8-27-2009

The Welfare Effects of International Remittance Income

Michael Milligan

Michael A. Milligan

Candidate

Economics

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

, Chairperson

The Welfare Effects of International Remittance Income

by

Michael Alan Milligan

B.S., Physics, State University of New York at Buffalo
M.A., Economics, University of New Mexico

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Economics

The University of New Mexico
Albuquerque, New Mexico

August, 2009

Acknowledgements

I am indebted to several persons in particular for their support of my research and their contributions to this dissertation. I thank Dr. Alok K. Bohara, professor of economics, my advisor and dissertation committee chair, for his valuable guidance throughout my time at the University of New Mexico, and especially for constant support and encouragement during the writing of this dissertation. I also thank Dr. Jennifer Thacher, committee member and professor of economics, for valuable comments and feedback which have helped me to improve my dissertation in many ways. I thank Dr. Don Coes, committee member and professor of economics, for his guidance and insight, and especially for his insights into the economic theory underlying the empirical work presented here. Finally, I thank Dr. Wendy Hansen, committee member and professor of political science, for her feedback, which has helped to improve this work. Of course, any errors in this dissertation are my own.

The Welfare Effects of International Remittance Income

by

Michael Alan Milligan

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Economics

The University of New Mexico
Albuquerque, New Mexico

August, 2009

THE WELFARE EFFECTS OF INTERNATIONAL REMITTANCE INCOME

By

Michael Alan Milligan

B.S., Physics, State University of New York at Buffalo

M.A, Economics, University of New Mexico

Ph.D., Economics, University of New Mexico

ABSTRACT

This dissertation explores the welfare effects of international remittance income, i.e., income earned by migrant workers and sent back to their home country. Remittance income has increased markedly in the last decade, particularly in the developing world. The primary purpose of this dissertation is to quantify the effects of this income on recipient countries.

Chapter 2 of this dissertation presents a study of how remittance income affects child welfare in Nepal using the 2003/2004 Nepal Living Standards survey. I examine how remittance income and non-remittance income affect child labor and child education. Specifically, I examine the probability that a child attends school; a child's educational attainment, given that the child attends school; the probability that a child labors; and the amount that a child labors, given that s/he does so. I find that while both income types positively and significantly impact child welfare, the effects of remittance income are much smaller than those of non-remittance income.

Chapter 3 presents an Engel curve analysis, in which I examine how remittance and non-remittance income affect consumption of various categories of goods in Nepal, again using the 2003/2004 Nepal Living Standards Survey. I use general additive models to allow remittance and non-remittance income to affect consumption nonparametrically and interactively and calculate elasticities of consumption for both remittance and non-remittance income. Confidence intervals for elasticities of consumption are calculated using a combination of bootrap methods and the method of Krinsky and Robb. I find that the elasticity of consumption is always much less from remittance than from non-remittance income.

Chapter 4 presents a macroeconomic analysis of how remittance income affects poverty in Eastern Europe and the former Soviet Union. I use World Bank poverty data on the region to examine how the rate, depth, and severity of poverty are related to GDP, inequality, and remittances in the period from approximately 1998-2003. The poverty data set has been collected and standardized by the World Bank and is an unusually good panel data set on poverty. I find that remittances have no significant impact on poverty in the region.

Throughout this dissertation, I find the effects of remittance income to be small. I posit that this is because of the way that remittances are transferred and used. Many remittances in the regions analyzed never enter the formal financial sector and are likely not used to increase permanent income. According to the permanent income hypothesis, income which does not impact permanent income will have smaller effects on consumption.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

I. Remittances and the Developing World

Remittances sent by migrant workers are an increasingly important means of wealth transfer from the developed to the developing world. In 1999, worldwide remittances were $127 billion, $78 billion of which was to developing countries (World Bank 2008; figures from this source reflect only officially reported remittances, and so are likely underestimates). In 2007, worldwide international remittances were $318 billion, $240 billion of which was to the developing world (ibid.)[1,2]. This dramatic increase in remittances in recent years, particularly to the developing world, is a macroeconomic phenomenon whose consequences are still not fully understood, despite a spate of remittance-related studies in the economic literature over the previous decade. The purpose of this dissertation is to examine the effects of international remittance income in the developing world. The dissertation contains two microeconomic analyses of international remittance income on Nepali households, in particular, on child welfare and consumption patterns. The dissertation also presents a macroeconomic analysis of remittances on poverty in Eastern Europe and the former Soviet Union.

[1] These figures are in 2006 US dollars.
[2] This increase is likely in part due to a transfer of remittances from informal to formal channels, such that official remittances increase even if actual remittances do not. However, the change is no doubt largely due to an increase in migration and greater ease of wealth transfer from globalization. It is unfortunately very difficult to determine the relative importance of these factors.

Remittances are, in some ways, more efficient than foreign aid (another wealth transfer mechanism from rich to poor countries). They represent direct transfers of wealth to needy households. They are much larger than foreign aid flows in many developing countries, can be larger than foreign direct investment, and may even exceed net exports in countries with very high remittance inflows. In addition, they may be countercyclical, tending to increase when conditions in the recipient country worsen (World Bank 2006); this enhances their potential role as a consumption smoothing mechanism.

II. Impact of Remittances

The impact of remittances on developing nations is not well understood. There is debate in the economics literature not only as to how much good remittances do, but if they tend to harm or help a recipient country. On a macroeconomic level, remittances are an important source of foreign currency, which should help to stabilize the balance of payments in countries which would otherwise have a large deficit. However, it has been argued (e.g. Kireyev 2006) that remittances can increase the trade deficit (1) if they are spent mostly on imports and (2) by appreciating the domestic currency, making exports less competitive; these could make remittance income welfare-decreasing in the long run. It has also been claimed (Keely and Tran 1989) that remittances actually increase rather than decrease international inequality, since rich countries benefit from poor countries' laborers, poor countries suffer increased inflation from the artificial influx of money, and migrants returning home to poorer countries are likely to have unrealistic employment aspirations and remain unemployed.

On a microeconomic level, the most intuitive effect of remittances is to increase consumption; though whether or not a given quantity of remittance income affects consumption in the same way as the same quantity of non-remittance income is open to debate, and a topic explored in some depth in Chapter 3 of this dissertation. However, some claim that remittances pose moral hazard problems and could cause the receiving family to work less or make riskier investments, and thus have a negative effect on GDP growth (Chami *et al.* 2003).

In some ways, the part of the world most in need of remittances is the least well-equipped to receive and use this income. The banking system is less developed, and as a consequence remittances are often sent through informal channels, where they are more prone to loss and theft. This also means that remittances are harder for recipient country governments to monitor and measure, which can make policy formulation difficult. Furthermore, financing migrants is often a costly endeavor, particularly for less educated households. Remittance-receiving families must often take out loans to finance a migrant, which mitigates the otherwise beneficial effects of remittances. Furthermore, poorer households may be unable or unwilling to use remittances for productive investments to increase long-term consumption. This could be because remittance income is needed immediately to finance basic consumption, because the family is uneducated about how to invest income, or because of the moral hazard problems discussed above. Since remittances tend to be sent by migrant workers planning to return home, this means that for most families, remittances are a temporary source of income and may not increase permanent income, and thus may do little to alleviate poverty or increase consumption in the long term.

III. Theoretical Background

Most remittances are sent by migrants planning at some point to return to their home countries. For example, studies of migrants to Nepal show that migrants to India tend to stay from a few months to a few years (Thieme 2003) while those to other areas tend to sign employment contracts for two to three years at most (Thieme and Wyss 2005). Eastern Europe and the former Soviet Union is harder generalize, though at least one study from the region (Pinger 2007) shows those migrants who plan to return whom remit more than those who do not. The upshot of these studies is that remittances are often only a temporary source of income.

All analyses in this dissertation focus in some way on consumption: on child welfare, which I argue is a form of household consumption, on household consumption of various categories of goods, and on poverty; whether or not a household is in poverty depends on its level of per capita consumption. Suppose that for each of *M* types of income Y_m and *J* categories of consumption c_j there is a marginal propensity to consume $MPC_{m,j}$, i.e.

$$c_j = \sum_{m=1}^{M} MPC_{j,m} Y_m \tag{1}$$

The focus of this dissertation is on the effects of remittance and non-remittance income, so I re-write Equation (1) as

$$c_j = MPC_{j,1} * (remittance\ income) +$$

$$MPC_{j,2} * (non - remittance\ income) \qquad (2)$$

According to Milton Friedman's permanent income hypothesis (Friedman 1957), consumption from permanent income is greater than that from temporary income. If remittances are considered as temporary income, we could then conclude that

$$MPC_{j,1} < MPC_{j,2} \qquad (3)$$

That is, the effects of remittance income on consumption are smaller than the effects of non-remittance income.

A similar conclusion can be reached from the life cycle hypothesis (Modigliani 1986), with the household considered as the unit of analysis. If the household finances a migrant, then the household's income will likely increase during the time when the migrant is sending remittances and will later decrease when the migrant returns (assuming none of the remittances are invested). Suppose, for example, that a household's non-remittance income y_1 is constant during and after migration, while remittance income y_2 is positive during migration and zero after. To smooth lifetime consumption, the household will then consume relatively little from remittance income, in order to use the savings later. (If, on the other hand, remittance income is productively invested such that non-remittance income y_1 expected to increase, even after the migrant's return, the household may consume more from the remittance income that remains after investing.) This again leads to the conclusion that the marginal propensity to consume from remittance income is less than that from non-remittance income.

Throughout this dissertation, the idea that remittances are more temporary than non-remittance income will be referenced and used as a theoretical justification for the empirical results presented. While remittance income will often be referred to as "temporary" income, in fact either the temporary income hypothesis or the life cycle hypothesis could justify the results presented here.

IV. Introduction to Data

For the analyses of Nepal in this work (Chapters 2 and 3), data comes from the 2003/2004 Nepal Living Standards Survey, which is Nepal's version of the World Bank's Living Standards Measurement Surveys. This is the second such survey done in Nepal; the first was the 1995/1996 Nepal Living Standards Survey. Data for the 2003/2004 survey is from 3,912 households from 326 Primary Sampling Units in Nepal (Central Bureau of Statistics 2004). The survey includes a household questionnaire covering consumption, income, assets, housing, education, health fertility, migration, employment, and child labor, and a community questionnaire to collect information on facilities, services, prices, and the environment (ibid.). Detailed information on the survey and methods used is found in Central Bureau of Statistics (2004).

Poverty inequality data for Eastern Europe and the former Soviet Union (analyzed in Chapter 4) is from a World Bank report on poverty in the region (Alam *et al.* 2005). Data there was collected from national household budget surveys in the period from 1997 to 2003. A detailed description data collection methods is given in Alam *et al.* (2005). Remittances and other national data is from the World Bank's World Development Indicators (World Bank 2008). The national remittance receipts data, though it is the best widely available macroeconomic data on remittances, may not reflect true receipts. This data reflects remittance receipts reported by countries' centralized banking agencies and likely underrepresents remittances sent outside the formal banking sector.

V. Purpose and Hypotheses

The purpose of this study is to examine the welfare effects of international remittance income at both the microeconomic and macroeconomic levels. These effects are compared with the effects of non-remittance income. (For the macroeconomic analysis in Chapter 3, remittance income is compared with GDP).

For policy makers and development organizations, it is important to know both *what* the effects of remittances are, and *why* remittances are having these effects. This dissertation focuses on only the first half of this problem, though each section contains discussion which, it is hoped, will help point the way towards answering the second part as well.

Both donor agencies and governments have recognized the importance of remittances, and programs have developed in recent years to enhance remittance flows (for example, through migrant training programs or efforts to strengthen international banking infrastructure). In order to determine if these programs are a priority or even beneficial, it is useful to know if remittances are going towards human capital investments, poverty alleviation, conspicuous consumption, or some other use; or, if a significant portion of remittances are used only to pay the costs associated with financing a migrant worker. Knowing how remittances are used, and why they are used this way, can help policy makers to decide if programs should focus on enhancing existing effects of remittances or on education programs to change the way in which remittances are used, or if programs should not focus on remittances at all.

One can theorize that remittances have positive or negative effects at either the microeconomic or macroeconomic level. Since remittances often go to poor households in need of income to finance immediate consumption (at least in the parts of the world analyzed in this dissertation), however, it seems more likely to suppose that, at least microeconomically, remittances have a positive net impact. Nonetheless, since remittance income is not usually permanent income, it seems likely that remittances would have less of an effect than non-remittance income at the microeconomic level, if consumers behave rationally and smooth their consumption over time. At the macroeconomic level, this could imply that remittances have little lasting impact on poverty.

VI. Contributions of this Dissertation

This dissertation contributes to the existing literature in several ways and, it is hoped, sheds light on some of the questions raised above. Chapter 2 presents a study of how remittance income affects child welfare in Nepal using the 2003/2004 Nepal Living Standards survey, a household survey data set. The point of view is taken that child welfare is a form of consumption. If a child works, for example, the family earns from the child's labor; for the child to not work, which I suppose to increase his/her welfare, the family must consume more of its resources from other sources. Similarly, if a child goes to school, this carries both direct and indirect costs. In Chapter 2, I examine how remittance income and non-remittance income affect child labor and child education. I examine how these income types affect: (1) the probability that a child will work; (2) the amount that a child works, given that he/she does so; (3) the probability that a child will have gone to school; (4) how much progress the child has made in school, given that he/she has some schooling. While a few other studies examine how remittance income affects child welfare, to my knowledge no other published study uses cross-sectional data to examine how remittance and non-remittance income affect these four child welfare metrics.

Chapter 3 presents an analysis which is in some ways more traditional: an Engel curve analysis, in which I examine how remittance and non-remittance income affect consumption of various categories of goods (food, education, health, select non-food, durables, and total consumption) in Nepal. This chapter is based on the same data set as Chapter 2. The econometric approach taken is in many ways more flexible than other published studies which try to answer this question. I use general additive models to allow remittance and non-remittance income to affect consumption nonparametrically and interactively, and, using bootstrap methods and the Krinsky-Robb sampling approach, construct a series of two-dimensional Engel curves, elasticities of consumption, and associated standard errors for both remittance and non-remittance income.

Chapter 4 presents an analysis of how remittance income affects poverty. This chapter differs in several ways from the preceding two chapters. Firstly, it is based on data at the macroeconomic (country) level, rather than microeconomic household survey data. Secondly, while the other two chapters focus on Nepal, Chapter 4 focuses on

Eastern Europe and the former Soviet Union. Data availability largely dictated the choice to study this part of the world. Macroeconomic analyses of poverty often suffer from a lack of appropriate country-level panel data. However, the World Bank has collected and standardized poverty measures, depth, and severity measures for many of the countries of Eastern Europe and the former Soviet Union for the period from approximately 1998 to 2003 (Alam *et al.* 2005). Because it is for a relatively homogeneous group of countries, has been standardized at considerable effort by the World Bank, contains data for multiple measures of poverty and multiple poverty lines, and contains inequality (Gini coefficient) data, this data set is more standardized and uniform than those used in the few other studies which examine the macroeconomic link between poverty and remittances. Chapter 4 presents an analysis of how poverty is affected by GDP, inequality, and remittance income.

Chapter 2: The Effects of International Remittance Income on Child Education and Child Labor in Nepal

I. Introduction

The welfare of children in Nepal has been a focus for both the Nepalese government and international and national NGOs, particularly since multiparty democracy was restored in 1990 (Baker and Hinton 2001). In particular, it is often a goal of Nepalese policy makers to increase children's school attendance and educational performance, and to reduce the number of children in the labor force. These goals are intertwined, since, ceteris paribus, a child who does not have to work will have more time to devote to school.

Children's education and child labor are important to study for several reasons. The most obvious may be the immediate compromise of a child's well-being if the child works rather than attends school. Children who labor have less time to devote to leisure and human capital development. Less educated children are likely to have fewer employment options when grown than their more educated counterparts. This has long-run negative consequences for the earning potential of the individual, which often has spillover consequences for the household. These effects also hinder the development of a thriving macroeconomy. It is thus important to understand how many factors, including remittances, affect child education and child labor.

This chapter focuses on the effects of household remittance income from international sources on children's education and on child labor. As the data used in this chapter reflects, many Nepali families receive remittance income from household members working abroad, particularly in India and the Middle East. Remittance income might be expected to have a different effect on education than income from other sources, such as wages or salaries. The remittance sender might have influence over the household's actions and spending patterns, so that receiving households may be constrained when deciding how to spend remittance income. Remittance income may be a more stable source of income than income earned in Nepal; this is particularly so for subsistence farmers, whose income is as often only as stable as the weather. However, remittance income is usually not a permanent income source—eventually, the sender will likely return home to Nepal (Thieme and Wyss 2005; Graner and Gurung 2003). Families may be less likely to base decisions such as whether to send their child to school or put the child to work on an income stream which is perceived as temporary.

The means of financing of migrant labor, the source of remittances, should also be considered. Many Nepalese households finance a household member's migration by taking out loans, and a significant portion of the remittances must go to pay off these loans (Ferrari *et al.* 2007). This would tend to mitigate the effects of remittance income. I examine the effects of remittance and non-remittance income on child welfare using two Heckman full information maximum likelihood regressions, one for education and another for labor. For the analysis of education, the dependent variables are a binary variable indicating whether or not the child has had formal schooling and, given that the child has schooling, the child's educational attainment. For the analysis of labor, the dependent variables are a binary variable indicating whether or not the child works in the year of the survey and, given that the child does so, the amount which the child works.

A few other studies have also analyzed how remittances affect some metric of child welfare (discussed in the next section). The primary contributions to the field from this chapter are that the methods presented allow the comparison of the effects of remittance and non-remittance income, in order to better assess the magnitude of the impact of remittances, and that I control for the possible endogeneity of remittance and non-remittance income. Furthermore, this analysis focuses on Nepal, which is a focus of human rights groups working to improve child welfare. It is also a country for which remittance income is an important and increasing source of income (CBS 2004).

This chapter is divided into seven sections. Section II reviews of some of the literature pertaining to child labor, education, and remittances. Section III describes the data used for analysis and presents summary statistics. Section IV describes the theoretical model to be analyzed. Section V presents estimation methods used. Section VI presents and discusses regression results. Section VII contains concluding remarks.

II. Literature Review

Modern research into child labor has been greatly influenced by a paper by Basu and Van (1998) outlining a link between low income and child labor. They established a theoretical microeconomic framework wherein the decision for a child to work was made by the household to help ensure the household's survival, and was not the result of selfish decisions by parents and employers. They argued that if parents could earn higher wages themselves, they would not send their children to work. This implies a strong connection between income and child welfare; this study is one way of analyzing this link.

