a lifecycle context.

This paper is not the first to study either remittances or inequality in rural Thailand. Paulson (2000) demonstrated that internal migration patterns in Thailand are consistent with a model where households select migration destinations based on an insurance motive. Paulson finds that households are more likely to send migrants to locations where rainfall variation, a key component of income risk in this agrarian setting, was less correlated with rainfall variation in the community of origin. Yang (2004) explicitly studied the link between income inequality and remittances by documenting high inequality in output between provinces, but relatively low inequality in income between provinces. Yang finds that remittance transfers are the equalizing force that explain these differences. In this paper I use more recent data and a longer panel dimension to study the role of remittances in explaining the dynamics of income inequality in these villages.

As was briefly alluded to above, inequality in rural Thailand has also been studied using data from the Townsend Thai Project itself. Pawasutipaisit and Townsend (2011) and Townsend (2013) establish that wealth inequality is falling among another set of Thai villages which are sampled on a monthly basis by the Project. By imposing an accounting framework on survey data on household income and expenditure, these papers show that differentials in savings between relatively poor and relatively rich household account for the convergence in net worth that they observe over the seven-year span of their data. Specifically, the authors find that poorer households save more relative to their richer counterparts, but also that poorer households earn relatively high returns on the assets which they do manage to accumulate. Together, these factors enable poorer households to close the gap with richer ones in terms of wealth. In contrast to what I find here (for income inequality in the annual series), the authors note that remittances and gifts play a relatively minor role in explaining the reduction in wealth inequality among the communities sampled in the monthly series. However, the strong relationship between high savings rates and high returns on assets to which the authors attribute falling wealth inequality are qualified by the important caveat that they are 'not robust to annual data' (Pawasutipaisit and Townsend, 2010: page 57). This chapter will demonstrate that in the annual data and in regard to income inequality, it is indeed remittances, specifically from the adult children of the heads of household who reside outside these villages of origin, which account for the decrease in income inequality over time.

3.2 The Data

As was described in chapter 1, the coverage of the Townsend Thai Data varies from year to year. Households from different survey regions may exhibit systematically different patterns of inequality. Therefore, to study the dynamics of income inequality, it is necessary to restrict the sample to only those 64 villages that are surveyed by the Townsend Thai Project in all fifteen years, i.e. the panel of 960 households. As mentioned in chapter 1, within this subsample, the year-on-year attrition rate reaches a maximum value of 6%, but is usually closer to 3%. Households that are subject to attrition from the sample are replaced by a randomly selected household from the same village, to keep the number of households in each cross section at 960. 609 of these households comprise the balanced panel that are interviewed in every period, without missing or obviously spurious[7] values for key variables of interest in any period. The main results in this paper will be informed by the full panel of 960 households, though in some select cases I use the balanced panel of 609 households.

This paper is about the evolution of income inequality, and so the key variable of interest is household income. I reiterate here that net income is the difference between the household's gross income and agricultural and business expenses over the last 12 months and that these numbers are revalued to allow for inflation using Bank of Thailand data. The summary statistics for the resulting income data are reported in the first row of table 3.1.

[7] Values that I have dropped for appearing spurious were reported in appendix 1.

Table 3.1: Summary Statistics

Variable	*Observations*	*Mean*	*Standard deviation*	*Minimum value*	*Maximum value*
Net household income	14,163	157,570.5	274,221	419.66	12,050,222
Individual monthly wages	2,929	10,561.66	7,805.43	136.43	85,744.91
Individual daily wages	6,695	180.63	59.82	11.99	1231.19
Year of birth of household head	14,244	1,949.17	13.304	1903	1989
Number of resident children	14,263	1.39	1.14	0	10
Remittances from children	9,570	21,374.46	43,037.53	0	1,096,907
Number of children living outside village	13,907	2.35	2.29	0	13
Number of children who remit	14,574	1.131124	1.561793	0	12

* *31.5147 Thai Baht = 1 U.S. Dollar on 31st Dec. 2011 (source: exchangerates.org.uk)*

Incomes in developing countries are notoriously difficult to measure. The Townsend Thai Project is designed to overcome some of the more common pitfalls of measuring income in these contexts. The survey records both ‘net income’ and the contribution of individual sources. The enumerator ensures that the latter add up to the former, providing a basic check on accuracy. Goods that are produced by the household for its own consumption are explicitly recorded as a part of income, as are gifts received by the household, addressing potentially important sources of underestimation. Despite the best efforts of the survey, measurement error is likely to remain a concern, as in most empirical work, and due care must be taken to think through the implications this may have for the analysis here. The presence of measurement error would lead me to an overestimate of income inequality, as the variance of these errors would be added to the true variance of underlying household income, when inequality in observed incomes is calculated. To study the evolution of income inequality over different stages of the lifecycle, it is therefore necessary to assume that the distribution of measurement error is independent of the age of the heads of household. It is difficult to see how a plausible violation of this assumption could generate the results presented in this paper: measurement error would need to be declining in age within, but not between, decade of birth cohorts.

Concerns about measurement error in household income are further allayed by recalling that the income data satisfy some reasonable properties that have been established elsewhere in this thesis. First, in the balanced panel, mean household income strongly predicts mean household consumption. Second, the downward trend in income inequality across households over time is also reflected in a downward trend in consumption inequality, though the magnitude of the decline in consumption inequality is less. Finally, the usual predictors of household income such as the level of education of the head of household, the number of income earners in the household, and the level of agricultural assets, all have significant predictive power. These observations lead me to conclude that though measurement error is almost always an issue, the income data in the Townsend Thai Project appear to be reasonably reliable.

The Townsend Thai Project also collects information on individual members of the household that will be relevant to the analysis that follows. Where household members are employed in jobs that pay either monthly or daily wages, these wage rates are recorded. Unfortunately, the 1997 data on individual wages appear to be inconsistent with the rest of the panel. In 1997 the data report 100 individuals as earning monthly wages that are less than 300 Baht a month. This compares with a total of 8 observations in all the remaining 14 years of the panel. This is either an error, or evidence that the 1997 sample is systematically different from other years. For this reason, I drop the 1997 data from the part of the analysis that is conducted on individual (as opposed to household) level data, in the analysis that follows. The second and third rows of table 3.1 show the summary statistics for monthly and daily wages, respectively. These figures have also been also inflated to 2011 Baht.

The year of birth of the head of household is of particular importance because it is with respect to this variable that I identify the lifecycle effects. A very small fraction of households report having multiple heads. These are dropped from the analysis in the years that they report having multiple heads. The resulting summary statistics for the years of birth of the heads of household are presented in the fourth row of table 3.1. The oldest head of household in the sample was born in 1903, while the youngest was born in 1989. The average head of household is older than the average household member by approximately 21 years.

As discussed previously, the literature suggests that in a rural, East Asian context, the dynamics of household composition may be crucially important to understanding the evolution of the distribution of household income over the lifecycle. The Townsend Thai Project collects detailed data on all individuals who either live in the household for at least six months out of the year, or who are in school and are financially supported by the household (so that those who are in school are counted as a part of the household of origin). The relationship of each individual to the head of household is recorded, and it is using these numbers that I trace the dynamics of the number of children cohabiting with their parents over the lifecycle. As the fifth row of table 3.1 demonstrates, on average each head has less than one-and-a-half of their children living in their household at a given time, though the standard deviation of this number is high, and as will be demonstrated in section 3.4, much of this variation is over the lifecycle of the head.

What makes the Townsend Thai data exceptionally suitable for investigating this research question is the section of the questionnaire dedicated to children of the head of household living outside the village. Along with some of the characteristics of these children, this section collects specific information on the amount remitted from these children to the household of origin. This permits the study of the intergenerational aspect of remittance transfers, without confounding the data with remittances from other sources, such as the spouse of the head of household, or extended family living outside the villages. The summary statistics of the real value of remittances from children in 2011 Baht are reported in the sixth row of table 3.1.

The amount of remittances received from non-resident children of the head of household is likely to depend on the number of such children. The seventh row of table 3.1 presents the summary statistics of the number of the head of household or the spouse of the head of household, who do not reside outside the village. The average number of such children in each household over the duration of the panel is 2.35. Section 3.5 will study whether differences in the number of non-resident children between richer and poorer households can help explain differences in remittance receipts.

Not all the children of the head of household who reside outside the village remit money to their households of origin. The final row of table 3.1 presents the summary statistics of the number of children from each household who remit a positive amount in a given year of the survey. The average household receives remittances from 1.13 children, substantially less than the average number of total children living outside the household.

Table 3.2: Remitter Status and Gender of Non-resident Children

Gender	*Non-remitters*	*Remitters*
Male	7,984	7,309
Female	7,340	8,983
Total	15,324	16,292

Table 3.2 tabulates the gender of children living outside the household against whether or not they remit money back to their parents' household. Approximately 48% of the children of the heads of household who live outside the village of origin are male, whereas the remaining 52% are female. The average woman living outside the village of the head of household is more likely to remit to their parent's households than the average man, with 55% of women remitting in a mean year, as opposed to only 48% of men. This pattern is also true of the amount of money remitted to the parents' household. On

average, female children of the head remit 12,461.27 Baht, whereas male ones remit 9,135.18 Baht (at 2011 prices). The difference is statistically significant, and a difference in means test rejects the null that the average man remits the same amount as the average woman, in favour of the alternative that women remit more, with a t-statistic of 8.79 with 19,448 degrees of freedom.

Unfortunately, there are some characteristics of children living outside the village of origin that I do not observe, which nevertheless are potentially relevant to understanding the role of remittances from children in reducing income inequality among their parents' households. The Townsend Thai Project does not collect information on either the reasons why these children choose to leave the village, or their earnings at their destination. I therefore cannot separate economic migrants from other migrants whose behaviour may be systematically different, such as those who migrate for marriage. Nor can I study potentially important and interesting questions about the share of income at destination that migrant children remit, and how this may vary with respect to observable characteristics of their parents' households.

Bearing in mind these characteristics (and limitations) of the data, I now proceed with studying the evolution of income inequality over this sample of Thai village households.

3.3 Inequality Over the Life-cycle

Figure 3.1 illustrates the evolution of income inequality in a balanced panel of 609 households, sampled by the Townsend Thai Project from 64 Thai villages between 1997 and 2011. Inequality, as measured by the standard deviation of the log of real income, is declining over the 15-year duration of the panel. The 95% confidence interval around the line of best fit shows that the decline is statistically significant. This is confirmed in table 3.3, which presents the results of a regression of a time trend and a constant term on the level of income inequality, which confirms that this decline is statistically significant at all conventional levels.

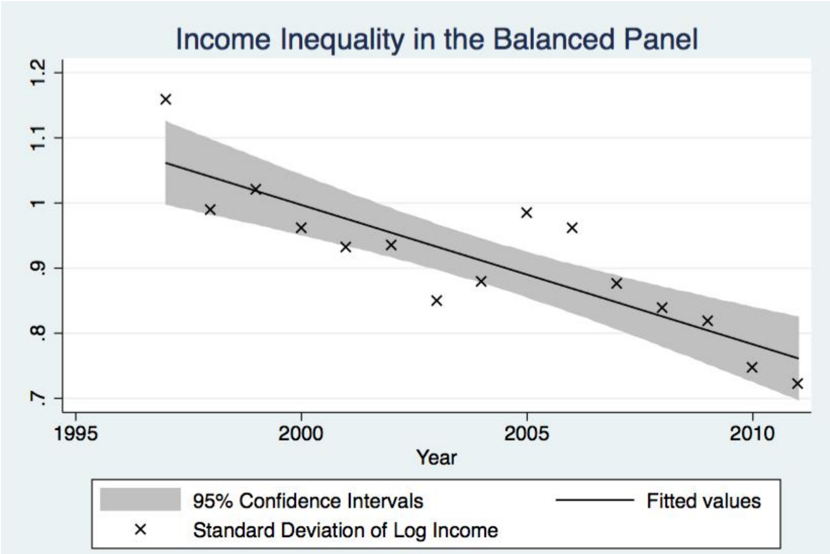


Figure 3.1

On the face of it, this pattern of inequality appears to disagree with what has been demonstrated by the literature to be a very robust prediction of the lifecycle model and the permanent income hypothesis. However, this depiction of the data can only be suggestive of declining inequality, because a number of important (and potentially offsetting) effects are presented together. Each cross section contains households headed by people who are drawn from different year of birth cohorts, and are at different stages of the lifecycle. There is a rich literature documenting that income inequality varies systematically between cohorts, and evolves over the lifecycle (Hall, 1978; Deaton and Paxson 1994; Blundell and Preston, 1998; and Dickens, 2000, among many others). Guided by this literature, I now turn to studying the evolution of income inequality within cohorts defined over the dates of birth of the heads of households.

Table 3.3: Decreasing Income Inequality	
	S.D. of log real income
Year	-0.0162***
	(-5.82)
Constant	33.52***
	(6.00)
N	15
t statistics in parentheses * p<0.10 ** p<0.05	

3.3.1 Household Income Inequality Declines Over the Lifecycle

In figure 3.2, I divide the sample into cohorts defined by the decade of birth of the head of household and follow the evolution of income inequality within each decade of birth cohort. With a view to enhancing cohort-year cell sizes, the analysis is based on all 960 of the households that are interviewed in each cross section, rather than the 609 which form the balanced panel. This choice opens the door to the possibility that the inequality dynamics that I will analyse here are influenced by selective attrition from the panel. To check that this is not the case, I will verify that the key insights of this paper are robust to following cohorts over the balanced panel, while retaining the larger, unbalanced panel as my primary and preferred sample. In the Townsend Thai data, the households in the survey are randomly chosen from the surveyed villages in the first wave, and so are representative of these villages in 1997. In the actual population new, younger households are continually replacing older ones, where this panel attempts to track the same households which are ageing over this period. Thus, even in the best case scenario of purely random attrition, the unbalanced panel will not be representative of the village economies in any year other than 1997.

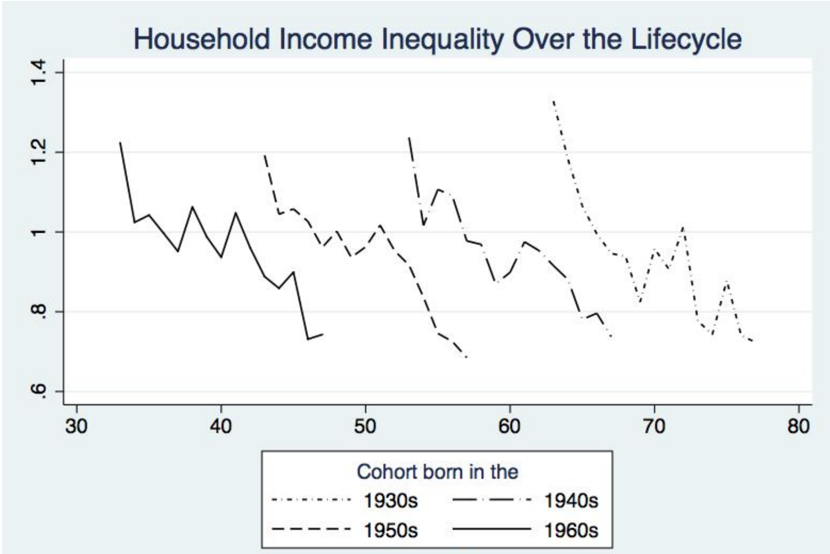


Figure 3.2

Due to the relatively small number of households in each cross section (approximately 960), I have not constructed a separate cohort for every observed year of birth of the head of household. Rather, I define a cohort by each decade of birth of the heads of household. I experiment with using a narrower range of ages to define each cohort, because this would potentially allow me to observe finer differences in inequality over the lifecycle, but doing so also reduces the number of households in each cohort-year cell, thereby increasing noise in the observations. As table 3.4 illustrates, constructing cohorts in this way yields reasonable cell sizes over the duration of the panel for households headed by cohorts born in the 1930s, 1940s, 1950s and 1960s.The ‘age’ of a cohort is short-hand for the number of years that have elapsed from the year that is at the centre of the range of birth years that defines that cohort. Figure 3.2 plots the evolution of income inequality between households headed by people from each of these four cohorts. The remaining cohorts, which are not well identified for the whole panel, are dropped from the analysis. As before, I use the standard deviation of the log of real income as my measure of inequality. These figures confirm that the decrease in income inequality observed in figure 3.1 is not driven by younger (and potentially less unequal) households replacing older ones as the panel progresses, but is a genuine (if somewhat surprising, from a theoretical point of view) feature of the lifecycle of Thai households. This general trend appears to hold true for every cohort for which I have a reasonable number of observations in each cohort-year cell. Furthermore, at a given age, younger cohorts tend to exhibit less income inequality than older ones. The level of income inequality at the beginning of the panel does not appear to vary systematically between cohorts, as it does in other studies (Blundell, Pistaferri and Preston, 2008 for example). It is possible that there is some heterogeneity in the rate at which inequality is declining: the oldest cohort may be experiencing a faster decline than others.

Table 3.4: Cohort – Year Cell Sizes for Household Income

Decade of birth:	*1900s*	*1910s*	*1920s*	*1930s*	*1940s*	*1950s*	*1960s*	*1970s*	*1980s*
Year									
1997	2	33	87	210	194	258	150	12	0
1998	2	27	88	208	195	253	148	17	0
1999	0	23	83	200	195	265	145	18	0
2000	0	20	84	196	203	273	148	19	0
2001	0	19	84	192	200	270	154	14	1
2002	0	17	86	181	202	269	158	21	2
2003	0	17	82	181	206	274	160	21	2
2004	0	16	74	183	210	270	161	26	2
2005	1	10	57	188	205	270	172	35	4
2006	2	9	51	181	208	276	174	37	3
2007	0	6	41	160	208	288	182	55	7
2008	0	5	29	138	209	288	208	63	8
2009	0	5	20	135	210	281	216	74	10
2010	0	4	18	120	207	286	223	83	11
2011	0	2	17	106	203	285	232	95	12

To pin down whether or not these observations are statistically significant, I now model the evolution of income inequality as an initial condition for each cohort, and a cohort-specific time trend. The objective is to test econometrically if inequality within each cohort is declining over time.

The model I estimate is:

$$\sigma_{ct} = \alpha + \beta_c \cdot t_0 + \gamma_c \cdot t + u_{ct} \qquad (3.1)$$

where σ_{ct} is the standard deviation of household income in cohort c in year t. The variable c is a set of cohort dummies. These are first interacted with a dummy variable that takes on a value of 1 in 1997, the first year of the panel, and a value of zero in all other years. The vector of coefficients β, will thus estimate of the level of initial income inequality in each cohort. The next term is an interaction between c, and a linear time trend, t. The vector of coefficients on this interaction, γ, is the key variable of interest in this part of the analysis: if inequality within a cohort is declining over time, then γ will be negative and significantly different from zero. The error term u_{ct} is assumed to have mean zero, and α is a constant.

An advantage of estimating an initial condition for each cohort, rather than a full cohort effect, is that I do not encounter the 'dummy variable trap' and can estimate initial inequality for all four cohorts. Thus identification is achieved by assuming that the time trends are linear within each cohort, permitting me to simultaneously model the evolution of inequality within each cohort over the entire duration of the panel (at the expense of omitting an overall time effect). The results are presented in table 3.5.

The estimated coefficients for the initial levels of inequality in the three younger cohorts are very similar. The 95% confidence intervals demonstrate that they are statistically indistinguishable from one another. Only households headed by the oldest cohort have significantly higher initial income inequality than households in other cohorts. With respect to the estimated cohort-specific time trends, for every cohort I reject the hypothesis that $\gamma = 0$ in favour of the alternative that $\gamma < 0$. Thus I conclude that income inequality is declining over time, within each group of households categorized by the cohort of birth of the head of household.

Table 3.5: Income Inequality Declines Within Cohorts		
Dependent variable	*S.D. of log real income*	
Sample used to construct cohorts	*Unbalanced Panel*	*Balanced Panel*
Initial Inequality for cohort born in 1930s	0.2253*** (13.7)	0.2839*** (13.17)
Initial Inequality for cohort born in 1940s	0.1315*** (8.03)	0.1042*** (4.82)
Initial Inequality for cohort born in 1950s	0.0881*** (5.48)	0.1359*** (6.38)
Initial Inequality for cohort born in 1960s	0.1164*** (7.19)	0.0655*** (2.98)
Time × born in the 1930s	-0.0264*** (-10.08)	-0.0223*** (-8.04)
Time × born in the 1940s	-0.0237*** (-11.68)	-0.0157*** (-5.39)
Time × born in the 1950s	-0.0250*** (-9.55)	-0.0176*** (-5.31)
Time × born in the 1960s	-0.0222*** (-8.63)	-0.0177*** (-6.28)
Constant	1.1293*** (63.26)	1.0389*** (43.93)
t statistics in parentheses, robust standard errors, * p<0.10, ** p<0.05, *** p<0.01. The initial conditions are identified as the coefficient on interactions between cohort dummies and dummy that is 'on' in 1997 and 'off' in all other years. Thus the 'omitted category' is the set of observations which are from all years other than 1997. Identification of the four within-cohort time trends is achieved by the assumption of linearity and the omission of an overall time trend.		

Figure 3.2 was suggestive of possible heterogeneity in the rate at which income inequality was declining within cohorts. Table 3.5, on the other hand indicates no statistically robust evidence of such heterogeneity. All estimated coefficients are within two robustly calculated standard errors of one another. To test these hypotheses formally, I conduct F-tests of pairwise comparisons between all estimated cohort-specific time trends. The results are presented in table 3.6. In every case I fail to reject the null hypothesis that each pair of trends is the same, against the alternative that they are not, even at the 10% level of significance. Thus income inequality is declining significantly within cohorts, and there is no evidence that the extent of the decline is different between cohorts.

To ensure that these inequality dynamics are not influenced by the attrition of households from the unbalanced sample, or the subsequent addition of replacement households, I re-estimate equation 3.1 following decade of birth cohorts of the heads of household over the balanced sample, of 609 households. Doing so yields the results presented in the second column of table 3.5. These results are very similar to those in the first column. The estimated coefficients are all of the same sign by and large, of similar magnitudes. The key result that inequality in household declines significantly over time for every decade of birth cohort carries over to this sample.

Table 3.6: F-tests for Differences between Cohorts of Time Trends in Household Income Inequality

	1930s	*1940s*	*1950s*
1940s	1.11 (0.2967)	.	.
1950s	0.23 (0.6370)	0.33 (0.5668)	.
1960s	2.19 (0.1453)	0.44 (0.5096)	1.18 (0.2833)

F-statistics distributed with (1, 51) degrees of freedom; p-values in parentheses.

3.3.2 Inequality in Individual Wages Over the Lifecycle

So far, this section has presented evidence of a significant decline in income inequality between households over the lifecycles of the heads of household. This convergence in the distribution of household income could be driven by convergence in the income streams of individuals who comprise the household, in ways that are exogenous to the household or by differences in the number and behaviour of earning adult members of the households, or by both of these factors. To distinguish between these two potentially important channels, I would ideally observe the contribution of each individual earning member to total household income. In practice, income within these rural Thai households cannot be unambiguously disaggregated, since much of household income is agricultural, where investments in agricultural assets are combined with the (potentially heterogeneous) unpaid labour of different members of the household before yielding income[8].

Given this problem, I focus on the 26% of working household members who are in a form of employment that pays either a daily or monthly wage. These household members are very likely to differ from those who are in non-wage paying employment, both in their observable and unobservable characteristics. As a result, it will not be possible to generalize findings based on these selected sub-samples to the broader population of household members. Nonetheless, it is the only check on dispersion in individual earnings that I can perform given the nature of these households and the available data.

I resist the temptation to simply multiply daily wages by 21 (the number working days in an average month) and pool the resulting values with the information available on monthly wages. Doing so would most likely underestimate the relative riskiness of daily wages, because there may be constraints on the number of days of work available to a daily wage earner. I recall here that some of the data on individual wages appeared to be miscoded or erroneous in 1997. For this reason, data on the year 1997 is dropped from the remainder of the analysis that is conducted on the individual (as opposed to household) income.

[8] Circumventing these issues of disaggregation, appendix 4 tracks the evolution of inequality in agricultural incomes over the lifecycle of the heads of household and finds that this type of income also fails to explain the convergence in the distribution of household income documented above.

Inequality in Monthly Wages

Monthly wage earners (including those with government jobs) are relatively rare in this sample of Thai village households, and on average I observe 251 individuals a year who earn monthly wages in their primary occupation (including those who are in government work). The resulting cell sizes (which are presented in appendix 5) lead me to expect that wage inequality will be reasonably well identified from when the cohort born in the 1980s reach 17 years of age to when the cohort born in the 1950s reach 53 years of age.

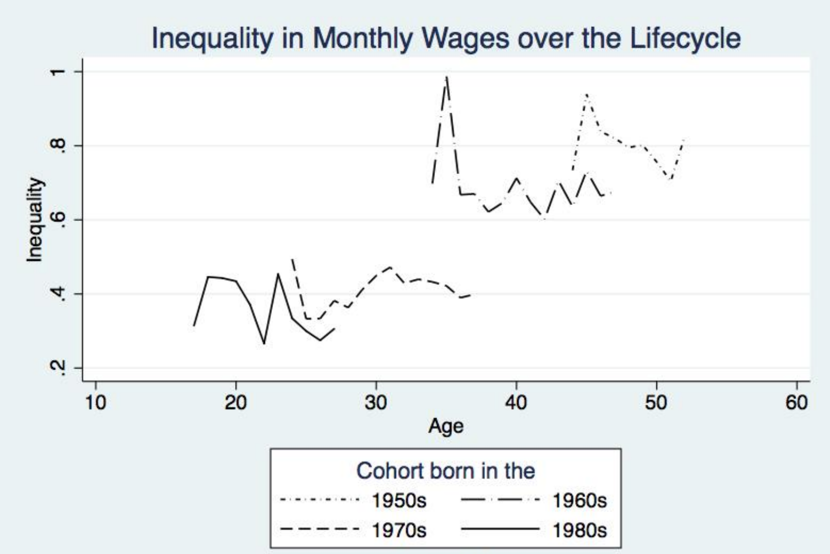


Figure 3.3

Figure 3.3 depicts the evolution of monthly wage inequality from these four cohorts. There is some evidence that monthly wages among older cohorts are more unequal. Importantly, however, over the part of the lifecycle for which we have reasonable cell sizes (between 17 and 52), inequality does not seem to systematically fall within cohorts. To test if inequality in monthly wages is decreasing within cohorts, I estimate the following model on the data:

$$\sigma_{ct} = \alpha + \beta c + \gamma_c \cdot t + u_{ct}, \tag{3.2}$$

using the standard deviation of monthly wages as the dependent variable. Figure 3.3 suggests that the levels of inequality in monthly wages may differ between cohorts. For this reason, I model them using cohort effects, rather than initial conditions, as was done in equation 3.1. I restrict the sample to individuals aged between 17 and 52, ages for which I have a reasonable number of observations in each cohort-year cell. The results of the regression are presented in the first column of table 3.7. There is some evidence that older cohorts are more unequal. This is consistent with the literature on wage dispersion over the lifecycle in developed countries, such as Dickens's (2000) study of UK wages, and Blundell, Pistaferri and Preston's (2008) study of US incomes. Importantly however, within-cohort time trends are all statistically indistinguishable from zero at the 95% level of significance.

Table 3.7: The Evolution of Inequality in Monthly and Daily Wages

Dependent Variables	*S.D. of Monthly Wages*	*S.D. of Daily Wages*
Cohort born in the 1940s		-0.0090 (-0.33)
Cohort born in the 1950s	0.4612*** (5.03)	0.1209*** (4.3)
Cohort born in the 1960s	0.3602*** (3.77)	0.1342*** (5.01)
Cohort born in the 1980s	0.0822 (0.98)	-0.0019 (-0.08)
Time × born in the 1940s		-0.0015 (-0.52)
Time × born in the 1950s	-0.0086 (-0.76)	-0.0100*** (-4.47)
Time × born in the 1960s	-0.0073 (-0.96)	-0.0108*** (-4.73)
Time × born in the 1970s	0.0022 (0.54)	-0.0032*** (-3.07)
Time × born in the 1980s	-0.0116* (-1.94)	-0.0031 (-1.62)
Constant	0.3922*** (8.79)	0.2735*** (32.76)
N	48	70
R-squared	0.8848	0.5955

t statistics in parentheses, robust standard errors, * p<0.10, ** p<0.05, *** p<0.0. There are insufficient observations of monthly wages to form reliable estimates for the cohort born in the 1940s. The cohort born in the 1970s is the omitted category in both the monthly and daily wage models. Identification of the within cohort time trends is achieved by the assumption of linearity and the omission of an overall time trend.

Thus inequality in monthly wages may explain part of the reduction in inequality over the balanced panel, as younger, more equal cohorts are replacing older ones. Nonetheless, they provide scant support for the hypothesis that the reduction in inequality in incomes *within* cohorts can be explained by a reduction in inequality of individual earnings.

Inequality in Daily Wages

The relationship between household income inequality and inequality in daily wages is less direct than that between income inequality and monthly wages because there are likely to be heterogeneous constraints (which are not directly observed in the data) to the number of days of paid work that daily wage labourers can find. Therefore, these results need to be interpreted with some care.

In each year from 1998 to 2011, on average I observe 522 daily wage earners in these households. Appendix 6 presents the sizes of the resulting cohort-year cells. The oldest cohort for which I have reasonable cell sizes over the duration of the panel was 53 years old in 1997 and the youngest was 13. Together, these 5 cohorts provide a synthetic view of a large part of the lifecycle of daily wage earners.

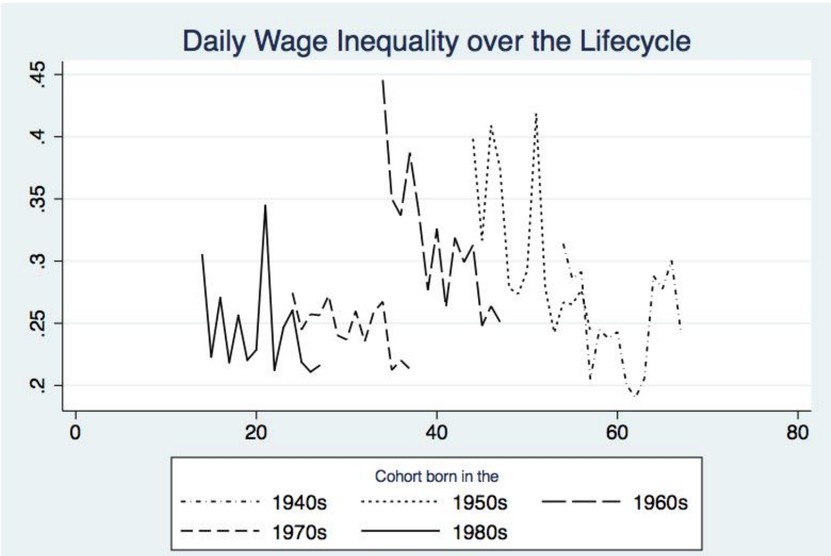


Figure 3.4

Figure 3.4 depicts how inequality in daily wages evolves within these 5 cohorts. The figure suggests that inequality may be decreasing for the middle cohorts, though any such pattern is accompanied by a great deal of noise. I test for declining inequality by re-estimating equation 3.2 using the standard deviation of daily wages as the dependent variable, and report the results in the second column of table 3.7. The declines in inequality for the cohorts born in the 1950s, 1960s and 1970s are indeed statistically significant, though the estimated declines for cohorts born in the 1940s and 1980s are not. Thus the evidence for declining inequality in daily wages is mixed. The data certainly do not exhibit the unambiguous decline observed in figure 3.2. Taken together with potential heterogeneity in the hours of work available to this group of earners, these results are unlikely to explain the entirety of the convergence in the distribution of household income documented above.

3.3.3 The Dynamics of Household Composition and Income Inequality

The presence of multiple generations of adults within a household could potentially drive the observed reduction in household income inequality. If more of the children of poorer households are likely to stay on in their parents' households throughout adulthood, these households would have a larger number of potential income earners, explaining the observed reduction in income inequality.

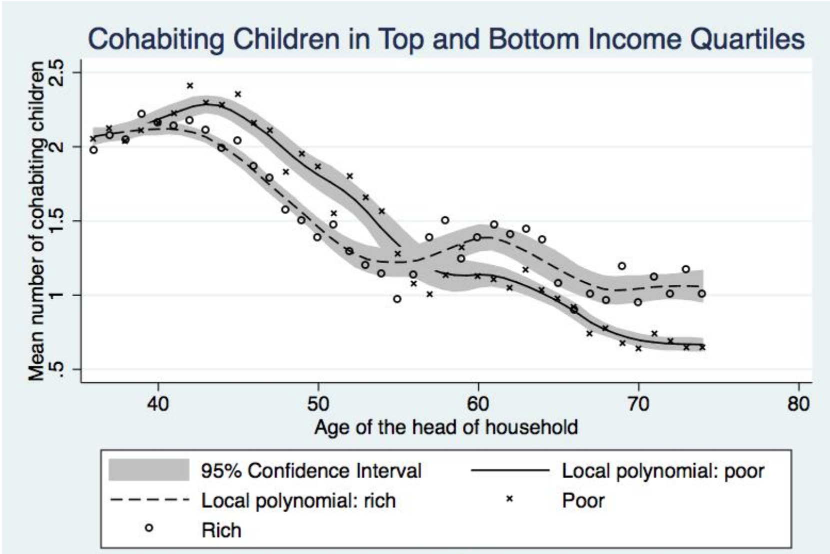


Figure 3.5

Figure 3.5 plots the evolution of the mean number of children of the heads of household who are resident in relatively rich and relatively poor households over the ages of the heads of household from when they are 35 to 75 years old. Each ‘x’ represents an estimated mean for the sub-sample of households who are in the bottom quartile of permanent income[9], that is households that are ‘relatively poor’, and each ‘o’ represents an estimated mean for the sub-sample of households that are in the top quartile, the ‘relatively rich’ ones. The local polynomials through these points (and their 95% confidence intervals) are also presented. The evidence suggests that poorer households have significantly more children than their richer counterparts, and these children start to exit their parents’ households later than their richer counterparts. Until their heads reach 52 years of age, households in the bottom quartile of permanent income have, on average approximately one-third of an extra child residing with them as compared with household heads in the top quartile of permanent income, and the difference is statistically

[9] ‘Relatively rich’ and ‘relatively poor’ are defined in a similar way to how they were used in the preceding chapter. I take the average level of real, equivalised consumption for each household over the duration of the panel as an observed proxy for permanent income. It should be noted here that in using this proxy for permanent income, I am restricting the data to that sample for which I have observations on consumption in every year, that is, the balanced panel of 609 households.

significant. Throughout their 50s, the heads of households in the bottom income quartile continue to experience the departure of their children from their households while exit of children from households in the top quartile appears to halt[10]. When heads in the top income quartile reach their 70s, on average they continue to have approximately one child resident with them, whereas those in the bottom income quartile have, on average, one-third fewer cohabiting children.

Under the maintained hypothesis that the only way in which the incomes of parents' households are supplemented by the productivity of their adult children is through cohabitation, these dynamics of household composition cannot account for the convergence in the distribution of household income documented above. If anything, in this view, the fact that the children from poorer households are more likely to leave their parent's household upon entering adulthood, so that these households have fewer resident children contributing to household income than their richer counterparts, would cause the distribution of household income to diverge over time, rather than to converge as it does in figures 3.1 and 3.2. But this interpretation neglects the possibility that children may continue to contribute materially to their parents' household, even after exiting the household. As Stark and Bloom (1985) observe, the family is not "an entity that is split apart as its independence seeking younger members move away in an attempt to dissociate themselves from familial and traditional bondage, regardless of the externalities thereby imposed upon their families." Indeed, figure 3.10 suggests poorer households have more adult children living outside the household, and thus a larger pool of potential remitters than their richer counterparts, especially later on in the lifecycle of the household heads. This finding provides further motivation to examine whether or not

[10] Indeed, from their mid-50s to their early 60s, the average number of cohabiting children for richer heads of households appear to increase, but this difference is not statistically significant.

remittances from the children of the head of household are a key driver of the reduction in inequality documented earlier in this section.

3.4 Do Remittances from Children Explain Falling Income Inequality?

The analysis so far has established that income inequality within a sample of Thai households is falling over time, both within a balanced panel and within cohorts of heads of household born in the same decade. I found scant evidence that the decline could be linked to falling inequality in individual incomes; or that it could be explained by the dynamics of children staying on in their parent's households, thereby supplementing household income. First, this section will establish that remittances from children account for the entirety of the convergence in the distribution of household income documented in the preceding section. Then, it will demonstrate that this result is not driven by differences in the dynamics of remittance receipts between villages. Finally, it will verify that the key findings of this paper – that income inequality is declining within decade of birth cohorts in the Thai sample, and that this decline is explained by the receipt of remittances – are robust to a range of different measures of inequality.

3.4.1 The Distribution of Income Without Remittances

I will now study the evolution of inequality over the lifecycle, for the component of household income that is not remitted by children. Simply subtracting the remittance contributions of children from household income (as I do here) clearly does not provide any information on the counterfactual where the children of heads of household did not migrate. Adams (1989) noted that if household members had not migrated and remitted, they would presumably have been in some other form of employment, potentially contributing to the income of the household of origin. McKenzie and Rapoport (2006) added that remittances also induce multiplier and other effects across the communities that receive them. The objective of the exercise in this section is not to compare observed income inequality with these counterfactuals, but rather to ask whether or not this particular transfer to the household explains the convergence in the distribution of household incomes documented in section 3.3. It is with the goal of answering this question that figure 3.6 traces the lifecycle profiles of income inequality among these households, after subtracting the contribution of remittances from the children of the head of household who do not reside with their parents.

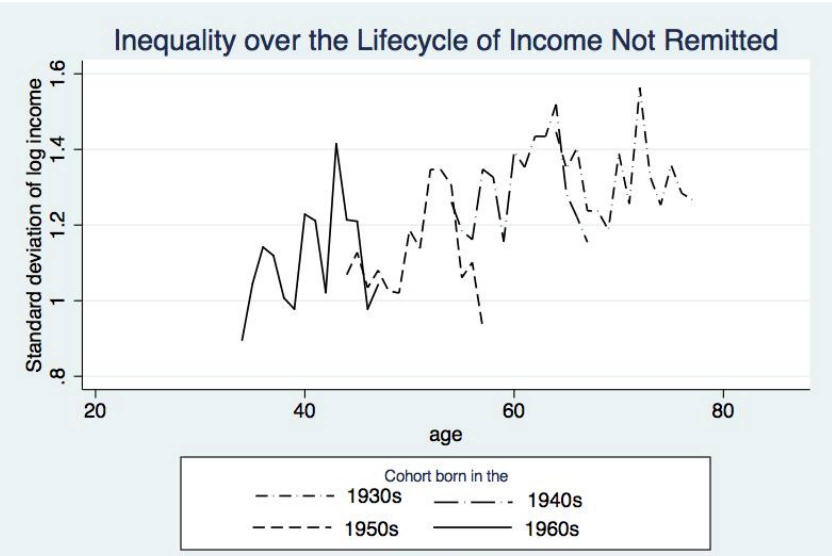


Figure 3.6

Figure 3.6 illustrates that inequality in the component of household income that is not remitted by children, does not appear to decrease over the lifecycle of the heads of household. To the contrary, the figure gives the impression that inequality in this component of household income may indeed be increasing over the lifecycle, in line with the standard narrative where shocks to income contain a permanent component (as was discussed in section 3.1). This particular visual representation of the data gives a strong impression of linearity in the relationship between inequality in this component of household income and the age of the household head. Guided by this characteristic of the data, I first test for rising inequality in this component of household income by assuming that income inequality varies linearly with age, that is:

$$\sigma_{ct}{}^{-r} = \alpha + \gamma a + \varepsilon_{ct} \qquad (3.3)$$

where the superscript "-*r*" emphasises that remittances from children have been subtracted from household income before calculating the standard deviation of the log, and all other variables are as previously defined. In this simple specification, inequality in each cohort in each year is regressed against the 'age' of the birth cohort of household head and a constant term. The estimated coefficients are presented in the first column of table 3.8. They suggest that on average, an increase in the age of the head of household of one year is associated with an increase in the standard deviation of this component of household income of 0.0078 log points. The t-statistic on this coefficient is 6.11 and is distributed with 54 degrees of freedom, so that in this basic specification I strongly reject the hypothesis that inequality in the component of household income that is not remitted by the children of the head of household, does not vary with age in favour of the alternative that it is increasing with age.

As noted earlier, my preferred specifications are based on cohorts constructed from the unbalanced panel, because this sample yields larger cohort-year cell sizes, and is therefore more likely to deliver reliable results. However, it is important to verify that none of the key insights of this paper hinge on this particular choice of sample. For this reason, I re-estimate equation 3.3, but using the cohorts constructed from the balanced panel of 609 households. The results are presented in the second column of table 3.5. The coefficient on age remains positive, and significantly different from zero. Indeed, the magnitude is slightly larger than in the model that was based on the unbalanced panel.

Table 3.8: The Evolution of Inequality in Income Not Remitted by Children of the Head of Household

	(1)	(2)	(3)	(4)
Sample used to construct cohorts	*Unbalanced panel*	*Balanced panel*	*Unbalanced panel*	*Unbalanced panel*
Dependent variable	*S.D. of income not remitted by children*	*S.D. of income not remitted by children*	*S.D. of income not remitted by children*	*S.D. of income not remitted by children*
Age of the head of household	0.00783*** (6.11)	0.0105*** (7.89)	0.00415 (1.05)	0.0119*** (3.50)
Cohort born in the 1940s			0.0179 (0.28)	
Cohort born in the 1950s			-0.116 (-1.14)	
Cohort born in the 1960s			-0.0936 (-0.71)	
Time × born in the 1930s				-0.0157 (-1.63)
Time × born in the 1940s				-0.00478 (-0.67)
Time × born in the 1950s				-0.0104* (-1.96)
Time × born in the 1960s				0.000105 (0.02)
_cons	0.781*** (10.75)	0.6141**** (8.39)	1.033*** (3.67)	0.623*** (3.80)
N	56	56	56	56
R-Squared	0.3662	0.4820	0.4272	0.3863

t-statistics in parentheses, robust standard errors, * p<0.10, ** p<0.05, *** p<0.01. The cohort born in the 1930s is the omitted, base category for the specification. Simultaneous identification of all four time trends is achieved by the assumption of linearity and the omission of an overall time trend.

Though these simple models have fairly high R-squares of 0.36 and 0.48 (that is, age and a constant term explain between one-third and almost one half of the observed variation in this component of income risk depending on the sample used to construct the cohorts), they neglect a number of potentially important and well-known sources of heterogeneity, such as the possibility that inequality varies systematically between cohorts.

Simultaneous identification of age and cohort effects is problematic, due to the dummy variables for each decade of birth cohort, by construction, capturing the average effect of the age-related variation for households headed by people born in that decade. Thus when estimated in the same regression, both the cohort effects and the age effects are likely to be biased towards zero. The point of the exercise here, however, is not to identify these effects, but rather to ascertain whether there is any evidence of heterogeneity between cohorts above and beyond that accounted for by the linear age effect. To allow for the level of inequality in this component to vary between cohorts, while retaining the assumption that it is linearly related to age, I estimate the following specification:

$$\sigma_{ct}{}^{-r} = \alpha + \lambda_c + \gamma a + \varepsilon_{ct}. \tag{3.4}$$

To estimate these results, I return to using the larger, unbalanced panel to construct the cohorts. The results are presented in the second column of table 3.8. Because the simultaneous estimation with the cohort effects biases the coefficient on age to zero, it is smaller in magnitude than in the first column and no longer statistically significant. The estimates do not provide any statistical evidence that the cohort effects are significantly different from one another. An F-test of the hypothesis that the cohort effects are jointly equal to zero yields an F (3, 51) statistic of 1.73, where the critical value is 2.79, for the 5% level of significance. These results suggest that the level of inequality in this component of household income does not vary systematically between cohorts and recommend the use of the simpler model from equation 3.3.

Despite the average levels of inequality not varying systematically between the cohorts, it is possible that the rate at which inequality increases within each cohort may differ from one cohort to another. To test this hypothesis, I estimate the following equation on the data:

$$\sigma_{ct}{}^{-r} = \alpha + \lambda_c \cdot t + \gamma a + \varepsilon_{ct}, \tag{3.5}$$

where cohort-specific time trends are estimated in addition to the linear age effects and the constant term. The results are presented in third column of table 3.8. The age effect is slightly larger in magnitude, but to that the one which was estimated from equation 3.3 and is statistically significant at the 1% level. There is some evidence that inequality in the component of household income that is not remitted by children is increasing less rapidly for the households headed by people born in the 1950s than for other cohorts. To interpret this coefficient, however, we must note that as time passes, the cohort is also ageing. Therefore, to test if inequality in this component of household income continues to decrease over time, I must test the null hypothesis that the sum of this coefficient and the age coefficient is equal to zero against the alternative that it is negative. This test yields an F statistic of 0.07 that is distributed with (1, 50) degrees of freedom, with a critical value of 4.03, leading me to fail to reject the null. Thus this regression provides no evidence that inequality in the component of household income that is not remitted by children is declining over the duration of the panel, for any of these decade of birth cohorts.

The statistical analysis that is summarized in table 3.8 thus supports the narrative that emerges from figure 3.6: when the contribution of remittances from the children of the head of household is removed, household income inequality increases over the lifecycle in a way that is explained by the standard model where household incomes are hit by a variety of shocks that are imperfectly correlated and have varying degrees of persistence. Thus this type of remittance accounts for the entirety of the convergence in the distribution of household income that was documented in section 3.3.

3.4.2 Robustness of the Results

Does variation between villages drive the results?

The households in the sample represent 64 different villages over a period of 15 years. Systematic differences between villages, such as the proximity to an urban centre or heterogeneity in the depth of available financial services, could conceivably predispose some villages to receiving a greater share of their income from remittances than others. It is therefore possible that the pattern of declining inequality documented above is driven by differences between villages, rather than between households and within villages. To learn whether or not this is the case, I regress household income on a fully interacted set of village and time fixed effects, and repeat the analysis on the residuals from that regression. That is, I estimate the econometric model:

$$y_{ivt} = \alpha + \vartheta_v \times \tau_t + \varepsilon_{ivt} \qquad (3.6)$$

and use the resulting coefficients to compute the vector of residuals, $\hat{\varepsilon}_{ivt}$. I then group these residuals into decade of birth cohorts and calculate the standard deviation of this residual within each cohort-year cell.

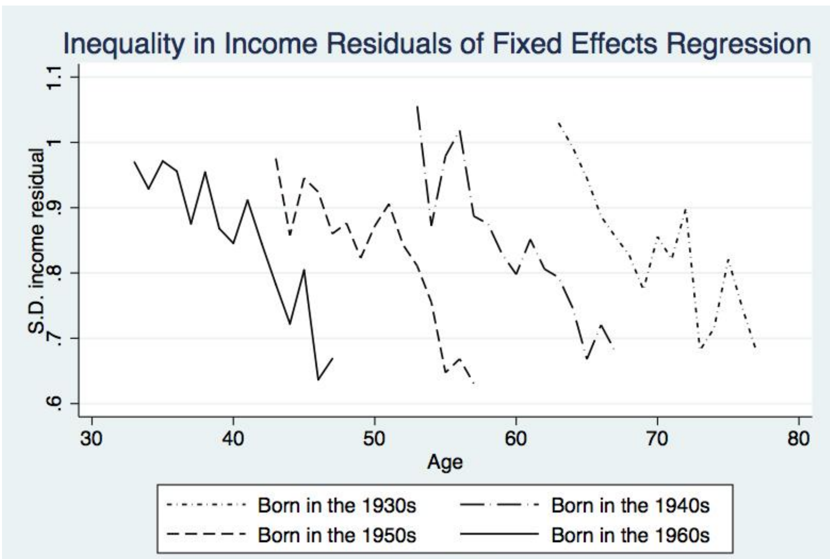


Figure 3.7

The resulting dynamics of the income residuals are presented in figure 3.7. This figure illustrates that income inequality within decade of birth cohorts of the heads of household is declining over time, even after the removal of all village-level income dynamics by means of the model in equation 3.6.

To verify that these declines in inequality are statistically significant, I re-estimate equation 3.1 using the standard deviation of the residuals from equation 3.6, rather than the standard deviation of reported net income, as the dependent variable. The results are presented in table 3.9. The key result that income inequality is decreasing over time within decade of birth cohorts of the heads of household is robust to the removal of all village-level income dynamics by means of the model in equation 3.6.

Table 3.9: Declining Inequality in Residual From Fixed Effect Regression

Dependent variable	*Income Residual*
Cohort born in the 1930s	0.0480*** (-3.09)
Cohort born in the 1940s	0.0743*** (-4.75)
Cohort born in the 1950s	-0.00528 (-0.34)
Cohort born in the 1960s	-0.0131 (-0.85)
Time × born in the 1930s	-0.0207*** (-8.37)
Time × born in the 1940s	-0.0212*** (-12.28)
Time × born in the 1950s	-0.0218*** (-9.64)
Time × born in the 1960s	-0.0197*** (-7.96)
_cons	1.002*** (-58.69)
N	60
R-Squared	0.7882

t statistics in parentheses, robust standard errors, * p<0.10, ** p<0.05, *** p<0.01. The initial conditions are identified as the coefficient on interactions between cohort dummies and dummy that is ‘on’ in 1997 and ‘off’ in all other years. Thus the ‘omitted category’ is the set of observations which are from all years other than 1997. Identification of the four within-cohort time trends is achieved by the assumption of linearity and the omission of an overall time trend.

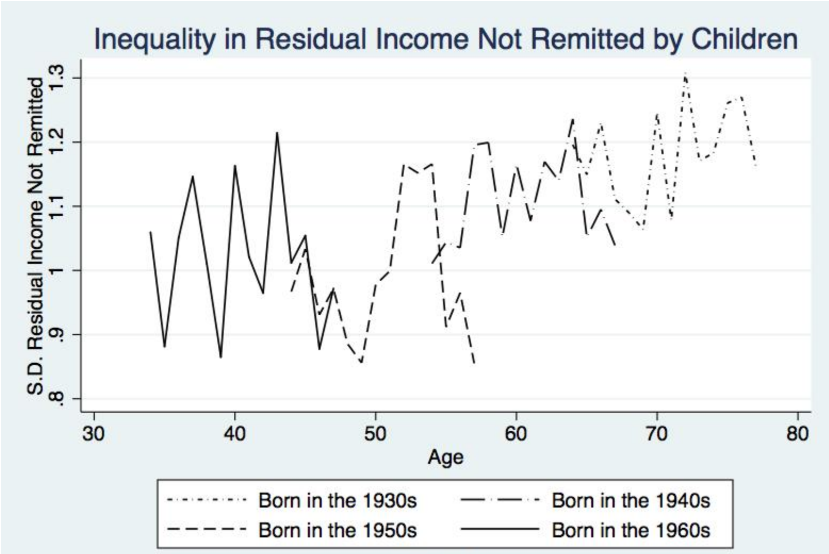


Figure 3.8

Figure 3.7 and table 3.9 have established that the declines in inequality documented in section 3.3 are not driven by village-level income dynamics. I now verify that inequality in the component of income that is not remitted by children increases over the lifecycle, after accounting for these village – time fixed effects. I re-estimate equation 3.6 using the component of income that is not remitted by children to purge this variable of all village-level income dynamics, and check whether the resulting residuals of income not remitted by children continue to increase in the standard way, as they do in figure 3.6.

Figure 3.8 plots the evolution of inequality in the residuals of the component of income that is not remitted by children, once the fully interacted set of village and time fixed effects have been removed. A univariate regression through the points in this graph yields a coefficient of 0.0066 that is significantly different from zero with a t-statistic of 7.69. Thus the finding that the differences between households in the receipt of remittances from their children accounts for the entirety of the convergence in the distribution of income observed in this panel is not driven by differences between villages in the propensity to attract remittances.

Other studies from the Townsend Thai Project have documented sharp differences between households residing in the relatively poor Northeastern region of Thailand and those residing in the relatively rich central region (Samphantharak and Townsend, 2017; Townsend 2013; Pawasutipaisit and Townsend, 2011), as was discussed in section 2.4.5. In light of that literature table 3.10 separates sampled households into two geographic groups and illustrates that the key insights of this paper – that inequality in total income is declining within decade of birth cohorts, but inequality in the component of income that is not remitted by the children of the head of household does not -- apply to households residing in both of these regions.

The first column of table 3.10 estimates equation 3.1 on the subset of households that are samples from the Northeastern region, whereas the second column estimates the same equation on households which were sampled from the central region. In both subsamples, I reject the null hypothesis that within cohort inequality is constant over the duration of the panel in favour of the alternative that it is declining at the 1% level of significance. The third column estimates equation 3.1 on the component of income that is not remitted by the children of the heads of household for households from the Northeastern region. For all cohorts, I fail to reject the null hypothesis that inequality remained constant over the duration of the panel. Finally, the fourth column estimates the equation on the component of income that is not remitted by the children of the heads of household on households from the central region. Here, there is some statistical evidence that inequality is actually rising for the oldest cohort. On the other hand, for the cohort born in the 1950s, there is some evidence that inequality in this component of income is falling, but if one is prepared to relax the standard of statistical significance to the 10% level. When compared with the robust declines documented in the the second column, the overall impression that remittances exert a downward pressure on inequality in these communities is confirmed, regardless of the region from which the households are sampled.

Table 3.10: Inequality Northeastern and Central Region				
Dependent variable	*Net income*	*Net income*	*Income not remitted*	*Income not remitted*
Region	*Northeastern*	*Central*	*Northeastern*	*Central*
Cohort born in the 1930s	0.202*** (7.68)	0.202*** (7.68)	0.281*** (5.49)	0.0388 (0.86)
Cohort born in the 1940s	0.151*** (5.73)	0.151*** (5.73)	0.0972* (1.93)	-0.129*** (-2.84)
Cohort born in the 1950s	0.0504* (1.92)	0.0504* (1.92)	0.204*** (3.99)	0.197*** (4.43)
Cohort born in the 1960s	-0.0984*** (-3.73)	-0.0984*** (-3.73)	-0.343*** (-6.71)	-0.666*** (-14.48)
Time × born in the 1930s	-0.0210*** (-5.70)	-0.0210*** (-5.70)	0.0124 (1.62)	0.0112** (2.17)
Time × born in the 1940s	-0.0293*** (-9.68)	-0.0293*** (-9.68)	0.00488 (0.57)	0.00771 (1.56)
Time × born in the 1950s	-0.0243*** (-6.51)	-0.0243*** (-6.51)	-0.00507 (-0.66)	-0.0121* (-1.96)
Time × born in the 1960s	-0.0201*** (-5.56)	-0.0201*** (-5.56)	-0.00829 (-1.10)	-0.00141 (-0.20)
_cons	1.092*** (37.89)	1.092*** (37.89)	1.132*** (20.13)	1.127*** (22.78)
N	60	60	60	60
R-Squared	0.6873	0.6551	0.2574	0.4418

t statistics in parentheses, robust standard errors, * p<0.10, ** p<0.05, *** p<0.01. The initial conditions are identified as the coefficient on interactions between cohort dummies and dummy that is 'on' in 1997 and 'off' in all other years. Thus the 'omitted category' is the set of observations which are from all years other than 1997. Identification of the four within-cohort time trends is achieved by the assumption of linearity and the omission of an overall time trend.