Several empirical studies have examined the link between poverty and child labor or education. Jensen and Nielsen (1997) analyze the activities of students in Zambia based on the assumption that for each child, households face a binary decision: to send the child to school or to engage the child in labor. They conclude that poverty was an important reason why children work rather than attend school, while higher head of household education and household savings and assets increase the probability of school attendance. Amin, Quayes, and Rives (2004) perform a similar analysis of determinants that a child would work in Bangladesh, and determine that poverty was the most important cause of child labor. Numerous studies worldwide have shown that household income is negatively correlated with child labor rates (Edmonds and Pavcnik 2005, and references therein).

Several studies have analyzed the influence of remittance income, as opposed to income from other sources, on spending, including spending on education. Stahl and Arnold (1986), in a survey of several studies of the effects of remittances on spending patterns in Asian countries, find that remittance income is more likely to be spent on food, durables, and housing, and less likely to be spent on investments like education, than income from other sources. In contrast, Adams (2005), using a 2000 household budget survey to perform a similar analysis on Guatemalan households, finds that remittance income is more likely to be spent on education than other sources of income. However, the effects of remittances on children's education may not be fully captured by an analysis of remittances on education spending, particularly in a country like Nepal where direct costs are usually small.

Of more direct relevance to this analysis are studies of how remittances effect educational attainment and child labor. Lopez Cordoba (2004, as cited in McKenzie 2005) find that 6- to 14- year-olds in Mexican municipalities which receive more remittances have higher literacy and school attendance rates. Yang (2006) takes advantage of exchange rate shocks to analyze how changes in real remittance levels affected remittance-receiving households in the Philippines, including investment in human capital. Among his conclusions are that increased real remittances are associated with more child schooling and less child labor.

Neither of these studies, however, attempt to compare the effects of remittance and non-remittance income, which makes it difficult to put into context the *magnitude* of the effects of remittance income. Cox and Ureta (2003) use a 1997 household survey to analyze and compare both types of income. They find that while both remittance income and non-remittance income contribute positively to school retention rates among 6- to 24-year-olds in El Salvador, remittance income contributes more than the same amount of non-remittance income. This conclusion differs from those of this study, though my different findings are not necessarily contradictory: Cox and Ureta analyze a different country, in different circumstances, and their data is from 6-7 years earlier. Moreover, they use different econometric methods than those used here, and in particular do not correct for the possible endogeneity of remittance and non-remittance income.

III. Theoretical Model

In practice, income is often not completely fungible—the source of a supply of income determines how households use it. This concept may have originated with Milton Friedman's permanent income hypothesis, according to which spending on consumption is taken only from permanent income, and not from temporary income. Suppose that household *h* maximizes utility *U*, which is function of a vector of quantities consumed $\{c_1, c_2, ..., c_j, ...c_J\} = \vec{c}$ and a vector of child welfare variables $\{w_1, w_2, ...w_k, ..., w_K\} = \vec{w}$, such that $\forall j, \frac{\partial U(\vec{c}, \vec{w})}{\partial c_j} > 0$ and $\forall k, \frac{\partial U(\vec{c}, \vec{w})}{\partial w_k} > 0$. Households maximize utility in two ways: by consuming, and by expending household resources to improve child welfare. These two means of increasing utility are often at odds; for example, by allowing a child to labor less, the child's welfare will increase; but household income will then decrease, and so must consumption. If a child devotes time and energy to school, then s/he will have less time to work, and less time to devote to household chores. Other household members must then divert time that could be spent earning income to doing these chores, and again income, and hence, consumption decrease.

Consumption and child welfare are both assumed to be positive functions of income. One can suppose that the way income is used to improve consumption and child welfare depends on how the income is obtained. For example, for a given household, consumption c_j of good *j* might be given by $c_j = \alpha_{0,j} + \sum_{m=1}^{M} MPC_{m,j} Y_m$, where $\alpha_{0,j}$ is a constant, Y_m are income amounts from M different sources and $MPC_{m,j}$ are coefficients (often assumed to be between zero and one). As discussed in Section II, there are several studies in the literature which examine propensities to consume from remittance income.

Here I analyze not consumption, but child welfare—another (opportunity) cost which may be paid with income. Suppose that the *k*th measure of child welfare w_k is a linear combination of M types of income and other household- and individual-level variables, such that

$$w_k = \alpha_{0,k} + \sum_{m=1}^{M} MPC_{m,k} Y_m + \beta_k \boldsymbol{x}_k \tag{4}$$

where $MPC_{m,k}$ are rates of improvement to child welfare metric w_k from income Y_m and β_k and $\boldsymbol{x}_k$ are vectors of coefficients and other explanatory variables, respectively. Since I am studying the effects of remittance and non-remittance income, I re-write Equation (4) stochastically as

$$w_k = \alpha_{0,k} + \alpha_{1,k} REMINC + \alpha_{2,k} NONREMINC + \beta_k \boldsymbol{x}_k + \varepsilon_k \tag{5}$$

where *REMINC* is log income received by the household from remittances from international sources, *NONREMINC* is log income received from all other sources (excluding income from child labor, as discussed in the previous section) and ε_k is an error term.

Because remittances are often sent by migrants who intend to return home, remittances can be a more temporary source of income than income from other sources. The permanent income hypothesis would then imply that remittance income would have a smaller effect on child welfare than non-remittance income: that is, $\alpha_{1,k} > \alpha_{2,k}$. Furthermore, some remittance income must often be used to pay back loans taken out to finance the costs of migration (Ferrari *et al.* 2007), which again could imply that $\alpha_{1,k} > \alpha_{2,k}$. There is also evidence that in Nepal, remittances are often used to finance the migration of other household members rather than for consumption (Graner and Gurung 2003; Thieme 2005).

I estimate Equation (5) using four child welfare measures w_k. The first is *LABORPART**, the unobserved probability of a child laboring; this probability is proxied by *LABORPART*, a dummy variable equal to one if the child worked in the year of the study and zero otherwise. I include time spent on family farms or doing household chores as work; household work can contribute significantly to a Nepali child's work load, as pointed out by Edmonds and Pacnvik (2005) and confirmed by NLSS 2003/2004 data. The second measure of welfare is *LABORHOURS*, the log of the number of hours a child worked in the year of the study, given that *LABORPART* is equal to one. The third measure is *SCHOOLING**, the unobserved probability of a child having had some schooling; this is proxied by *SCHOOLING*, a dummy variable equal to one if the child was currently attending school or had successfully completed at least one year of school in the year of the study and zero otherwise. The fourth measure of child welfare is *EDINDEX*, the child's educational attainment given that *SCHOOLING* equals 1. *EDINDEX* is the adjusted ratio of the number of years of schooling completed by the child and the child's age (an index similar to that used by Ruan *et al.* [2009]), i.e., the number of years of schooling the child successfully completed, plus five, divided by the child's age.

IV. Data and Summary Statistics

The data used for this chapter are from the 2003 Nepal Living Standards Survey, conducted from April 2003 to April 2004 by Nepal's Central Bureau of Statistics (CBS). The survey follows the World Bank's Living Standards Measurement Survey methodology (CBS 2004). The survey included 3912 households and 21531 individuals, of which 6478 individuals were between the ages of 5 and 16 (the age group of focus for this chapter).

I calculate from this survey certain statistics pertaining to my sample, i.e. children aged 5 to 16. Of the sample, 22% of children had never attended school. In the year of the survey, 31% both attended school and worked (including unpaid work, such as household chores); 16% worked and did not attend school; 41% went to school and did not work; while 12% did neither. Thus, 78.1% of the sample had some schooling, and 47% labored. Table 1 reports the primary reasons why children not in school never attended or left school, as reported by the household.

Table 1 Primary Reasons for Never Attending or Leaving School, Children 5-16

Reason	For never attending (%)	For leaving (%)
Parents did not want	20.1	10.5
Too expensive	19.5	16.1
Not willing to attend	15.5	
Had to help at home	13.9	25.1
Too far away	4.7	2.5
School not present	1.4	
Disabled	1.3	
Education not useful	1.3	
Poor academic progress		26.8
Completed desired schooling		3.4
Further schooling not available		2.5
Moved away		1.4
Environment of school not good		1.1
Other reasons	22.2	10.5

*Author's calculations using data from the Nepal Living Standards Survey, 2003/2004.

Aggregate household income was calculated by summing reported income from various sources, namely, revenue from agriculture and livestock operations, rental income, remittance income, income from enterprise, wage income, and other sources (such as investment in stocks or bonds). Also included in income was the value of agricultural products produced by the household for self-consumption, and, for those who owned their homes, the opportunity cost of not renting the home to others (as estimated by the household). Income was adjusted using regional price indices. This real aggregate income was divided into two categories: that from international remittances ("remittance income") and that from all other sources ("non-remittance income"). The quantity one was added to each type of income (because of the many households with zero remittance income), and the natural logs of these quantities were used as explanatory variables in the child welfare estimations.

Child welfare is a determinant of household income in one important way: the more a child labors, the more income he or she earns. Since this analysis focuses on how income affects child welfare, rather than how child welfare affects income, income from child labor is not included in the definition of income used here. The income of children age 16 or less is excluded from non-remittance household income.[3] Also, 114 children

---

[3] To exclude the value of agricultural products produced by the children, I assumed that the proportion of income generated by the child was equal to the proportion of hours worked by the child; e.g., if child labor accounted for half of total household hours of work in agriculture, then half of the income from agriculture, livestock, and the household's consumption of its own production was excluded from the income aggregate.

from households with negative non-remittance income (often due to losses incurred by enterprises) were removed from the sample, since the log of non-remittance (plus one) was undefined for these observations.

Table 2 contains descriptions of variables used in the analyses presented in this chapter. Those statistics pertaining to households apply to those households in the relevant sample, i.e., households with children between the ages of 5 to 16. Those statistics pertaining to individuals apply to children ages 5 to 16.

It is probably not the case that a child is as productive as an adult per unit of time, but the data available did not allow for a more precise determination of production by children.

Table 2 Summary Statistics (Child Welfare Analysis)

Variable	Description	mean (s.e.)
At household level (n = 2755):		
NONREMINC[4]	Log of real household income plus 1, less income from remittances from international sources or from household members age 16 or less	10.506 (.020)
REMINC	Log of real household income from remittances from international sources plus 1	1.797 (.079)
HEADUNMARRIED	Dummy = 1 if head of household is unmarried	.109 (.006)
HEADMIGRATED	Dummy = 1 if head of household migrated to current residence	.422 (.010)
HEADFEMALE	Dummy = 1 if head of household is female	.181 (.008)
HEADAGE	Age of head of household	44.517 (.255)
HEADEDUC	Years of schooling successfully completed by head of household	2.855 (.079)
CASTE1	Dummy = 1 if head of household is of Magar, Tamang, Rai, Gurung, or Limbu caste or ethnicity	.205 (.008)
CASTE2	Dummy = 1 if head of household is of Kami, Damai, Dholi, or Sarki caste or ethnicity	.083 (.006)

[4] Mean real non-remittance, non-child labor income is 104,276.5 Nepalese Rupees, with a standard error of 38,413.76. This statistic is heavily influenced by outliers; if the nine households with such income over 1,000,000 rupees are dropped, the mean is 59,045.24 with a standard error of 1,357.599. Removing outliers does not significantly change the important conclusions of this paper. For the subset of households in my sample who receive remittances (n=471), the mean amount of real remittances received is 50582.08 Nepalese Rupees, with a standard error of 3794.006. *LABORHOURS* is only defined when *LABORPART* = 1 (n = 2987), and *EDINDEX* is only defined when *SCHOOLING* = 1 (n = 4968).

CASTE3	Dummy = 1 if head of household is of Tharu, Yadav, Brahmin Terai, Thakur, or Hazam caste or ethnicity	.103 (.006)
CASTE4	Dummy = 1 if head of household is of Newar caste or ethnicity	.065 (.004)
CASTE5	Dummy = 1 if head of household is Muslim	.060 (.005)
*CASTE^*	Dummy = 1 if head of household does not fall into categories covered by above five caste/ethnicity dummy variables, and is not Brahmin or Chhetry	.211 (.008)
HH_SIZE	Number of people in household	6.295 (.055)
SUBS_AG	Dummy = 1 if subsistence agriculture is one of the head of household's occupations	(.786) (.008)
RURAL	Dummy = 1 if household is located in rural area	.860 (.006)
MOUNTAIN	Dummy = 1 if household is located in mountain ecological zone	.069 (.004)
HILL	Dummy = 1 if household is located in hill ecological zone; Terai (plains) ecological zone is unspecified	.428 (.010)
LANDVALUE	Log of the value of the land owned by household plus 1	9.558 (.101)
FINANCIAL	Financial sophistication of the household, proxied by the number of financial instruments the household owns (including savings accounts, fixed deposit accounts, stocks/shares, provident funds, pensions, commission fees, and instruments reported as "others")	1.082 (.011)
LOANS	Number of outstanding loans owed by the household	1.116 (.023)
At child level (n = 6365)		
LABORPART	Dummy =1 if the child labored in the year of the survey (including wage-earning labor, household work, and work in household businesses, including subsistence agriculture)	.487 (.007)
LABORHOURS	Natural log of the number of hours the child worked in the past year, plus one	6.236 (.022)
SCHOOLING	Dummy = 1 if the child has successfully completed at least one year of school or is currently in school	.759 (.006)
EDINDEX	Child's educational attainment, proxied by the number of years the child has successfully completed plus five (the age at which schooling usually starts), divided by the child's age	.880 (.003)

FEMALE	Dummy = 1 if child is female	.483 (.007)
AGE	Child's age	10.388 (.046)

*Author's calculations using data from the Nepal Living Standards Survey, 2003/2004.

V. Econometric Models and Estimation Methods

The labor dependent variables, *LABORPART* and *LABORHOURS*, are examined jointly using a Heckman full information maximum likelihood regression, where *LABORPART* is the selection variable and *LABORHOURS* the outcome variable. The education dependent variables, *SCHOOLING* and *EDINDEX*, are examined jointly in another Heckman full information maximum likelihood regression, where *SCHOOLING* is the selection variable and *EDINDEX* the outcome variable. The probability of a child working or attending school is assumed to have a probit relationship with the relevant explanatory variables, and the expected value of the amount labored or educational performance is assumed a linear combination of the relevant explanatory variables.

Given these relationships Equation (5) can be rewritten for both the labor and education models:

Labor Model:

$$LABORPART^{*} = \alpha_{0,LP} + \alpha_{1,LP}REMINC$$

$$+\alpha_{2,LP}NONREMINC + \beta_{LP}\boldsymbol{x}_{LP} + \varepsilon_{LP} \quad (6a)$$

$$LABORPART = 1 \text{ if } LABORPART^{*} > 0 \text{ and } 0 \text{ otherwise} \quad (6b)$$

$$LABORHOURS = \alpha_{0,LH} + \alpha_{1,LH}REMINC + \alpha_{2,LH}NONREMINC$$

$$+\beta_{LH}\boldsymbol{x}_{LH} + \varepsilon_{LH} \text{ observed only if } LABORPART = 1 \quad (6c)$$

$$(\varepsilon_{LP}, \varepsilon_{LH}) \sim N(\begin{bmatrix} {\sigma_{LH}}^{2} & \rho_{L}\sigma_{LH} \\ \rho_{L}\sigma_{LH} & 1 \end{bmatrix}) \quad (6d)$$

where the subscripts *LP* and *LH* indicate quantities pertaining to the estimation of *LABORPART* and *LABORHOURS*, respectively, ${\sigma_{LH}}^2$ is the variance of ε_{LH}, and ρ_L is the correlation between ε_{LP} and ε_{LH}.

Education Model:

$$SCHOOLING^* = \alpha_{0,SC} + \alpha_{1,SC}REMINC +$$

$$\alpha_{2,SC}NONREMINC + \beta_{SC}\boldsymbol{x}_{SC} + \varepsilon_{SC} \qquad (7a)$$

$$SCHOOLING = 1 \text{ if } SCHOOLING^* > 0 \text{ and } 0 \text{ otherwise} \qquad (7b)$$

$$EDINDEX = \alpha_{0,ED} + \alpha_{1,ED}REMINC + \alpha_{2,ED}NONREMINC$$

$$+\beta_{ED}\boldsymbol{x}_{ED} + \varepsilon_{ED} \text{ observed only if } SCHOOLING = 1 \qquad (7c)$$

$$(\varepsilon_{SC}, \varepsilon_{ED}) \sim N(\begin{bmatrix} {\sigma_{ED}}^2 & \rho_E\sigma_{ED} \\ \rho_E\sigma_{ED} & 1 \end{bmatrix}) \qquad (7d)$$

where the subscripts *SC* and *ED* indicate quantities pertaining to the estimations of *SCHOOLING* and *EDINDEX*, respectively, ${\sigma_{ED}}^2$ is the variance of ε_{ED} and ρ_E is the correlations between ε_{SC} and ε_{ED}, respectively.[5] These welfare equations are estimated at

[5] If $\boldsymbol{X}_{LP}$ represents all the regressors for *LABORPART* (unity, *REMINC*, *NONREMINC*, and $\boldsymbol{x}_{LP}$) and b_{LP} all associated coefficients, and $\boldsymbol{X}_{LH}$ and b_{LH} all regressors and associated coefficients for *LABORHOURS*, the log likelihood function to be maximized is the sum over observations:

the individual child level. Observations are weighted with household-level sample weights included in the survey data.

To ensure that that the primary equations (the estimations of child welfare) are identified, I exclude an independent variable from each child welfare equation which is found in the co-estimated welfare equation. It is difficult to justify theoretically that some variables might influence, for example, the probability that a child would attend school, but not how well a child does in school; or how much a child labors, but not the probability that a child will labor. However, such assumptions are necessary for the identification of the child welfare equations. I choose to include a dummy variable indicating whether or not the head of household is female in the select equations (dependent variables *LABORPART* and *SCHOOLING*), but not the outcome equations (dependent variables *LABORHOURS* and *EDINDEX*); and a dummy variable indicating whether or not the head of household has migrated to his/her current location in the

---

$$\ln \mathrm{L} = \sum_{LABORPART=0} \ln(1 - \Phi\,(b_{LP}\boldsymbol{X}_{LP})) + \sum_{LABORPART=1} \ln \varphi(\frac{LABORHOURS - b_{LH}\boldsymbol{X}_{LH}}{\sigma_{LH}}) + \sum_{LABORPART=1} \ln \Phi(\frac{b_{LP}\boldsymbol{X}_{LP} + \rho_L \frac{LABORHOURS - b_{LH}\boldsymbol{X}_{LH}}{\sigma_{LH}}}{\sqrt{1-{\rho_L}^2}})$$

where $\varphi(\cdot)$ is the normal probability density function and $\Phi(\cdot)$ is the normal cumulative density function. The log likelihood function for the education estimations is exactly analogous.

outcome equations, but not the select equations. This may seem arbitrary, but some specification choice of this nature is needed. The conclusions of this chapter, particularly the primary conclusion that non-remittance income has a significantly more positive effect on child welfare than remittance income, are robust to alternative specifications. Results of these alternative regressions are available upon request.