Are the Results Sensitive to the Measure of Inequality Used?

This paper has used the standard deviation of the logarithm of income to measure inequality, because this is the measure used by much of the literature on the evolution of inequality and the dynamics of income processes (Blundell, Pistaferri and Preston, 2008; Dickens, 2000, among others) from which this paper draws some inspiration. Much of that literature deals with the residuals from regressions of various observable characteristics on the log of income, so that the standard deviation of these residuals arises naturally as a measure of income dispersion. The broader inequality literature however, has employed a range of different measures of inequality, depending on the nature of the particular research question being addressed. Dalton (1920) made the point that a researcher's choice of inequality measure also implies a choice over a social welfare function. Rather than place ad-hoc restrictions on an implied social welfare function, I now demonstrate that the key insights of this paper are robust to a range of different measures of inequality and are therefore not overly sensitive to the choice of social welfare function being considered. I use the statistical package developed by Jenkins (2015) to calculate these different inequality measures and plot their evolution within decade of birth cohorts over the age of the cohort in figure 9. The panels in the left column of figure 3.9 depict the evolution of net household income, whereas those panels in the right column depict the evolution of income that is not remitted by children. The results for different inequality measures are presented in different rows.

The Gini Coefficient

I start with one of the most commonly used measures of inequality, the Gini coefficient. Graphically, the Gini coefficient is the area between the line of perfect equality and the Lorenz curve, divided by the area under the line of perfect equality[11]. It is well known that compared to the standard deviation of the log (which is most sensitive to transfers near the bottom of the distribution), the Gini coefficient is more sensitive to transfers around the mode of the income distribution.

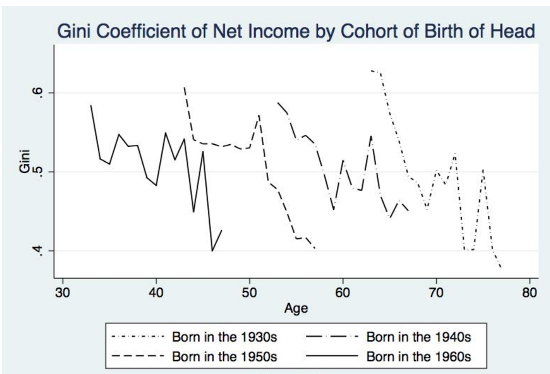


Figure 3.9a

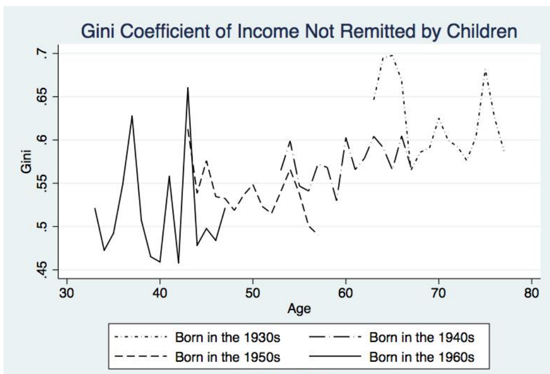


Figure 3.9b

[11] Consider a graph where the y-axis measures the cumulative share of income, and the x-axis measures the cumulative share of the population. If the population were ranked in ascending order of income, then the locus of points traced out by the resulting graph is a Lorenz curve. In this case, the line of perfect equality is the 45-degree line.

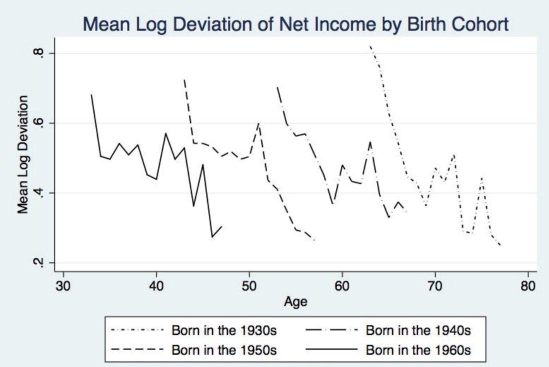


Figure 3.9c

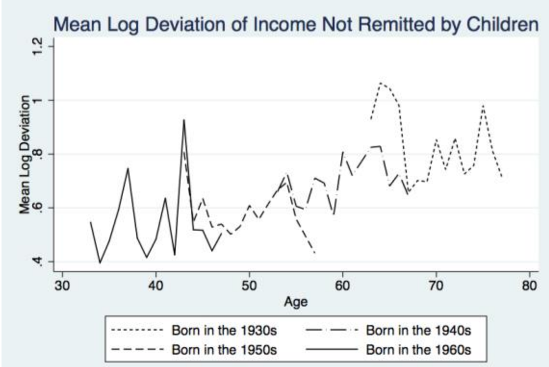


Figure 3.9d

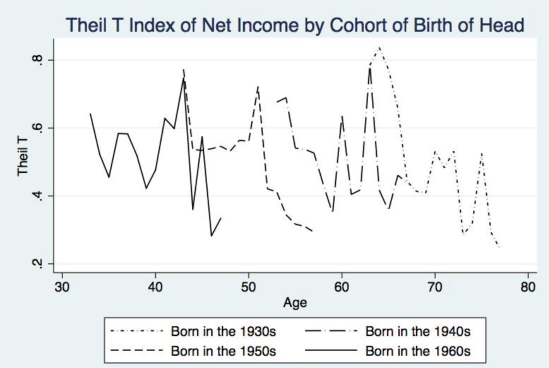


Figure 3.9e

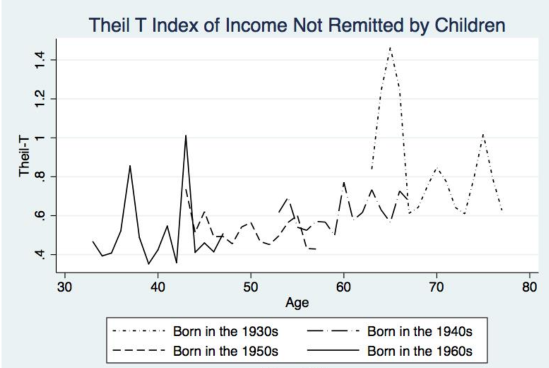


Figure 3.9f

Table 3.11: Inequality in Net Income within Cohorts Using Different Measures

Dependent variable	*Gini*	*Mean Log Deviation*	*Theil-T*
Cohort born in the 1930s	0.0603*** (6.43)	0.217*** (10.61)	0.157*** (4.83)
Cohort born in the 1940s	0.0175* (1.85)	0.0969*** (4.76)	0.0423 (1.30)
Cohort born in the 1950s	0.0377*** (4.00)	0.120*** (5.90)	0.144*** (4.44)
Cohort born in the 1960s	0.0131 (1.42)	0.0732*** (3.65)	0.00681 (0.21)
Time × born in the 1930s	-0.0118*** (-7.00)	-0.0232*** (-7.92)	-0.0224*** (-5.23)
Time × born in the 1940s	-0.00916*** (-7.30)	-0.0194*** (-7.44)	-0.0164*** (-3.31)
Time × born in the 1950s	-0.00999*** (-7.37)	-0.0211*** (-7.78)	-0.0210*** (-5.18)
Time × born in the 1960s	-0.00868*** (-4.97)	-0.0181*** (-5.50)	-0.0155*** (-2.78)
_cons	0.579*** (55.89)	0.626*** (28.06)	0.650*** (18.33)
N	60	60	60
R-Squared	0.6812	0.7425	0.4524

t statistics in parentheses, robust standard errors, * p<0.10, ** p<0.05, *** p<0.01. The initial conditions are identified as the coefficient on interactions between cohort dummies and dummy that is 'on' in 1997 and 'off' in all other years. Thus the 'omitted category' is the set of observations which are from all years other than 1997. Identification of the four within-cohort time trends is achieved by the assumption of linearity and the omission of an overall time trend.

The evolution of income inequality (by decade of birth cohorts of the heads of household in the sample), as measured with the Gini coefficient, is presented in panel (a) of figure 3.9. It is evident from the figure that on this measure, inequality continues to exhibit a declining trend within each cohort. I re-estimate equation 3.1 using the Gini coefficient as the dependent variable and present the resulting coefficients in the first column of table 3.11. This table confirms that the declines are statistically significant within each cohort. To the right of figure 3.9, panel (a), in panel (b) I plot the evolution of income inequality as measured by the Gini coefficient for that component of income that is not remitted by children. As was the case when the standard deviation of the log was used to measure inequality, overall, inequality in this component of income appears to increase over the lifecycle. The only exception is within the cohort born in the 1950s, for which inequality continues to decline, albeit at a much slower rate than in panel (a). Re-estimating equation 3.3 using the Gini coefficient of the component of income that is not remitted by children, in place of the standard deviation of the log yields the coefficients in the second column of table 3.12. The coefficient on age in this regression confirms that the observed overall increase in inequality with the age of the head of household is statistically significant, under this alternative measure of inequality.

Table 3.12: Inequality in Income Not Remitted and Age of the Head Using Different Inequality Measures

Measure	*Gini*	*Mean Log Deviation*	*Theil-T*
Age	0.00286*** (5.54)	0.00837*** (6.31)	0.00909*** (4.54)
_cons	0.405*** (13.36)	0.205*** (2.74)	0.129 (1.22)
N	60	60	60
R-Squared	0.3611	0.3912	0.2502

t statistics in parentheses, robust standard errors, * p<0.10, ** p<0.05, *** p<0.01.

Generalized Entropy Index

This paper has not yet answered whether it is differences in the receipt of remittances from children between or within villages that drive the observed reduction in inequality. Despite their intuitive transparency, neither the standard deviation of the log, nor the Gini coefficient allow inequality to be decomposed in the way that would be necessary to answer such a question. Shorrocks (1980) identifies the set of inequality measures that satisfies this "additive decomposability" property (in addition to the other properties that are necessary to be considered an appropriate measure of inequality). The set of inequality measures Shorrocks (1980) identifies is the Generalized Entropy Index. Two notable members of this family are the mean logarithmic deviation (sometimes called the "Theil-L index"), which is given by the formula:

$$L = \frac{1}{N}\sum_{i=1}^{N} \ln(\frac{\bar{y}}{y_i}), \qquad (3.7)$$

and the Theil Index (sometimes referred to as the "Theil-T index"), which is defined as:

$$T = \frac{1}{N}\sum_{i=1}^{N} \frac{y_i}{\bar{y}} \ln(\frac{y_i}{\bar{y}}), \qquad (3.8)$$

where $\bar{x}$ is the mean value of x in the sample, N is the sample size.

These two measures differ in the region of the distribution of income to which they are most sensitive to transfers: the mean logarithmic deviation is more sensitive to transfers at the bottom of the distribution than the Theil Index.

The second row of figure 3.9 illustrates that the key insights of this paper are robust to using the mean logarithmic deviation to measure inequality. Panel (c) continues to give the impression that inequality in net household income is declining. This decline is confirmed as being statistically significant in the second row of table 3.11, which reports the coefficients that result from re-estimating equation 3.1 using the mean logarithmic deviation as the dependent variable. Panel (d) plots the evolution of inequality in income not remitted by children over synthetic cohorts using this measure of inequality. The overall impression that inequality in this component of household income is increasing linearly with age is also robust to this measure of inequality. Despite the possible presence of some heterogeneity within cohorts (inequality in the beginning of the panel for the cohort born in the 1930s is fairly high when this measure is used), the overall increase in inequality with age is statistically significant, as is demonstrated in the second column of table 3.11, which reports the coefficients yielded by re-estimating equation 3.3 using this measure of inequality.

The pair of diagrams in the third row of figure 3.9 use the Theil (T) coefficient to measure inequality. The declines in inequality in net household income appear noticeably less pronounced than before. Nonetheless, when I re-estimate equation 3.1 for this measure of inequality, I continue to formally reject the hypothesis that inequality remains constant over time, in favour of the alternative that it is decreasing within each cohort. The relevant coefficients are reported in the third column of table 3.10. Inequality in household income not remitted by children again appears to increase linearly using this measure of inequality, too. Again, there is some evidence of heterogeneity between cohorts, though as the third column of table 3.12 demonstrates, the overall increase is statistically significant.

As noted above, the mean logarithmic deviation is more sensitive to transfers near the bottom of the income distribution, compared to the Theil (T) index. Thus differences in the evolution of income inequality between panels (c) and (e) may be informative about the part of the income distribution where inequality reduction is occurring. Graphically, the decline in inequality in net household income appears to be more pronounced for the mean logarithmic deviation, than the Theil (T) index, though formally we cannot compare the declines in this way. Nonetheless, this suggests that the transfers that are leading to inequality reduction are disproportionately affecting the bottom of the income distribution. This impression is corroborated when I use the Atkinson's (1970) inequality index, and vary the "inequality aversion parameter", which determines that index's sensitivity to transfers at the bottom of the distribution, though I do not report those results here for the sake of brevity. The declines in inequality in net income are more pronounced when the inequality aversion parameter is assigned a value of 2, as compared with when it is assigned a value of 1, supporting the narrative that inequality reduction is greater in the left tail of the income distribution.

This section has demonstrated that the reduction in income inequality documented in the preceding section can be fully explained by remittances from the adult children of the heads of household. Then, it demonstrated that the inequality reducing effect of remittances was not driven by heterogeneity between villages in the propensity to receive remittances, but was rather a feature of the distribution of income between households, within villages. It went on to verify that these results are not unduly sensitive to the choice of inequality measure. Indeed, to the extent that the results differed between inequality measures, these differences suggested that inequality was being reduced by transfers affecting income the bottom of the distribution.

3.5 The Distribution of Remittances

The preceding section had established that the remittances from adult children of the heads of household were driving a reduction in income inequality among this panel of Thai village households. This section will study the characteristics of the distribution of remittances from children, between households and over time, that enable it to reduce inequality.

3.5.1 The Distribution of Remittances Over Time

Figure 3.10 shows that the mean of the real value of remittance received by households in the panel from the children of the heads has been increasing from 1997 to 2011. These figures are of course, the outcomes of a variety of different forces. The children of those households which continue to be headed by the same person between 1997 and 2011, are ageing and so more of them enter the pool of potential remitters to their parents' household. In other households, either death or retirement may cause the head of household from an older generation to be replaced by one from a younger generation. In such cases, the remitting children of the former head may no longer be obliged to provide support to what may have become their sibling's household. Due to such heterogeneity, it is important to decompose remittance receipts into the familiar lifecycle components.

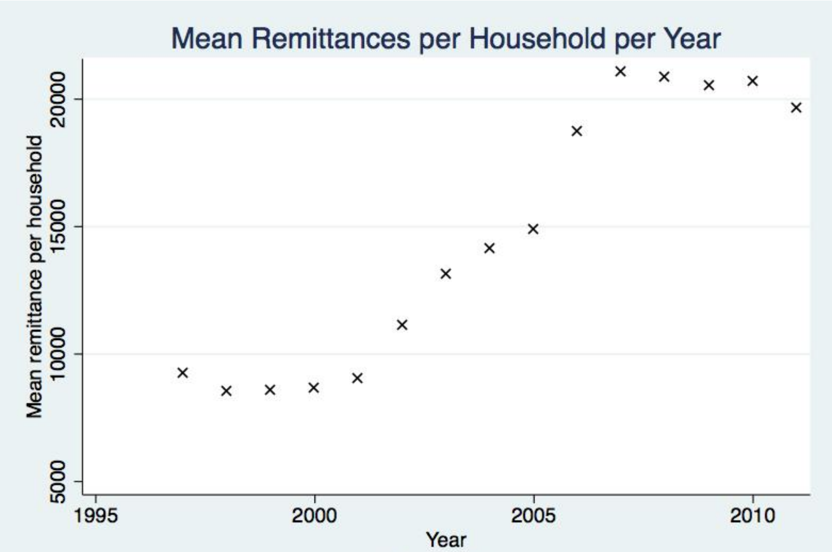


Figure 3.10

Figure 3.11 plots the proportion of household income that is accounted for by remittances from children, over the lifecycles of the heads of household. Unsurprisingly, remittances from children form an especially large proportion of the incomes of the households headed by the elderly. The figure illustrates that remittances from children start to gain importance as household heads enter their early 40s. Their share in household income continues to increase until the heads of household reach their late 50s or early 60s. On average these transfers continue to account for approximately 30% of household income for the remainder of the lifecycle of the heads of household. Appendix 7 reconciles the dip in the proportion of household income derived from children's remittances in the last four years of the panel, illustrated in figure 3.11, with the lack of any corresponding dip in the overall level of remittances illustrated in figure 3.10. It finds that these are driven by rapid growth in rural incomes continuing while remittances remain at their peak.

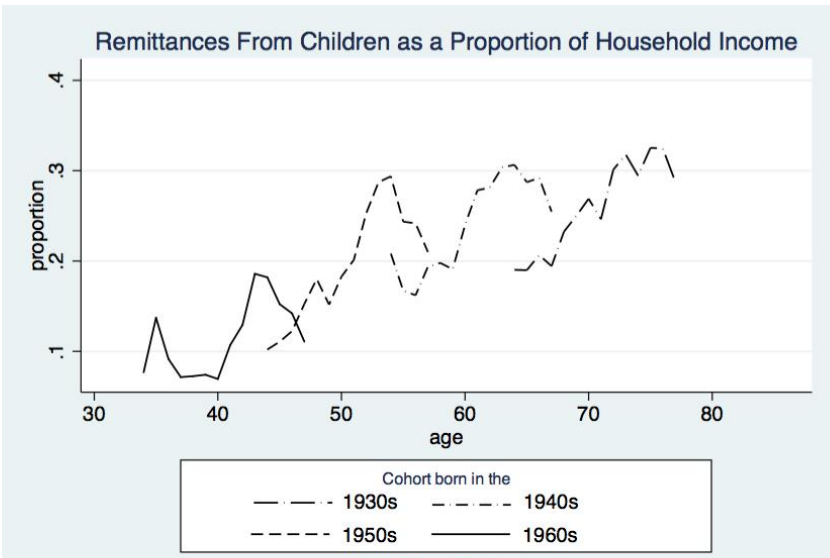


Figure 3.11

The importance of remittances from children in the later stages of the lifecycle of the heads of household has implications for lifecycle savings: children can act as a substitute to 'hump shaped' lifecycle savings not only by living in the same household as their parents throughout adulthood as, for example, documented by Deaton (1989), but also by remitting money to their parents even in the case that they do not reside with their parents. Thus remittances not only help insure households from potential covariate risk that would otherwise be uninsurable as noted by Stark and Bloom (1985), but also against low productivity later in the lifecycle. As a result, the elderly may continue to receive a great deal of financial support from their children, even though the number of adults living together in East Asian households has been declining (Lo, 1987, cited in Deaton 1989; Deaton and Paxson 1995).

Having understood how remittances, on average are distributed over time and over the lifecycle of the heads of household, I now study the distribution of remittances from children across households with differing levels of permanent income, where I use the average over time of each, individual household's log of real, equivalised consumption as a proxy for permanent income. The strengths and potential weaknesses of this approach were discussed in section 2.4.4, of the preceding chapter.

3.5.2 Remittances across Households by Decile of Permanent Income

The Proportion of Income Accounted for by Remittances by Decile of Permanent Income

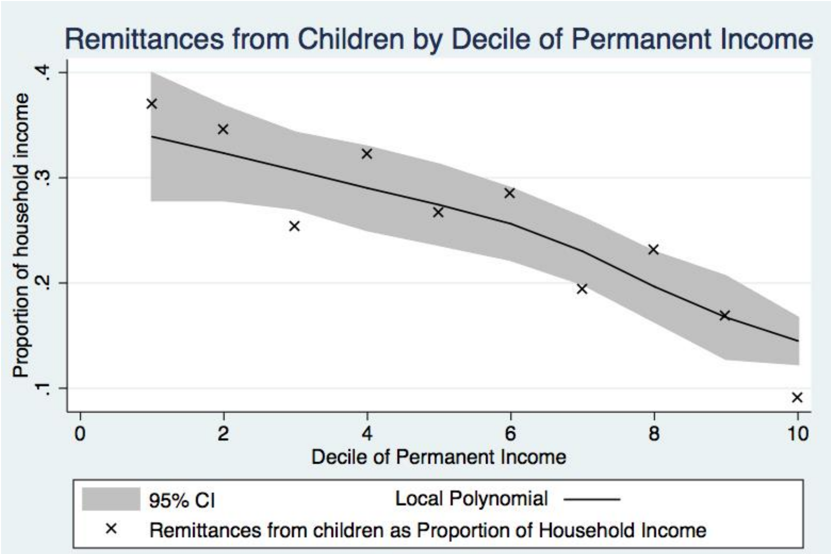


Figure 3.12

Figure 3.12 plots remittance receipts from the children of the head of household as a fraction of net household income, at each decile of household permanent income. For the first decile, remittances account for over 35% of household income; for the fifth decile this declines to a little over 25%, whereas for the tenth decile they account for less than 10%. I have plotted an estimated local polynomial and the associated 95% confidence bands over these points. The decline in the fraction of household income that is accounted for by remittances from children with a household's level of permanent income is clearly statistically significant.

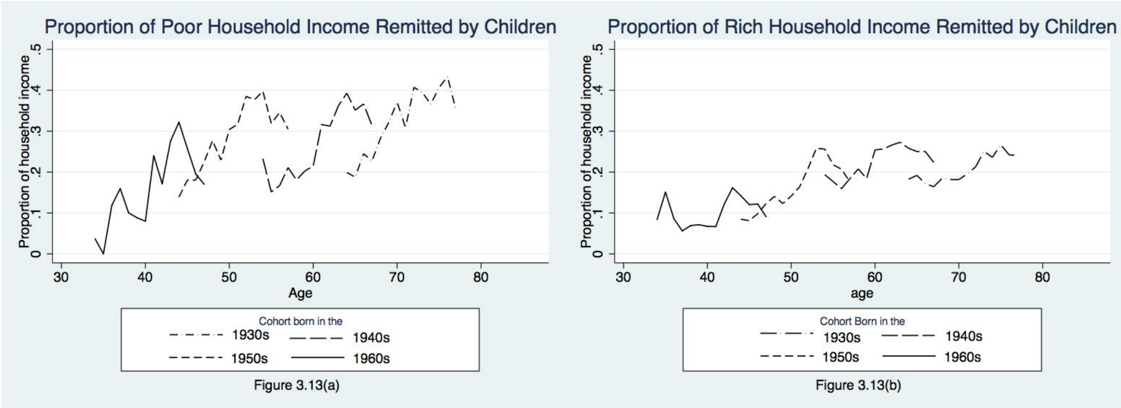


Figure 3.13(a)

Figure 3.13(b)

The fact that in this sample, the heads of poorer households are older on average (a finding that is discussed in chapter 2), raises concerns that this figure may be conflating lifecycle effects with those that vary across the distribution of income. For this reason, it is important to make comparisons not just over levels of income permanent income, but within cohorts of birth. If data permitted, I would study the evolution of remittance receipts from children over the lifecycles of the heads of household at each decile of the distribution of household permanent income. However, the need to maintain reasonably large cohort-year cell sizes restricts me to breaking the sample into two parts: households with above and below median permanent income.

Figures 3.13a and 3.13b illustrate how the proportion of household income that is accounted for by remittances from children evolves over the lifecycle for relatively poor and relatively rich households, respectively. On average, the proportion is higher for poorer households and over the course of the panel, it grows much more rapidly for these households. For richer households, remittances from children do not account for more than 30% of household income for any group. In contrast, poorer households headed by anyone older than 50 received over 30% of their household income in the form of remittances from children in every year after 2004. These differences are striking and all indications are, that they would be even more striking at the extremes of the distribution of permanent income, if the data permitted me to observe them.

The Total Number of Children by Decile of Permanent Income

The larger share of remittances in the incomes of relatively poor households may be associated with higher rates of fertility in these households. Indeed, figure 3.5 (which plotted the number of cohabiting children over the lifecycle of the heads of household) found that richer and poorer households had similar numbers of children until their heads reached their late 30s. After this point, however, the number of cohabiting children in poorer households continued increase until these heads reached their mid 40s, while for richer households it did not. These descriptive statistics suggest that poorer households exhibit higher rates of fertility overall, than their richer counterparts.

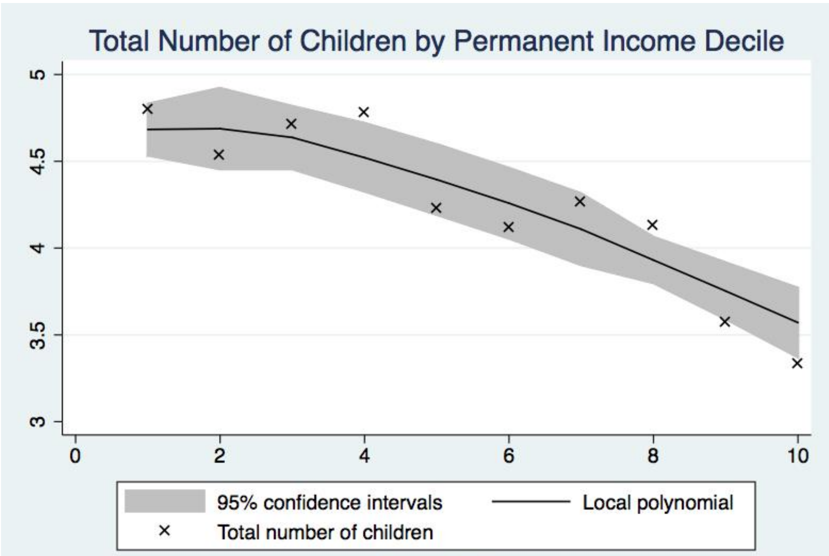


Figure 3.14

This is confirmed in figure 3.14, which plots total number of children of these heads of household by decile of permanent income. I also plot a local polynomial and the associated 95% confidence intervals. A household in the bottom decile of permanent income can be expected to have approximately 4.7 children, whereas one in the top decile of permanent income may be expected to have only 3.6 children. Thus the households in the poorest decile have, on average, more than one additional child compared to those in the richest decile. The plotted 95% confidence intervals illustrate that this difference is statistically significant. A larger number of children constitutes a larger pool of potential remitters. However, for differential fertility to explain the inequality reducing role of remittances, the number of observed remitters (as opposed to the number of observed children) need to be higher among poorer households. I now turn to establishing whether or not this is the case.

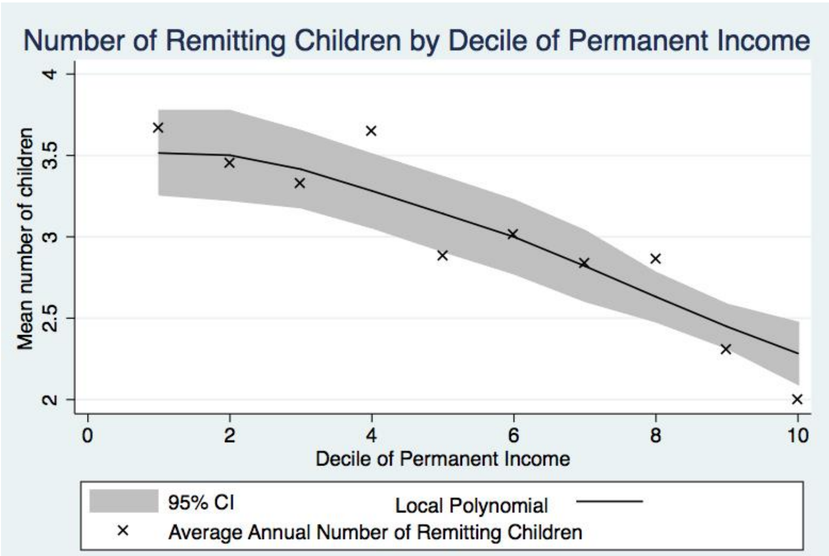


Figure 3.15

The 'Quantity' and 'Quality' of Remitters by Decile of Permanent Income

The redistributive effect of remittances from children on the income distribution of their parents' households may be due to the fact that poorer households have a larger number of children to support them later in the lifecycle, differences in the amount remitted by each child, or a combination of these forces. Figure 3.14 demonstrated that poorer households had a larger number of children overall, but figure 3.5 illustrated that fewer of these children tended to reside with their parents throughout adulthood. Thus the number of potential remitters to these households is larger than it is for their richer counterparts.

Figure 3.15 plots the average of the actual number of different children from which a household receives remittances in each year, by decile of household permanent income. The differences are economically (and statistically) significant: the typical household in the 1st or 2nd decile of the income distribution receives remittances from approximately 3.5 children, as compared with households in the 9th and 10th decile, which typically receive remittances from less than 2.5 children. Thus differences in the number of remitting children across the distribution of household permanent income may indeed be an important channel through which remittances cause the distribution of income to converge over time.

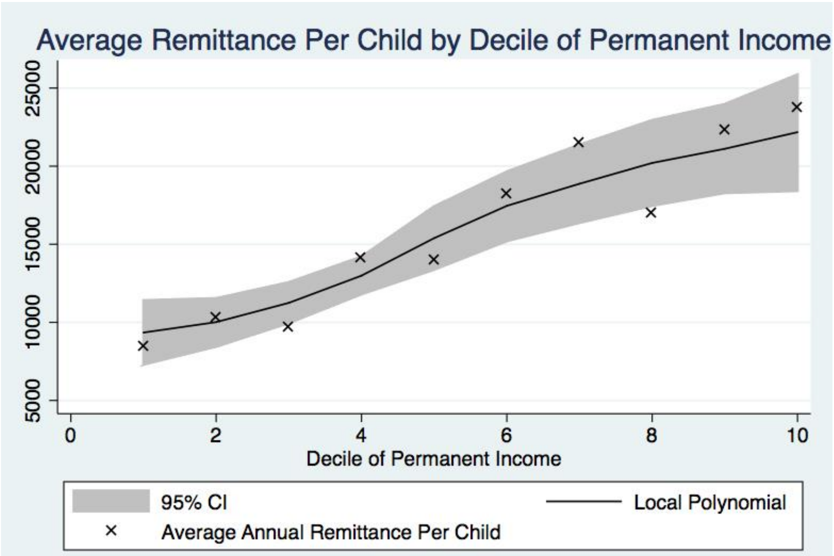


Figure 3.16

I now attempt to ascertain whether there are differences in the ‘quality’ of the average remitter, over the income distribution. Figure 3.16 plots the average real amount remitted by each child in each year, by decile of household permanent income. It is clear that on average, each child from a household with relatively high permanent income remits more per year than one from a household with relatively low permanent income. A linear regression of a household's percentile in the income distribution on the mean amount remitted by each child suggests that, on average, a 1 percentage point improvement in a household's percentile in the income distribution, is associated with an increase in the amount of money remitted by each child annually of 148.4 Thai Baht, at 2011 prices. The null hypothesis that this coefficient is equal to zero is rejected, in favour of the alternative that it is positive, with a t-statistic of 12.17. Thus the convergence in the distribution of household income documented in section 3.3 cannot be explained by this absolute measure of 'quality' of remitters. However, it is not the absolute value of remittances that is important in explaining the evolution of income inequality, but the value relative to household income.

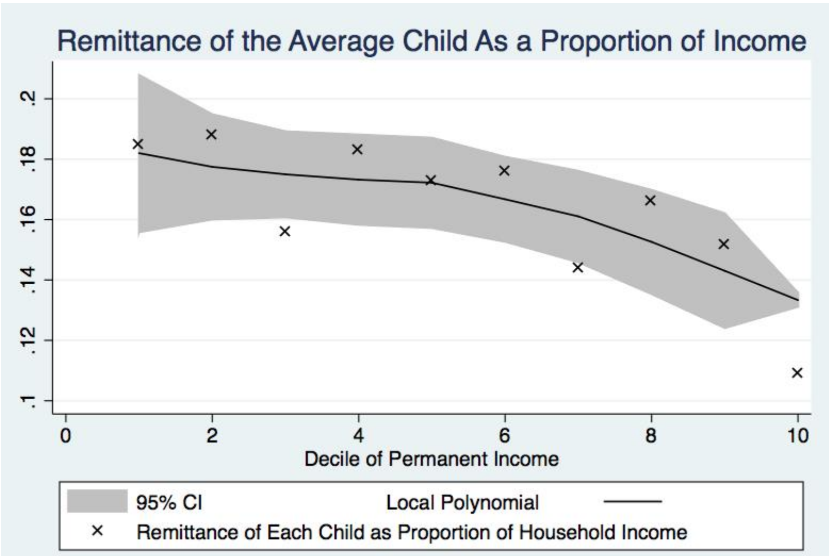


Figure 3.17

Figure 3.17 shows that remitters from poorer households tend to outperform their richer peers, when the average amount remitted each year by each remitter is expressed as a proportion of net household income, though the confidence intervals around the estimated local polynomial do not suggest that the difference is statistically significant. On the other hand, a linear regression of each household's percentile in the income distribution on the mean proportion of household income that is remitted by each child (and a constant) is able to discern a statistically significant difference, because such a test has higher statistical power due to the fact that it uses all the variation in the available data, as opposed to figure 3.17 which collapsed the variation into deciles. The estimated coefficient in this regression suggests that a 10 percentage point improvement in a household's percentile in the distribution of permanent income is associated with a 0.537% reduction in the proportion of household income that is transferred by the average remitter. The null hypothesis that this effect is equal to zero is rejected with a t-statistic of -4.86. As such, differences in the relative magnitude of transfers from individual remitters do help explain the reduction in income inequality documented in the preceding section.

Inequality and the Gender of Migrant Children

Section 3.3 noted that female adult children of the heads of these Thai households on average remit more than their male children. This could be associated with the inequality reducing effect of remittances in a number of ways. On the one hand, if the generosity of female remitters, relative to their male peers were constant across the distribution of income, then poorer households producing a greater number of female migrants may be a channel through which remittances from children of the head of household reduce inequality. On the other hand, if the daughters from richer households, who presumably earn more than their poorer counterparts, are more likely to select into migration outside the village, then this would explain the finding that on average women remit more, but in a way that is not associated with falling inequality among the households of origin.

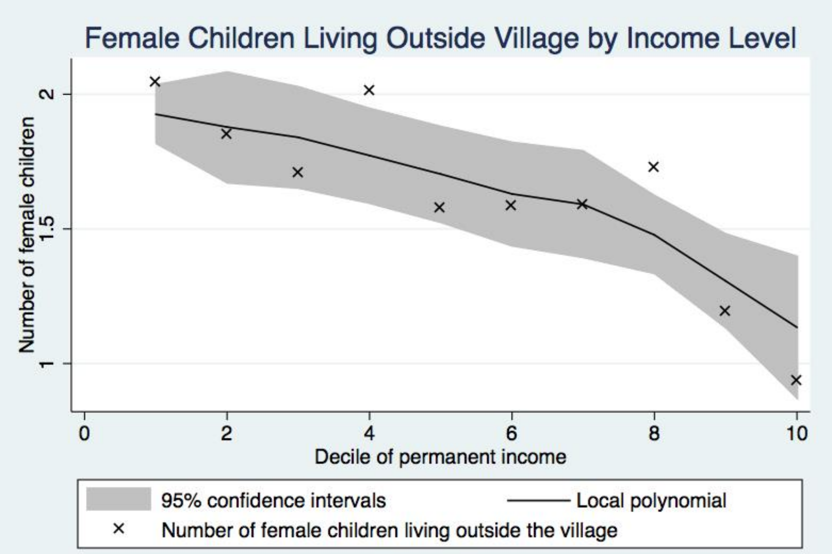


Figure 3.18

Figure 3.18 plots the association between the number of female children of the head of household who reside outside the village by decile of permanent income. The 95% confidence intervals around the local polynomial suggests that poorer households do indeed have significantly more female children who live outside the village of origin than their richer counterparts. It is still possible that these women migrate for reasons that are unrelated to remittances, for example, because of the structure of the marriage market. Unfortunately, the information collected in the Townsend Thai Project does not catalogue the stated reason for leaving the village. Thus I cannot distinguish between women who left the village to seek employment and those who left due to marriage. I can however, observe whether or not this greater number of female migrants among poorer households is also reflected in a greater number of female remitters to those households.

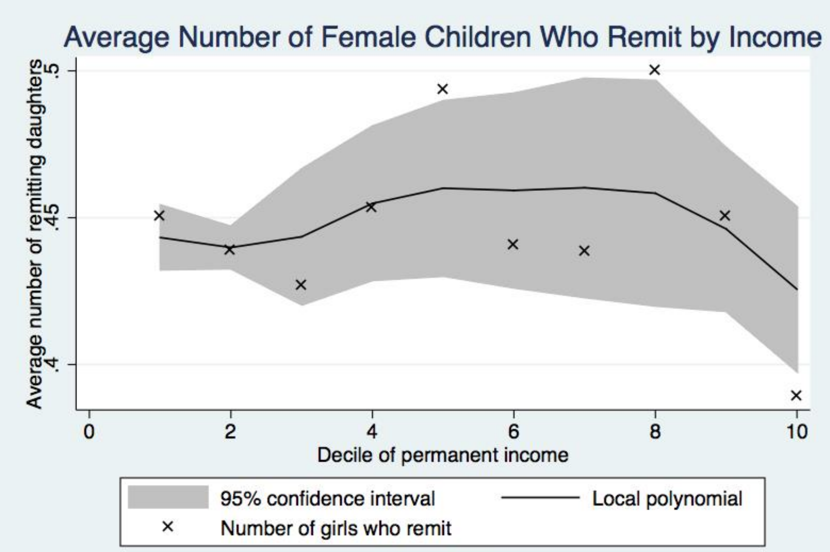


Figure 3.19

Figure 3.19 plots the average number of female remitters by the decile of permanent income of the household of origin. The scatter points do not appear to display any clear pattern, so that despite producing a greater number of female migrants, poorer households do not produce more female remitters. The remaining women migrants are likely to be absorbed by the marriage market, though the available data do not allow me to rule out the possibility that they are unsuccessful economic migrants.

Thus while the average woman remits more to the household of her parents than her average male sibling, poorer household are not any more likely to produce female remitters in a way that would explain the convergence of inequality documented in this paper.

The analysis has thus established five aspects of the distribution of remittances which can explain their ability to substantively reduce household income inequality among these rural Thai households:

1. They constitute a large share of household income: the average of the proportion of net household income that is accounted for by remittances is approximately 25%.
2. They increase in importance later on in the lifecycle of the heads of household, accounting for 10% of household income, on average, when heads are in their 40s, a figure which rises to approximately 30%, by the time heads reach their 50s.
3. They comprise a larger proportion of the incomes of poorer households, accounting for approximately 35% of the income of households in the bottom decile of the income distribution, and less than 10% for those in the highest.
4. The heads of poorer households receive remittances from a significantly larger number of children. Households in the bottom decile of the distribution of permanent income receive on average receive remittances from 3.5 children, whereas those in the top do so from and average of fewer than 2.5 children.
5. Among the relatively poor, the average amount remitted annually by each child, constitutes a greater proportion of household income than it does among the relatively rich.

3.6 Conclusions

This paper has documented and analysed declining income inequality in a panel of Thai households with reference to the permanent income hypothesis and lifecycle theory. It noted on theoretical grounds that decreasing inequality was unlikely to result from a convergence in the distribution of incomes that was exogenous to the households. To the extent that the data permitted, it verified that the incomes of individual household members diverged in the standard way. It speculated that the number of children who continued to live with their parents throughout adulthood was likely to vary endogenously with household income. If children from poorer households were more likely to stay on in their parent's households thereby supplementing household income, this choice might explain the observed reduction in inequality, but the paper failed to find any support for this hypothesis in the data.

The paper went on to observe that cohabitation was not the only strategy available to parents wishing to secure support from their children, later on in the lifecycle. It showed that village level income and remittance dynamics were not sufficient to explain the decline in inequality, and that these declines were robust to a number of different measures of inequality. Furthermore, the paper showed that differences in remittances received from children over the distribution of income account for the entirety of the reduction in income inequality observed between these households. The effectiveness of remittances in reducing income inequality in this context is associated with five features of the distribution of remittances from children between households: first, they constitute a large fraction of household income; second, they account for a larger proportion of the incomes of poorer households; third, their share of household income increases later on in the lifecycle of the recipients households causing inequality to fall as households age over the panel; fourth, poorer households receive remittances on average from a larger number of children; and finally, the average child from a poorer household remits a larger fraction of household income per year than one from a richer household.

This paper's primary contribution is to the literature on the effect of migration on income inequality in migrant producing areas. The perspective however, has been different from that of much of that literature: rather than ask whether migration 'caused' a decline income inequality, the paper has sought to establish that the observed reduction in inequality is explained by the effect of migrant's remittances.

This paper has highlighted the importance of remittances from children in insuring the income streams of their parents against reduced productivity later in the lifecycle. The quality of the insurance yielded by this strategy is striking – it appears to be sufficient to reverse the typical increases in inequality that have been documented in a wide variety of countries, as predicted by lifecycle theory.

This evolution of the household as an institution that pools income from multiple adults to insure itself against productivity declines of individual members, is likely to further enhance the income growth of younger generations. Where cohabitation required high productivity members to live within the same household and therefore work within geographical proximity to that household, exposing their income streams to many of the same sources of risk as other household members, supporting parental households through remittances allows high productivity members of the household the choice to be geographically mobile. If there is significant spatial heterogeneity in the returns to labour, this mobility may well enhance the earning capacity of these members.

These findings have important implications for policy. Retirement benefits and other support from the government was found to be extremely rare among these Thai villages. Insofar as a wider provision of these services may crowd out these private transfers, their effect on the welfare of the elderly may be muted. Indeed, given that the private transfers are sufficient to actually engender a reduction in inequality, the sign of effect on the welfare of some of the elderly may even be debatable. On the other hand, elderly headed households that are unable to draw support from their children, including remittances, would indeed be made unambiguously better off by the introduction of such policies.

Chapter 4

Income Inequality and the Extensive and Intensive Margins of Remittances

Introduction

The preceding chapter demonstrated that income inequality was declining over time in a panel of households surveyed by the Townsend Thai Project, and demonstrated that the receipt of remittances from the adult children of the head of household accounted for the documented decline in inequality. Descriptive analysis presented in that chapter identified two characteristics of the distribution of remittances across households that may be associated with the ability of these transfers to reduce income inequality among recipient households. First, poorer households were more likely to receive remittances than their richer counterparts; and second, conditional on receiving remittances, transfers accounted for a larger proportion of household income in poorer households than in richer ones. The objective of this paper is to ascertain the relative importance of each of these margins in explaining the inequality reducing role of remittances.

To achieve this objective, this chapter models both the likelihood of receiving remittances and the proportion of remittances in household income as functions of observable household characteristics. Using these econometrics, the paper constructs counterfactual distributions of income where one of these margins is allowed to vary at a time, while the other is held constant. Depending on the measure of inequality used, differences in the likelihood of receiving remittances at different points of the income distribution account for between 52% and 55% of the reduction in inequality that can be explained econometrically by these margins. Conversely, differences in the proportion of remittances in household income explain between 45% and 48% of these reductions.

The key econometric issue that is addressed in this chapter is the potential endogeneity of a subset of the explanatory variables. In regressions that do not make allowances for endogeneity, the levels of household, agricultural and business assets (which will be defined in greater detail below), are found to be significant predictors of remittance receipts. By using village-level information on historical migration episodes as instruments for current remittance levels, I am able to identify the effect of remittances on asset levels, thereby addressing the initial endogeneity problem. After I have addressed endogeneity, asset levels no longer retain any significant explanatory power, suggesting that their predictive power in the single equation models was spurious. I also use the lagged number of households affected by a drought in each village as an alternative instrument for current remittances to check that these results are not unduly sensitive to the choice of instrument. Doing so does not cause a model that includes assets to substantively outperform one that does not.

The remaining, exogenous explanatory variables allow me to model the extensive and intensive margins of remittance receipts at the household level. By comparing these estimates across the distribution of income, I find that these margins do indeed vary systematically with income. I identify this systematic variation and exploit it to construct the counterfactual distributions of income described above.

These models are only capable of explaining the part of the effect of remittances on income inequality that is related to the observable characteristics of households. Some of the inequality reducing effect of remittances will be driven by unobservables, and will therefore remain unexplained by the counterfactual distributions of income that I construct. The performance of these models in explaining the inequality reducing effect of remittances will depend on the measure of inequality that is used to assess them. Rather than select an arbitrary measure of inequality (and implicitly, an arbitrary social welfare function) I assess these models using a range of different measures of inequality. The models perform best for the Theil (T) index, for which they can explain 47.9% of the observed decline in inequality, and they perform worst for the standard deviation of the log, for which they can only explain 4.7% of the measured decline. For the mean logarithmic deviation, and the Gini coefficient, the models are able to explain 17.2% and 24.9% of the measured declines in inequality, respectively.

The chapter is organised as follows. The next section details some fundamental challenges to identifying the effect of remittances on the incomes of recipients. Section 4.2 describes the data that will be used. Section 4.3 discusses the identification strategy that will be used to address potential endogeneity between remittance receipts and some of the observable household characteristics. Section 4.4 presents an econometric model of the probability with which a household receives remittances. This is complemented by the analysis in Section 4.5, where I model the proportion of household income received from each remitter. Section 4.6 conducts sensitivity analysis using an alternative instrument. In Section 4.7 I use the household-level predictions from the preceding models to construct counterfactual distributions of income which allow me to assess the extent to which the extensive and intensive margins of remittance receipts account for the inequality reducing effect of remittances, for a range of different measures of inequality. Section 4.8 concludes.

4.1 The Economics of the Counterfactual

The identification of reliable counterfactuals in the context of remittances and the incomes of recipient households may be impeded by a number of factors. First, past migration decisions may affect both the costs of current migration and the current wealth and income prospects of a household, so that observed household characteristics are not exogenous to the migration process. Second, households select into migration, so that those which do not receive remittances do not provide an appropriate counterfactual for those that do. Third, a household is likely to change its behaviour in response to the receipt of remittances. Not all of these issues can be comprehensively addressed using observational data, though I do adopt some measures to mitigate their effects.

Mckenzie and Rappoport (2007) address the endogeneity of household and community characteristics using information on historical migration between the US and communities in rural Mexico. The Townsend Thai data collects information on village migration histories which I am able to use to instrument for current migration propensities to address this source of endogeneity.

McKenzie et al. (2010) very effectively illustrate the issue of selection, using random assignment via a ballot which New Zealand authorities used to allocate a fixed, but oversubscribed quota of migration opportunities to Tongans. Not only do the authors use this random assignment to identify the effect of migrating from a poor country to a rich one on the earnings of migrants, but they also compare these effects to those that would have been estimated by non-experimental methods. By interviewing people who applied for the ballot and a random selection from the general population, the authors find that potential migrants are positively self-selected on observable characteristics such as education and past income. Furthermore, by comparing estimates from the lottery with those yielded by non-experimental techniques, the authors find that the latter consistently overstate the effect of migration on earnings. They interpret these results as evidence that migrants also positively self-select on unobservables such as ability.

In contrast to McKenzie et al. (2010), this chapter focuses on assessing the effect of remittances on income inequality, so that biases in the estimated effect of remittances on the level of income will pose a threat to my results insofar as these biases affect households differentially by income level. The large share of the inequality reducing effect of remittances that I am unable to explain on the basis of observable characteristics is an indication that unobserved heterogeneity in the Thai sample plays an important role in the relationship between remittance receipts and household income.

Finally, the possibility that the members of remittance receiving households may change their behaviour, say in labour market participation decisions, in response to receiving remittances is also difficult to address in the absence of random assignment. What is more, chapter 2 demonstrated that richer and poorer households in this sample significantly differ in the number of jobs held by the head of household and the number of active participants in the labour force. These systematic differences in labour market outcomes between households at different points in the income distribution suggest that these households will, very likely, differ in their behavioural responses to remittance receipts. This poses a potentially serious challenge to the identification of the effect of remittance receipts on income inequality. Indeed, McKenzie et al. (2010) argue that this challenge can only be robustly addressed using random assignment.

Cognizant of these potential limitations, I proceed with analysing the inequality reducing effect of remittances among these Thai households with the available data.

4.2 The Data

This chapter will again utilize observations from the Townsend Thai Project (Townsend, 2011) that are drawn from the 64 villages which were surveyed in all 15 years of the project. It will model the probability that households receive remittances, and the proportion of household income that is accounted for by remittances. To do so effectively requires the use of a relatively large number of conditioning variables. For this reason, I use the full sample of (approximately) 960 households that are interviewed in each period,

Table 4.1: Summary Statistics

	Observations	*Mean*	*Standard Deviation*	*Min*	*Max*
Whether or not the household receives remittances	14395	0.509	0.500	0	1
Real value of remittances from children	14397	14,924.240	37,361	0	1,096,907
Real value of net household income[a]	14393	155,449.100	273,349	-1,165,068	12,050,222
Raw proportion of remittances in net income	14389	0.170	0.407	-19.075	10.370
True proportion of remittances in income for remittance recipients	7320	0.334	0.301	0	1
Age of head of household	14365	54.906	13.135	16	103
Head earns monthly or government wages	12754	0.064	0.245	0	1
Head is a rice farmer	14397	0.317	0.317	0	1
Head farms crops other than rice	14397	0.078	0.269	0	1
Head raises livestock	14397	0.065	0.246	0	1
Head is a blue collar professional	14397	0.089	0.285	0	1
Head is a white collar professional	14397	0.030	0.170	0	1
Head in jobs that vary	14397	0.110	0.313	0	1
Head is inactive	14397	0.113	0.316	0	1
Head has other occupation	14397	0.154	0.360	0	1
Household owns a business	12754	0.699	0.459	0	1
Whether the household is headed by a female	14397	0.306	0.461	0	1
Years of schooling of the head	14325	4.410	3.268	0	19
Number of non-working age members	14394	1.769	1.295	0	13
Total number of adult children	14394	2.904	2.455	0	14
Mean years of schooling of all adult children	11408	8.812	3.859	0	19
Household assets	14397	50,366.110	114,075	0	1,616,400
Agricultural assets	14397	28,119.150	98,468	0	2,194,591
Business assets	14397	30,569.620	171,872	0	4,277,450

[*] *31.5147 Thai Baht = 1 U.S. Dollar on 31st Dec. 2011 (source: exchangerates.org.uk)*

rather than the 609 households that comprise the balanced panel.

The first row of table 4.1 presents the summary statistics for a dummy variable that takes on a value of 1 if a household receives remittances in a particular period, and 0 if it does not. Half the household-year observations report receiving remittances from the children of the head of household.

Naturally, I calculate the proportion of remittances in household income by dividing the total amount of remittances received from the children of the head of household by the total amount of household income. The summary statistics for the real value of remittances from children, inflated to 2011 Baht, are presented in the second row of table 4.1[1], whereas those for net household income appear in the third row.

The fourth row presents the summary statistics for the proportion of remittances calculated by dividing the amount of remittances from children by net household income. It is important to note here that because of the way net household income is calculated, the share of remittances is not necessarily bounded between 0 and 1. Because the agricultural or business expenses of some households are so great and their incomes are so small, net income in a period can fall below zero. In the sample, there are 120 households with negative net income. While realizations of negative net income are perfectly plausible in this setting, such observations confound the interpretation of the ratio of remittances to income as a simple proportion. Since such an interpretation is crucial the exercise I plan on doing here, I address this issue in the following way.

42 of these households report receiving some remittances from their children. Because these households are wholly dependent on remittances, I assign a value of 1 to the proportion of remittances in household income. The remaining 78 of these households do

[a] These summary statistics are calculated inclusive of those households that do not receive remittances.

not report receiving remittances from children, and so I assign a share of 0 to the proportion of remittances in their income.

In some circumstances, net household income (including remittances) is positive, but less than the value of remittances. These cases lead to the observations in the fourth row of table 4.1 which are greater than unity. In the sample there are 221 such households. Because these households are effectively wholly dependent on remittances in these periods, I assign the share of remittances in their incomes to equal 1.

The summary statistics for the share of remittances in household income (for households that receive remittances) that arises after these adjustments have been made are presented in the fifth row of table 4.1. The average of the share of remittances in household income is 33.37%, among households that receive remittance.

A number of observable household characteristics are likely be associated with the ability of the heads of households to secure remittances from their children. It has been reported elsewhere in this thesis that the likelihood of receiving remittances varies systematically over the lifecycle of the head of household, so it will be important to condition on the age of the head of household. The summary statistics for the age of the head are presented in the sixth row of table 4.1. The average head is approximately 55 years old, while the youngest is 16 and the oldest is 103.

The type of employment contract the head is in may be related to their retirement strategies and so may influence their ability to attract remittances from their children. Therefore, I use a dummy variable for whether or not the head earns a monthly wage in their primary occupation. According to row 7, approximately 6% of heads are in this type of wage contract.

The occupation of the head may also influence retirement strategies. For this reason, I include a set of dummies for different occupation types. These categories include rice farming, farming a crop other than rice, raising livestock (including cows, pigs, poultry, shrimp and fisheries), blue collar professionals (working for example, in construction, as a rice miller, a factory worker or a mechanic), white collar professionals (including clerical workers, those in managerial positions, nurses, teachers, etc.), jobs that vary over the year, heads who are inactive, and those who have some other primary occupation. I note here, that there are 1,625 household-year observations where the head of the household is 'inactive'. Unsurprisingly, inactivity of the head of household perfectly predicts the receipt of remittances from children. These observations would frustrate the computation of the intended counterfactual distributions of income, and so need to be dropped from the empirical models.

Ownership of a business is very common among the sampled households. To ascertain whether or not this phenomenon is related to remittance receipts, I include a dummy variable for business ownership among my regressors. The summary statistics for this variable are also presented in table 4.1

Female headed households may exhibit different propensities to attract remittances than their male headed counterparts. I capture this using a dummy variable, for which I also present the summary statistics in table 4.1. The education level of the head of household may also be an important determinant of whether a household receives remittances or not. It is plausible that even after conditioning on the type of work and the type of employment contract, high levels of education are associated with different retirement strategies. For this reason, I use the years of education of the household head as an explanatory variable, and also present the summary statistics of this variable in table 4.1.