Both log non-remittance income (plus one) and log remittance income are likely to be endogenous functions of some of the same variables which determine child welfare. To instrument these variables, I included all household-level explanatory variables used to estimate the child welfare equations, as well as instruments which seemed unlikely to influence child welfare other than through their effects on income. These additional instruments were *FINANCIAL*, a measure of financial sophistication proxied by the number of financial instruments used by the family (which may be particularly relevant as an indicator of non-remittance income), and *LOANS*, a measure of the number of loans taken out by the family (which may be particularly relevant as a measure of remittance income, since many families take out loans in order to finance a migrant's travel [Ferrari *et al.* 2007]). Non-remittance income was instrumented using an ordinary least squares regression, while remittance income was instrumented with a tobit regression, since many households reported zero remittance income. These instrumenting estimations were done at the household level with household level sample weights. Hausman specification tests were used to confirm that endogenization of remittance and non-remittance income was necessary for consistent estimation results.

Appendix A contains a proof that the system of equations estimated in this chapter is identified under these conditions. All statistical analysis presented in this chapter was done using the Stata statistical analysis program.

VI. Estimation Results and Discussion

Regression results are presented in Table 3, and marginal effects of independent variables on the binary select variables are presented in Table 4.

Table 3 Full Information Maximum Likelihood Estimation of Child Welfare Measures

	Labor		Education	
Variable	*LABORPART*	*LABORHOURS*	*SCHOOLING*	*EDINDEX*
NONREMINC	-.264***	-.367***	.276***	.052***
	(.094)	(.096)	(.093)	(.012)
REMINC	-.060*	-.061***	.065**	.010***
	(.033)	(.015)	(.027)	(.002)
HEADUNMARRIED	-.019	.080	-.243***	-.017
	(.086)	(.068)	(.073)	(.011)
HEADMIGRATED		.030		.017***
		(.05)		(.006)
HEADFEMALE	.004		.051	
	(.174)		(.141)	
HEADAGE	-.009***	-.0005	-.001	.0007***
	.002	(.002)	(.002)	(.0002)
HEADEDUC	-.032***	.005	.030***	.007***
	(.009)	(.009)	(.008)	(.001)
CASTE1	.109*	.137**	-.176***	-.048***
	(.056)	(.056)	(.054)	(.007)
CASTE2	-.007	.175**	.030	-.043***
	(.084)	(.073)	(.079)	(.010)
CASTE3	.011	.388***	-.208**	-.045***
	(.085)	(.097)	(.081)	(.011)
CASTE4	-.062	.042	.015	-.0002
	(.094)	(.101)	(.099)	(.011)
CASTE5	-.132	.742***	-.496***	-.061***
	.101	(.113)	(.134)	(.023)
*CASTE^*	.093	.458***	-.426***	-.060***
	(.066)	(.076)	(.064)	(.009)
HH_SIZE	-.006	.027**	-.057***	-.010***
	(.011)	(.011)	(.009)	(.002)
SUBS_AG	.253***	.002	-.008	.007
	(.082)	(.083)	(.081)	(.013)
RURAL	.398***	.078	-.350***	-.040***
	(.071)	(.080)	(.070)	(.009)
MOUNTAIN	.326***	.056	-.082	-.011
	(.075)	(.084)	(.070)	(.010)
HILL	.104*	0.265***	.025	.008
	(.056)	(.067)	(.056)	(.007)
LANDVALUE	.006	-.019***	.033***	.005***
	(.005)	(.006)	(.005)	(.001)
FEMALE	.548***	0.451***	-.310***	-.024***
	(.039)	(.040)	(.041)	(.006)
AGE	.243***	0.139***	.027***	-.023***
	(.007)	(.007)	(.006)	(.0009)

Constant	-.123 (1.00)	7.73*** (.986)	-1.86* (.973)	.520*** (.125)
ρ	.147*** (.025)		.893*** (.034)	
log likelihood	-8633817		-1265225	
pseudo-R^2	.162		.521	
N	6364		6364	

NONREMINC* and *REMINC* were instrumented as described in Section V. The quantity ρ is the correlation between co-estimated equations. Numbers in parentheses are heteroskedasticity-consistent-standard errors. *, **, and * indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 4 Marginal Effects on Binary Child Welfare Variables				
	LABORPART		*SCHOOLING*	
Variable	Marginal effect	Standard error	Marginal effect	Standard error
NONREMINC	-.105***	.038	.083***	.028
REMINC	-.024**	.069	.019**	.008
HEADUNMARRIED	-.008**	.034	-.078***	.025
HEADAGE	-.004	.0009	-.0003	.0006
HEADFEMALE	.001**	.069	.015	.042
HEADEDUC	-.013***	.003	.009***	.003
TAMAGURALI	.044*	.022	-.055***	.017
DAKASA	-.003	.033	-.009	.023
TERAICASTE	.004	.034	-.066**	.027
NEWAR	-.025	.037	.005	.030
MUSLIM	-.052	.040	-.170***	.049
OTHERCASTE	.037	.026	-.139***	.022
HHSIZE	-.003	.004	-.017***	.003
SUBSAG	.010***	.032	-.002	.024
RURAL	.154***	.026	-.095***	.017
MOUNTAIN	.129***	.029	-.026	.022
HILL	.042*	.022	.008	.017
LANDVALUE	.002	.002	.010***	.001
FEMALE	.215***	.015	-.094***	.012
AGE	.097***	.003	.008***	.002

*Standard errors are heteroskedasticity-consistent. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Marginal effects and associated standard errors for continuous variables *LABORHOURS* and *EDINDEX* are identical to the coefficients and standard errors reported for these variables in Table 3.

Several conclusions can be drawn from these results. Firstly, children in rural areas are more likely to work than in urban areas, even when one controls for the child's household being subsistence farmers and other factors. Children in rural areas are also less likely to attend school, and perform more poorly when they do so. This is probably due to schools being more inaccessible in rural areas, as well as a more traditional rural mindset regarding the role of a child and the importance of a Western-style education. When a household's caste is statistically significant, the children in that household work more, go to school less, and do worse in school than the unspecified castes, Brahmin and Chhetry. This is not surprising, since Brahmin and Chhetry are traditionally the most privileged castes in Nepalese society. Children from larger families are no more likely to work than those from smaller families, but when they do, they work longer hours. They are also less likely to go to school and do not do as well in school as children from smaller families. This is probably the result of the household's resources being divided between more household members. Female children work more and go to school less, and with less success, than their male counterparts; this is as one would expect in a male-dominated society like Nepal. Unsurprisingly, children in households with more educated heads of household are less likely to join the labor force, are more likely to go to school, and perform better in school.

It is interesting to note that these socioeconomic effects occur even when controlling for disparate amounts of household income. One might think, for example, that privileged caste children have higher welfare simply because their families tend to be better off financially. This analysis provides evidence of the effects of socioeconomic characteristic on child welfare independent of the effects of these characteristics on income.

The focus of this study is the effects of income on child welfare. Both remittance and non-remittance income contribute positively and significantly to child welfare: increased household income implies that a child is less likely to work, will work fewer hours if he or she does work, is more likely to go to school, and does better in school if he or she attends. However, the coefficients and marginal effects for these variables are very different. I find that a given amount of remittance income contributes much less to child welfare than the same amount of income from other sources.

To formalize some of these conclusions, I constructed four hypotheses, one for each measure of child welfare, to determine if the income coefficients $\alpha_{1,k}$ and $\alpha_{2,k}$ in 2quation (2) are significantly different. Each of these hypotheses was tested with a Wald test with the null hypothesis $\alpha_{1,k} = \alpha_{2,k}$. (Two-tailed Wald tests are used to avoid assuming *a priori* that remittance income has greater or lesser effects than non-remittance income.) The results of these tests are presented in Table 5. In all four cases the coefficients for remittance and non-remittance income were found to be statistically significantly different at the 10% level or better.

Table 5 Wald Tests of Child Welfare Hypotheses

Hypothesis	Chi-squared (p-value)	Remark
H1: $\alpha_{1,LABORPART} = \alpha_{2,LABORPART}$	3.51 (.0608)	H1 is rejected at the 10% level; remittance income and non-remittance income do not have significantly different effects on the probability that a child will labor
H2: $\alpha_{1,LABORHOURS} = \alpha_{2,LABORHOURS}$	11.24 (0.0008)	H2 is rejected at the .1% level; remittance income and non-remittance income have significantly different effects on the amount that a child labors
H3: $\alpha_{1,SCHOOLING} = \alpha_{2,SCHOOLING}$	4.34 (0.0372)	H3 is rejected at the 5% level; remittance income and non-remittance income have significantly different effects on the probability that a child has had some schooling
H4: $\alpha_{1,EDINDEX} = \alpha_{2,EDINDEX}$	13.24 0.0003	H4 is rejected at the .1% level; remittance income and non-remittance income have significantly different effects on a child's educational attainment

Though the effects of remittance income are small compared to the effects of non-remittance income, remittance income unambiguously enhances child welfare in Nepal. This implies that policies which enhance remittance income will also enhance child welfare. This may even be an area of focus for NGOs and advocacy groups interested in enhancing child welfare. Of course, it must be remembered that this chapter deals only with remittances and not migration as a whole; while remittance income certainly enhances child welfare, the effects of having adult males—perhaps with children—away from home for several years is well beyond the scope of this analysis.

VII. Conclusions

I find that both remittance and non-remittance income contribute positively and significantly to all measures of child welfare I have analyzed, though with significantly different magnitudes. I also analyzed several individual- and household- level socioeconomic variables. I find that a children who live in rural areas, who are female, who are not members of a privileged caste, who come from large households, or who come from households with less educated heads of household, tend to work more and achieve less in school than their counterparts.

That the marginal effects from remittance income are of lesser magnitude than the effects from non-remittance income does not change the conclusion that remittances positively contribute to child welfare. Were the Nepalese government or another agency to enhance the flow of remittances into Nepal, children would benefit. However, the differences in remittance and non-remittance income coefficients indicate that this should not necessarily be done at the expense of developing income sources within Nepal.

It is worth briefly discussing the challenges to migration and remittance transmission in Nepal, and how these might relate to the results presented here. Many families finance a household member's migration with loans from private organizations (Ferrari *et al.* 2007). Some of the remittance income received by the household is likely used to pay back these loans. If the Nepali government were to train and sponsor its citizens to work abroad, helping to free them from dependence on these loans, more remittance income could be used to improve household welfare, including child welfare. Of course, this is more easily said than done for a developing nation like Nepal.

There may be other ways to change the way that remittances affect child welfare. Remittances' positive effects on child welfare could potentially be enhanced by facilitating the flow of remittances to Nepal. There is much room for improvement in the system through which remittances are sent. Remittances are often sent through informal networks or hand-carried by the migrant or a friend; very few Nepali migrants remit through the formal banking system (Ferrari *et al.* 2007; Graner and Gurung 2003; Thieme 2005). Establishing a more formal and reliable transmission network would allow more remittances to reach their targets, and would amplify their effects on Nepalese households. Integration of Nepal's banks with the global banking community, for example, would allow remittance earners to deposit their earnings in an account directly accessible by their intended recipients. A program in the Nepali government to certify or license remittance couriers might also lead to a more efficient transmission system. Besides reducing losses of remittances en route, these changes could also increase remittance flows by increasing senders' confidence that their money would reach its intended source.

Furthermore, if more remittances are sent through the formal system, Nepali households may better be able to leverage remittances—often a temporary income source—to increase permanent income. If remittances are sent through the formal sector, they may be more likely to stay in the formal sector; for example, banks have more opportunities to market financial services to remittance senders and recipients. Investment opportunities for remittance recipients may thus increase if more remittances are sent through formal channels. If remittances are used to increase permanent income, rather than only temporary income, then their welfare-enhancing effects may be significantly increased.

Chapter 3: Consumption from Remittance and Non-Remittance Income in Nepal: A Semiparametric Analysis

I. Introduction

International remittances are an important stream of income for many developing countries. Their effect on household consumption patterns is a phenomenon on which no consensus has been reached. This study attempts to shed light on this phenomenon by constructing Engel curves and associated consumption elasticities for both remittance and non-remittance income for households in Nepal.

The effects of remittances on consumption patterns are important for several reasons. According to traditional macroeconomic theory, if remittances are saved rather than immediately consumed, their macroeconomic effects are enhanced because they are available for on-lending and to banks and firms for investment. For governments or donor agencies assessing policy priorities, it is also important to know if remittances are used as investments in human capital (such as education), for consumption to meet basic needs and alleviate poverty, or for luxury items. Finally, if food budget share is used as a measure of household welfare or poverty, then it is important to know how remittance income affects this metric.

Others have used surveys wherein the respondent is directly asked to what use he/she puts remittance income to attempt to determine the effects of remittances on spending patterns (for example, Arrehag *et al.* 2005; SECO 2007). However, how a respondent feels that remittance receipts are being spent does not necessarily reflect the impact of remittance income on spending. The use of more extensive household survey data to construct Engel curves as functions of remittance and non-remittance income is promising; this is the approach taken by Adams (2005), but his econometric approach differs significantly from that presented here; most notably, he uses parametric methods.

The use of semiparametric techniques to construct Engel curves is now rather common. However, to my knowledge these techniques have not been used to define the effects of remittance income versus income from other sources. This analysis is a natural extension of modern consumption analysis to explore the microeconomic impacts of remittances.

This chapter is divided into six sections. Section II gives an overview of the role of international remittances to Nepal. Section III presents the econometric theory and techniques used in the analysis presented here. Section IV describes the data and variables used for analysis and presents summary statistics. Section V presents regression results with a focus on consumption elasticities. Section VI contains concluding remarks and policy recommendations.

II. International Remittances to Nepal

Remittance income (money sent by migrant workers back to their home countries) has increased significantly in recent years. In 1999, worldwide remittances were $127 billion, $78 billion of which was to developing countries (World Bank 2008); in 2007, worldwide international remittances were $318 billion, $240 billion of which was to the developing world (ibid.).[6,7]

Remittances to Nepal have also increased markedly in recent years. In Nepal, 17.9% of households received remittances from international sources in 2003/2004, compared to 9.8 % in 1995/1996 (author's calculations based on 1995/1996 and 2003/2004 Nepal Living Standards Survey data). For the median remittance-receiving family, remittances accounted for 48.1% of household income in 2003/2004, compared to 28.9% in 1995/1996 (ibid.).

In 2003/2004, nearly two-thirds of those sending remittances to Nepal worked in India, with most of the rest living in the Middle East (ibid.). However, less than a third of remittances to Nepal come from India; the Middle East is the largest source of remittances to Nepal (ibid.). The vast majority of international remittances to Nepal are sent through informal channels. Approximately 79.0% are hand-carried (by the migrant or another person) while only 10.8% are sent through formal financial institutions (ibid.). Financing a migrant is often expensive, and remittances must often go towards paying these costs. According to a 2007 World Bank report (Ferrari *et al.* 2007), "large shares of

---

[6] These figures are in 2006 US dollars.

[7] This increase is likely in part due to increased use of formal rather than informal remittance transfer channels, so that reported remittances increase even if actual remittances do not. However, the change is no doubt largely due to an increase in migration and greater ease of wealth transfer due to globalization. It is unfortunately very difficult to determine the relative importance of these factors.

remittances are used to repay loans (most likely incurred during the immigration process) reducing the potential impact of remittances on household welfare." Many remittances are also used to finance the migrations of other household members (Graner and Gurung 2003). This implies that remittance income in Nepal will tend to have a lower impact on consumption than non-remittance income.

That a migrant sends remittances implies a continuing vested interest in the home country. It is likely that many or most of those who send remittances to Nepal plan to return to Nepal. Remittance income is therefore often a temporary, rather than permanent source of income. For example, studies of migrants to Nepal show that migrants to India tend to stay from a few months to a few years (Thieme 2003) while those to other areas tend to sign employment contracts for two to three years at most (Thieme and Wyss 2005). According to Milton Friedman's permanent income hypothesis (Friedman 1957) or the life cycle hypothesis (Modigliani 1957), a source of income which is expected to be of relatively short duration has less of an impact on current consumption than longer-lasting supplies of income. There are, therefore, both practical and theoretical reasons for supposing that the elasticity of consumption from remittance income will be smaller than that from non-remittance income in Nepal.

III. Data and Variables Analyzed

Data for this analysis come from the 2003/2004 Nepal Living Standards Survey, a study carried out jointly by Nepal's Central Bureau of Statistics and the World Bank following the Living Standards Measurement Methodology developed at the World

Bank. Households were selected from 326 primary sampling units using Probability Proportional to Size (PPS) sampling (CBS 2004). The survey used a two-stage stratified sampling scheme to select a nationally representative sample of 3912 households (ibid.).

I consider "remittance income" (x_2) to refer to per capita remittances received by the household from international sources (sources outside Nepal). "Non-remittance income" (x_1) includes per capita household income from all other sources, including wages, the value of home-produced consumption goods, the rental value of the home for home-owners, profits from investments, rent received, and *remittance income from domestic sources* (which is not the focus of this study). The terms "remittance" and "non-remittance" income are thus somewhat inaccurate and used for convenience. Procedures for constructing income totals largely, but not exactly, follow CBS (2004). Income aggregates were adjusted using regional price indices to account for differing prices within Nepal. Since I use logged income as explanatory variables, 86 households with negative non-remittance income (often due to losses incurred by enterprises) were dropped from the analyses presented here, making the effective sample size n = 3826.

Construction of per capita household consumption aggregates (y) also largely follows CBS (2004), with a few exceptions. My consumption categories are: food (which does not include alcohol, coffee, or other products primarily used as drugs), durable goods[8] (examples of which include furniture, dishware, and appliances), housing

---

[8]I consider expenditure on durable goods to be the amount spent on durable goods by the household. Nepal's Central Bureau of Statistics calculates durables expenditure as the depreciation of goods owned by the household, arguing that actual expenditure is more

(including rent or, for homeowners, the rental value of the home, and home improvements), education, health care (both Western and traditional), and select non-food consumption (a heterogeneous category including goods and services not included elsewhere, including expenditures for special events such as weddings and funerals). Total consumption is the sum of consumption from all these categories. Unity was added to both income and consumption aggregates before natural logs were taken in order to define observations of value zero. Consumption and income are expressed in Nepalese rupees.

Other explanatory variables are head of household characteristics, household location dummy variables, and household composition variables. I also control for the health of the household members. To account for the possibility that remittance-receiving families differ from non-remittance receiving families in a way not otherwise captured in my model, I also include dummy variables for households receiving remittances from India and from outside of India. The main results of this chapter, especially elasticity results, are not significantly changed by excluding these remittance dummy variables.