Household demographics may be crucially important for explaining remittance receipts. Households with high fertility rates have a larger pool of potential remitters, and so may be more likely to receive remittances. Households with a large number of members who are not of working age may be more dependent on remittances than those with fewer dependents. The pool of potential remitters to a household will be determined by the total number of adult children of the head of household, including both children who live within the household and those who live away from the household. For the purposes of the chapter, I take ‘adult’ to mean over 18 years of age. The mean level of education among these adult children may also be an important determinant of remittances. The summary statistics of these variables are also presented in table 4.1.

Finally, the stocks of assets available to the household may influence remittance behaviour among children within that household. Households that are particularly asset-rich, may retire on the returns to these assets and therefore demand fewer remittances from their children. Alternatively, these assets may provide a bequeath motive for children to remit to their parents. The Townsend Thai project collects information on three categories of a household’s asset holdings. These are the stock of durable goods owned by the household, which I shall refer to as ‘household assets’, and the stocks of business and agricultural assets owned by the household. The summary statistics for these three types of assets are also presented in table 4.1.

4.3 Endogeneity of Household Characteristics

The structural characteristics of households, such as the age of the head, the type of job the head is employed in and the number of children in the household are arguably exogenous to both whether or not a household receives remittances, and the proportion of remittances in net household income. However, the stocks of household assets (i.e. durable consumption goods), agricultural assets and business assets are likely to vary in response to the receipt of remittances, biasing estimates that do not explicitly allow for this endogeneity.

Assets may vary in response to remittance receipts in a number of ways. On the one hand, households may use remittances to finance purchases of different categories of assets. In this case, the single-equation estimates will be biased upwards. Alternatively, households may finance migration by liquidating assets, leading to a negative contemporaneous correlation between asset levels and remittance receipts, when in fact high pre-existing levels of assets actually facilitate migration. Households of origin may also see remittances as a substitute to future income from productive assets, yielding a possible incentive to divest from business and agricultural assets in the presence of remittances. These channels could potentially bias the coefficient on assets downwards.

Another channel through which asset levels may be endogenous to remittances is through the household's history of remittance receipts. Households that receive remittances today are likely to have done so in the past as well. These past remittances constitute a part of the history of household income, on which current asset holdings will depend. These dynamics could potentially cause estimates of the effect of current assets on current remittances to be biased and inconsistent. Counterfactual distributions of income must therefore be constructed in a manner that is robust to these potential biases.

To solve this problem using an instrumental variable would require an instrument that is correlated with current asset levels, but uncorrelated to remittances through any other channel. The descriptive statistics from the preceding chapter, which provide the motivation for this chapter, suggest that the level of household income may itself affect the probability of receiving remittances. The search for an external instrument that is correlated with asset holdings, but not with household income is unlikely to bear fruit, because each of these variables is likely to be affected by the same processes as the other. As such, another approach is required.

The preceding paragraphs have recognized why it is problematic to directly establish causation from assets to remittance receipts. The literature can, however, offer some guidance on establishing causation in the reverse direction. McKenzie and Rapoport (2007) use historical village-level migration to instrument for current migration levels in their ground breaking study on the relationship between remittances and income inequality in migrant-sending areas. In this chapter, I will use information on village-level historical migration to identify the effect of remittances on the stocks of different categories of assets. I will use the parameter estimates yielded by this model to predict what household assets would have been, if households had not received any remittances from their children. Assuming that the effect of remittances on assets was correctly identified by means of the instrument, these predicted asset levels will be purged of the endogeneity that posed a threat to identification of the single-equation model.

4.3.1 Village Histories

The 1997 wave of the Townsend Thai data reports information on historical migration from and to each village, that was collected from interviews with a village key informant, typically the village headman. Each informant is asked if in the history of the village there has ever been a period when three or more households moved away from that village, together. I use this information to construct a dummy variable for whether or not the village has experienced historical 'out-migration'. In addition, each key informant is also asked if there was ever a period when three or more households moved *to* that village together, which I use to construct a dummy variable for historical 'in-migration'. Of the

64 villages in my sample, 12 (approximately 19%) report experiencing historical out-migration, whereas 25 (approximately 39%) report experiencing historical in-migration. I note here that these instruments are fixed, not only within villages, but also over time. Therefore, in all the subsequent analysis, the reported standard errors are clustered by village and by time period, wherever appropriate, unless otherwise stated.

Applying the logic of McKenzie and Rapoport's (2007) migration network model to these Thai villages suggests that instances of three or more households leaving a village together to settle in another region, are likely to reduce the costs of future migration from the village of origin to that destination, through a variety of channels. Information about transportation, housing, and employment prospects are likely to flow more freely between origin and destination. The extension of kinship networks to the place of destination may entail a reduction in the cost of housing and board, in the early stages of migration, when migrants are at their most vulnerable. These forces may reduce the risks and costs associated with migration. For these reasons, historical migration out of a particular village is likely to affect current village-wide migration rates, and consequently, the likelihood that a particular household in that village receives remittances from their children. Importantly, there is no obvious channel through which the current asset levels of households that remain at the place of origin would be affected by this type of historical migration, except through the migration process.

It is possible that similar network effects are engendered when three or more households migrate *to* a village in the sample, though this narrative remains to be established in the literature. Such households would presumably maintain links with their communities of origin, and if employment prospects were to arise in those communities, these households may be able to facilitate the diffusion of this information to their current village of residence. This narrative however, is less compelling than the preceding one because it involves a reversal in the direction of migration flows. Such a reversal implies either that the attractiveness of the historical origin and destination labour markets had switched, relative to one another, after the occurrence of the historical migration; or that the historical migration episode was not economically motivated.

To test statistically if either of these variables are relevant instruments for household remittance receipts, that is, if they significantly predict the amount of remittances received by households, I regress each of them on a remittances received by households in a pooled OLS model of all years in the panel, estimating:

$$r_{ivt} = c + \alpha m_v + \varepsilon_{ivt}, \qquad (4.1)$$

where *r* is the amount of remittances received by a household from the children of the head, inflated to 2011 Baht; m is whether or not a village has experienced at least one episode of historical migration (I run separate regressions for historical episodes of in- and out-migration), ε is a mean-zero error term, and c and α are parameters to be estimated. As before, the subscript *i* refers to individual households, *v* refers to villages and *t* refers to the time period.

Table 4.2: Results of Univariate Regressions of Potential Instruments on Real Remittance Receipts from Children

	Without clustered errors $F(1,12668)$	*Village X time clustered errors* $F(1, 959)$
Historical inward migration	11.82	7.02
Historical outward migration	18.22	11.56
Remittance from people other than head's children	0.21	0.36
Critical value for α*= 0.01 of* $F(1, \infty) = 6.635$		

The F-statistics for tests of the null hypotheses that historical in-migration and historical out-migration are independent of income are reported in the first column of table 4.2.

Both of these F-statistics are greater than 10, the critical value that is conventionally used to differentiate between strong and weak instruments. Because the instrument does not vary over time, or within villages, a standard F-test may overstate its statistical relevance. I therefore repeat this diagnostic, after clustering errors by village and calendar year. Despite the fact that clustering somewhat diminishes the magnitude of the resulting F-statistics, I continue to strongly reject the null hypothesis that historical migration is unrelated to remittance receipts at the 1% level of significance. I therefore conclude that both historical in-migration to a village, and historical out-migration are relevant instruments for current remittance receipts.

In addition to remittances received from the children of the head of household who live outside the village, the Townsend Thai Project also collects information on remittances received from friends and family who are not children of the head of household. The presence of such remitters may be indicative of whether or not the children of the head of household have access to a migrant network, something which may also be used to instrument for remittance receipts. Unfortunately, only approximately 6% of the households receive remittances from members other than the children of the head of household. The first column of the third row of table 4.2 presents the F-statistic that results from a univariate regression of whether or not a household receives remittances from someone other than a child of the head of household in 1997, on remittances received from children of the head in all years of the panel. It is evident from this result that this variable fails the ‘relevance’ criterion necessary to be of use as an instrument for remittances from the children of the head of household.

Historical migration in to and out of these villages, will therefore allow me to identify an arguably causal effect of remittance receipts on household asset holdings. I now turn to estimating this relationship.

4.3.2 Identifying the Effect of Remittances on Assets

The objective of this section is to construct a measure of assets for each household in each period that is free from the influence of remittance receipts on asset holdings. To ensure that the resulting predicted assets are not correlated with the error term in the final stage, I model assets as a function of all other observable characteristics which I treat as exogenously determined, in addition to (instrumented) remittances. These include characteristics of the head of household (sex, age, occupation and type of employment contract) and household demographic variables (number of dependents, the number of adult children and the education level of children).

Among the 14,271 household-year observations, 3,124 report owning no household assets, while 7,117 report holding no agricultural assets and 9,616 report holding no business assets. Because the distribution of these dependent variables are bounded below by zero, estimation strategies such as two stage least squares, which assume a dependent variable that is distributed over an unbounded support, will not be suitable to model these asset holdings. Instead, I use a Tobit model (Tobin, 1958), as generalized by Amemiya (1978) to accommodate endogenously determined dependent variables. In the first stage, I use historical migration and the exogenous characteristics of the household to model the receipt of remittances.

$$r_{ivt} = X_{ivt}\beta + \gamma m_v + \theta t + \delta t^2 + \varepsilon_{ivt}, \tag{4.2}$$

where β, γ, θ, and δ are parameters to be estimated, t is a time trend, and the other terms remain as they have been previously defined. The resulting maximum likelihood estimates are used to predict remittance receipts for each household-year observation,

$$\hat{r}_{ivt} = X_{ivt}\hat{\beta} + \hat{\gamma} m_v + \hat{\theta} t + \hat{\delta} t^2, \tag{4.3}$$

where $\hat{r}$ is the predicted level of remittances while $\hat{\beta}$ and $\hat{\gamma}$ are the parameters estimated from equation 4.2. In the second stage, I use these predicted remittances, which are constructed to be free of the influence of reverse causation from assets to remittance receipts, to identify the effect of remittances on assets. That is, I estimate the Tobit model:

$$a^{*j}{}_{ivt} = X_{ivt}\beta' + \mu\hat{r}_{ivt} + \theta' t + \delta' t^2 + \varepsilon_{ivt}, \qquad (4.4)$$

where the 'j' superscript is used to differentiate between the three classes of household, agricultural and business assets. The primes on the parameters are used to differentiate these coefficients from those that were estimated from equation (4.2). Finally, $a*$ is an underlying latent variable such that observed assets, a take on the value:

$$a = \begin{cases} a^* \; if \; a^* > 0 \\ 0 \;\; if \; a^* \leq 0 \end{cases} \qquad (4.5)$$

The error term, ε_{ivt}, is assumed to be normally distributed and the model is estimated using the ivtobit package in STATA. The resulting coefficients for household, agricultural and business assets are presented in the first, third and fifth columns of table 4.3. As mentioned before, the standard errors are clustered at the level of individual villages and time periods. For comparison, the un-instrumented, single equation estimates for household, agricultural and business assets are presented in the second, fourth and sixth columns of the table, respectively. Instrumenting makes a substantial difference to the estimated coefficients on remittances for all categories of assets, suggesting that endogeneity was indeed a problem (something that is confirmed by a Wald test, discussed below). The estimated coefficient on remittances for household assets is small, but positive and statistically significant. The coefficient on remittances for agricultural assets is positive, but insignificant, whereas the one for business assets was negative and statistically significant, but also relatively small in magnitude.

After instrumenting for the potentially endogenous explanatory variables, I reject the null hypothesis that remittances have no effect on asset levels in favour of the alternative that remittances reduce asset holdings at the 5% level of significance for all three of these asset categories. For household and agricultural assets, a 1 Baht increase in remittances is associated with between a 4 and 5 Baht decline in asset holdings, at the sample mean, other things remaining equal. The effect for business assets is stronger, with the same increase in remittances reducing holdings of business assets by over 9 Baht, on average and other things remaining equal (in the un-instrumented specification, a 1 Baht increase in remittances always leads to less than 1 Baht of a difference in holdings of a given asset category).

Let ‘underlying assets’ describe the level of assets that would have prevailed in a particular village if that village had not experienced an episode of historical migration. As mentioned above, the coefficients on un-instrumented remittances are always relatively small, whereas those on instrumented remittances are always relatively large and negative. This suggests that the level of underlying assets in those villages that have experienced historical migration are lower than the observed level of assets. Thus historical migration episodes tend to increase the measured level of assets in these villages.

Though these reduced form results need to be interpreted with caution, I note that the negative instrumented coefficients of the effect of remittances on assets (of which the effect on business assets is the strongest) are consistent with the narrative that households may liquidate productive assets because remittances provide a substitute to the future stream of income that could have been earned by combining these assets with the labour that is foregone when adult children migrate.

The final row of table 4.3 presents the Chi-squared statistics for a Wald test of the null hypothesis that remittances are exogenously determined from assets. For all asset categories, the statistic is greater than the 5% critical value, leading us to reject the null in favour of the alternative that remittances are indeed endogenously determined with assets. Because of the inherently heteroskedastic nature of the Tobit model, I am unable to perform an over-identifying restrictions test in this setting, despite having more instruments than endogenous variables for each asset category. For the purpose of running an over-identifying restrictions test, I repeat the analysis in a 2SLS setting. For the sake of brevity, I do not report those coefficients here, but I note that the Chi-squared statistics for the Sargan tests, for household, agricultural and business assets are 0.002, 0.264 and 0.077, respectively. The 5% critical value of this distribution with 1 degree of freedom is 3.84. Thus for all asset categories, I fail to reject the null hypothesis that the instruments are uncorrelated with the error term and that the model is correctly specified.

Using these estimated coefficients, I now predict the expected value of each category of assets for each household in each time period, assuming that remittances were zero, holding all other observable characteristics constant. That is, I predict:

$$\widehat{a^j}_{it}|_{r_{it}=0} = \max\{0, \ (X_{ivt}\hat{\beta}'^j + \widehat{\theta'^j}t + \hat{\delta}'^j t^2)\}, \tag{4.6}$$

where the estimated parameters are those from equation 4.4, which were presented in table 4.3. These predicted assets are, by construction free of the influence of remittances on assets. Therefore, they can be used to address the endogeneity problem that is likely to confound identification of the effect of assets on a sample household’s probability of receiving remittances.

Table 4.3: Testing the Effect of Remittances on Asset Holdings

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	*Household assets*	*Household assets*	*Agricultural assets*	*Agricultural assets*	*Business assets*	*Business assets*
Model	*Tobit*	*Instrumented Tobit*	*Tobit*	*Instrumented Tobit*	*Tobit*	*Instrumented Tobit*
Remittances (or instrument)	0.0961** (2.10)	-4.329*** (-4.38)	0.00574 (0.14)	-4.355** (-2.30)	-0.288** (-2.28)	-9.458*** (-3.28)
Sex	-0.0811*** (-2.93)	-0.171*** (-3.63)	-0.0302 (-0.65)	-0.119 (-1.63)	-0.518*** (-4.58)	-0.698*** (-4.85)
Age of head	0.0142 (1.19)	0.121*** (3.44)	0.0797*** (4.04)	0.184*** (3.30)	0.0112 (0.26)	0.232** (2.29)
Age of head squared	-0.0000917 (-0.93)	-0.000952*** (-3.22)	-0.000683*** (-4.16)	-0.00153*** (-3.34)	-0.000323 (-0.91)	-0.00211** (-2.49)
Head's years Of schooling	0.0651*** (7.99)	0.0403*** (3.33)	-0.0125** (-2.23)	-0.0368*** (-2.69)	0.0616*** (3.89)	0.00947 (0.36)
Monthly or govt. wages	-0.116 (-1.02)	-0.162 (-1.05)	-0.365*** (-3.05)	-0.411*** (-2.59)	-1.185*** (-3.89)	-1.267*** (-3.37)
Business owner	0.147*** (2.76)	0.188** (2.24)	0.821*** (9.76)	0.862*** (7.79)	2.211*** (10.40)	2.305*** (9.02)
Crops other than rice	0.201*** (3.38)	-0.0788 (-0.79)	0.275*** (2.99)	-0.000684 (-0.00)	-0.553*** (-2.64)	-1.128*** (-3.92)
Livestock	0.154*** (3.77)	0.0644 (0.84)	0.181** (2.39)	0.0929 (0.91)	1.674*** (10.00)	1.479*** (7.03)
Blue collar worker	0.245*** (4.78)	0.0750 (0.81)	-0.548*** (-7.15)	-0.714*** (-5.50)	2.285*** (12.43)	1.937*** (7.92)
White collar worker	1.017*** (4.99)	0.809*** (3.57)	0.255* (1.91)	0.0381 (0.20)	2.386*** (6.49)	1.944*** (4.41)
Other work	0.274*** (5.78)	0.0510 (0.62)	-0.0994 (-1.23)	-0.322** (-2.48)	1.251*** (6.45)	0.797*** (3.17)
Work that varies	0.0897 (1.52)	-0.0934 (-0.93)	-0.388*** (-4.26)	-0.566*** (-4.15)	0.000123 (0.00)	-0.355 (-1.22)
Non working age	0.0256** (2.46)	0.110*** (4.03)	0.109*** (5.96)	0.191*** (4.29)	0.227*** (4.35)	0.400*** (4.53)
Schooling of children	0.0674*** (3.27)	0.164*** (4.07)	0.209*** (7.10)	0.306*** (5.14)	0.337*** (4.88)	0.536*** (4.54)
Schooling of children squared	0.000720 (0.65)	-0.00112 (-0.62)	-0.00647*** (-4.84)	-0.00838*** (-3.94)	-0.00753** (-2.54)	-0.0113** (-2.48)
Number of Adult children	0.0453*** (5.90)	0.185*** (4.92)	0.0000373 (0.00)	0.137** (2.17)	0.0991*** (3.74)	0.387*** (3.77)
Time	-0.146*** (-9.05)	-0.116*** (-4.94)	-0.0215 (-1.11)	0.00742 (0.27)	0.178*** (3.76)	0.241*** (3.77)
Time squared	0.00462*** (4.57)	0.00566*** (3.66)	-0.000120 (-0.09)	0.000927 (0.50)	-0.00133 (-0.44)	0.000751 (0.18)
_cons	-0.840** (-2.31)	-4.585*** (-4.08)	-4.418*** (-6.45)	-8.110*** (-4.22)	-8.326*** (-5.86)	-16.12*** (-4.69)
N	9871	9871	9871	9871	9871	9871
Wald test of exogeneity		20.18		5.37		10.31

Assets and remittances in 100,000 Baht
t statistics in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

4.4 The Probability of Receiving Remittances

In this section, I econometrically characterise the likelihood that a particular household will receive remittances. In the baseline specification, I use a simple Probit to model the likelihood of remittance receipts as a function of observed household characteristics. I then go on to address the potential endogeneity of assets by substituting the predicted values from equation 4.6, for the observed values. The resulting parameter estimates are then used to construct counterfactual distributions of income, where the likelihood of receiving remittances is allowed to vary over the distribution of income, but the proportion of net income that is accounted for by remittances is held constant, at the sample mean.

In the Probit model, whether a household receives remittances in a particular period or not, which I denote by the dummy variable, d_{it}, is assumed to be a function of an underlying latent variable, d_{it}^*:

$$d_{it}^* = X_{it}\beta + \delta t + \theta t^2 + \varepsilon_{it}, \qquad (4.7)$$

where d^* is an unobserved latent variable such that:

$$d_{it} = \begin{cases} 1 \; if \; d_{it}^* > 0 \\ 0 \;\; otherwise \end{cases} \qquad (4.8)$$

The other variables remain as they have been defined above. The results of this baseline specification are presented in the first column of table 4.4.

In this naïve specification, where I do not address the endogeneity of asset levels, relatively large asset holdings appear to discourage the receipt of remittances. Households that are headed by women appear to be more likely to receive remittances. The lifecycle variables, such as the age of the head of the household (and its square) and the number of adult children of the head of household are highly significant and have the expected signs. Better educated heads of household are less likely to receive remittances, potentially explaining part of the inequality reducing effect of remittances. Other things remaining equal, a larger number of non-working age members does not significantly affect the likelihood that a household receives remittances.

Due to reasons discussed in the preceding section, assets are likely to be endogenously determined with remittances. If so, these coefficients will be biased and inconsistent. To address this issue, I replace observed assets for a given household in a given time period with their predicted values from equation 4.6. These predicted values are, by construction, unaffected by the endogeneity that is likely to confound the naïve estimates that were presented in the first column. Specifically, I assume a latent variable of the form:

$$d_{it}^{*} = X_{it}\beta + \alpha^{j}\widehat{a^{j}}_{it}|_{r_{it}=0} + \delta t + \theta t^{2} + \varepsilon_{it}, \qquad (4.9)$$

where α is a parameter to be estimated, and other variables are as before (with the exception that observed household, agricultural and business assets are no longer included among the Xs). Two distinct approaches may be used to handle the standard errors in this application and each has its own advantages and disadvantages. Because one of the explanatory variables is constructed on the basis of an instrument that does not vary between villages or over time periods, a case can me made for proceeding with clustering standard errors at the village and time period levels. On the other hand, because the value of these variables are predicted, rather than observed, a case can be made for bootstrapping the standard errors (unfortunately, the available statistical software does not enable me to bootstrap within clusters). For this reason, I present two sets of results for this regression, with the only difference being the way the standard errors are calculated. In the second column of table 4.4, the standard errors are clustered at the village and time-period level, whereas in the third column, they are bootstrapped. These results however, suggest that there is no material difference between these two methods of estimating the standard errors as the degree of statistical significance varies only slightly, for only two of the conditioning variables.

Purging assets of their endogenous component reverses the signs of the coefficients on household assets and agricultural assets and greatly increase their magnitude. Despite the increase in magnitude, the coefficients remain statistically indistinguishable from zero, due to the larger errors associated with the two-stage model on which the prediction is based. In contrast to these coefficients, the sign on the coefficient for business assets remains the same, and the magnitude is somewhat diminished. The coefficient on the predicted value of this asset too, is statistically indistinguishable from zero, at conventional levels of significance.

It appears then, that the apparent explanatory power of assets on the extensive margin of remittance receipts was through reverse causation – possibly because remittances are seen as a substitute to the future stream of payoffs from these assets.

4.4.1 Modelling the Probability of Remittance Receipts

Section 4.2 presented statistical evidence that the levels of different categories of assets did indeed vary endogenously with the level of remittances received by households. After addressing this endogeneity however, the estimated coefficients for the effect of the levels of different categories of assets on the probability of a household receiving remittances was not statistically distinguishable from zero. This begs the question, how important are assets in explaining the likelihood of a household receiving remittances?

The pseudo R-squared from the Probit model that utilized the predicted values of assets that were designed to be free of endogeneity was 0.1469. This was based on a sample of 9513 household-year observations and 21 regressors (without counting the constant term).

Table 4.4: The Probability of Receiving Remittances				
	(1)	(2)	(3)	(4)
Variables:	*Observed Assets*	*Predicted assets*	*Predicted assets*	*Core variables*
Standard errors	*Analytical*	*Clustered errors*	*Bootstrapped errors*	*Analytical*
Agricultural assets	-0.0980*** (-6.44)	0.431 (1.47)	0.431 (1.18)	
Business assets	-0.0606*** (-6.41)	-0.0467 (-0.54)	-0.0467 (-0.55)	
Household assets	-0.0949*** (-6.48)	0.602 (1.40)	0.602 (1.37)	
Sex of head	0.0713** (2.17)	0.154*** (2.69)	0.154*** (3.24)	0.0681** (2.01)
Age of head	0.130*** (9.48)	0.0441 (1.19)	0.0441 (1.30)	0.123*** (7.87)
Age of head squared	-0.000964*** (-8.25)	-0.000287 (-0.97)	-0.000287 (-1.06)	-0.000920*** (-7.03)
Head’s years Of schooling	-0.0405*** (-6.98)	-0.0549*** (-3.10)	-0.0549*** (-2.88)	-0.0520*** (-10.13)
Monthly or govt. wages	0.280*** (3.14)	0.444*** (4.39)	0.444*** (3.60)	0.127* (1.87)
Business owner	-0.0208 (-0.39)	-0.297*** (-2.59)	-0.297** (-2.51)	-0.230*** (-5.33)
Crops other than rice	-0.338*** (-6.58)	-0.384*** (-5.00)	-0.384*** (-5.02)	0.0336 (0.56)
Livestock	-0.137** (-2.36)	-0.207** (-2.36)	-0.207*** (-2.64)	0.237*** (3.72)
Bluecollar worker	-0.361*** (-6.37)	-0.233 (-1.24)	-0.233 (-1.18)	
Whitecollar worker	-0.436*** (-3.72)	-0.873*** (-2.98)	-0.873*** (-2.97)	
Other work	-0.358*** (-7.75)	-0.332*** (-3.35)	-0.332*** (-3.06)	
Work that varies	-0.148** (-2.17)	-0.0311 (-0.28)	-0.0311 (-0.30)	
Non working age	0.00572 (0.48)	-0.0877** (-2.53)	-0.0877*** (-2.60)	
Schooling of children	0.0408** (2.10)	-0.0858* (-1.76)	-0.0858* (-1.65)	
Schooling of children squared	-0.00187** (-2.06)	-0.000418 (-0.27)	-0.000418 (-0.23)	
Number of Adult children	0.178*** (18.04)	0.0707 (1.44)	0.0707 (1.47)	0.178*** (16.88)
Time	0.0920*** (6.96)	0.155*** (3.94)	0.155*** (4.09)	0.0985*** (6.51)
Time squared	-0.00414*** (-4.69)	-0.00721*** (-3.66)	-0.00721*** (-4.07)	-0.00375*** (-3.61)
Rice farmer				0.490*** (12.08)
_cons	-4.606*** (-11.34)	-1.956* (-1.80)	-1.956* (-1.88)	-4.498*** (-9.97)
N	9513	9513	9513	9513
Assets and remittances in 100,000 Baht t statistics in parentheses ="* p<0.10 ** p<0.05 *** p<0.01"				

I also estimate the same model on only a core group of clearly exogenous variables. To do so, I drop not only the three asset categories as conditioning variables, but also the number of non-working age members of the household, the mean years of education of the adult children of the head of household (and its square) and the dummy variables for whether or not the head of household owns a business, is in a blue-collar job, is in a white-collar job, is in jobs that vary over the course of the year, or in some ‘other’ kind of job. This leaves me with a group of core variables that include the sex of the head of household, his or her age (and age square) and years of schooling, the type of agriculture the household is dependent on (rice, crops other than rice, livestock), whether or not the head of household is in work that pays a monthly wage, the number of adult children of the head of household and a time trend and its square. Estimating this model using these 11 regressors on the same 9513 observations as the previous one, yields a pseudo R-squared of 0.1465. That is, almost halving the number of regressors leaves the pseudo R-squared unchanged to three significant figures.

Furthermore, the eleven core variables are quite plausibly exogenous to the receipt of remittances. Most of these are structural characteristics of the household that are unlikely to vary in response to the receipt of remittances.