As explained in the next section, to instrument remittance and non-remittance income I included several instruments not used in the primary regressions. The first such instrument is household financial sophistication, proxied by the number of certain types

---

properly considered investment than consumption (CBS, 2004). While this argument is valid, I feel that actual expenditure more accurately reveals how different types of income influence durable goods consumption.

of financial instruments used by the household. This seems to be an important explainer of both remittance and non-remittance income; households more familiar with financial investment opportunities may also be more familiar with migration opportunities. Also included were instruments related to the gender composition and marital status of the household: the proportion of household members which are female, a dummy variable indicating whether or not the head of household is female, and a dummy variable indicating whether or not the head of household is married. These variables are relevant to explaining non-remittance income because Nepal is still largely a male-dominated society, and women often do not have the earning potential that men do. They are also relevant to explaining remittance income, since the vast majority of migrants from and remittance senders to Nepal are men, many of whom leave their wives or households to earn and send remittances (Graner and Gurung 2003; Thieme and Wyss 2005). I also include head of household age (in years), head of household education (in years), and a dummy indicating whether or not the head of household had migrated to his/her current location, which may affect both household non-remittance earnings and awareness of migration opportunities. As a final instrument I include the distance (in hours of travel) from the household to the nearest paved road. Access to transportation infrastructure to, for example, sell farm products, seems an obvious explainer of non-remittance income. Furthermore, since many Nepali migrants are recruited by recruiting agents (Graner and Gurung 2003; Thieme and Wyss 2005), proximity to transportation may reflect how readily these agents can reach Nepali households. Our main results are robust to alternative selections of instruments; results of alternative specifications are available upon request. Summary statistics for all variables used as instruments and in the primary regression appear in Table 1.

Table 6 Summary Statistics (Consumption Analysis)

variable	Description	mean (standard error)
Consumption		
TOTAL	natural log of total household per capita consumption plus one	9.441 (.765)
FOOD	natural log of total food per capita consumption plus one	8.720 (.523)
DURABLES	natural log of total durables per capita consumption plus one	3.442 (2.731)
HOUSING	natural log of total housing per capita consumption plus one	6.663 (1.979)
EDUCATION	natural log of total education per capita consumption plus one	.571 (3.007)
HEALTH CARE	natural log of total health care per capita consumption plus one	5.094 (2.311)
NON-FOOD	natural log of total select non-food per capita consumption plus one	8.041 (1.044)
Income		
Log non-remittance income ($\ln x_1$)	natural log of real household per capita annual income plus 1, less income from remittances from international sources	8.973 (3.152)
Log remittance income ($\ln x_2$)	natural log of real household per capita annual remittance income from international sources plus 1	1.410[9] (3.152)
Household composition		
AGE0TO4	number of household members aged 0 to 4 years	.701 (.901)
AGE5TO16	number of household members aged 5 to 16 years	1.733 (1.499)
AGE17TO49	number of household members aged 17 to 49 years	2.413 (1.432)
AGE50PLUS	number of household member aged 50 years or more	.775 (.726)
PROP_FEMALE	proportion of household members which are female	.501 .186
Head of household characteristics		
HEADUNMARRIED	dummy = 1 if head of household is unmarried	.143 (.350)
HEADMIGRATED	dummy = 1 if head of household migrated to	.428

[9] For the subsample with remittances >0 (616 observations), $\ln x_2$ has a mean of 8.179 and a standard error of 1.502.

	current residence	(.495)
HEADFEMALE	dummy = 1 if head of household is female	.194 (.396)
HEADAGE	age of head of household	45.54 (14.160)
HEADEDUC	years of schooling successfully completed by head of household	2.910 (4.156)
CASTE1	dummy = 1 if head of household is of Magar, Tamang, Rai, Gurung, or Limbu caste or ethnicity	.210 (.407)
CASTE2	dummy = 1 if head of household is of Kami, Damai, Dholi, or Sarki caste or ethnicity	.080 (.272)
CASTE3	dummy = 1 if head of household is of Tharu, Yadav, Brahmin Terai, Thakur, or Hazam caste or ethnicity	.096 (.295)
CASTE4	dummy = 1 if head of household is of Newar caste or ethnicity	.077 (.267)
CASTE5	dummy = 1 if head of household is Muslim	.050 (.218)
CASTE6	dummy = 1 if head of household does not fall into categories covered by above five caste/ethnicity dummy variables, and is not Brahmin or Chhetry	.205 (.404)
SUBS_AG	dummy = 1 if subsistence agriculture is at one of the head of household's occupations	.759 (.428)
Health		
ILLNESSES	proportion of household to have suffered an injury or illness in year of study	.349 (.306)
Financial		
ASSETS	natural log of total household assets plus one	10.389 (2.868)
INDIA_REM	dummy = 1 if household received remittances from India	.112 (.316)
NONINDIA_REM	dummy = 1 if household received international remittances, none from India	.060 (.238)
FINANCIAL	financial sophistication of the household, proxied by the number of financial instruments the household owns (including savings accounts, fixed deposit accounts, stocks/shares, provident funds, pensions, commission fees, and instruments reported as "others")	1.102 (.572)
Location		
RURAL	dummy = 1 if household is located in rural area	.835 (.372)
MOUNTAIN	dummy = 1 if household is located in	.075

	mountain ecological zone	(.264)
HILL	dummy = 1 if household is located in hill ecological zone; Terai (plains) ecological zone is unspecified	.449 (.497)
WESTERN	dummy = 1 if household is in Western region of Nepal	.206 (.404)
MIDWESTERN	dummy = 1 if household is in Midwestern region of Nepal	.120 (.325)
FARWESTERN	dummy = 1 if household is in Far Western region of Nepal	.070 (.254)
EASTERN	dummy = 1 if household is in Eastern region of Nepal; Central zone is unspecified	.248 (.432)
ROAD_DISTANCE	distance (in hours travel time[10]) to nearest paved road	9.051 23.972

*Author's calculations based on the 2003/2004 Nepal Living Standards Survey. Means and standard errors are weighted with survey sample weights.

[10] Means of travel varies between individuals.

IV. Econometrics

A. Preliminaries and the problem of endogeneity

Numerous empirical studies use semiparametric techniques to construct Engel curves in order to avoid specifying a relationship between expenditure and income. My primary contribution to the existing semiparametric Engel curve literature is that I analyze the effects of two different types of income (that from international remittances and that from all other sources) and analyze the elasticities of consumption from these income types. To my knowledge, only one other paper constructs Engel curves for remittance and non-remittance income (Adams 2005), and this study uses traditional parametric analysis.

In general, I wish to determine

$$\widehat{\ln y} = \mathrm{E}(\ln y \,|\, \ln x_1, \ln x_2, \boldsymbol{s}, \boldsymbol{z}), \tag{8}$$

where $\mathrm{E}(q|\boldsymbol{Q})$ represents the expected value of quantity q given quantities $\boldsymbol{Q}$, y is per capita household consumption[11] of a certain category of good, x_1 is per capita household income not from remittances from international sources (non-remittance income), x_2 is per capita household income from remittances from international sources (remittance

[11] It is common in Engel curve analyses to use budget share as the primary dependent variable, and consumption (considered to proxy income) as an explanatory variable; my choices of variables are more appropriate for comparing the effects of two types of income.

income), and $\boldsymbol{s}$ and $\boldsymbol{z}$ are other, non-income household-level socioeconomic variables. Variables $\boldsymbol{z}$ are not included in the primary estimations of consumption. I also compensate for the endogeneity of log remittance and non-remittance income, $\ln x_1$ and $\ln x_2$. I allow the logs of x_1 and x_2 to be functions of linear combinations of instruments $\boldsymbol{s}$ and $\boldsymbol{z}$:

$$\mathrm{E}(\ln x_1) = \tau_1(\boldsymbol{\gamma_{1x}s} + \boldsymbol{\gamma_{1z}z}) \quad (9)$$

$$\mathrm{E}(\ln x_2) = \tau_2(\boldsymbol{\gamma_{2x}s} + \boldsymbol{\gamma_{2z}z}), \quad (10)$$

where $\boldsymbol{\gamma_{1x}}$, $\boldsymbol{\gamma_{1z}}$, $\boldsymbol{\gamma_{2x}}$, and $\boldsymbol{\gamma_{2z}}$ are vectors of coefficients and $\tau_1(\cdot)$ and $\tau_2(\cdot)$ are functions which depend upon the methods of instrumentation used.

Because $\ln x_1$ had significant outliers, and the residuals resulting from ordinary least squares regression were heteroskedastic, a robust regression method using an M-estimator with fitting by iterated re-weighted least squares as in Huber (1981) was used to predict $\mathrm{E}(\ln x_1)$. A tobit estimation was used to instrument $\mathrm{E}(\ln x_2)$ to take account of the many observations of value zero. Several other means of instrumentation were explored (including OLS to predict $\mathrm{E}(\ln x_1)$ and a Heckman maximum likelihood estimation to predict $\mathrm{E}(\ln x_2)$); using these alternative methods did not significantly change results.

Substitution of instrumented values of explanatory variables into a semiparametric regression equation does not generally yield consistent results. I control for the endogeneity of income variables using the control function method described in Newey *et al.* (1999) (and less rigorously, but more accessibly, in Blundell and Powell 2001). (This method has been used in several other semiparametric Engel curve analyses. A few examples are Blundell *et al.* 2098, Blundell *et al.* 2003, and Gong *et al.* 2005.)

The control function method relies upon the calculation of the residuals v_1 and v_2 from the instrumenting Equations (9) and (10):

$$v_1 = \ln x_1 - \mathrm{E}(\ln x_1 \,|\boldsymbol{s}, \boldsymbol{z}) \tag{11a}$$

$$v_2 = \ln x_2 - \mathrm{E}(\ln x_2 \,|\boldsymbol{s}, \boldsymbol{z}) \tag{11b}$$

These residuals are then used as regressors in the estimation of Equation (8).

Since the focus of this chapter is on the effects of income, only income variables enter the model nonparametrically. Re-writing Equation (8) more precisely, the goal then is to estimate

$$\widehat{\ln y} = \mathrm{E}(\ln y \,|\, \ln x_1, \ln x_2, \boldsymbol{s}, \boldsymbol{z}) = \mathrm{E}(\mathrm{f}(\ln x_1, \ln x_2) + \boldsymbol{\beta_s s} + \mathrm{g}(v_1, v_2)), \tag{12}$$

where $\boldsymbol{\beta_s}$ is a vector of coefficients, $\mathrm{f}(\cdot)$ and $\mathrm{g}(\cdot)$ are functions to be determined, and $\mathrm{E}(g(v_1, v_2)) = 0$. Note that this equation is considerably more general than, say,

$$\widehat{\ln y} = \mathrm{E}(\mathrm{f}_1(\ln x_1) + \mathrm{f}_2(\ln x_2) + \boldsymbol{\beta_s s} + \mathrm{g}_1(v_1) + \mathrm{g}_2(v_2)), \tag{13}$$

which does not allow interaction between x_1 and x_2.[12] I also suppose that the residuals from the estimation of Equation (12) to be normally distributed with mean zero:

$$(\ln y) \sim \mathrm{N}(\mathrm{E}(\ln y), \sigma^2) \tag{14}$$

B. Semiparametric estimation techniques

[12] To allow interaction between remittance and non-remittance income is to allow that the way remittance income affects consumption is affected by the level of non-remittance income and vice versa.

To estimate Equation (12) I use general additive model spline techniques as described in Wood (2003) and Wood (2006) with the mgcv program package (Wood 2008) for the R statistical computing program (R Development Core Team 2008). I use penalized thin-plate regression splines because they are computationally efficient and because spline methods are compatible with the control method approach described above (Newey *et al.* 1999).

Estimation of Equation (12) yields a predicted functional relationship:

$$\widehat{\ln y} = \hat{f}(\ln x_1, \ln x_2) + \hat{\boldsymbol{\beta}}_s \boldsymbol{s} + \hat{g}(v_1, v_2) \tag{15}$$

Some discussion of the form of the nonparametric functions $\hat{f}(\cdot)$ and $\hat{g}(\cdot)$ and how they are estimated will facilitate explanations of regression results. Brevity necessitates an explanation which is somewhat qualitative and imprecise; a much more thorough explanation is to be found in Wood (2003) or Wood (2006), from which the following brief explanation is distilled.

The functions $\hat{f}(\cdot)$ and $\hat{g}(\cdot)$ are a linear combination of several basis functions and their coefficients. Let $\boldsymbol{\beta}$ collectively represent these coefficients and $\boldsymbol{\beta}_s$, the coefficients for the linear terms $\boldsymbol{s}$.

Calculation of predicted functions must balance the predictability and the smoothness of the resultant function. (It is possible to construct a function which fits the sample perfectly, but this would be an extremely "wiggly" and unaesthetic function and probably of very little use for out-of-sample predictions.) The problem of estimating $\widehat{\ln y}$ can be given as:

$$\min \left\| \ln y - \widehat{\ln y} \right\|^2 + \lambda_{\hat{f}} J_{\hat{f}}(\hat{f}) + \lambda_{\hat{g}} J_{\hat{g}}(\hat{g}) \tag{16}$$

with respect to the estimated values of all model coefficients $\widehat{\boldsymbol{\beta}}$. Here, J.(·) are functions penalizing the wiggliness of function · and the smoothing parameters $\lambda_{\hat{f}}$ and $\lambda_{\hat{g}}$ (collectively written $\boldsymbol{\lambda}$) determine to what extent one values smoothness over data fitting. The estimated values of these smoothing parameters $\hat{\boldsymbol{\lambda}}$ are determined via generalized cross validation.

Variances of estimated coefficients are calculated using the Bayes' law and an assumed prior distribution. Define

$$\hat{f}(\ln x_1, \ln x_2) + \widehat{\boldsymbol{\beta}}_s \boldsymbol{s} + \hat{g}(v_1, v_2) = \hat{h}(\boldsymbol{w}|\boldsymbol{\beta}), \quad (17)$$

where $\boldsymbol{w}$ is the observed data. Then

$$\hat{h}(\boldsymbol{\beta}|\boldsymbol{w}) \propto \hat{h}(w|\boldsymbol{\beta})\hat{h}(\boldsymbol{\beta}), \quad (18)$$

where $\hat{h}(\boldsymbol{\beta})$ is a prior distribution which favors smooth models over wiggly ones and gives equal weight to models of equal smoothness; for details see Wood (2006).

C. Nonlinear functions of estimated parameters

Most of the analysis in this chapter focuses on the effects of non-remittance and remittance income, that is, the $\hat{f}(\ln x_1, \ln x_2)$ term from Equation (15). I also calculate elasticities ε as functions of remittance and non-remittance income:

$$\varepsilon_i(\ln x_1, \ln x_2) = \frac{\partial \hat{f}(\ln x_1, \ln x_2)}{\partial \ln x_i}, i \in 1,2 \quad (19)$$

Distributional properties of these derived quantities depend on the distributional properties of coefficients and the smoothing parameters; I re-write Equation (15) as

$$\hat{f}(\ln x_1, \ln x_2) + \widehat{\boldsymbol{\beta}}_s \boldsymbol{s} + \hat{g}(v_1, v_2) = \hat{h}(\boldsymbol{w}|\boldsymbol{\beta}, \boldsymbol{\lambda}) \quad (20)$$

to make this explicit. Confidence bands for these derived quantities can be obtained by the method of Krinsky and Robb (1986), that is, by resampling $\hat{\mathrm{h}}$ a large number of times assuming a multivariate normal distribution for $\boldsymbol{\beta}$, calculating the desired derived quantities for each sampling iteration, and then taking appropriate quantile measurements for these resultant quantities. However, this method, using the distribution $\hat{\mathrm{h}}(\boldsymbol{\beta}|\boldsymbol{w}, \hat{\boldsymbol{\lambda}})$, ignores that the coefficients $\hat{\boldsymbol{\beta}}$ are calculated given smoothing parameters $\hat{\boldsymbol{\lambda}}$, and is only valid if $\boldsymbol{\lambda}$ is known with certainty (which it is generally not).

To correct for this I use a variation of the method of Krinsky and Robb as given in Wood (2006), which is much less computationally intensive than pure bootstrapping. The primary regression Equation (15) is estimated using thin plate regression splines and $\hat{\boldsymbol{\lambda}}^0$ determined via generalized cross validation. Over *nboot* iterations (19 in my case; only a small number of bootstrap iterations are needed since the primary variables of interest are functions of $\boldsymbol{\beta}$, not $\boldsymbol{\lambda}$) random deviates are simulated with mean and variance defined by the fitted values and residuals from the regression; these deviates are used to create a bootstrap response vectors $\boldsymbol{w}^k$, $k \in [1, nboot]$. For each $\boldsymbol{w}^k$, Equation (15) is again estimated with bootstrap smoothing parameter estimates $\hat{\boldsymbol{\lambda}}^k$ determined by generalized cross validation. Equation (15) is then estimated with $\boldsymbol{w}$ using smoothing parameters specified at $\hat{\boldsymbol{\lambda}}^k$ to obtain bootstrap coefficient estimates $\hat{\boldsymbol{\beta}}^k$ and the associated variance-covariance matrix $\hat{\boldsymbol{V}}^k$. For each $\hat{\boldsymbol{\beta}}^k$ and $\hat{\boldsymbol{V}}^k$ a large (in my case 1000-observation) multivariate normal distribution is sampled. The aggregate of these multivariate normal distributions yield a distribution of $\hat{\boldsymbol{\beta}}$ which is unconditional on $\hat{\boldsymbol{\lambda}}$. For each estimated value of $\hat{\boldsymbol{\beta}}$ (of which there are 19000 in my case) elasticities were calculated. The mean of this distribution of elasticities gives the estimated elasticity ε_i. The 2.5th and 97.5th percentiles of the elasticity distribution yield 95% credible intervals.

V. Estimation Results

A. Parametric Components

Regression results for the parametric components of the estimation of Equation (15) for different categories of consumption good are presented in Table 7, as well as statistical significance estimates for the nonparametric components.

Table 7 Estimation Results for Consumption Equations (10)

	Total	Food	Durables	Housing	Education	Health care	Select non-food
Income:							
$f(\ln x_1, \ln x_2)$	***	***	***	***	***	*	***
Control function							
$g(v_1, v_2)$:	***	***	***	***	***		***
Household composition:							
AGE0TO4	-.063***	-.049***	.220***	-.080**	-.477***	.175***	-.055***
	(.011)	(.008)	(.064)	(.037)	(.060)	(.053)	(.019)
AGE5TO16	-.006	-.032***	.210***	.040*	1.062***	.017	.035***
	(.007)	(.005)	(.038)	(.022)	(.035)	(.031)	(.011)
AGE17TO49	.007	-.033***	.198***	.045**	.371***	.085***	.048***
	(.006)	(.004)	(.033)	(.019)	(.031)	(.028)	(.010)
AGE50PLUS	-.031***	-.025***	-.221***	.086***	-.271***	.152***	-.085***
	(.009)	(.007)	(.050)	(.029)	(.047)	(.042)	(.015)
Head of household characteristics:							
CASTE1	-.224***	-.160***	-.395***	-.372***	-.877***	-.679***	-.143***
	(.023)	(.017)	(.128)	(.074)	(.120)	(.107)	(.038)
CASTE2	-.133***	-.116***	-.008	-.331***	-.524***	-.262*	-.045
	(.031)	(.023)	(.176)	(.102)	(.164)	(.147)	(.052)
CASTE3	-.106***	-.054**	-.019	-.309***	-1.214***	-.418***	-.057
	(.033)	(.025)	(.188)	(.109)	(.175)	(.157)	(.055)
CASTE4	-.054*	-.171***	-.324**	.510***	-.312**	-.353***	-.132***
	(.029)	(.022)	(.164)	(.095)	(.153)	(.137)	(.048)
CASTE5	-.093**	-.030	-.428*	-.387***	-2.146***	-.569***	-.146**

	(.042)	(.032)	(.240)	(.139)	(.224)	(.200)	(.070)
CASTE6	-.112***	-.077***	.050	-.302***	-1.201***	-.207*	-.074*
	(.025)	(.019)	(.140)	(.081)	(.131)	(.117)	(.041)
SUBS_AG	-.032	-.042*	.706***	-.049	.475***	.204	.142***
	(.032)	(.024)	(.182)	(.105)	(.170)	(.152)	(.070)
Health:							
ILLNESSES	.044*	.033*	-.141	-.210**	-.062	2.241***	.018
	(.026)	(.019)	(.145)	(.084)	(.135)	(.121)	(.043)
Financial:							
ASSETS	.072***	.058***	.096**	.182***	.083*	.035	.068***
	(.008)	(.006)	(.047)	(.027)	(.044)	(.039)	(.013)
INDIA_REM	-.531	-.320**	-.141	-.976	-.090	-.088	-.855*
	(.414)	(.173)	(1.368)	(.711)	(1.387)	(.672)	(.495)
NONINDIA_ REM	-.598	-.431**	-.588	-.612	.990	.483	-.865*
	(.426)	(.185)	(1.409)	(.741)	(1.419)	(.768)	(.520)
Location:							
RURAL	-.164***	-.023	.214	-.725***	-.638***	.051	-.108***
	(.025)	(.019)	(.140)	(.081)	(.131)	(.117)	(.041)
MOUNTAIN	.176***	.277***	-.201	.169*	-.333**	- 1.003***	.134***
	(.030)	(.023)	(.172)	(.099)	(.160)	(.143)	(.050)
HILL	.184***	.182***	.039	.383***	.068	-.355***	.215***
	(.022)	(.017)	(.125)	(.073)	(.117)	(.104)	(.037)
WESTERN	.064***	.153***	.083	.002	.317***	.063	.051
	(.023)	(.017)	(.126)	(.073)	(.118)	(.106)	(.038)
MIDWESTERN	-.146***	-.092***	.979***	.017	-.459***	-.450***	-.150***
	(.026)	(.020)	(.150)	(.087)	(.140)	(.125)	(.044)
FARWESTERN	-.184***	-.174***	.681***	.412***	-.165	-.286*	-.190***
	(.034)	(.025)	(.192)	(.111)	(.179)	(.160)	(.056)

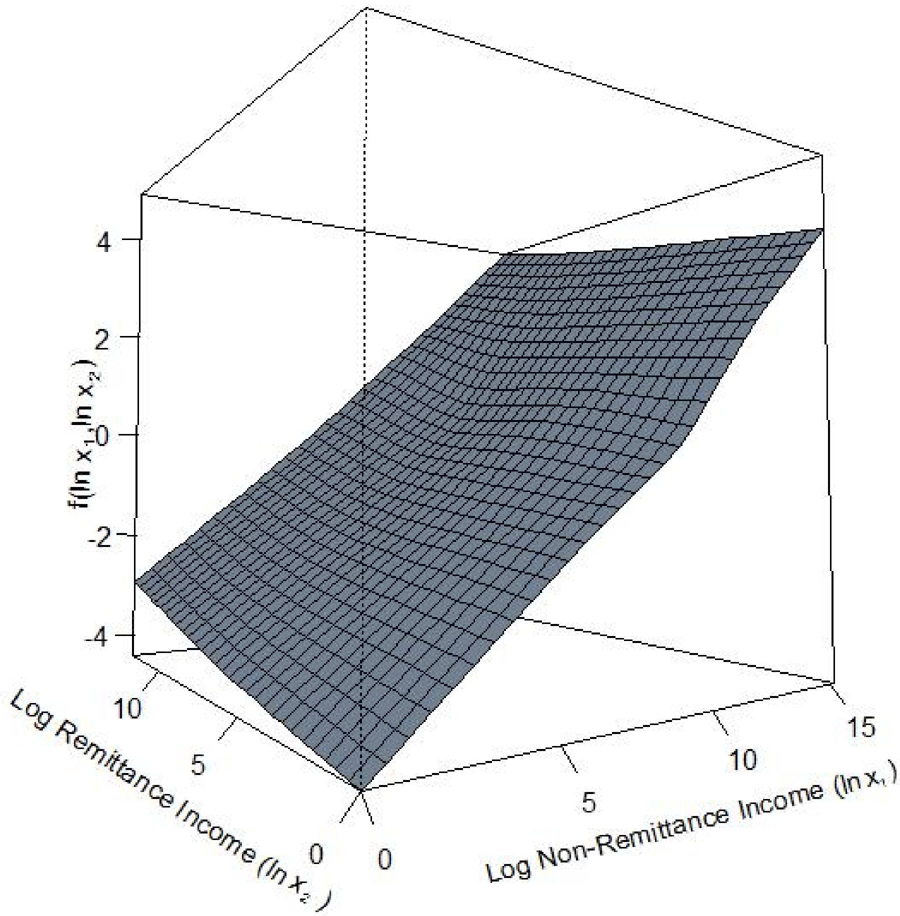
EASTERN	-.100***	.023	.116	-.376***	-.028	-.098	-.174***
	(.022)	(.016)	(.118)	(.069)	(.110)	(.099)	(.035)
intercept	9.061	8.352***	1.114**	5.313***	1.627***	3.937***	7.513***
	(.097)	(.058)	(.443)	(.251)	(.423)	(.340)	(.141)
Adjusted R^2	.700	.603	.151	.429	.410	.164	.519
n	3826	3826	3826	3826	3826	3826	3826

*The symbols *, **, and *** represent significance at the 90%, 95%, and 99% levels, respectively. Standard errors follow estimated coefficients in parentheses.

B. Engel Curves for Total Consumption

Consider the function $\hat{f}(\ln x_1, \ln x_2)$ where the independent variable in Equation (15) ($\ln y$) is log per capita total household consumption. The estimated function itself is shown in Figure 1.

Figure 1 Income Dependent Component of Estimated Total Consumption Function[13]



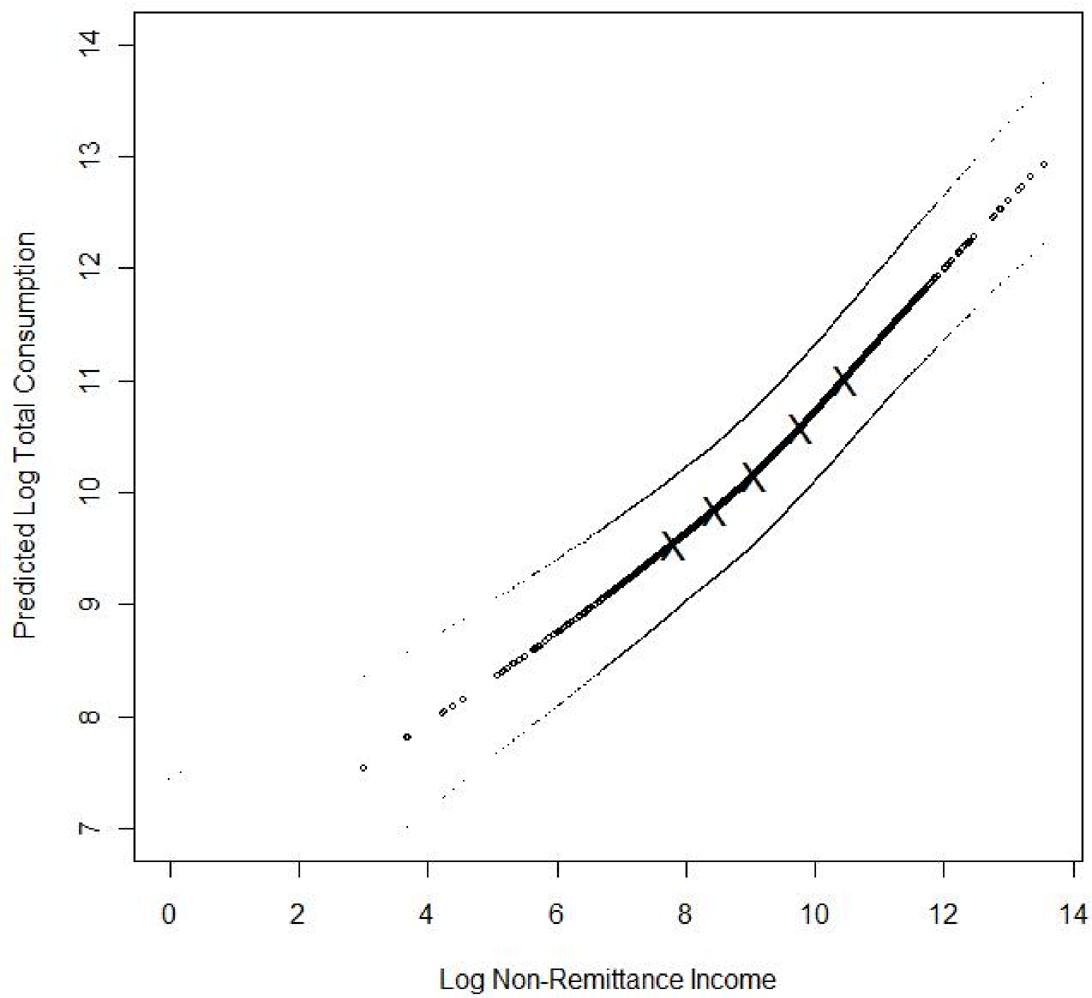
[13] The income-dependent component $\hat{f}(\ln x_1, \ln x_2)$ of the estimated consumption function as a function of log non-remittance income ($\ln x_1$) and log remittance income ($\ln x_2$).

Though this graph is quite informative in the strictest sense of the word, it is difficult for most to draw conclusions from it by visual inspection. The graph would be yet harder to interpret if credible interval surfaces were added.

One can construct something which looks more like a familiar Engel curve graph by holding $\ln x_1$ or $\ln x_2$ constant. Figure 2 shows log total consumption as a function of log non-remittance income with remittance income held constant at the median of positive observed remittance income values ($\ln x_2 = 8.23$, or $x_2 = 3737.56$ Nepalese rupees), where $\ln x_1$ is allowed to range through the observed values of non-remittance income.[14] I also set the values of other socioeconomic variables equal to their means, $\boldsymbol{s} = \bar{\boldsymbol{s}}$, for this and later analyses.

[14] In this and future graphs in this chapter with $\ln x_1$ as the abscissa, five outlying observations are excluded since they greatly extend the domain of the graph.

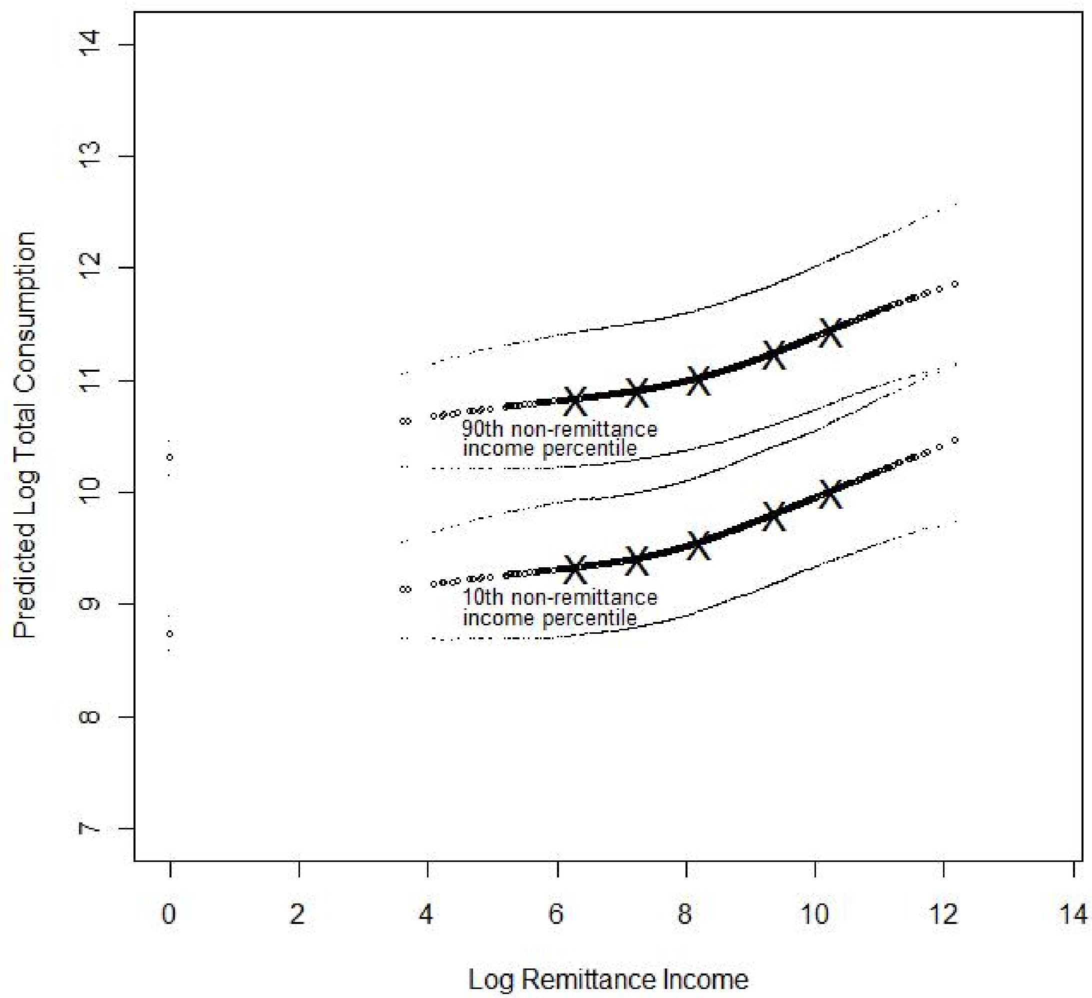
Figure 2 Log Total Consumption versus Log Non-Remittance Income[15]



[15] Predicted log total consumption as a function of the observed values of log non-remittance income $\ln x_1$, with remittance income $\ln x_2$ constant at the median of observed non-zero values and socioeconomic variables $\boldsymbol{s}$ at mean observed values. 95% credible bands are shown. "X" marks indicate the 10th, 25th, 50th, 75th, and 90th percentiles of $\ln x_1$.

Figure 3 shows a pair of Engel curves for log consumption versus log remittance income, with non-remittance income fixed at the 10th and 90th percentiles. In this case, the remittance income Engel curve has much the same shape for different levels of non-remittance income, though this was not supposed *a priori*. The relatively narrow credible band where $\ln x_2 = 0$ is a result of the many observations of zero remittance income.

Figure 3 Log Total Consumption versus Log Remittance Income[16]



16 Predicted log total consumption as a function of the observed values of log remittance income $\ln x_2$, with socioeconomic variables $\boldsymbol{s}$ held at mean values and non-remittance income $\ln x_1$ held constant at the 10th and 90th percentiles of observed values. 95% credible intervals are shown. "X" marks indicate the 10th, 25th, 50th, 75th, and 90th percentiles of the non-zero values of $\ln x_2$.

The mathematically inclined reader may note from Figures 2 and 3 that many of the predicted consumption values are larger than predicted total income values. This is not an artifact of my methods or my income and consumption calculations: though Nepal's Central Bureau of Statistics income and consumption calculations differ slightly from my own, they report consumption to be higher than income for all but the richest three population deciles (CBS 2004).

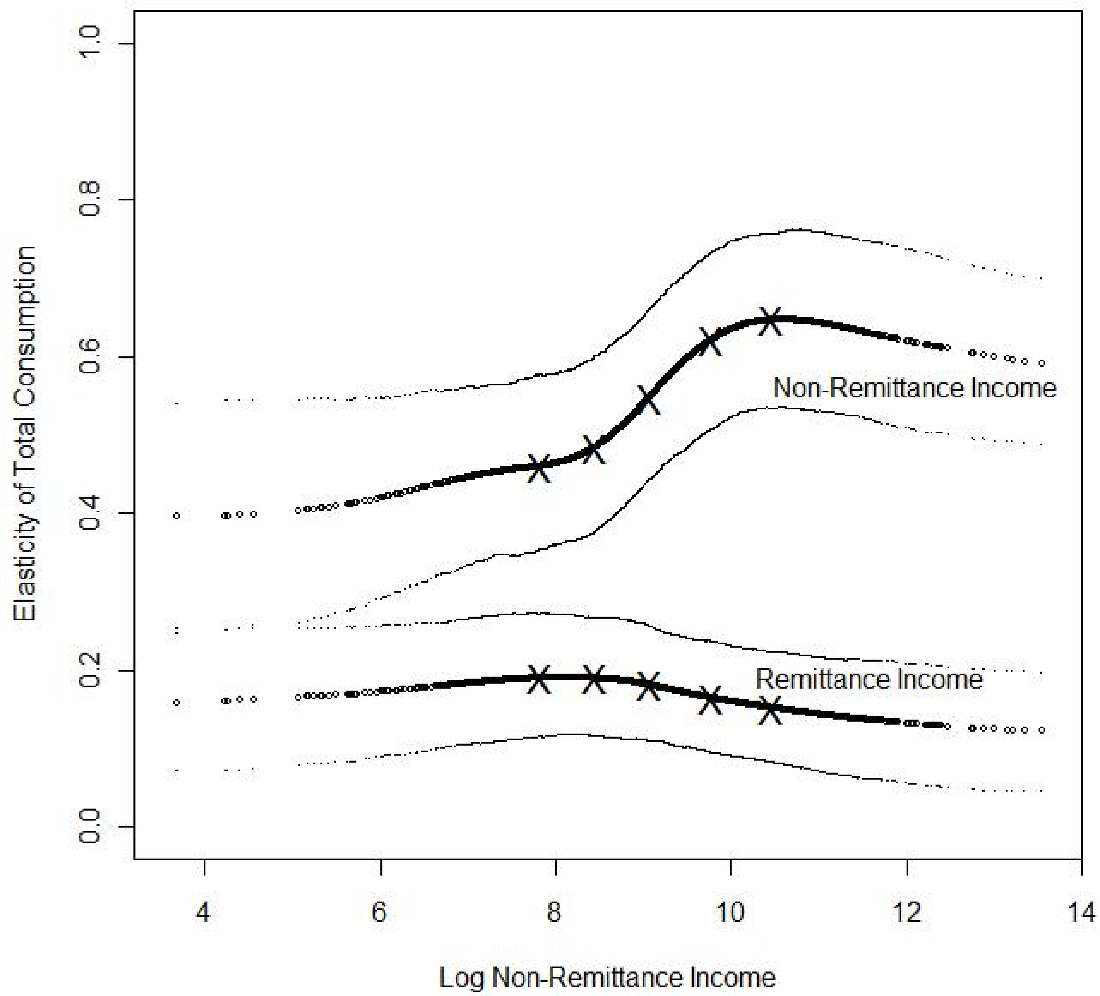
Several explanations exist. Both income and consumption are calculated based on figures reported by the households themselves. It may be that consumption is generally overestimated by households, or income underestimated; the latter is perhaps more likely, as reported consumption is often considered a more accurate measure of income than reported income in developing nations. It is also possible that families consume beyond what their income would normally allow, due to loans, gifts, or even illegitimate activities.