The claim that a small subset of the variables account for the vast majority of the explained variation in the likelihood with which a household receives remittances can be tested formally by minimizing the Akaike (1973) or Bayesian (Schwarz 1978) information criteria, which not only compare the goodness of fit of these models, but also include a penalty for increasing the number of covariates. The Akaike information criterion for the model with the full set of regressors is 10,931.26, whereas that for the core set of regressors is 10,916.8, recommending use of the simpler model. A similar conclusion is reached by minimizing the Bayesian information criterion, which takes on a value of 11,088.79 for the richer model, but 11,002.73 for the simpler model.

These information criteria help inform a discussion of the amount of variation explained by additional regressors, but this is not the only reason to include additional covariates in a regression. If an omitted covariate is correlated with other covariates, then the omission can bias coefficient estimates of the included regressors. The fourth column of table 4.4 presents the estimated coefficients on the core group of regressors discussed above. It is apparent from these results that no significant bias is incurred by the exclusion of the variables which measure asset levels, the education levels of adult children and the substitution of a number of the occupation categories for the existing base group (rice farmers). These considerations lead me to conclude that the predictive power of assets in the naïve specification was spurious, and driven by endogeneity. Furthermore, no discernable bias is incurred by omitting these variables as regressors. For this reason, I proceed by using the stripped down specification, with only the core variables to predict the likelihood of a household receiving remittances.

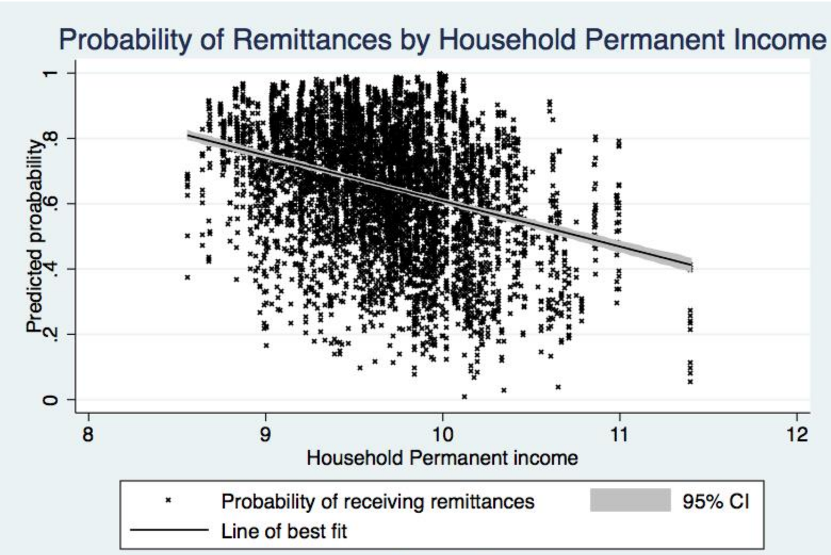


Figure 4.1

Figure 4.1 plots the probabilities of receiving remittance for each household in each year against the permanent income[2] of that household over the duration of the panel. The scatter points suggest that poorer households are substantially more likely to receive remittances than their richer counterparts. The line of best fit and the associated 95% confidence intervals confirm that this pattern is statistically significant.

In this section, I have modelled the extensive margin of remittance receipts at the household level, using a core set of observable household characteristics. Section 4.5 will model the intensive margin, also at the level of individual households. In section 4.6, I will study how the extensive and intensive margins of remittance receipts vary over the distribution of income. This will allow me to decompose the inequality reducing effect of remittances into three components: a part that is explained by the extensive margin of remittance receipts, a part that is explained by the intensive margin, and a part that remains unexplained by these econometrics.

4.5 The Proportion of Remittances in Household Income

The preceding section has established that poorer households are indeed significantly more likely to receive remittances from their adult children than their richer counterparts. In this section I will complement that analysis by modelling the proportion of household income that is accounted for by remittances.

[2] Permanent income is calculated as it was in the preceding two essays, by taking the mean of real, equivalised consumption over the 15 year duration of the panel, for each household.

In the baseline specification, I will use a Tobit model to for the proportion of remittances in household income as a function of the full set of observable household characteristics that were identified in section 2. This specification, too, however is likely to suffer from the same endogeneity problem that was encountered when modelling the likelihood that a household receives remittances – namely that assets will not only determine remittance contributions, but may also be a function of those contributions. To address this issue, I will again rely on the predicted asset levels from equation 4.6, which were constructed to be free of this endogenous component.

In the baseline specification, I assume:

$$s_{it}{}^{*} = X_{it}\beta + \delta t + \theta t^{2} + \varepsilon_{it}, \tag{4.10}$$

where ε is a mean-zero, normally distributed error term and s^{*} is an unobserved latent variable such that:

$$s_{it} = \begin{cases} 1 \; if \; s_{it}{}^{*} > 1 \\ 0 \; if \; s_{it}{}^{*} < 0 \\ s_{it}{}^{*}, otherwise. \end{cases} \tag{4.11}$$

where *s* is the share of remittances in household income and all other variables are as they were defined previously. The naïve estimates are presented in the first column of table 4.5.

Over all, the narrative that emerges from this benchmark specification for the proportion of remittances in household income is very similar to that which characterised the probability of a household receiving remittances. High levels of assets appear to be associated with a reduction in the share of remittances in household income. Female headed households receive a significantly larger share of their household income in the form of remittances.

The variables that are associated with the lifecycle of the head of household are found to be key determinants of the proportion of remittances in household income. The share of remittances in household income increases with the age of the head of household, though the sign of the second order term suggests that the rate of increase diminishes somewhat with age. The better educated a head of household is, the smaller is the share of remittances from children in net income. Having an additional adult child significantly increases the share of remittances in household income.

As mentioned above, the coefficients from this baseline specification are likely to be biased and inconsistent because a number of the conditioning variables are likely to vary endogenously with the proportion of remittances in household income. I now use the values predicted by equation 4.6, which I constructed to be free of influence of remittances on assets, in place of observed assets to model the proportion of remittances in household income. That is, the latent variable in equation 4.10 is now specified as being:

$$s_{it}^* = X_{it}\beta + \alpha^j \widehat{a^j}_{it}|_{r_{it}=0} + \delta t + \theta t^2 + \varepsilon_{it}, \qquad (4.12)$$

where the Xs no longer include the observed level of household, agricultural and business assets. The resulting estimates are presented in the second and third columns of table 4.5. As was discussed in the preceding section, I present standard errors that are clustered at the village-time level in the second column, whereas bootstrapped standard errors are presented in the third column. Again, the choice of standard error does not meaningfully alter inference.

Replacing observed assets with their predicted counterparts from equation 4.6 has similar effects on the coefficients of these assets as it did when modelling the probability of receiving remittances. The coefficient for business assets diminishes slightly and loses significance. That of household assets reverses sign and increases greatly in magnitude, but remains statistically indistinguishable from zero. The coefficient on agricultural assets also changes sign and increases in magnitude, but here, it is statistically different from

Table 4.5: The Proportion of Remittances in Income				
	(1)	(2)	(3)	(4)
Endogenous Variables	*Observed Assets*	*Predicted assets*	*Predicted assets*	*Not included*
Standard errors	*Analytical*	*Clustered errors*	*Bootstrapped errors*	*Analytical*
Agricultural assets	-0.0376*** (-7.94)	0.154* (1.75)	0.154 (1.61)	
Business assets	-0.0242*** (-8.09)	-0.0193 (-0.82)	-0.0193 (-0.86)	
Household assets	-0.0302*** (-7.01)	0.135 (1.20)	0.135 (1.20)	
Sex of head	0.0405*** (4.77)	0.0607*** (4.17)	0.0607*** (4.04)	0.0413*** (4.73)
Age of head	0.0389*** (10.72)	0.0153 (1.51)	0.0153 (1.44)	0.0368*** (7.88)
Age of head squared	-0.000296*** (-9.69)	-0.000107 (-1.32)	-0.000107 (-1.26)	-0.000281*** (-7.25)
Head's years Of schooling	-0.0126*** (-7.85)	-0.0150*** (-3.07)	-0.0150*** (-2.67)	-0.0145*** (-10.85)
Monthly or govt. wages	-0.0460* (-1.85)	0.00140 (0.06)	0.00140 (0.05)	-0.0623*** (-3.69)
Business owner	-0.0400*** (-2.78)	-0.126*** (-4.01)	-0.126*** (-3.35)	-0.101*** (-8.46)
Crops other than rice	-0.118*** (-8.42)	-0.137*** (-6.58)	-0.137*** (-6.09)	0.0135 (0.82)
Livestock	-0.0217 (-1.51)	-0.0340 (-1.44)	-0.0340* (-1.72)	0.112*** (6.63)
Blue collar worker	-0.136*** (-8.68)	-0.0814 (-1.47)	-0.0814 (-1.43)	
White collar worker	-0.101*** (-2.99)	-0.202** (-2.56)	-0.202*** (-2.61)	
Other work	-0.0984*** (-7.84)	-0.0847*** (-2.88)	-0.0847*** (-2.90)	
Work that varies	-0.0748*** (-4.15)	-0.0346 (-1.11)	-0.0346 (-1.10)	
Non working age	0.00132 (0.44)	-0.0252*** (-2.70)	-0.0252** (-2.42)	
Schooling of children	0.0185*** (3.60)	-0.0195 (-1.40)	-0.0195 (-1.27)	
Schooling of children squared	-0.000754*** (-3.12)	-0.000187 (-0.40)	-0.000187 (-0.37)	
Number of Adult children	0.0371*** (16.19)	0.00962 (0.75)	0.00962 (0.67)	0.0358*** (15.35)
Time	0.0296*** (8.35)	0.0454*** (4.18)	0.0454*** (4.88)	0.0313*** (6.67)
Time squared	-0.00133*** (-5.70)	-0.00208*** (-3.82)	-0.00208*** (-4.22)	-0.00119*** (-3.76)
Rice farmer				0.170*** (15.22)
_cons	-1.341*** (-12.20)	-0.597** (-1.96)	-0.597* (-1.82)	-1.298*** (-9.44)
N	9513	9513	9513	9513
Assets and remittances in 10,000 Baht t statistics in parentheses ="* p<0.10 ** p<0.05 *** p<0.01"				

zero at the 10% level of significance. These changes suggest that assets did indeed vary endogenously with the proportion of remittances in household income in ways that seriously biased the coefficients on these assets in the baseline model.

Using the predicted value of assets has substantive implications for some of the other coefficients estimated in this model. Notably, the coefficients on the variable for the age of the head of household are smaller in magnitude but still statistically distinguishable from zero at the 10% level with bootstrapped standard errors. The dummy variable for those who earn monthly or government wages is also much smaller in magnitude and no longer statistically distinguishable from zero, as is the coefficient on the number of adult children of the head of household.

4.5.1 Assessing Models using Information Criteria

Using observed asset levels in the model that did not account for endogeneity suggested that assets were important predictors of the proportion of remittances in household income (as the results in the first column of table 4.5 illustrate). Addressing the endogeneity problem between asset levels and remittance receipts reduced much of the predictive power of assets. Only agricultural assets were found to predict the proportion of remittances in household income at the 10% level of significance (but not at the 5% or 1% levels).

As in section 4.3, I use the Akaike (1973) and Bayesian (Schwarz 1978) information criteria to select between these models. The Akaike information criterion for the model with the full set of regressors is 8,768, whereas the Bayesian information criterion is 8,933. These statistics for the model with only the core set of variables are 8,648 and 8,748 respectively. Minimizing the criteria recommends the use of the simpler model.

A crucial difference between the Tobit results here and the Probit results presented before is that in the Tobit model, the level of agricultural assets is found to have some predictive power, whereas in the Probit model it did not. Therefore, I also compute information criteria scores for a model where the predicted level of agricultural assets is used in addition to the 11 core variables. The Akaike information criteria for this model is evaluated at 8650 whereas the Bayesian information criteria is 8,757. Both of these values are greater than the corresponding values for the model with only the core variables, leading me to prefer the simpler model.

These considerations lead me to conclude that the predictive power of assets in the naïve model was spurious and driven by the endogeneity of these variables. I therefore use the model with only the core group of exogenous variables as regressors to predict the proportion of remittances in household income.

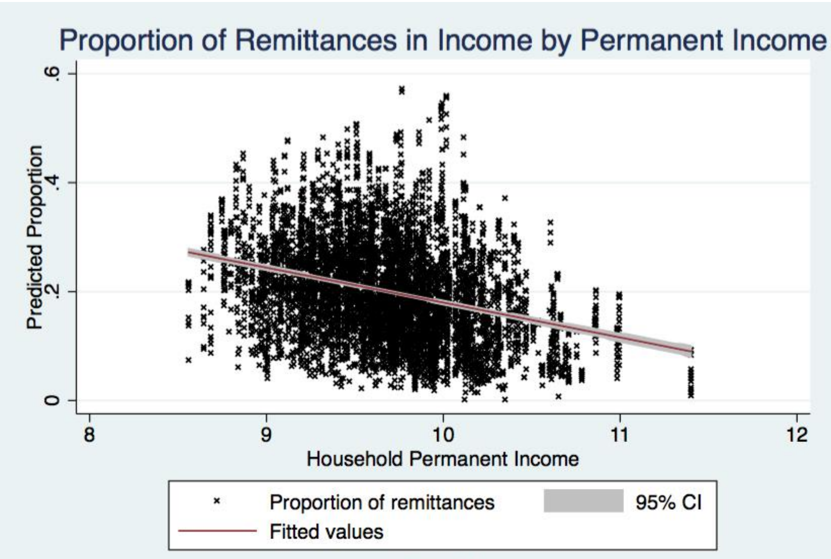


Figure 4.2

Figure 4.2 plots the predicted proportion of remittances in the income of each household in each year, against the permanent income of that household over the duration of the panel. The scatter plot supports the descriptive evidence from the preceding chapter, which observed that the share of remittances in the incomes of poorer households is larger than those of richer households. A line of best fit and the associated 95% confidence intervals confirm that this pattern is statistically significant.

I now turn to testing if the conclusions of this section and the one that preceded it are unduly sensitive to the choice of instrument.

4.6 Sensitivity Analysis: Drought as an Alternative Instrument for Remittance Receipts

The preceding analysis suggests that making econometric corrections for the endogeneity of assets in a model of remittance receipts, rids them of their explanatory power. That finding may be sensitive to the choice of instrument used to identify the effect of remittances on asset holdings. Specifically, historical migration episodes may be driven by processes which also affect historical asset levels. As asset levels are known to display high persistence, this may lead to a violation of the exclusion restriction and hence invalidate the instrumental variable. Though the results of an overidentifying restrictions test suggest this is not the case, it may nonetheless be useful to verify the main implications of the preceding sections with an alternative instrument. In the second chapter of this thesis I demonstrated that drought was a significant covariate shock to incomes in these villages. I now use the proportion of households in a village that are affected by drought[3] as an instrument for the amount of remittances received by village

[3] Chapter 2 contains a detailed description of how this variable was constructed, and the descriptive statistics of this variable.

households, with a view to constructing counterfactual underlying assets as was done in the preceding sections.

Within the period that a village is affected by a drought, households may well run down their assets to cover the shortfall in income. Thus asset levels may still be endogenously determined with the current prevalence of drought. For this reason, I use the lagged proportion of households in each village affected by drought as my instrument for the level of remittance receipts.

A univariate regression of the proportion of households in each village affected by drought on the level of remittances a household received by a household yields an F-statistic of 25.2, well over the standard F-statistic of 10 that is the general benchmark to establish instrument relevance. The use of clustered standard errors, to account for the fact that the regressor is measured at the village level reduces this to 17.97, but this is still well over the threshold value of 10. I therefore conclude that the instrument is relevant.

Table 4.6 presents the results from re-estimating equation 4.4 using the lagged value of drought as in instrument, as opposed to village migration histories where were used earlier. The results in this table should be compared with those presented in table 4.3. In contrast to those results, the estimates from the instrumented Tobit presented here are not able to discern a statistically significant effect of remittances on assets. Using these new estimates, I again compute what assets would have been in the counterfactual where remittance receipts were equal to zero, using equation 4.6. Using these ‘underlying’ assets (which are constructed to be free of the estimated influence of remittances on assets) in place of observed assets allows me to identify the effect of remittances on assets, that is arguably free of reverse causation. The results of re-estimating equation 4.9 using underlying assets as identified by the drought instrument are presented in table 4.7.

Table 4.6: Testing the Effect of Remittances on Asset Holdings Instrumenting with Lagged Drought

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	*Household assets*	*Household assets*	*Agricultural assets*	*Agricultural assets*	*Business assets*	*Business assets*
Model	*Tobit*	*Instrumented Tobit*	*Tobit*	*Instrumented Tobit*	*Tobit*	*Instrumented Tobit*
Remittances (or instrument)	0.0961** (2.10)	-2.501 (-0.89)	0.00574 (0.14)	-3.028 (-0.77)	-0.288** (-2.28)	5.656 (0.60)
Sex	-0.0811*** (-2.93)	-0.117 (-1.53)	-0.0302 (-0.65)	-0.0731 (-0.67)	-0.518*** (-4.58)	-0.342 (-1.28)
Age of head	0.0142 (1.19)	0.100 (1.11)	0.0797*** (4.04)	0.203 (1.58)	0.0112 (0.26)	-0.178 (-0.60)
Age of head squared	-0.0000917 (-0.93)	-0.000793 (-1.09)	-0.000683*** (-4.16)	-0.00168 (-1.63)	-0.000323 (-0.91)	0.00122 (0.50)
Head's years Of schooling	0.0651*** (7.99)	0.0441** (2.06)	-0.0125** (-2.23)	-0.0373 (-1.28)	0.0616*** (3.89)	0.100 (1.47)
Monthly or govt. wages	-0.116 (-1.02)	-0.139 (-0.94)	-0.365*** (-3.05)	-0.443*** (-2.71)	-1.185*** (-3.89)	-1.216*** (-3.14)
Business owner	0.147*** (2.76)	0.106 (1.51)	0.821*** (9.76)	0.770*** (7.28)	2.211*** (10.40)	2.228*** (8.66)
Crops other than rice	0.201*** (3.38)	0.0284 (0.12)	0.275*** (2.99)	0.139 (0.41)	-0.553*** (-2.64)	-0.0544 (-0.07)
Livestock	0.154*** (3.77)	0.0184 (0.22)	0.181** (2.39)	0.118 (0.91)	1.674*** (10.00)	1.747*** (5.72)
Blue collar worker	0.245*** (4.78)	0.119 (0.79)	-0.548*** (-7.15)	-0.648*** (-3.05)	2.285*** (12.43)	2.543*** (5.11)
White collar worker	1.017*** (4.99)	0.712*** (2.85)	0.255* (1.91)	0.219 (0.83)	2.386*** (6.49)	2.821*** (4.28)
Other work	0.274*** (5.78)	0.0758 (0.42)	-0.0994 (-1.23)	-0.242 (-0.92)	1.251*** (6.45)	1.598** (2.55)
Work that varies	0.0897 (1.52)	-0.0776 (-0.53)	-0.388*** (-4.26)	-0.529** (-2.51)	0.000123 (0.00)	0.173 (0.34)
Non working age	0.0256** (2.46)	0.0835 (1.22)	0.109*** (5.96)	0.190** (1.96)	0.227*** (4.35)	0.110 (0.47)
Schooling of children	0.0674*** (3.27)	0.138* (1.72)	0.209*** (7.10)	0.291** (2.56)	0.337*** (4.88)	0.219 (0.81)
Schooling of children squared	0.000720 (0.65)	-0.00122 (-0.59)	-0.00647*** (-4.84)	-0.00795*** (-2.77)	-0.00753** (-2.54)	-0.00579 (-0.88)
Number of Adult children	0.0453*** (5.90)	0.128 (1.40)	0.0000373 (0.00)	0.0982 (0.78)	0.0991*** (3.74)	-0.0834 (-0.28)
Time	-0.146*** (-9.05)	-0.0975 (-1.38)	-0.0215 (-1.11)	0.0689 (0.66)	0.178*** (3.76)	-0.0126 (-0.05)
Time squared	0.00462*** (4.57)	0.00324 (1.22)	-0.000120 (-0.09)	-0.00370 (-0.92)	-0.00133 (-0.44)	0.00631 (0.71)
_cons	-0.840** (-2.31)	-3.736 (-1.14)	-4.418*** (-6.45)	-8.783* (-1.86)	-8.326*** (-5.86)	-1.571 (-0.14)
N	9871	9871	9871	9871	9871	9871

Assets and remittances in 100,000 Baht
t statistics in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

These results should be juxtaposed with those of table 4.4 which provide analogous results using the migration history instrument.

Here, also there are substantive changes in the estimated coefficients. Whereas in table 4.4, none of the instrumented coefficients were statistically significant, here, (counterfactual) agricultural assets have a strong, positive effect on predicted remittances and business assets have a strong negative effect. Household assets remain insignificant. As before, however, there are indications that these results may not be as informative as they appear. The pseudo R-squared of the instrumented regression with the full set of regressors, is 0.1507, while that of a regression which restricts attention to the core regressors (discussed in section 4.5) yields an R-squared of 0.1487. The Bayesian information criteria for the full set of regressors including (counterfactual) assets is 11041.34, where that of the core set of regressors is 10984.48. Minimizing this information criterion therefore recommends use of the simpler model. However, the Akaike information criteria for the full model is 10883.81 and that of the core model is 10891.4. Thus the two information criteria recommend different models.

The substantial differences between the results yielded by the two instruments raises concerns about the accuracy with which either identifies the effect of assets on the probability of receiving remittances. Furthermore, the mixed results from the information criteria suggest that again, no significant explanatory power is added by the inclusion of these potentially mis-identified regressors. Together, these findings confirm the conclusions of section 4.4 which proceeded to model the extensive margin of remittance receipts using just the core set of plausibly exogenous regressors.

I now establish similar results using counterfactual asset holdings identified using the drought instrument for the proportion of remittances in household income, that is, the intensive margin of remittance receipts.

Table 4.7: The Probability of Receiving Remittances (Drought as Instrument)				
	(1)	(2)	(3)	(4)
Variables:	*Observed Assets*	*Predicted assets*	*Predicted assets*	*Core variables*
Standard errors	*Analytical*	*Clustered errors*	*Bootstrapped errors*	*Analytical*
Agricultural assets	-0.0980*** (-6.44)	0.532** (2.29)	0.532** (2.25)	
Business assets	-0.0606*** (-6.41)	-1.055*** (-6.90)	-1.055*** (-8.57)	
Household assets	-0.0949*** (-6.48)	0.358 (0.87)	0.358 (0.81)	
Sex of head	0.0713** (2.17)	0.0608 (1.33)	0.0608 (1.34)	0.0681** (2.01)
Age of head	0.130*** (9.48)	0.00572 (0.17)	0.00572 (0.15)	0.123*** (7.87)
Age of head squared	-0.000964*** (-8.25)	-0.0000184 (-0.07)	-0.0000184 (-0.06)	-0.000920*** (-7.03)
Head's years Of schooling	-0.0405*** (-6.98)	-0.0193 (-1.07)	-0.0193 (-0.97)	-0.0520*** (-10.13)
Monthly or govt. wages	0.280*** (3.14)	0.151 (1.41)	0.151 (1.49)	0.127* (1.87)
Business owner	-0.0208 (-0.39)	0.310** (2.37)	0.310*** (2.62)	-0.230*** (-5.33)
Crops other than rice	-0.338*** (-6.58)	-0.467*** (-7.67)	-0.467*** (-7.46)	0.0336 (0.56)
Livestock	-0.137** (-2.36)	0.245*** (2.60)	0.245*** (3.15)	0.237*** (3.72)
Bluecollar worker	-0.361*** (-6.37)	0.506*** (3.24)	0.506*** (3.41)	
Whitecollar worker	-0.436*** (-3.72)	-0.0486 (-0.17)	-0.0486 (-0.16)	
Other work	-0.358*** (-7.75)	0.0955 (1.01)	0.0955 (1.05)	
Work that varies	-0.148** (-2.17)	0.228** (2.35)	0.228** (2.47)	
Non working age	0.00572 (0.48)	-0.0516* (-1.72)	-0.0516 (-1.57)	
Schooling of children	0.0408** (2.10)	-0.0432 (-1.00)	-0.0432 (-0.99)	
Schooling of children squared	-0.00187** (-2.06)	-0.000324 (-0.25)	-0.000324 (-0.28)	
Number of Adult children	0.178*** (18.04)	0.0966*** (2.85)	0.0966** (2.44)	0.178*** (16.88)
Time	0.0920*** (6.96)	0.107*** (2.97)	0.107*** (2.91)	0.0985*** (6.51)
Time squared	-0.00414*** (-4.69)	-0.00299* (-1.82)	-0.00299* (-1.91)	-0.00375*** (-3.61)
Rice farmer				0.490*** (12.08)
_cons	-4.606*** (-11.34)	-0.839 (-0.81)	-0.839 (-0.73)	-4.498*** (-9.97)
N	9513	9513	9513	9513

Assets and remittances in 100,000 Baht
t statistics in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

Table 4.8: The Proportion of Remittances in Income (Drought as Instrument)

	(1)	(2)	(3)	(4)
Endogenous Variables	*Observed Assets*	*Predicted assets*	*Predicted assets*	*Not included*
Standard errors	*Analytical*	*Clustered errors*	*Bootstrapped errors*	*Analytical*
Agricultural assets	-0.0376*** (-7.94)	0.193*** (2.76)	0.193*** (2.87)	
Business assets	-0.0242*** (-8.09)	-0.307*** (-7.25)	-0.307*** (-6.73)	
Household assets	-0.0302*** (-7.01)	0.111 (0.99)	0.111 (0.98)	
Sex of head	0.0405*** (4.77)	0.0399*** (3.36)	0.0399*** (3.30)	0.0413*** (4.73)
Age of head	0.0389*** (10.72)	-0.000354 (-0.04)	-0.000354 (-0.04)	0.0368*** (7.88)
Age of head squared	-0.000296*** (-9.69)	0.00000656 (0.09)	0.00000656 (0.09)	-0.000281*** (-7.25)
Head's years Of schooling	-0.0126*** (-7.85)	-0.00629 (-1.25)	-0.00629 (-1.39)	-0.0145*** (-10.85)
Monthly or govt. wages	-0.0460* (-1.85)	-0.0731*** (-2.61)	-0.0731** (-2.49)	-0.0623*** (-3.69)
Business owner	-0.0400*** (-2.78)	0.0298 (0.83)	0.0298 (0.85)	-0.101*** (-8.46)
Crops other than rice	-0.118*** (-8.42)	-0.159*** (-9.35)	-0.159*** (-11.46)	0.0135 (0.82)
Livestock	-0.0217 (-1.51)	0.0831*** (3.46)	0.0831*** (3.10)	0.112*** (6.63)
Blue collar worker	-0.136*** (-8.68)	0.118*** (2.84)	0.118*** (3.07)	
White collar worker	-0.101*** (-2.99)	-0.00429 (-0.06)	-0.00429 (-0.06)	
Other work	-0.0984*** (-7.84)	0.0306 (1.19)	0.0306 (1.23)	
Work that varies	-0.0748*** (-4.15)	0.0397 (1.55)	0.0397* (1.79)	
Non working age	0.00132 (0.44)	-0.0203** (-2.37)	-0.0203** (-2.58)	
Schooling of children	0.0185*** (3.60)	-0.0128 (-1.03)	-0.0128 (-1.13)	
Schooling of children squared	-0.000754*** (-3.12)	-0.000149 (-0.38)	-0.000149 (-0.37)	
Number of Adult children	0.0371*** (16.19)	0.0112 (1.23)	0.0112 (1.21)	0.0358*** (15.35)
Time	0.0296*** (8.35)	0.0335*** (3.25)	0.0335*** (3.58)	0.0313*** (6.67)
Time squared	-0.00133*** (-5.70)	-0.000982** (-2.06)	-0.000982*** (-2.62)	-0.00119*** (-3.76)
Rice farmer				0.170*** (15.22)
_cons	-1.341*** (-12.20)	-0.121 (-0.40)	-0.121 (-0.43)	-1.298*** (-9.44)
N	9513	9513	9513	9513

Assets and remittances in 10,000 Baht
t statistics in parentheses
* p<0.10 ** p<0.05 *** p<0.01

Table 4.8 presents the results obtained by re-estimating equation 4.12 using counterfactual asset holdings constructed by using the drought instrument. These results should be compared to those in table 4.5 which estimates 4.12 using migration histories to instrument for current migration. Here too, the effect of (counterfactual) agricultural assets on the share of remittances in household income is positive and significant and those of business assets are significant and negative. Household assets remain insignificant, as they were in table 4.5.