Some minimum level of consumption is necessary for a household to survive. To take a look at the extreme poor in my sample, for the 146 households which report per capita income between 0 and 2000 rupees, mean per capita consumption is 7639 rupees (excluding 10 outliers with abnormally high consumption). A positive, rather than a zero, intercept for log consumption as a function of log income is theoretically and empirically justifiable.

C. Semiparametric Elasticities

I calculate elasticities of consumption from changes in both remittance and non-remittance income using Equation (19). Figure 4 shows these elasticities as a function of $\ln x_1$, again with $\ln x_2$ held constant at the median of its positive observed values.

Figure 4 Elasticity of Total Consumption with Respect to Remittance and Non-Remittance Income[17]



[17] Elasticity of total consumption with respect to non-remittance income and remittance income as a function of log non-remittance income $\ln x_1$, with log remittance income $\ln x_2$ held constant at the median of observed positive values. 95% percent credible bands are shown. "X" marks indicate the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of $\ln x_1$.

That the elasticity of consumption is significantly less than one for both types of income is not surprising given a positive intercept for consumption versus income. Some earned income is likely offset by a reduction in the activities which caused reported consumption to exceed reported income. (For example, higher income households might be less likely to receive donor aid or gifts from other households). For non-remittance income, the results presented here show a gradual increase in the elasticity of consumption until approximately the 90th percentile (still not a wealthy household by Western standards), followed by a slow decline (perhaps due to increased savings). The lower elasticity of total consumption for remittance income is a topic addressed in the next section. Though it is less aesthetically appealing, it is almost as informative, and much more compact, to look at elasticity calculations in table form. Tables 8 and 9 give the elasticities of consumption with respect to remittance and non-remittance income, respectively, for various categories of goods, for certain representative percentiles of non-remittance income. (The results for total consumption correspond to the "X" points in Figures 2 and 3.)

Table 8 Elasticities of Consumption with Respect to Remittance Income, by Percentiles of Non-Remittance Income

Non-Remittance Income Percentile	Total	Food	Durables	Housing	Education	Health Care	Select Non-Food
10	.190***	.065***	.278**	.177***	.278*	.067	.234***
	(.041)	(.022)	(.141)	(.072)	(.152)	(.076)	(.054)
25	.190***	.063***	.261*	.180***	.269*	.067	.236***
	(.039)	(.019)	(.138)	(.071)	(.145)	(.074)	(.053)
50	.181***	.061***	.245*	.187***	.244*	.066	.228***
	(.038)	(.016)	(.135)	(.068)	(.143)	(.074)	(.051)
75	.65***	.059***	.231*	.190***	.197	.064	.210***
	(.036)	(.0016)	(.121)	(.067)	(.134)	(.073)	(.048)
90	.52***	.0058***	.223*	.192***	.152	.061	.193***
	(.036)	(.15)	(.116)	(.067)	(.130)	(.070)	(.048)

*For these calculations non-remittance income is held constant at the median observed value. Standard errors follow estimated elasticities in parentheses. *, **, and *** represent statistically significant difference from zero at the 90%, 95%, and 99% levels, respectively.

Table 9 Elasticities of Consumption with Respect to Non-Remittance Income, by Percentiles of Non-Remittance Income

Non-Remittance Income Percentile	Total	Food	Durables	Housing	Education	Health Care	Select Non-Food
10	.460***	.341***	1.577***	.681***	1.039***	.404***	.702***
	(.059)	(.031)	(.237)	(.146)	(.232)	(.175)	(.083)
25	.484***	.344***	1.609***	.715***	1.035***	.406***	.719***
	(.059)	(.029)	(.240)	(.135)	(.238)	(.174)	(.080)
50	.547***	.347***	1.634***	.755***	1.042***	.409***	.761***
	(.056)	(.029)	(.244)	(.127)	(.241)	(.173)	(.079)
75	.621***	.350***	1.642***	.782***	1.076***	.411***	.807***
	(.056)	(.029)	(.245)	(.127)	(.236)	(.171)	(.079)
90	.646***	.351***	1.625***	.788***	1.093***	.413***	.818***
	(.056)	(.029)	(.241)	(.133)	(.243)	(.171)	(.081)

*For these calculations remittance income is held constant at the median observed positive value. Standard errors follow estimated elasticities in parentheses. *, **, and *** represent statistically significant difference from zero at the 90%, 95%, and 99% levels, respectively.

This analysis supports Engel's law, that the elasticity of food consumption is always less than one (with respect to both remittance and non-remittance income). That the elasticity of durables consumption with respect to non-remittance income is uniformly greater than one indicates that durables are luxury goods, as is often shown or assumed to be the case in Engel curve analyses.

In Tables 8 and 9, elasticity of consumption from remittance income is shown to be always less than that from non-remittance income. I formally compared the two types of elasticities using t-tests[18] to compare the means of the generated samples of elasticities for the two types of income at the 10th, 25th, 50th, 75th, and 90th percentiles, for all of categories of consumption. For all non-remittance income percentiles and consumption quantities represented in Tables 8 and 9, I strongly rejected the null hypotheses that the elasticity from non-remittance income is equal to the elasticity from remittance income ($p<10^{-4}$).

It is worth comparing the results presented here to those of the only other published study (to my knowledge) which is similar to this one, Adams (2005). Adams finds that remittance-receiving households spent less on the margin on food, consumer goods, and durables, but spend more on the margin on education, health, and housing. Thus, our findings are similar vis-à-vis food, durables, and select non-food consumption but differ vis-à-vis the other categories. Since I use quite different

---

[18] The validity of these tests depends upon the normality of the distributions to be tested. Because of the methods I use, my simulated distributions should be normal; Jarque-Bera tests confirm that this is the case.

methods, and I analyze Nepal while Adams analyzes Guatemala, I do not consider these results to be contradictory.

D. Specification Tests

Finally, I use F-tests to determine whether, in this case, semiparametric methods provide advantages over less general specifications. These less general specifications are those where $\ln x_1$ and $\ln x_2$ (and associated errors v_1 and v_2) enter the model parametrically and linearly, where these terms enter the model parametrically and quadratically, and where these terms enter the model nonparametrically in an additively separable form (Equation 13). Results of these tests are summarized in Table 10. In all cases parametric specifications were strongly rejected. F-tests indicate that (for the data analyzed here) the nonparametric model has little if any advantage over a model additively separable in income when the consumption category being analyzed is durables, housing, education, or select non-food. However, I reject at the 10% level the additively separable model when analyzing total consumption and food consumption, and at the 5% level the additively separable model when analyzing health care consumption; for these classes of goods, using a model additively separable in income may yield inconsistent estimated functions.

Table 10 F-Test P-Values for Specification Tests of Alternative Models versus the Model: $E(\ln y) = \hat{f}(\ln x_1, \ln x_2) + \hat{\boldsymbol{\beta}}_s \boldsymbol{s} + \hat{g}(v_1, v_2)$							
alternative model	total cons.	food	durables	housing	education	health care	select non-food
linear:	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$
quadratic:	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$	$<10^{-16}$
semiparametric additively separable:	.088	.098	.519	1.000	1.000	.049	.244

The linear (in income) model is $E(\ln y) = \hat{\beta}_{x1} \ln x_1 + \hat{\beta}_{x2} \ln x_2 + \hat{\beta}_{v1} v_1 + \hat{\beta}_{v2} v_2 + \hat{\boldsymbol{\beta}}_s \boldsymbol{s}$, the quadratic (in income) model $E(\ln y) = \hat{\beta}_{x1} \ln x_1 + \hat{\beta}_{x2} \ln x_2 + \hat{\beta}_{x1sq} (\ln x_1)^2 + \hat{\beta}_{x2sq} (\ln x_2)^2 + \hat{\beta}_{x12} \ln x_1 \ln x_2 + \hat{\beta}_{v1} v_1 + \hat{\beta}_{v2} v_2 + \hat{\beta}_{v1sq} {v_1}^2 + \hat{\beta}_{v2sq} {v_2}^2 + \hat{\beta}_{v12} v_1 v_2 + \boldsymbol{\beta}_s \boldsymbol{s}$, and the semiparametric additively separable (in income) model $E(\ln y) = \hat{f}_1(\ln x_1) + \hat{f}_2(\ln x_2) + \hat{\boldsymbol{\beta}}_s \boldsymbol{s} + \hat{g}_1(v_1) + \hat{g}_2(v_2)$.

VI. Conclusions and Policy Implications

The results presented here imply that, for each analyzed consumption category, the elasticity of consumption is significantly less from remittance than from non-remittance income. The difference is particularly marked for durables and health care consumption.

I will discuss briefly effects of remittances and the possible effects of foreign aid, since these two means of transferring wealth from rich nations to poor are often compared. Firstly, the relatively low elasticities of consumption from remittance income, even for very poor households, indicate that care must be taken when deciding how much benefit a given amount of remittance income is to impoverished nations. My results imply that elasticities with respect to remittance income are highest for durables, education, and select non-food consumption and lowest for food and health care consumption. On the other hand, one might suppose aid programs to primarily focus on, for example, food, education, and health care. If this or a similar supposition is correct, then this study indicates that remittance income is hardly a substitute for foreign aid. Still, remittance income seems to be an important source of funds for education, particularly for families with low levels of non-remittance income.

I now try to shed some light on why elasticities of consumption from remittance income are so low. Two possibilities seem relevant: that remittances are saved or that they are spent on expenses not captured by the measures of consumption used here, such as loan payments. There is reason to think that the latter is the larger factor. Migrant workers from Nepal often take out loans to finance their migration.

Large shares of remittances are often used to repay loans, which are generally incurred during the migration process (Ferrari *et al.* 2007). Remittances are often also used to finance other migrants rather than for consumption (Graner and Gurung 2003).

Ferrari *et al.* (2007) make several recommendations to enhance the consumption increasing effects of remittances. A government or aid-organization backed viable loan scheme for migrants could help reduce the proportion of remittances which goes towards paying off loans. A financial literacy program for migrants would allow migrants to choose more efficient remittance transfer mechanisms and bring more remittances into formal transfer networks. Changes in legal and regulatory obstacles to formal remittance transfer mechanisms in the Nepal-India corridor would also help to bring more remittances into the formal financial sector.

The merits of formal versus informal remittance transfer mechanisms may seem an issue tangential to this analysis, but this is not necessarily the case. If remittances were used to draw recipients into the formal financial sector, this would lead to more financially savvy consumers who could better convert a temporary income stream (remittances) into a permanent one (productive investments). If remittances are perceived to increase permanent income, then their effects on consumption could be enhanced.

Besides the implications of this study for the specific case of international remittances to Nepal, this chapter presents a robust and flexible way to compare the uses of remittance and non-remittance income (or any number of different income types) at the microeconomic level using household survey data. Because I do not specify an artificial functional relationship between income and consumption and I allow for the possible endogeneity of income, I feel that the methods presented here hold significant potential for future Engel curve analyses.

Chapter 4: Remittances and Poverty in Eastern Europe and the Former Soviet Union

I. Introduction

The study of the impact of remittances from international sources on developing countries is very much a developing field. In particular, few studies exist examining the effect of remittance income on poverty. This chapter presents a macroeconomic study of the impact of remittance income on poverty during the late 1990s and early 2000s in many of the Eastern European and Central Asian nations which composed the former Soviet bloc[19].

Remittances are an important source of income in much of this region. According to the World Bank, Europe and Central Asia received an estimated $62 billion in remittances in 2008 (Ratha *et al.* 2008). The two largest recipients of remittances in proportion of GDP are in this region: Tajikistan received remittances equivalent to 45.5% of its GDP and Moldova received remittances equivalent to 38.3% of its GDP in 2008 (Ratha *et al.* 2008). The largest source of remittances to the region in 2008 was the European Union; this is in part due to the recent depreciation of the dollar relative to the euro (Ratha *et al.* 2008).

---

[19] Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyz Republic, Latvia, Lithuania, Macedonia, Moldova, Poland, Romania, Serbia and Montenegro, and Ukraine.

The macroeconomic effects of remittances are ambiguous from a theoretical perspective. The most obvious effects are that remittances increase the income of recipient families, which should increase consumption, which should lead to increased welfare and a stronger economy. Remittances are also a source of foreign currency and, in this sense, contribute positively to the recipient nation's balance of payments.

However, there are reasons to suppose that, at least at the macroeconomic level, remittances are not welfare enhancing. If remittances are spent primarily on imports, then they could fuel a trade imbalance which would make the recipient country more vulnerable to certain shocks (for example, a change in host country worker permit regulations which caused a sudden drop in migrant workers). Remittances may also create a moral hazard problem – recipient households may work less, and thus there would be less than a one-to-one increase in consumption for a given increase in remittances.

The means by which remittances are transferred and the financial sophistication of recipient families may also affect the effectiveness of remittances. According to the permanent income hypothesis, remittance income will have an impact on consumption to the extent that they are perceived to increase *permanent* income. Remittance senders who plan to return to their home countries likely remit more than their counterparts who do not, since (1) returning migrants are likely to benefit to some degree from the remittance they have sent, while senders who do not plan to return remit primarily for altruistic reasons, and (2) permanent migrants likely use part of their income to purchase property and for other large costs which temporary migrants do not have to incur. A study from the region (Pinger 2007), in Moldova, shows that migrants who plan to return to their home countries remit 30% more than their permanent counterparts. This indicates that most remittances are sent by those who do not plan to work abroad abruptly, but rather to return to their home countries after several years of work. This indicates that remittances are often a temporary income source at the microeconomic level.

The degree to which remittances affect permanent income depends upon the ways that remittances can be leveraged into productive investments. In many parts of the world, this is done most effectively through the financial sector. However, banks have not achieved significant penetration in many rural and poor areas in much of Eastern Europe and the former Soviet Union. This can be seen by the prevalence of informal transfer mechanisms for sending remittances. For example, in Albania, 81% of remitters use informal channels (Gedeshi 2002); in Armenia, over 80% of remitters never use formal channels (PA Consulting Group 2006); in Moldova, 42% (European Bank for Reconstruction and Development) to 50% (Orozco 2007) use informal mechanisms or hand carry the money; in Tajikistan, a large and unknown amount of remittances are sent through the informal sector (Kireyev 2006). Estimates of informal flows to Serbia range from 50% (Martinez *et al.* 2006) to 80% (SECO 2007). This is the case despite the fact that costs of formal transfers are not much higher than costs of informal transfers (PA Consulting Group 2006; Quillin *et al.* 2007; Roberts and Banaian 2004; Orozco 2007). While some migrants use Hawala-type networks (that is, informal organized money transfer networks) to send remittances (Roberts and Banaian 2004; PA Consulting Group 2006), quite often they are sent through informal couriers such as bus drivers (bus drivers are particularly used for remittances to Serbia: Martinez *et al.* 2006; SECO 2007) or hand-carried by the migrant (e.g., Aretha *et al.* 2005; PA Consulting Group 2006).

That so few send remittances through formal channels indicates that remittances often do not enter the formal financial sector. This implies that households investment and savings opportunities from remittance income are limited, which impairs households' ability to leverage remittance income into a more permanent income source. This implies that the effect of remittance income on consumption may be small.

II. The Measurement of Poverty

Since this chapter deals with poverty, it will be helpful to clarify a few basic concepts in how poverty is defined. Perhaps the most basic measure of poverty is the *rate of poverty*: the number of persons whose consumption (or income) is below an arbitrary poverty line. For this analysis, poverty lines used are \$2.15 US and \$4.30 per capita of consumption per day (discussed in more detail below), though \$1 US per capita per day is also often used in other studies.

One problem with examining only the rate of poverty can be illustrated with the following thought experiment: if all the consumption from one half of the impoverished population were redistributed to the other half then, in most scenarios, the poverty rate would decrease. However, few would consider this change to be an improvement. An alternative measure of poverty, which does not decrease given said change, is the *depth of poverty*: the difference between the per capita consumption of an average impoverished person and the poverty line, as a percentage of the poverty line.

Poverty depth is lacking from a theoretic point of view in that it ignores the generally accepted concavity of the utility function with respect to consumption. The *severity of poverty,* a poverty measure which allows for a nonlinear relation between consumption and utility, is the mean square of the distance between per capita consumption of impoverished individuals and the poverty line. The rate, depth, and severity of poverty are all examined in this study.

III. Literature Review and the Contribution of this Study

A few other studies examine the remittance-poverty link at the macroeconomic level. Three such studies are discussed in this section, with a focus on four key elements: 1) how data, particularly poverty data, is collected; 2) the estimation techniques used to analyze the determinants of poverty; 3) which instruments, if any, are used to account for the endogeneity of remittances; and 4) what these studies found the impact of remittances on poverty to be.

Adams and Page (2005) is probably the study most similar to this one. Remittance data for this study come from the International Monetary Fund's *Balance of Payments Statistics Yearbook*s. Poverty and inequality data were collected from various household data sets, and the collection of this data is itself a significant contribution to existing literature. Their data allow for observations of 71 low- and middle-income countries, sometimes with multiple observations (at different years) for the same country. They use ordinary least squares analysis to estimate poverty as a function of GDP, Gini coefficient, international remittances[20], and in some regressions also include dummy variables indicating region of the world. They also allow for the endogeneity of remittance income by instrumenting remittance income as a function of distance of the recipient country from the United States, Europe, and the Persian Gulf (important world remittance sources), the percent of the population over age 25 with a secondary education, a measure of government stability, and explanatory variables from the primary regression. They consistently find that remittances significantly (at the 10% level) and negatively affect the rate of poverty in a nation, and find that a 10% per capita increase in per capita official international remittances will lead on average to 3.5% decline in the share of people below the poverty line. They also find that remittance income significantly and negatively affects the poverty gap (poverty depth) and squared poverty gap (poverty severity).

Acosta *et al.* (2008) look at the link between remittances and poverty and inequality in Latin America. Rather than directly use poverty as a dependent variable, they examine the impact of remittances on GDP growth and on inequality, and construct the change in poverty from these results based on the Foster, Greer, and Thorbecke family of poverty measures (Foster *et al.* 1984). Acosta *et al.* gather data for 59 industrial and developing countries over the period 1970-2000. Remittance data is from

---

[20] In separate analyses also presented in Adams and Page (2005), Adams and Page substitute migration for remittances as an explanatory variable. Here are addressed only their conclusions regarding remittances, not migration, since they relate most directly to this study.

World Economic Outlook (2005), which is again primarily drawn from IMF's *Balance of Payment Statistics Yearbook*s. Both GDP growth and inequality are modeled as functions of education levels, capital prices, and remittances. Unlike Adams and Page (2005), Acosta *et al.* take advantage of the panel nature of their data and use Arellano and Bond's general method of moments estimator. This technique allows them to account for the endogeneity of remittance income in a sophisticated manner: they incorporate both lagged levels of explanatory and dependent variables and other exogenous variables as instruments for remittances. These exogenous variables are the per capita GDP of the top 10 migrant-receiving countries in the world weighted by the inverse of the distance from these countries to the remittance-receiving countries in their sample and the real GDP per capita of the five OECD countries that are the top recipients of migrants for each remittance-receiving country, weighted by the share of migration from the corresponding country to each of these five OECD countries.

Acosta *et al.* find that remittances have a positive impact on both GDP growth and inequality for the non-Latin American countries in their sample, which should have contrary impacts on poverty. For Latin American countries, they find that remittances have a positive impact on GDP growth and a negative impact on inequality, which should reduce poverty. They calculate the poverty elasticity with respect to remittances and find the impact on poverty to be negative (i.e. increased remittances reduce poverty) over a range of reasonable values of the Gini coefficient and a supposed poverty level. Unfortunately, they do not calculate standard errors for these elasticities and so it can be stated whether or not this impact is statistically different from zero[21]. It is also worth pointing out that according to their definitions, the impact of remittance income on poverty is effected only through its impact on GDP growth and inequality levels; the effect of using remittances to directly finance consumption is not considered. Acosta *et al.* do not examine the impact of remittances on poverty depth or severity in the macroeconomic portion of their study.[22]

To the author's knowledge there is only one other serious published macroeconomic study of the impact of remittances on poverty, in World Economic Outlook (2005). They use a sample of 90 developing and developed nations from 1970 to 2003 and regress poverty rate against the log of the remittance/GDP ratio, log average

---

[21] Acosta *et al.* also include a nuanced microeconomic analysis of remittances and poverty in 10 Latin American countries, including a counter-factual study where they compare the impact of remittance income with the impact of income that would be earned had the migrant not yet migrated. They do provide confidence intervals for the elasticity of poverty with respect to remittance income in this analysis, and determine that the impact is negative and significant. Further discussion of these results would be a digression from this review of macroeconomic analyses of the effect of remittances on poverty.

[22] They do examine poverty depth and severity in their microeconomic analyses, and in most cases find that remittances negatively impact both measures of poverty, though they do not report the statistical significance of this impact.

income, and the Gini coefficient[23]. Remittance data is from the IMF's *Balance of Payments Yearbook*. Remittances are endogenized as a function of the distance between the remittance receiving country and the country which contains the largest share of its migrants and a dummy variable indicating whether or not these two nations share a common language. They find a negative, statistically significant relationship between remittances and poverty, though the impact is small: a 2.5 percent increase in the remittances/GDP ratio results in less than .5% reduction in poverty rate.

Reliable panel poverty data are often difficult to obtain: Adams and Page (2005) construct their own data set from various sources, and Acosta *et al.* (2007) avoid the problem by analyzing remittances' effects on GDP growth and inequality and extrapolating poverty impacts from these results. This study uses poverty data from a World Bank study of poverty in Eastern Europe and the former Soviet Union (Alam *et al.* 2005). Using this data set presents several advantages. The World Bank and the International Bank for Reconstruction and Development have taken considerable effort to standardize this poverty data, so that one can have more confidence in the comparability of data points than one would in data pulled from heterogeneous sources. Moreover, the data relates to countries which have, to a large extent, a shared historical and cultural background, at least for the latter half of the twentieth century. This significantly reduces the confounding effects of country-specific differences for which it is difficult to fully

[23] Unfortunately, the work is not targeted to empirical economists, and the description of data sources and methods used is not sufficient to present a much more detailed analysis of the paper than what is offered here.

account in studies which pull data from around the world, especially if one allows for the possibility that these differences manifest themselves other than through fixed or random effects.

The analysis presented here contributes to existing literature in several other ways. This chapter presents an analysis of a subject which few papers have treated, using data and techniques which have not been used in other published studies of this topic. Unlike Adams and Page (2005), who used ordinary least squares analysis[24], this study takes advantage of the panel nature of the data and uses standard fixed- and random-effects analyses, which are described in more detail in the following section.[25] Like Adams and Page (2005), but unlike Acosta *et al.* (2007), this study analyzes the rate, depth, and severity of poverty. This study also presents an analysis of Eastern Europe and the former Soviet Union; it is not known *a priori* that what holds true in one part of the world necessarily holds true in another. In fact, this study concludes that remittance income does not significantly affect poverty in the region, which is different from the results of the other studies discussed here.

---

[24] This is not meant as a criticism: since for most countries in their data set, data was available for only one year, ordinary least squares may have been the most appropriate analysis method.

[25] Acosta *et al.* use a GMM estimator, as discussed above. This study does not, primarily because the technique requires observations to occur in consecutive years, which resulted in a substantial number (>50%) of the already somewhat small number of observations to be dropped.

IV. Data Sources and Variables Analyzed

Poverty data for this study comes from Alam *et al.* (2005)[26], who collate the data from the World Bank's Europe and Central Asia (ECA) household survey archive. The poverty line used is $2.15 US of consumption per capita per day, adjusted for inflation (to account for temporal variation in prices) and purchasing power parity (to account for spatial variation in prices). This figure is more appropriate than the $1 of consumption a day often used in poverty analyses, since while the region of analysis is in the developing world, it does have a significantly higher standard of living than many low-income countries. This figure does not include durables or health services. Goods consumption achieved through non-market means, such as home-produced food, is included in this measure. Also presented are results using a poverty line of $4.30 US of consumption per capita per day. According to Alam *et al.* (2005), persons with consumption under this level, while not necessarily impoverished, are vulnerable to poverty if adverse economic shocks occur. Alam *et al.* (2005) includes measures of poverty rate, poverty depth, and

---

[26] Alam *et al.* (2005) also includes poverty data for Columbia, Turkey, and Vietnam, in order to compare Eastern European and former Soviet Union poverty levels with poverty levels in developing nations outside this region. These countries are not included in analyses presented here. Russia is also excluded, because it is unique in several ways from the other countries: its economy is much larger than the others', and it is both a significant source and destination of remittances.

poverty severity. Inequality data (Gini coefficients) is also taken from Alam *et al.* (2005).

Remittance data are taken from the World Bank's World Development Indicators (World Bank); this data is taken from the International Monetary Fund's *Balance of Payments Statistics Yearbook*s, the same source used by other, similar studies cited above. This data is not perfect; it includes only official remittance figures reported by national central banks[27], and thus likely underrepresents the actual flow of remittances into the countries of interests. Unfortunately, no other source of macroeconomic remittance data is available for the region, and one can do little but what others have done, which is to acknowledge this shortcoming. To improve records of remittance transfers is a goal, if not always a priority, of governments and development organizations in the region. Per capita GDP data (in year 2005 PPP adjusted dollars) is also taken from the World Development Indicators, and is again based on the IMF's *Balance of Payments Statistic*s yearbooks.

The data used is appealing from an econometric point of view for several reasons. The countries analyzed share, to some extent, a common historical and cultural background, which mitigates the distorting effects of country-specific differences in the underlying regressed equations (though our models do allow for fixed and random effects). However, the data is rich in that the countries studied have widely varying poverty levels, levels of inequality, remittances, and GDPs. Table 11 contains summary

[27] This is *not* quite the same as reporting only remittances sent through official channels, though other studies have imprecisely interpreted the IMF data as such.

statistics for the variables used in this chapter. Table 12 lists values of these variables for the countries analyzed, averaged over the appropriate time frame.

Table 11 Summary Statistics (Poverty Analysis)

Variable	Mean (Standard error)
Per Capita GDP (2005 US dollars)	6786.478 (487.756)
Per Capita Remittances (2005 US dollars)	173.445 (33.081)
Gini Coefficient	.310 (.005)
Poverty Rate (poverty line = $2.15 US/day)	.225 (.032)
Poverty Depth (poverty line = $2.15 US/day)	.079 (.013)
Poverty Severity (poverty line = $2.15 US/day)	.087 (.022)
Poverty Rate (poverty line = $4.30 US/day)	.53 (.036)
Poverty Depth (poverty line = $4.30 US/day)	.23 (.024)
Poverty Severity (poverty line = $4.30 US/day)	.19 (.032)

*Author's calculations based on Alam *et al.* (2005) and World Bank (2008). Because different countries may provide different numbers (years) of observations, some countries are weighted more than others in mean and standard error calculations. Per capita GDP and per capita remittances are adjusted for inflation and for purchasing power parity.

Table 12 Country Level Data (Poverty Analysis)									
Country (Years)	*Per capita GDP*	*Per capita remittances*	*Gini coefficient*	*Poverty rate* ($2.15/ day)	*Poverty depth* ($2.15/ day)	*Poverty severity* ($2.15/ day)	*Poverty rate* ($4.30/ day)	*Poverty depth* ($4.30/ day)	*Poverty severity* ($4.30/ day)
Albania (2002)	4711	777	.32	.24	.05	.02	.71	.28	.14
Armenia (1999-2003)	2707	295	.31	.56	.17	.07	.92	.48	.29
Azerbaijan (2002-2003)	3195	84	.18	.05	.01	.00	.72	.18	.07
Belarus (1998-2002)	5822	80	.30	.06	.01	.00	.35	.10	.04
Boznia and Herzegovina (2001, 2004)	5322	1255	.28	.05	.01	.00	.38	.10	.04
Bulgaria (1995, 2001, 2003)	7470	369	.31	.06	.02	.07	.30	.09	.04
Estonia (2000-2003)	12275	32.36	.33	.04	.01	.00	.27	.08	.03
Georgia (1997-2003)	2428	199	.39	.49	.20	.11	.84	.45	.29
Hungary (1998-2002)	13584	68	.25	.00	.00	.00	.16	.03	.01
Kazakhstan (2001-2003)	6748	47	.33	.26	.07	.26	.70	.29	.70
Kyrgyzstan (2000-2003)	1569	31	.29	.74	.28	.74	.97	.59	.97
Latvia (2002-2003)	10485	157	.35	.03	.01	0	.18	.05	.02

Lithuania (1998-2003)	10081	45	.31	.04	.01	.00	.27	.07	.03
Macedonia (2002-2003)	6763	224	.37	.04	.01	.00	.24	.07	.03
Moldova (1998-2003)	1571	249	.35	.65	.27	.15	.92	.55	.37
Poland (1998-2002)	11349	94	.31	.02	.00	.02	.25	.06	.02
Romania (1998-2003)	7291	18	.28	.16	.04	.04	.65	.23	.11
Serbia and Montenegro (2002)	7265	959	.29	.06	.01	.01	.42	.12	.05
Ukraine (2002-2003)	4551	28	.27	.02	.00	.00	.27	.07	.03

*Author's calculations based on Alam *et al.* (2005) and World Bank (2008). Per capita GDP and per capita remittances are adjusted for inflation and for purchasing power parity. Quantities are expressed in 2005 U.S. dollars (where applicable).

V. Econometric Methods

The basic model assumed is a variant of the poverty growth model proposed by Ravillion (1997) (Adams and Page [2005] assume a nearly identical model). Rather than use ordinary least squares analysis, I suppose that poverty levels in each nation are affected by country-specific fixed effects which manifest themselves as constant differences in poverty between countries. The fixed effects model for country *i* at time *t* is

$$\ln(poverty)_{it} = \alpha_i + \beta_1 \ln(GDP)_{\mathrm{it}} + \beta_2 \ln(gini)_{\mathrm{it}}$$
$$+\beta_3 \ln(remittance)_{\mathrm{it}} + \varepsilon_{it} \qquad (21)$$

where *poverty* is the poverty measure in question, *GDP* is per capita GDP, *gini* is the Gini coefficient, *remittance* is per capita remittance income, and $\varepsilon_{it} \sim \mathrm{N}(0, \sigma_\varepsilon{}^2)$. The term α_i can be considered unobserved temporally constant spatial heterogeneity.

The fixed effects model is perhaps the most commonly used and intuitively appealing model for analyzing panel data sets such as ours. For completeness, though, and to test the robustness of results, also presented are results from ordinary least squares analysis:

$$\ln(poverty)_{it} = \alpha + \beta_1 \ln(GDP)_{\mathrm{it}} + \beta_2 \ln(gini)_{\mathrm{it}}$$
$$+\beta_3 \ln(remittance)_{\mathrm{it}} + \varepsilon_{it} \qquad (22)$$

Ordinary least squares estimation ignores the panel nature of the data and these results are presented primarily for comparison with other studies, such as Adams and Page (2005).

Also presented are results under the assumption of a random effects model. The random effects model further assumes that the covariance between country-specific effects and explanatory variables is zero, such that

$$\ln(poverty)_{it} = \alpha + \beta_1 \ln(GDP) + \beta_2 \ln(gini))$$
$$+\beta_3 \ln(remittance) + v_i + \varepsilon_{it} \quad (23)$$

where $v_i \sim N(0, \sigma_v^2)$. Specification tests (the artificial regression approach described by Arrelano 1983) indicate that the more parsimonious fixed effects model is preferred to the random effects model for our specifications; however, results assuming random effects relationships are presented as well for completeness.

I also corrected for the possible endogeneity of remittances by instrumenting log remittance income $\ln(remittance)$ with instruments $\ln(GDP)$ and $\ln(gini)$. It was also necessary to include instruments which are plausible explainers of remittance level and do not appear in the primary regression equations. These instruments are the one- and two-year lagged values of remittance receipts and a proxy for the GDP of developing countries, i.e., the logged GDP of the United States, Germany, and Russia (perhaps the three most important sources of remittances for the region) weighted by inverse squared distance of these countries to the remittance receiving countries. These instruments are similar to those used in other macroeconomic studies of remittances.[28] Regression of

[28] Alternative regressions were done including a measure of secondary school participation, as in Adams and Page (2005), and a measure of trade openness (the sum of imports and exports as a proportion of GDP). While including this additional instrument did not significantly affect our results, results from these regressions are not presented

Equations (21), (22), and (23) was done with standard two-stage fixed effects, two-stage instrumental variables, and two-stage random effects estimators. The statistical analysis program Stata was used for these estimations.

In order to define observations of value zero, .01 was added to poverty and inequality measures and 1 to other quantities before natural logs were taken. Omission of this step does not significantly change estimation results, though it does significantly reduce viable sample size.

VI. Estimation Results

Tables 13 and 14 contain estimation results for different econometric models and different measures of poverty. The R^2 values indicate that the model used is a good predictor of poverty rate and poverty depth, though it is a much less strong predictor for poverty severity.

---

here, since inclusion of this variable necessitated dropping 16 observations from the sample.

Table 13 Estimates of the Effects of Explanatory Variables on Poverty with Poverty Line at $2.15 US

Variable	*Poverty rate*			*Poverty depth*			*Poverty severity*		
	OLS	Fixed Effects	Random Effects	OLS	Fixed Effects	Random Effects	OLS	Fixed Effects	Random Effects
Per Capita	-.230***	-.186***	-.219***	-.107***	-.115***	-.108***	-.117***	-.103***	-.109***
GDP	(.022)	(.049)	(.027)	(.011)	(.041)	(.012)	(.031)	(.032)	(.027)
Per Capita	-.005	.008	.012*	.0006	.002	.0005	.004	.004	.004
Remittances	(.044)	(.007)	(.007)	(.021)	(.006)	(.006)	(.062)	(.004)	(.005)
Gini	1.275***	1.653***	1.633***	.656***	1.231***	.900***	-.186	1.097***	1.016***
Coefficient	(.392)	(.423)	(.350)	(.189)	(.357)	(.202)	(.552)	(.276)	(.259)
Intercept	1.840***	1.303***	1.553***	.810***	.721**	.745***	1.11**	.646**	.708***
	(.318)	(.422)	(.246)	(.153)	(.357)	(.123)	(.447)	(.276)	(.230)
R^2	.847	.817	.817	.842	.821	.838	.364	.292	.302
n	67	67	67	67	67	67	67	67	67

*All variables (except the intercept) are expressed as logged quantities. Standard errors are in parentheses. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. Log per capita remittances is instrumented as described in section IV.

Table 14 Estimates of the Effects of Explanatory Variables on Poverty with Poverty Line at $4.30 US

Variable	*Poverty rate*			*Poverty depth*			*Poverty severity*		
	OLS	Fixed Effects	Random Effects	OLS	Fixed Effects	Random Effects	OLS	Fixed Effects	Random Effects
Per Capita GDP	-.229*** (.026)	-.182*** (.054)	-.243*** (.032)	-.184*** (.017)	-.151*** (.038)	-.184*** (.021)	-.175*** (.040)	-.127*** (.037)	-.150*** (.032)
Per Capita Remittances	.003 (.053)	.007 (.008)	.023*** (.008)	.0003 (.035)	.006 (.005)	.013** (.006)	.018 (.081)	.005 (.005)	.008 (.006)
Gini Coefficient	.067 (.469)	-.021 (.469)	.168 (.385)	.705** (.306)	.987*** (.327)	.993*** (.273)	.307 (.714)	.994*** (.322)	.306*** (1.013)
Intercept	2.345*** (.380)	1.942*** (.468)	2.331*** (.282)	1.586*** (.248)	1.202*** (.326)	1.438*** (.193)	1.501** (.579)	.955*** (.322)	1.127*** (.279)
R^2	.768	.761	.736	.847	.825	.816	.450	.478	.474
n	67	67	67	67	67	67	67	67	67

*All variables (except the intercept) are expressed as logged quantities. Standard errors are in parentheses. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. Log per capita remittances is instrumented as described in section IV.