The fourth column of the table presents the coefficients from the model estimating the intensive margin of remittance receipts with only the core set of explanatory variables. The Akaike information criteria for this model takes on a value of 8648.03, whereas that of the full model is 8705.17. The Bayesian information criteria for the model with the core variables is 8748.27, and that of the full model is 8869.86. Thus both information criteria recommend use of the core model with plausibly exogenous regressors, confirming the conclusions of sections 4.4 and 4.5.

I now turn to studying how the intensive and extensive margins of remittance receipts vary over the distribution of household income, so as to understand the relative importance of each of these margins in explaining the inequality reducing effect of remittances.

4.7 Accounting for Inequality on the Extensive and Intensive Margins

So far, the analysis in this chapter has been conducted at the level of individual households, though the key research question pertains to the distribution of income across households. I start this section by extrapolating from the preceding household-level results to characterise how the extensive and intensive margins of remittance receipts vary over the distribution of income. Having done so, I will move on to identifying the share of the reduction in inequality that can be explained econometrically, and to subsequently decomposing that share into parts that are explained by the extensive and intensive margins of remittances.

I characterise the estimated probability of a household receiving remittances, $\hat{d}$ (from equation 4.7) as a function of the real value of the component of household income that is not remitted by children using a univariate Tobit model, because the household-level predictions of the likelihood of receiving remittances vary continuously between zero and one. I assume a latent variable

$$\hat{d}_{it}^{\ *} = c + \varphi y^{-r}{}_{it} + \varepsilon_{it} \tag{4.13}$$

such that

$$\hat{d}_{it} = \begin{cases} 1\ if\ \hat{d}_{it}^{\ *} > 1 \\ 0\ if\ \hat{d}_{it}^{\ *} < 0 \\ \hat{d}_{it}^{\ *}, otherwise, \end{cases} \tag{4.14}$$

where φ is the parameter to be estimated and all other terms remain as they have been defined before. The estimates yielded by this regression are presented in the first column of table 4.9. The probability of receiving remittances does indeed vary significantly over the distribution of income, and the likelihood of remittance receipts is higher for households that are lower down the income distribution. The resulting estimates for the way in which the probability of receiving remittances varies over the distribution of household income (that was not remitted by children) is illustrated in figure 4.3. I denote these distribution level predictions as $\widehat{pr}$.

Table 4.9: The Extensive and Intensive Margins as Functions of Income not Remitted by Children

	Probability of receiving remittances	*Proportion of remittances in income*
Income not remitted[1]	-0.8882***	-0.5293***
	(-10.05)	(-15.09)
Constant	0.5270***	0.1911***
	(196.52)	(173.70)
N	9512	9512

[1]*In10,000,000 Baht*
t statistics in parentheses
* p<0.10, ** p<0.05,*** p<0.01"

Heterogeneity on the intensive margin of remittance receipts over the distribution of household income is characterised in an exactly analogous way, only with the estimated probability of receiving remittances $\hat{d}$, replaced by the estimated share of remittances in household income $\hat{s}$, in equations 4.13 and 4.14. The estimates from this model are presented in the second column of table 4.9. The share of remittances in household income is also found to decline significantly as the level of household real income that was not remitted by children increases. Figure 4.4 illustrates the resulting predicted values of how the share of remittances in household income varies over the distribution of this component of income. In the following analysis, I refer to these distribution-level predicted shares as $\widehat{sh}$. These shares are illustrated in figure 4.4.

Using these estimates for the way in which the likelihood of receiving remittances and the share of remittances in household income varies over the distribution of income, I will now construct three counterfactual distributions of income. In the first, I allow the extensive margin of remittance receipts to vary over the distribution of household income,

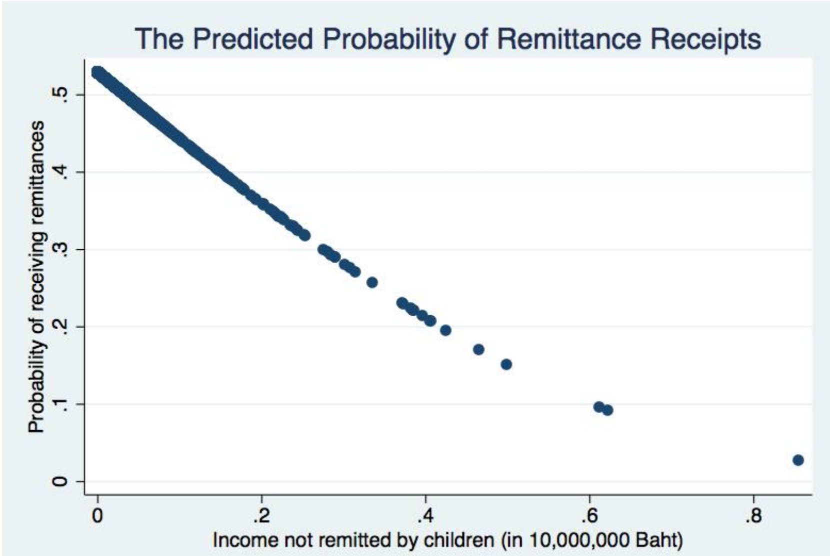


Figure 4.3

using the values illustrated in figure 4.3, while holding the share of income remitted to each household constant at the sample mean of 33.4% (from table 4.1). That is, I compute:

$$\widehat{y_{it}} = y_{it}^{-r} \cdot (1 + 0.3337 \cdot \widehat{pr}(d = 1|y_{it}^{-r})), \tag{4.15}$$

where $\widehat{pr}$ is the probability of a household receiving remittances conditional on its observed level of income that was not remitted by children. In what follows, I will refer to this as 'the extensive margin' distribution.

The second counterfactual income distribution that I generate is where I allow the intensive margin of remittance receipts to vary as a function of income that was not remitted by children, as it illustrated in figure 4.4, while holding the likelihood with which each household receives remittances at the sample mean of 0.4832. I do so by computing:

$$\hat{\hat{y}}_{it} = y_{it}^{-r}(1 + 0.4832 \cdot \widehat{sh}(y_{it}^{-r})). \tag{4.16}$$

I will refer to this distribution as the 'intensive margin' distribution.

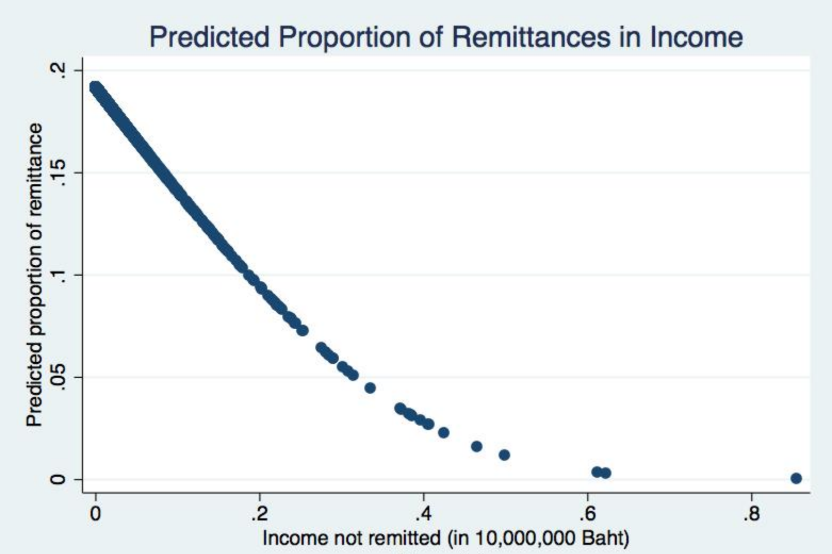


Figure 4.4

Finally, I construct a counterfactual distribution where I allow both of these to margins to vary simultaneously. That is, I calculate:

$$\hat{\hat{\hat{y}}}_{it} = y_{it}^{-r}(1 + 0.3337 \cdot \widehat{pr}(d = 1|y_{it}^{-r}))(1 + 0.4832 \cdot \widehat{sh}(y_{it}^{-r})). \quad (4.17)$$

I refer to this as the 'explained' distribution, because it captures the total shift in the distribution of income that I am able explain econometrically by a combination of the extensive and intensive margins of remittance receipts.

I now turn to assessing the degree of inequality in each of these counterfactual distributions of income for each year of the panel. I also compare measured inequality in each of these distributions to measured inequality in the distribution of income not remitted by children, and in the distribution of net household income, inclusive of remittances from children. I do so for the same range of inequality measures that I used in the preceding chapter[4], namely the standard deviation of the log, the Gini coefficient, the Theil (T) coefficient and the mean logarithmic deviation. The extent to which I am able to explain the reduction in measured inequality varies depending on the measure that

[4] The definitions of each of these measures and a brief description of their relevant properties can also be found in chapter 3.

is used, but I am able to make some general comments on the extent to which the extensive and intensive margins of remittance receipts account for the explained reduction in inequality.

Table 4.10 presents the Theil index, evaluated for each of these distributions for each year of the panel. The first column measures inequality in the component of income that is not remitted by children and the second does so for the 'extensive margin' distribution. Comparing the values reported in these two columns, I find that the second column does indeed exhibit less inequality than the first, in all years of the panel. The third column presents the Theil index evaluated on the counterfactual distribution where only the intensive margin is allowed to vary. This distribution also exhibits less inequality in every year than the component of income that is not remitted by children. A comparison of the levels of inequality reported in the second and third columns reveals that on average, the intensive margin accounts for a slightly larger reduction in inequality than the extensive margin.

Table 4.10: Thiel (T) Index for different Income Distributions

Year	*Income not remitted*	*Extensive margin counterfactual*	*Intensive margin counterfactual*	*Counterfactual where both vary*	*Observed net income*	*Proportion of observed decline explained*
1997	0.7630	0.7437	0.7459	0.7273	0.7034	0.5995
1998	0.7543	0.7239	0.7320	0.7036	0.6942	0.8445
1999	0.7636	0.7185	0.7357	0.6941	0.7025	1.1383
2000	0.7940	0.7542	0.7650	0.7279	0.7286	1.0108
2001	0.6558	0.6344	0.6389	0.6188	0.5961	0.6193
2002	0.6275	0.6159	0.6171	0.6059	0.5461	0.2651
2003	0.6380	0.6201	0.6225	0.6053	0.5668	0.4591
2004	0.6913	0.6661	0.6703	0.6465	0.5908	0.4462
2005	0.7137	0.6874	0.6918	0.6670	0.6039	0.4260
2006	0.5933	0.5784	0.5807	0.5666	0.4745	0.2244
2007	0.6222	0.6051	0.6075	0.5913	0.4933	0.2397
2008	0.5494	0.5404	0.5412	0.5325	0.4259	0.1370
2009	0.5950	0.5774	0.5799	0.5632	0.4780	0.2722
2010	0.5119	0.5003	0.5015	0.4904	0.4021	0.1959
2011	0.5118	0.4969	0.4987	0.4844	0.4235	0.3110
Average percentage of decline explained		54.79%	45.20%	47.93%		

Table 4.11: Mean Logarithmic Deviation for different Income Distributions

Year	*Income not remitted*	*Extensive margin counterfactual*	*Intensive margin counterfactual*	*Counterfactual where both vary*	*Observed net income*	*Proportion of observed decline explained*
1997	1997	0.7423	0.7313	0.7325	0.7218	0.6693
1998	1998	0.6949	0.6818	0.6846	0.6721	0.6114
1999	1999	0.6663	0.6503	0.6555	0.6405	0.5955
2000	2000	0.6850	0.6687	0.6726	0.6571	0.6078
2001	2001	0.6263	0.6161	0.6178	0.6080	0.5369
2002	2002	0.6624	0.6551	0.6558	0.6485	0.5350
2003	2003	0.6011	0.5912	0.5923	0.5827	0.5057
2004	2004	0.6755	0.6628	0.6646	0.6523	0.5084
2005	2005	0.7081	0.6950	0.6970	0.6845	0.5323
2006	2006	0.6686	0.6606	0.6616	0.6538	0.4571
2007	2007	0.6732	0.6637	0.6649	0.6557	0.4447
2008	2008	0.6336	0.6271	0.6277	0.6213	0.4126
2009	2009	0.6098	0.5996	0.6009	0.5910	0.4243
2010	2010	0.5430	0.5353	0.5360	0.5285	0.3723
2011	2011	0.5107	0.5013	0.5024	0.4933	0.3794
Average percentage of decline explained		54.07%	45.93%	17.20%		

Table 4.12: Gini Coefficient for different Income Distributions

Year	*Income not remitted*	*Extensive margin counterfactual*	*Intensive margin counterfactual*	*Counterfactual where both vary*	*Observed net income*	*Proportion of observed decline explained*
1997	0.6071	0.6026	0.6031	0.5987	0.5844	0.3666
1998	0.5919	0.5864	0.5876	0.5822	0.5683	0.4107
1999	0.5798	0.5730	0.5752	0.5687	0.5580	0.5120
2000	0.5938	0.5869	0.5886	0.5820	0.5700	0.4964
2001	0.5693	0.5647	0.5655	0.5611	0.5435	0.3185
2002	0.5762	0.5730	0.5733	0.5701	0.5390	0.1626
2003	0.5642	0.5596	0.5602	0.5557	0.5318	0.2616
2004	0.5746	0.5690	0.5698	0.5643	0.5288	0.2253
2005	0.5835	0.5779	0.5787	0.5733	0.5349	0.2097
2006	0.5546	0.5510	0.5514	0.5478	0.4959	0.1158
2007	0.5638	0.5595	0.5601	0.5559	0.4985	0.1212
2008	0.5494	0.5464	0.5466	0.5437	0.4831	0.0861
2009	0.5523	0.5475	0.5481	0.5434	0.4921	0.1475
2010	0.5253	0.5214	0.5218	0.5179	0.4632	0.1185
2011	0.5177	0.5129	0.5134	0.5086	0.4688	0.1861
Average percentage of decline explained			54.11%	45.89%	24.92 %	

Unsurprisingly, when both margins are allowed to vary, inequality declines even further. These results are reported in the fourth column of the table. Comparing these measures to those in the fifth column (which reports inequality in net household income inclusive of observed remittances) reveals that inequality in actual net income was typically lower than in this counterfactual. Thus while the econometrics presented in this chapter have been able to explain some of the decline in inequality that was caused by remittance receipts, they have not been able to explain it all (the exceptions are the years 1999 and 2000 where inequality is higher in observed net income than in this counterfactual). The final column of the table computes the percentage of the observed decline in inequality that these models have been able to explain, showing that there is substantial variation in the performance of these models from year to year.

In the final row of table 4.10 I present some percentages which summarize the relative success of each of the counterfactuals[5]. On average, the share of the decline in inequality that is explained by the extensive margin is 54.8%, and the corresponding share for the intensive margin is 45.2%. Allowing both margins to vary simultaneously on average explains 47.9% of the measured decline in the Theil index between the distributions of income not remitted by children, and net income.

Table 4.11 presents the corresponding results where the mean logarithmic deviation (sometimes called the Theil-L index) is used to measure inequality. Again, from year to year there is some variation in the extent to which allowing the extensive and intensive margins to vary simultaneously explains the reduction in this measure of inequality. On average however, the models perform reasonably well explaining 17.2% of the observed

[5] Not all of the inequality measures presented here are additively decomposable. As a result, percentages calculated in the obvious way do not add up to 100. Instead, I compute the ‘total decline’ as the sum of the declines in measured inequality between income not remitted and the extensive and intensive margins, and express the declines in each margin as a percentage of this total.

decline in inequality. What remains consistent is that the extensive margin explains slightly more of the decline in inequality than the intensive margin, this time accounting for 52.9% of the decline.

A similar narrative emerges from table 4.12, which presents the Gini coefficients for these distributions for the 15 years of the panel. This time, on average I am able to explain 24.9% of the observed reduction in inequality. The intensive margin again accounts for slightly more than half the explained decline in inequality, approximately 52.9%.

Table 4.13 uses the standard deviation of the log of income to measure of inequality. On this measure, a great deal of the observed decline in inequality remains unexplained, with the counterfactual where both the extensive and intensive margins are both allowed to vary explaining only 4.6% of the observed decline in inequality. Within the explained decline in inequality however, the results continue to be very consistent with the preceding tables. The extensive margin is again slightly more important than the intensive margin, explaining 52.4% of the reduction.

Table 4.13: Standard Deviation of the Log for different Income Distributions

Year	*Income not remitted*	*Extensive margin counterfactual*	*Intensive margin counterfactual*	*Counterfactual where both vary*	*Observed net income*	*Proportion of observed decline explained*
1997	1.2307	1.2268	1.2272	1.2233	1.1569	0.0997
1998	1.1899	1.1862	1.1867	1.1830	1.0781	0.0617
1999	1.1593	1.1558	1.1564	1.1529	1.0731	0.0745
2000	1.1415	1.1375	1.1380	1.1340	1.0454	0.0777
2001	1.1277	1.1241	1.1244	1.1208	0.9949	0.0521
2002	1.2014	1.1979	1.1982	1.1948	1.0230	0.0371
2003	1.0813	1.0771	1.0775	1.0733	0.9557	0.0636
2004	1.2180	1.2137	1.2141	1.2099	0.9598	0.0314
2005	1.2483	1.2441	1.2445	1.2404	1.0010	0.0318
2006	1.2949	1.2916	1.2919	1.2886	0.9625	0.0189
2007	1.2640	1.2602	1.2605	1.2567	0.9116	0.0207
2008	1.2381	1.2345	1.2348	1.2313	0.8938	0.0199
2009	1.1622	1.1576	1.1581	1.1536	0.8759	0.0301
2010	1.1026	1.0982	1.0986	1.0942	0.8364	0.0313
2011	1.0353	1.0302	1.0307	1.0256	0.8299	0.0473
Average percentage of decline explained			52.69%	47.30%	4.65 %	

4.8 Conclusion

This chapter has sought to understand the relative importance of the extensive margin and the intensive margin of remittance receipts in explaining the ability of remittances to reduce inequality. To do so, it modelled the likelihood of a household receiving remittances and the proportion of remittances in a household's income as functions of observable household characteristics.

Some potentially endogenous regressors appeared to have significant explanatory power in single-equation models. Addressing the endogeneity however, revealed that the apparent power of these variables in explaining remittance receipts was spurious. The chapter therefore went on to model the extensive and intensive margins of remittance receipts using a core group of plausibly exogenous variables. I used the resulting estimates to construct counterfactual distributions of income where I allowed either the probability of receiving remittances, or the proportion of remittances in household income, or both of these margins to vary.

I then used a variety of measures of inequality to assess the extent to which these econometrics explained measured reductions in inequality, and the relative importance of the extensive and intensive margins of remittance receipts in driving these reductions. For the Theil (T) coefficient the extensive and intensive margins together explained approximately 47.9% of the observed reduction in inequality. The performance of these models were somewhat diminished when the mean logarithmic deviation, which is more sensitive to transfers near the bottom of the distribution than the Theil index, was used to measure inequality. On that measure, the models explained approximately 17.2% of the observed reduction in inequality. For the Gini index and the standard deviation of the log, the models were able to explain 24.9% and 4.6% of the observed reduction respectively.

Again, the models performed better on the measure that was less sensitive to transfers near the bottom of the distribution.

For all measures of inequality, the decline in inequality was slightly larger on the extensive margin than the intensive margin, with the former accounting for between 52 and 55% of the explained decline depending on the measure of inequality used.

Chapter 5

Conclusion

The recurrent themes in this thesis have been inequality and insurance. In each of the three preceding chapters, I have studied some particular aspect of the behaviour of rural, Thai households with respect to these themes, with detailed panel data that has been made available by the Townsend Thai Project.

The second chapter of this thesis studied the extent to which these households depend on the smoothness of their income streams to satisfy their insurance needs. Descriptive evidence suggested that in rural Thailand, richer households derive a greater share of their insurance needs from the smoothness of their income streams, while poorer households depend more heavily on traditional consumption smoothing. These descriptive statistics were inconsistent with the existing narratives of income smoothing which have typically studied this insurance strategy among the poor and vulnerable. The chapter showed formally, using Morduch's (1994) analytical model of income smoothing that if the income streams of richer households offer greater insurance possibilities then (holding all other things including credit constraints constant) these rich households will depend more heavily on their income streams to satisfy their insurance needs than their poorer counterparts. I used the incidence of drought as a source of exogenous variation in the income streams of these households. I failed to reject the null hypothesis that income streams were fully insured against this covariate shock for households that were headed

by people from younger cohorts, by people with more than a primary education, or by people who were in jobs that paid a monthly wage. These characteristics were more prevalent among richer households, and so the income streams of richer households offered better insurance possibilities. I was unable to reject the hypothesis that households which enjoyed above median permanent income over the duration of the panel had income streams that were fully insured against the covariate shock. In contrast, the effect of drought on income was double the average effect for the whole sample when I restricted attention to households which exhibited below median permanent income.
Chapter 2 thus contributes to this literature by re-interpreting the existing theory to allow for the possibility that the income streams of richer households offer better insurance possibilities, and by documenting empirical results from rural communities in a middle-income country that support this re-interpretation. It remains to be established whether these insights are peculiar to rural Thailand or if they are of broader relevance. The increasing availability of high quality panel data from a number of low- and middle-income countries means that future research can convincingly answer this question. If these insights are more broadly applicable, then the increased prevalence of salaried jobs in many rapidly industrializing parts of the world may enhance workers' welfare, not only by increasing their mean wages, but also by fulfilling a valuable insurance function. Welfare evaluations that neglect this insurance function are likely to underestimate the welfare gains from industrialization. Future work may also be directed at carefully quantifying these welfare gains so as to better inform the policies that many governments undertake to foster industrialization.

The third chapter of this thesis documented and analysed declining income inequality in a panel of Thai households with reference to the permanent income hypothesis and lifecycle theory. I found no evidence of convergence in individual earnings, or differences in the propensity of adult children to cohabitate with their parents between richer and poorer households that could plausibly explain the decline in inequality within decade of birth cohorts of the heads of household. Rather, I found that declining inequality was explained by differences in the receipt of remittances from the children of the head of household, which supplemented the incomes of the ageing parents. Over the duration of the panel the average of the proportion of remittances in the incomes of households in the bottom decile of permanent income was over 35% whereas that of households in the top decile of permanent income was less than 10%. The strong lifecycle component of this inequality reducing transfer causes it to become more and more prevalent as households age over the duration of the panel. The chapter also presented some descriptive evidence that poorer households were both more likely to receive remittances, and that remittances accounted for a larger share of their household income than it did for richer households.

The main contribution of this chapter is to study the link between remittance receipts and income inequality from a lifecycle perspective. The robust decrease in income inequality, even between households grouped by the cohort of birth of their heads is a striking finding that contrasts with the results of the usual studies of the evolution of inequality over the lifecycle. These descriptive results can be fruitfully extended in future research by implementing matching methods on the observable characteristics of households which could, in principle, justify a stronger causal interpretation. Furthermore, the narrative that younger generations of adults are responsible for providing material support to their ageing parents is germane to a number of East Asian countries. Widespread, rapid industrialization and urbanization experienced by many of these countries over the last quarter century are likely to have created similarly high rates of reliance on remittances among the family members of rural to urban migrants that remain in the communities of origin. Future research may therefore also be directed at ascertaining whether remittances engender similar declines in lifecycle inequality in these other East Asian countries.

Chapter 4 of this thesis studied the inequality reducing impact of remittances more formally, by modelling the extensive and intensive margins of remittances as functions of observable household characteristics. It noted that a household's economic position was in part determined by its migration history, so that some current household characteristics such as asset holdings might be endogenous to the migration process. These variables appeared to explain substantial variation on both of these margins in single-equation models. The paper addressed endogeneity by explicitly modelling the effect of remittances on assets holdings, and found that doing so relieved assets of any substantive explanatory power over the extensive and intensive margins of remittance receipts. Therefore, I proceeded by using a core set of exogenous characteristics to model these margins.

Using these estimates, I identified the degree to which the likelihood of receiving remittances and the share of remittances in household income varied over the distribution of income. These estimates informed the construction of three counterfactual income distributions: 1. where the extensive margin of remittances were allowed to vary over the distribution on income, but the intensive was held fixed at the sample mean; 2. where the intensive was allowed to vary but the extensive margin was held at the sample mean; and 3. where both margins were allowed to vary simultaneously. The last counterfactual distribution allowed me to quantify the share of the inequality reducing effect of remittances which these econometrics were able to explain, whereas the first two allowed me to measure the relative importance of the extensive and intensive margins in accounting for these explained reductions. I evaluated inequality for these distributions for a variety of different measures. Though the share of net explained inequality varied from measure to measure, the proportions that were explained by the extensive and intensive margins remained remarkably consistent. The extensive margin always accounted for slightly more than half (between 52% and 55%) of the explained reduction in inequality with the intensive margin accounting for the rest.

While this chapter addressed the potential endogeneity of various household characteristics to the migration process, it was unable to adequately account for behavioural responses that remittance receipts may engender among recipients. Given the observational nature of these data, matching remittance recipients to non-recipients appears at first to be an attractive option. But while there is an established literature on using matching methods to identify average treatment effects, I do not know of any such methods that are suitable to identify differential treatment effects across a distribution, as is the objective here. Indeed, McKenzie et al. (2010) suggest that there may be no way to comprehensively address this issue except with the use of experimental data.

This thesis has demonstrated inequalities that affect households among these villages in rural Thailand are substantive and pervasive. It has documented that the welfare of the relatively poor in these communities is more affected by a lack of insurance opportunities than was previously understood. The income streams of these poor village households bear the brunt of uninsurable covariate shocks. But these poor also benefit disproportionately from the receipt of remittances from their migrant offspring. The support that poor ageing heads of household receive from their children is striking in that it is sufficient to reverse the increase in inequality that is typically observed within birth cohorts over the lifecycle.

Bibliography

Adams, R.H. Jr, (1989) "Worker Remittances and Inequality in Rural Egypt," *Economic Development and Cultural Change*, University of Chicago Press, vol. 38(1), pages 45-71, October

Aiken, L. S., & West, S. G. (1991) *Multiple regression: Testing and interpreting interactions.* Newbury Park: Sage.

Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B.N.; Csáki, F., *2nd International Symposium on Information Theory*, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, p. 267-281.

Alem, M., & Townsend, R. M. (2014). "An evaluation of financial institutions: Impact on consumption and investment using panel data and the theory of risk-bearing". *Journal of Econometrics*, 183(1), 91-103.

Amemiya, T. (1978) "The Estimation of a Simultaneous Equation Generalized Probit Model" *Econometrica*, Vol. 46 No. 5 pp.1193-1205

Ando, A., and F. Modigliani (1963) "The Life-Cycle Hypothesis of Saving: Aggregate Implications and Tests" *American Economic Review*, Vol. 53 No. 1 pp. 55-84

Atkinson, A. B. (1970) "On the measurement of inequality" *Journal of Economic Theory*, 2(3), 244-263.

Attanasio, O. P. & Pavoni, N. (2011) "Risk sharing in private information models with asset accumulation: Explaining the excess smoothness of consumption" *Econometrica*, 79(4), 1027-1068.

Attanasio, O. and G. Weber (1995) "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy*, 103(6):1121–57.

Attanasio, O. and G. Weber (2010) "Consumption and Saving: Models of temporal Allocation and Their Implications for Public Policy." *Journal of Economic Literature*, Vol 48 pp 693-691.

Arrow, K. J., (1951) "An Extension of the Basic Theorems of Classical Welfare Economics." in *Proceedings of the Second Berkley Symposium on Mathematical Statistics and Probability*, J. Neyman (ed.) Berkeley and Los Angeles: University of California Press, pp 507-532.

Arrow, K. J. and G. Debreu, (1954) "Existence of an Equilibrium for a Competitive Economy." *Econometrica* Vol. 22 No. 3 pp. 265-290.

Banerjee, A., X. Meng, N. Qian, (2010) “The life cycle model and household savings: micro evidence from urban China.” Available from Http://econ.yeale.edu/nq3/.

Becker, G., N. Tomes (1979) “An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility”, *Journal of Political Economy,* 87, 1153-1189.

Blundell, R., L. Pistaferri and I. Preston (2008) “Consumption Inequality and Partial Insurance.” *American Economic Review*, Vol. 98 No. 5 pp. 1887-1921.

Blundell, R. & I. Preston (1998) "Consumption Inequality and Income Uncertainty," *The Quarterly Journal of Economics*, Oxford University Press, vol. 113(2), pages 603-640.

Bonke, T., G. Corneo & H. Luthen, (2015) "Lifetime Earnings Inequality in Germany," *Journal of Labor Economics*, University of Chicago Press, vol. 33(1), pages 171 - 208

Cai, F. and Giles, J. and Meng, X., (2006) “How Well Do Children Insure Parents Against Low Retirement Income? An Analysis Using Survey Data from Urban China.” *Journal of Public Economics,* Vol. 90 No. 12.

Campbell, J., & Deaton, A. (1989). “Why is consumption so smooth?” *The Review of Economic Studies*, 56(3), 357-373.

Carter, M.R. and C. Barrett (2006) “The Economics of Poverty Traps and Persistent Poverty: an Asset-Based Approach” *Journal of Development Studies,* Vol. 42 No. 2 pp. 178-199.

Chatterjee, A. & Singh, A. & Stone, T. (2015) "Understanding Wage Inequality in Australia," *Working Papers 2015-06*, University of Sydney, School of Economics.

Chiappori, Pierre-André, Krislert Samphantharak, Sam Schulhofer-Wohl, and Robert M Townsend. (2011) *Heterogeneity and Risk Sharing in Village Economies.* NBER Working Paper no. 16696

Chiappori, Pierre-Andre, Krislert Samphantharak, Sam Schulhofer-Wohl and Robert Townsend (2013) “Portfolio Choices and Risk Sharing in Village Economies,” Federal Reserve Bank of Minneapolis Research Department Working Paper 706.

Chiappori, Pierre-Andre, Krislert Samphantharak, Sam Schulhofer-Wohl and Robert Townsend, (2014) “Heterogeneity and Risk Sharing in Village Economies,” *Quantitative* Economics 5(1).

Cochrane, J.H., (1991) “A Simple Test of Consumption Insurance.” *Journal of Political Economy,* Vol. 5 No. 1 pp. 957-976

Dalton, H. (1920) “The Measurement of the Inequality of Incomes”, *The Economic Journal,* Vol. 30, No. 119 pp. 348-361.

Deaton, A., (1989) "Saving in Developing Countries: Theory and Review," Papers 144, Princeton, Woodrow Wilson School - Development Studies

Deaton, A. (1991) “Savings and Liquidity Constraints.” *Econometrica*, Vol. 59 No. 5, pp. 1221-1248.

Deaton, A. (1992) “Saving and Income Smoothing in Cote d’Ivoire.” *Journal of African Economies*, Vol. 1 No. 1 pp. 1-24

Deaton, A. and C. Paxson, (1994) “Intertemporal Choice and Inequality”, *The Journal of Political Economy* Vol 102, No. 3 pp. 437 -467.

Deaton, A. and C. Paxson, (1995) “Saving, Inequality and Aging: an East Asian Perspective” *Asia-Pacific Economic Review*, Vol 1.

Debreu, G., (1952) “A Social Equilibrium Existence Theorem.” *Proceedings of the National Academy of Sciences* Vol. 38, No. 10 pp. 886-893.

Dercon, S. (2002)“Income risk, Coping Strategies, and Safety Nets”, *The World Bank Research Observer*, 17 (2), 141-166.

Dercon, S. (2004) “Growth and shocks: evidence from rural Ethiopia.” *Journal of Development Economics* Vol. 74 pp. 309-329.

Dercon, S. (2006) “Vulnerability: A Micro Perspective.” QEH Working Paper Series No. 149, University of Oxford.

Dercon, S. and P. Krishnan, (1996) “Income Portfolios in Rural Ethiopia and Tanzania: Choices and Constraints”, *Journal of Development Studies*, 32 (6), 850-875.

Dickens, R. (2000) "The Evolution of Individual Male Earnings in Great Britain: 1975-95," *Economic Journal*, Royal Economic Society, vol. 110(460), pages 27-49, January.

Felkner, J. S., & Townsend, R. M. (2011). “The geographic concentration of enterprise in developing countries,” *The Quarterly Journal of Economics*, 126(4), 2005.

Friedman, M. (1957) *A Theory of the Consumption Function*, Princeton University Press.

Giné, X., & Townsend, R. M. (2004) “Evaluation of financial liberalization: a general equilibrium model with constrained occupation choice,” *Journal of Development Economics*, 74(2), 269-307.

Gutierrez, F.H. (2014) “Acute Morbidity and Labour Market Outcomes in Mexico: Testing the Role of Labor Contracts as an Income Smoothing Mechanism”, *Journal of Development Economics*, 110:1-12.

Hall, R. E. (1978) “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence”. *Journal of Political Economy*, 86(6), 971-87.

Hall, R. E. and F. S. Mishkin (1982) “The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households.” *Econometrica* Vol. 50 No. 2 pp. 461-481.

Jaccard, J. and R. Turrisi (2003) *Interaction Effects in Multiple Regression.* Sage.

Jappelli, T. and L. Pistaferri (2010) “Does Consumption Inequality Track Income Inequality in Italy?” *Review of Economic Dynamics,* Vol 13 No. 1.

Jenkins, S. P. (2015). “INEQDECO: Stata module to calculate inequality indices with decomposition by subgroup,” *Statistical Software Components*.

Kaboski, J. P., & Townsend, R. M. (2005) Policies and Impact: An Analysis of “Village-Level Microfinance Institutions*” Journal of the European Economic Association,* 3(1), 1-50.

Kaboski, J. P., & Townsend, R. M. (2011) “A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative,” *Econometrica,* 79(5), 1357-1406.

Kaboski, J. P., & Townsend, R. M. (2012). “The impact of credit on village economies,” *American Economic Journal: Applied Economics*, 4(2), 98-133.

Kochar, A. (1999) “Smoothing Consumption by Smoothing Income: Hours-of-Work Responses to Idiosyncratic Agricultural Shocks in Rural India” *The Review of Economics and Statistics,* 81(1), 50-61.

Kotlikoff, L. J. and A. Spivak (1981) “The Family as an Incomplete Annuities Market” *Journal of Political Economy* Vol. 89.

Kremer, M. &, D.L. Chen, (2002) “Income Distribution Dynamics with Endogenous Fertility”, *Journal of Economic Growth* 227-258

Lam, D. (1986) “The Dynamics of Population Growth, Differential Fertility, and Inequality”, *American Economic Review,* 76(5): 1103-16.

Lipton, M. (1980) "Migration from rural areas of poor countries: The impact on rural productivity and income distribution," *World Development,* Elsevier, vol. 8(1), pages 1-24, January.

Mare, R. D. (2011) “A Multigenerational View of Inequality” *Demography*, Volume 48, Number 1, Page 1

Mazzocco, M., and S. Saini (2012) "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences." *The American Economic Review* Vol. 102 No. 1 pp. 428-468.

McKenzie, D. & Rapoport, H. (2007) "Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico," *Journal of Development Economics,* Elsevier, vol. 84(1), pages 1-24, September.

McKenzie, D., Stillman, S., & Gibson, J. (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.

Mincer, J (1962) “Labor Force Participation of Married Women: A Study of Labor Supply.” *Aspects of Labor Economics,* Princeton University Press.

Morduch, J. (1994) "Poverty and Vulnerability," *American Economic Review,* American Economic Association, vol. 84(2), pages 221-25.

Morduch, J. (1995) "Income Smoothing and Consumption Smoothing," *Journal of Economic Perspectives*, American Economic Association, vol. 9(3), pages 103-114, Summer.

Nelson, Julie A. (1994) “On Testing for Full Insurance Using Consumer Expenditure Survey Data: Comment.” *Journal of Political Economy,* 102(2): 384–94.

OECD (1982) *The OECD List of Social Indicators*, Paris.

Oliveira, J. (2016) “The value of children: Inter-generational support, fertility and human capital” *Journal of Development Economics* Vol. 120 No. 1.

Paulson, A. L. (2000) “Insurance Motives for Migration: Evidence from Thailand”, Kellogg School, Northwestern University.

Paulson, A. L., & Townsend, R. (2004) “Entrepreneurship and financial constraints in Thailand,” *Journal of Corporate Finance,* 10(2), 229-262.

Piketty, T. (2013) *Capital in the Twenty-First Century*, Harvard University Press.

Poggi, C. (2015) *Credit Availability and Internal Migration: Evidence from Thailand*. Presented at the 2015 PhD. conference at the University of Sussex.

Rosati N. (2003) “How has economic inequality evolved over the past two decades? A look at the Italian experience”. *Research in Economics*, 57, 93-122

Rosenzweig, M.R. and H.P. Binswanger (1993) “Wealth, Weather Risk and the Profitability of Agricultural Investments”, *The Economic Journal*, Vol 103, No. 416 pp. 56-78.

Rosenzweig, M. R. & K. Wolpin, (1993) "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India," *Journal of Political Economy,* University of Chicago Press, vol. 101(2), pages 223-44, April.

Samphantharak, K. and Robert M. Townsend (2013) "Risk and Return in Village Economies.”, Working paper, 2017

Schulhofer-Wohl, S. (2011) “Heterogeneity and tests of risk sharing,” *Journal of Political Economy*, 119(5), 925-958.

Schwarz, Gideon E. (1978), "Estimating the dimension of a model", *Annals of Statistics*, 6 (2): 461–464.

Shorrocks, A. F. (1980), “The class of additively decomposable inequality measures,” *Econometrica*, 48(3), 613-625.

Stark, O. and D. E. Bloom (1985) “The New Economics of Labor Migration” *The American Economic Review* Vol. 75, No. 2, 173-178.

Stark, O., Taylor, J.E. & Yitzhaki, S. (1986) "Remittances and Inequality," *Economic Journal*, Royal Economic Society, vol. 96(383), pages 722-40, September.

Tobin, J. (1958) “Estimation of Relationships for Limited Dependent Variables”, *Econometrica,* Vol 26. No1 pp 24-36.

Townsend, R. M. (1994) “Risk and Insurance in Village India.” *Econometrica* Vol. 62, No. 3 pp. 539-591.

Townsend, R. M. (2011) “Townsend Thai Project Household Annual Resurvey 1997 – 2011” http://hdl.handle.net/1902.1/10673UNF:3:bN2yS4jbisVzbRVs8ZF0xg== Robert M. Townsend;Murray Research Archive [Distributor] V2 [Version]
Udry, Christopher. (1994) “Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria.” *Review of Economic Studies*, Vol. 61 No. 3 pp. 495–526.

Townsend, R. M. (2013) "Accounting for the Poor", *American Journal of Agricultural Economics* 95(5): 1196-1208

Willis, R. J (1979) “The Old Age Security Hypothesis and Population Growth.” mimeo, State University of New York at Stony Brook.

World Bank (2016) *Country Overview: Thailand.* Updated September 2016. Available at: http://www.worldbank.org/en/country/thailand/overview

Yamada, T. (2009) "Income Risk, Consumption Inequality, and Macroeconomy in Japan," *Global COE Hi-Stat Discussion Paper Series* gd08-041, Institute of Economic Research, Hitotsubashi University.

Yang, Liu (2004) “Unequal Provinces But Equal Families? An Analysis of Inequality and Migration in Thailand,” Working paper.

Appendix 1: Deleted observations

Table A1: Deleted observations			
Household ID:	*year*	*Expenditure*	*Income*
072136660911	1998	430	99999997952
072840610910	1997	33603.13	116951000
493141600924	2010	0	12000
532938650940	1999	5.88E+10	764.7059
493141650962	1999	2.78E+10	63097.22
493141600902	1998	3.23E+10	34370.97
492838650923	1998	7.32E+10	50243.9
492624650927	1999	2.78E+10	22666.67
272931620968	1999	4.88E+10	24817.07
072436540977	1998	4.17E+10	12500

Appendix 2: Robustness of differential insurance over the income distribution

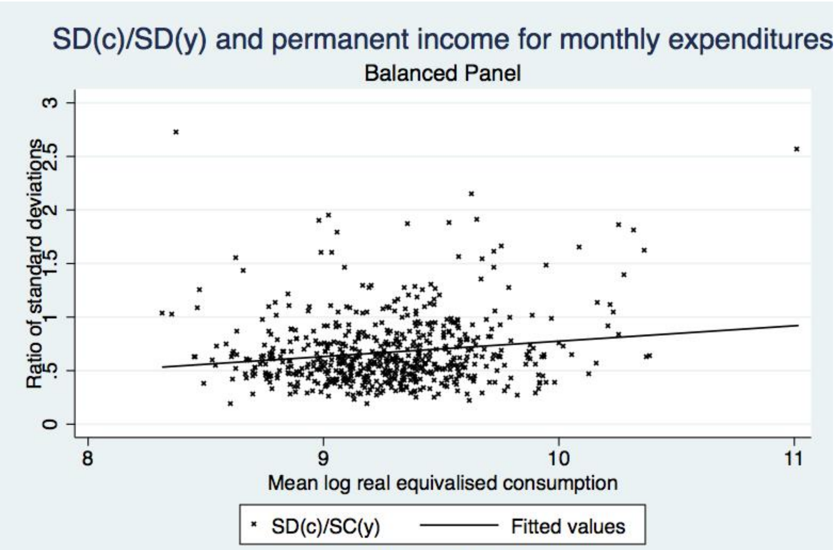


Figure A2.1

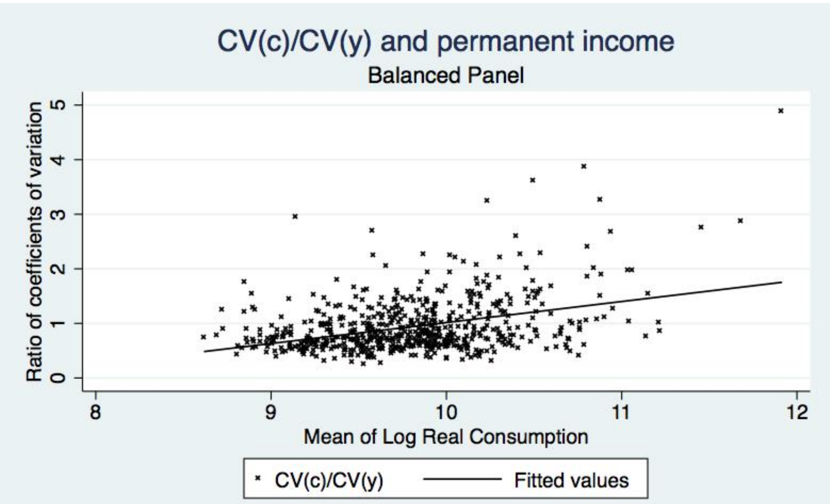
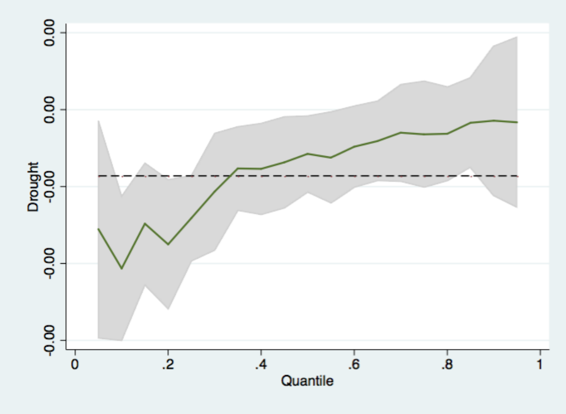
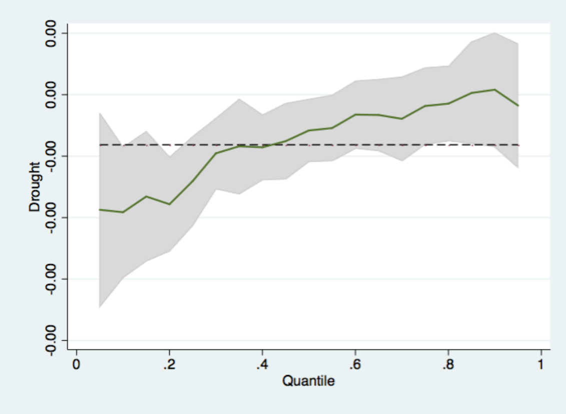


Figure A2.2

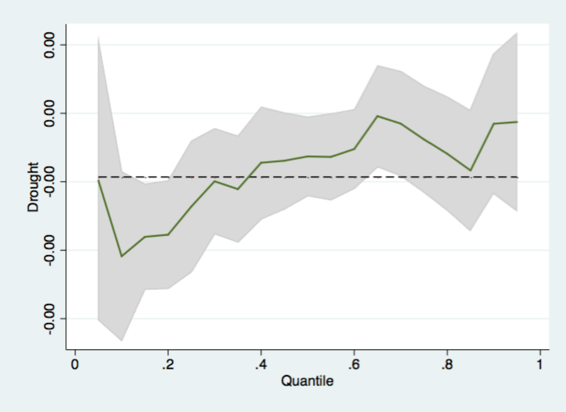
Appendix 3: Quantile regressions of drought on income by characteristic



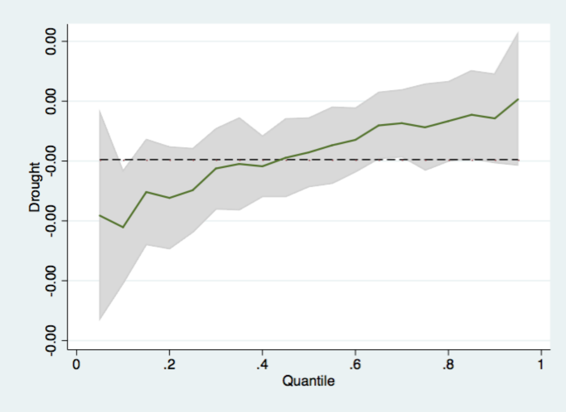
(a) Entire sample



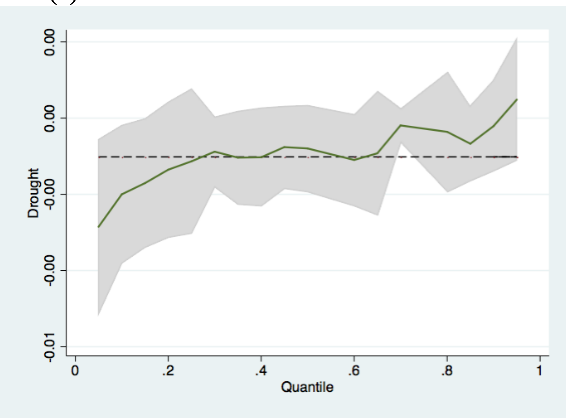
(b) Not earning monthly wages



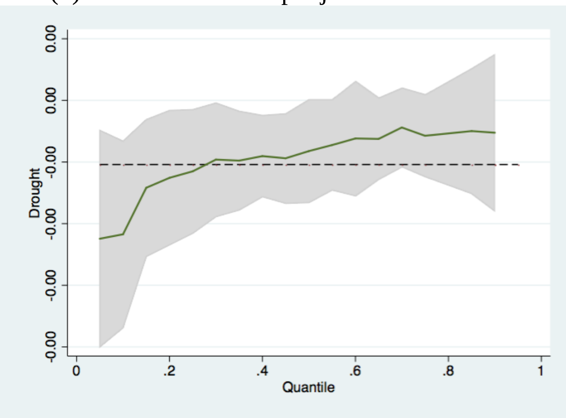
(c) At least two breadwinners



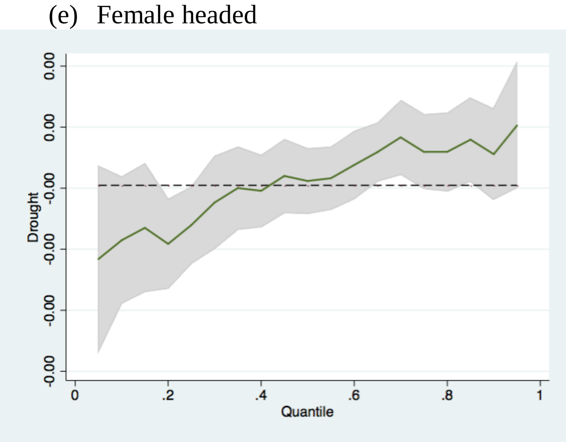
(d) Head holds multiple jobs



(e) Female headed



(f) Business owner



(g) Involved in agriculture

Notes:

1. Line linking quantile estimates
2. Shaded area is the 95% confidence interval

Appendix 4: Inequality in agricultural profits over the lifecycle

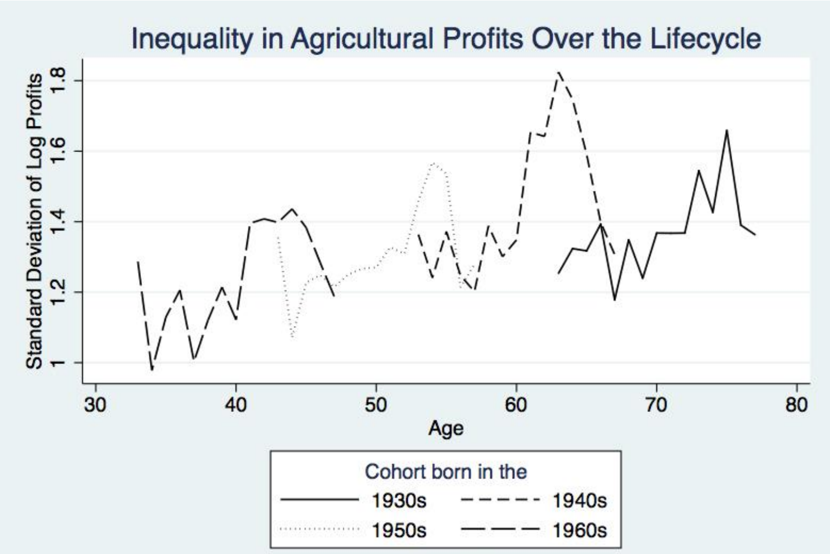


Figure A4

Appendix 5: Cohort-year cell sizes for monthly wage earners

Table A 5: Cohort – Year Cell Sizes for Monthly Wage Earners

Decade of birth	*1900*	*1910*	*1920*	*1930*	*1940*	*1950*	*1960*	*1970*	*1980*
1997	0	0	0	0	0	0	0	0	0
1998	0	1	1	9	14	48	57	74	16
1999	0	1	2	6	13	46	53	84	21
2000	0	0	2	4	15	45	62	75	18
2001	0	0	3	5	10	45	59	86	23
2002	0	0	2	5	12	41	58	81	39
2003	0	0	0	3	9	37	61	65	31
2004	0	0	1	5	8	33	62	71	46
2005	0	0	1	6	7	34	61	78	37
2006	0	0	1	5	6	32	45	69	46
2007	0	0	1	4	6	29	33	49	37
2008	0	0	0	4	8	26	38	54	55
2009	0	0	0	3	8	23	38	51	65
2010	0	0	0	3	5	25	40	61	70
2011	0	0	0	3	2	23	36	52	81

Appendix 6: Cohort-year cell sizes for daily wage earners

Table A 6: Cohort – Year Cell Sizes for Daily Wage Earners									
Decade of birth:	*1900*	*1910*	*1920*	*1930*	*1940*	*1950*	*1960*	*1970*	*1980*
1997	0	0	0	0	0	0	0	0	0
1998	0	1	6	29	49	72	106	105	37
1999	0	0	5	29	61	86	105	130	45
2000	0	0	7	23	62	79	101	129	65
2001	0	0	4	17	62	64	101	134	83
2002	0	0	4	22	55	53	93	115	92
2003	0	0	4	17	56	57	112	111	109
2004	0	0	4	17	52	66	101	123	126
2005	0	0	6	19	56	84	107	101	126
2006	0	0	1	15	50	72	100	99	121
2007	0	0	1	16	50	89	129	129	141
2008	0	0	1	23	61	101	138	136	140
2009	0	0	1	7	41	81	120	118	120
2010	0	0	1	6	35	74	103	94	133
2011	0	0	1	5	41	76	101	93	142

Appendix 7: Real income growth over the lifecycle

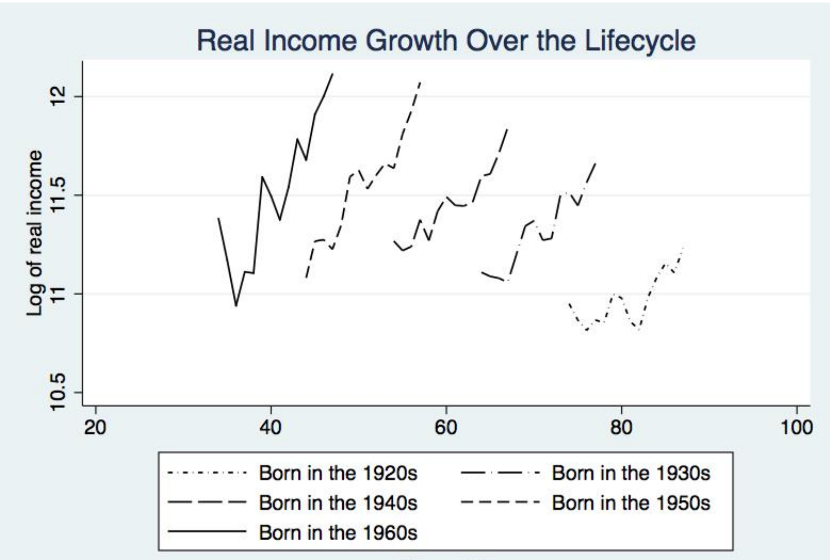


Figure A7