The results shown in Table 13, where the poverty line is $2.15 US, are remarkably consistent: with only one exception, the effect of remittances on poverty is insignificant at the 10% level. The one exception is a random effects estimation of poverty depth, where remittances are actually found to increase poverty. Similarly, when the poverty line is taken to be $4.30 US (Table 14), remittances have no statistically significant effect on poverty in most regressions. In those cases when the effect is significant (random effects estimations of poverty rate and poverty depth), the effect is again positive, the opposite of what one might expect. Particularly since specification tests indicate that fixed effects estimations are to be preferred to random effects estimations, the balance of information strongly indicates that remittances have no statistically significant effects on poverty. Certainly, one cannot argue that remittances reduce poverty based on these results.

Alternative regressions were also done including additional instruments for remittance income (a measure of secondary school participation and a measure of trade openness, i.e. the sum of imports and exports as a fraction of GDP), using poverty rate rather than log poverty rate as the dependent variable, and with data averaged over time to yield a single observation for each country. The result that remittances do not significantly reduce poverty is robust to these alternative specifications. Results of these specifications are available upon request.

VII. Conclusions

This study finds that remittance income has no significant impact on poverty levels in Eastern Europe and the former Soviet Union. This is contrary to the results found in the studies cited above (Adams and Page 2005, Acosta *et al.* 2008, World Economic Outlook 2005) and, perhaps, to intuition; one might expect income from any source to reduce poverty. Nonetheless, the underlying relationship between remittances and poverty well may be different for the region analyzed here and for the other parts of the world. Furthermore, the data and methods presented here are, in some ways, an improvement over those used in other studies.

The most obvious explanation for the small impact of remittances on poverty is that remittances are not reaching the impoverished in Eastern Europe and the former Soviet Union. It may be that in order to finance a migrant, a family must already have achieved a certain level of consumption; and thus few remittance-senders come from impoverished homes. Even when remittances do reach the impoverished, they may not have lasting or significant impacts on consumption, for those reasons discussed in the introduction.

Regardless, the conclusion remains the same: ceteris paribus, increased remittance income does not lead to lower poverty levels in Eastern Europe and the former Soviet Union. Policy makers may wish to examine ways to increase remittance flow to the poorest in the region, such as through work training programs to increase the capacity of poor households to produce remittance senders, or through international banking education programs to allow poorer households to transfer and invest remittances in a more informed manner. While these changes could improve the impact of remittance income, this study also suggests that focusing on international remittances to reduce poverty in the region could yield few results in the short term, and it may be best for government and development agencies to focus efforts elsewhere.

The World Bank projects that the current economic downturn will result in a marked reduction of the growth rate of remittances. The estimated growth rate of remittances to Europe and Central Asia was 5% in 2008, compared to 31% in 2007 and 26% in 2006 (Ratha *et al.* 2008). In 2009, the growth rate of remittance income for the region is expected to be -.1% (ibid.) (compared to -.9% for all developing countries). This reduction is primarily due to reduced wages and employment opportunities in remittance source countries. However, this study indicates that the direct effects of this reduction in remittance growth on the poor should be minimal. That remittances have a small impact on poverty is good in that it indicates that the impoverished in the region analyzed here are not susceptible to remittance shocks.

Chapter 5: Conclusion

I. Summary of Dissertation

I have presented three studies in which I analyze the impacts of remittance income from international sources. The first two analyses are on the sub-national level and use data from the 2003/2004 Nepal Living Standards Survey. Of these, the first is an analysis of how remittance income and non-remittance income affect child education and child labor in Nepal. The second is a semiparametric Engel curve analysis of the effects of remittance income and non-remittance income on consumption. The third analysis presented here is a cross-national study of how GDP, inequality, and remittances affect poverty levels in Eastern Europe and the former Soviet Union.

In each study, the conclusions are similar: that the welfare-enhancing effects of remittance income are small or nonexistent. In the first study, I find that while both non-remittance and remittance income positively contribute to child welfare, the effects of remittance income are much smaller than the effects of non-remittance income. In the second study, I find that though both remittance and non-remittance income increase consumption, the elasticities of consumption from remittance income are much smaller than those from non-remittance income. In the final study, which is of poverty in Eastern Europe and the former Soviet Union, I find no evidence that remittance income reduces poverty levels in that region.

One reason that remittances have a small impact on welfare may be that they are not leveraged to increase permanent income. In the areas studied in this dissertation, remittances are often sent through informal channels and many individuals are unfamiliar with the banking system. This makes it more difficult to use remittance income for productive investments. For many households, then, remittances represent only a temporary increase in income. According to Milton Friedman's permanent income hypothesis, remittances would then have only small consumption effects on recipient households. If more remittances could be brought into the formal financial sector, or if some other way were found to make it easier to channel remittance income towards productive investments, their effects on consumption could increase. There is also reason to believe that the debt burden taken on by households supporting a migrant worker may mitigate the otherwise beneficial effects of remittances, particularly in Nepal.

The economics literature on the effects of remittance income on child welfare or on consumption is somewhat sparse, and published results are not uniform. As such, it cannot definitively be said whether the microeconomic analyses in this dissertation agree with existing literature: results are consistent with some studies and inconsistent with others. There are only a few other studies of the effects of remittances on poverty, but the results of these studies are consistent: that remittances significantly reduce poverty in recipient countries. The conclusions of Chapter 4 of this dissertation are different: remittances do not reduce poverty in Eastern Europe and the former Soviet Union. Since I analyze a different part of the world than that analyzed in previous studies, this difference is not difficult to resolve: remittances may indeed reduce poverty in other parts of the world than that studied here. In particular, it may be the case that in parts of the world where remittances are more likely to enter the formal financial sector, they are more likely to have significant effects on consumption by the poor and therefore poverty.

II. Opportunities for Future Research

Several opportunities for future research to expand upon and clarify the results in this dissertation seem apparent. In Nepal, it would be interesting to further explore how and to what extent financing a migrant imposes a debt burden, and if households fully understand this debt burden when they assume it. This could be explored by surveys in which respondents are asked these questions or by an analysis of the debt burden of recipient and non-recipient households. Along the same lines, one could examine if and how migrants' remittances change with time; it is possible that receipts could decrease as time away from Nepal increases (or perhaps increase as the time to return to Nepal draws near), which may affect the net benefit of a household's financing of a migrant.

In Eastern Europe, microeconomic analyses might explain why remittances do not reduce poverty in these countries. That remittances do not reduce poverty implies that they do not increase consumption significantly for poor households. Engel curve analyses similar to Chapter 3 of this dissertation would reveal how remittances affect consumption for rich and poor households, and could shed further light on the results presented in this dissertation. For many countries in the region, World Bank Living Standards Surveys (analogs of the Nepal data used in chapters 2 and 3) are available.

III. Remittances and the Current Economic Downturn

Currently, the world is experiencing a significant economic downturn, and it is not known how long this period of depressed GDP and higher unemployment will last or how severe it will be. Nonetheless, some of the downturn's effects on remittances can, perhaps, be anticipated. The downturn is likely to depress both wages and employment opportunities for remittance senders; this affect will depress remittances along with world GDP. However, remittances may be counter-cyclical with respect to recipient country GDP (World Bank 2006; Ratha *et al.* 2008). When conditions in recipient countries worsen, more migrants may leave to try to find better opportunities elsewhere. Existing migrants may be inclined to send more when they know that the quality of life of recipient households has decreased. These factors tend to *increase* remittance receipts in the face of a downturn, making remittances an important consumption-smoothing source of income in adverse economic conditions. The effects of a worldwide economic downturn on international remittances are thus theoretically ambiguous.

Though the effects that this downturn will have on remittances in the long term cannot be predicted with certainty, some of the short- and medium-term effects are already manifesting. According to a World Bank report (Ratha *et al.* 2008), international remittances flow growth decreased sharply in the third quarter of 2008. International remittances to developing countries grew 6.7% in 2008, compared to 16% in 2007 (ibid.). The sharpest deceleration occurred in Europe and Central Asia (the focus of Chapter 4 of this dissertation), where growth was 31% in 2007 and 5% in 2008. In contrast, remittance growth remains strong in South Asia (the focus of Chapters 2 and 3 of this dissertation): remittance growth increased from 11% in 2007 to 16% in 2008 (ibid.)

The World Bank has projected remittance growth rates to the developing world for 2009 and 2010 (ibid.). They predict that while remittances to the developing world will be -.9% in 2009, it will quickly rebound to 6.1% in 2010; Europe and Central Asia and South Asia are also expected to face insignificant or negative remittance receipt growth rates in 2009 followed by a growth rate above 2008 levels in 2010. Remittances are thus expected to remain resilient in the medium term. Furthermore, since the economic downturn is anticipated to be especially sharp in the United States and Europe, recipient countries which primarily rely on other sources of remittances should be less affected; this applies to Nepal, which receives most remittances from India and the Middle East, and Central Asia and parts of Eastern Europe, for which Russia is the most important remittance source country.

In summary, remittance growth is anticipated to slow sharply in 2009 for the world and areas analyzed in this dissertation. However, remittances are in general a relatively non-volatile foreign source of income. In the medium-term, remittance growth rates are expected to be stable and return to approximately 2008 levels.

IV. Remittance Transfers and Migration in the Regions of Interest

A. Nepal

Migration from Nepal and the transfer of remittances to Nepal often occur in conditions that are in many ways suboptimal. Most migration from Nepal is to bordering India, and migration between these countries is generally not formally regulated or documented (Thieme and Wyss 2005). Migration to India takes considerably less capital than to other destinations and jobs are generally low-paying (Thieme and Wyss 2005). However, migrants must still often buy these jobs from a predecessor or inherit them from family members, which can contribute to significant migration costs (Thieme 2003). Seed capital estimates for migration to India range from about $10 US (Thieme 2003) to about $208 US (Thieme and Wyss 2005). Nearly all Nepalese migrants to India hand carry their remittances or send them with a friend or relative; use of the formal banking sector is virtually nil (Thieme and Wyss 2005).

Migrants going to areas outside of India face much higher capital costs. Those wishing to work in the Middle East almost always work through manpower agencies who charge from about $500 to $1000 to arrange a two to three year working contract for a generally menial job (Graner and Gurung 2003). Remittances from the Middle East are generally hand-carried but are sometimes sent through informal money transfer networks known as hundis or hawalas (Thieme 2003). There are few ways to transfer remittances through the formal sector.

Migrants from Nepal do go to other areas, such as Western Europe, Hong Kong, and Japan. Jobs in these locations are generally more coveted than in, for example, India and the Middle East, and have higher startup costs (Graner and Gurung 2003). Still, because of migrants' and recipients' limited knowledge of banks and lengthy transfer times (indicating the underdevelopment of Nepal's international banking infrastructure), migrants to Western Europe are the only group of which a significant number transfer remittances via the banking sector (Thieme and Wyss 2005). Money transfer operators have considerably shorter transfer times and are low cost, but are still not used by most migrants (Ferrari *et al.* 2007).

Because few remittances enter the formal sector, families have limited options when attempting to put remittances towards productive investments. A large share of remittances goes towards paying loans incurred during the migration process (Ferrari *et al.* 2007). A large share also goes to a productive investment with which the recipient family has experience: the financing of another migrant (Thieme and Wyss 2005). This is often perceived by households to be the best investment to which they have access.

B. Eastern Europe and the Former Soviet Union

It is, of course, harder to make general statements for so large an area as Eastern Europe and the former Soviet Union than it is for Nepal. However, some statements seem to be true for most of the region. Like in Nepal, remittances in Eastern Europe and the former Soviet Union are often sent through informal channels; Chapter 4 of this dissertation presents evidence that this is the case.

V. Initiatives to Bring Remittances to the Formal Sector

In this dissertation it has been argued that increasing the amount of remittances which flow in the formal sector would amplify their welfare-enhancing effects. Recently, several governmental, non-governmental and private organizations have explored ways to channel more remittances to the formal sector. Even if these initiatives are undertaken by for-profit organizations, if these programs help migrants and recipients to transform remittances from a temporary to a permanent income stream, they should be welfare-enhancing.

Several banks have developed programs to bring remittances into the formal banking sector and to help migrants and recipients to invest and productively use remittance income. Banco Solidario in Ecuador has targeted senders and recipients of remittances from Spain to Ecuador (CARANA Corporation 2004; Orozco 2004). They allow senders to control how much recipients can withdraw at a given time. They use remittance transfer services to draw senders and recipients to other services such as insurance, credit lines, home loans, and savings accounts. They use a variety of media to market specifically to the remittance market and make transfers easier by allowing clients to make transactions through a point of sales network.

Banco Salvadoreño, El Salvador has also developed programs to target the remittance market (Orozco 2004). They advertise through a variety of media and have collaborative agreements with money transfer organizations in the U.S. to increase their presence and market services in the U.S. They also allow recipients to use the future stream of remittances as collateral and to take out loans of up to 80% of the amount of remittances received in the past six months.

Wells Fargo has a number of initiatives to take advantages of the large flow of remittances from Mexico to the United States (FDIC 2004; Orozco 2007). They engage in financial literacy campaigns and work with a non-governmental organization to provide access to literature explaining basic financial tasks to remittance recipients. They were also the first U.S. bank to accept Mexico's Matricula Consular as valid identification for banking, which helped them to tap into the large pool of undocumented migrants in the U.S.

Companies outside of the banking sector are also working to capitalize on and help consumers to productively use remittance flows. Construmex is a Mexican home-building company which partners with U.S. money transfer organizations to allow remittance senders to automatically make loan payments for home improvement or home construction from their remittances. Construmex's large physical presence in Mexico allows for the construction of homes throughout most of the country. An initiative such as this seems particularly relevant for areas with underdeveloped mortgage markets, which includes the areas analyzed in this dissertation.

Credit unions (CUs) and microfinance institutions (MFIs) could also help to bring more remittances into the formal sector, particularly for demographics typically underserved by traditional financial institutions. Credit unions and microfinance institutions often target the poor and those in rural communities. However, at least in the developing world, most do not have the capacity to receive remittances. There are several reasons why this is the case (Hamilton and Orozco 2006): there are high costs associated with complying with international money transfer regulations; remittance transfers involve more complicated liquidity and risk management than usual MFI and CU activities; MFI and CU staff lack training to deal with remittances; and a rather large critical mass of transfers is necessary for remittance transfers to be financially feasible.

Several initiatives exist to enable CUs and MFIs to process remittances. Organizations behind these initiatives generally provide technical assistance, help MFIs and CUs to negotiate with money transfer operators to minimize costs, and/or help MFIs and CUs to market other financial services to remittance customers. The World Council of Credit Unions has worked to integrate credit unions in Latin America and the United States into IRnet, its remittance transfer network (Orozco 2004). The Multilateral Investment Fund has a number of projects to help MFIs in Latin America to process remittances and cross-sell related services (IADB 2009).

APPENDICES

Appendix A: Identification of Child Welfare Equations

Appendix A provides a proof that systems of Equations (6) and (7) in Chapter 2 of this dissertation are identified under the assumptions given in that chapter. Note that notation used here differs from that used in the main body of the chapter.

Let y_1* be the probability of a child having schooling or not (or, the probability of a child laboring or not); y_2 be a child's scholastic attainment, given that the child has some schooling (or, the amount a child labors, given s/he does so); y_3* be the intensity of the amount of remittance income received; and y_4 be amount of non-remittance income received. Consider a four-equation model:

$$y_1{}^* + a_{1,3}y_3 + a_{1,4}y_4 = b_{1,1}x_1 + \beta_1\mathbf{z}_1 + \alpha_1\mathbf{z}_2 + u_1 \quad (24a)$$

$$y_2 + a_{2,3}y_3 + a_{2,4}y_4 = b_{2,2}x_2 + \beta_2\mathbf{z}_1 + \alpha_2\mathbf{z}_2 + u_2 \quad (24b)$$

$$y_3{}^* = b_{3,1}x_1 + b_{3,2}x_2 + \beta_3\mathbf{z}_1 + \gamma_1\mathbf{z}_3 + u_3 \quad (24c)$$

$$y_3 = b_{4,1}x_1 + b_{4,2}x_2 + \beta_4\mathbf{z}_1 + \gamma_4\mathbf{z}_3 + u_4 \quad (24d)$$

where $\boldsymbol{x}_1$ represents a column vector of independent household-level variables (which does not include independent household-level variables x_1, x_2, or $\mathbf{z}_3$ but does include an intercept term, i.e. unity), $\mathbf{z}_2$ a column vector of different, individual-level variables, $\mathbf{z}_3$ a column vector of variables different from those in $\mathbf{z}_1$ or $\mathbf{z}_2$, α_i, β_i, and γ_i the corresponding coefficient row vectors for the *i*th equation, *a* and *b* terms are scalar coefficients, and *u* terms are residuals. Independent variables x_2 (*HEADMIGRATED*) and $\mathbf{z}_3$ (*FINANCIAL* and *LOANS*) are excluded from the first equation, and x_1 (*HEADFEMALE*) and $\mathbf{z}_3$ from the second equation. Individual-level variables are excluded from the third and fourth equations, since these regressions are done at the household level. Rearranging these equations and writing them in matrix form yields:

$$\begin{bmatrix} 1 & 0 & a_{1,3} & a_{1,4} & -b_{1,1} & 0 & -\beta_1 & -\alpha_1 & 0 \\ 0 & 1 & a_{2,3} & a_{2,4} & 0 & -b_{2,4} & -\beta_2 & -\alpha_2 & 0 \\ 0 & 0 & 1 & 0 & -b_{3,1} & -b_{3,2} & -\beta_3 & 0 & -\gamma_3 \\ 0 & 0 & 0 & 1 & -b_{4,1} & -b_{4,2} & -\beta_4 & 0 & -\gamma_4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ x_1 \\ x_2 \\ \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \boldsymbol{z}_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \tag{25a}$$

or

$$\boldsymbol{Aw} = \boldsymbol{u} \tag{25b}$$

where

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & a_{1,3} & a_{1,4} & -b_{1,1} & 0 & -\beta_1 & -\alpha_1 & 0 \\ 0 & 1 & a_{2,3} & a_{2,4} & 0 & -b_{2,4} & -\beta_2 & -\alpha_2 & 0 \\ 0 & 0 & 1 & 0 & -b_{3,1} & -b_{3,2} & -\beta_3 & 0 & -\gamma_3 \\ 0 & 0 & 0 & 1 & -b_{4,1} & -b_{4,2} & -\beta_4 & 0 & -\gamma_4 \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ x_1 \\ x_2 \\ \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \boldsymbol{z}_3 \end{bmatrix},$$

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \tag{22c}$$

Since I assume that only the first and second equations have correlated residuals, let the covariance matrix $\boldsymbol{\Sigma} = E(\boldsymbol{uu}')$ be of the form

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & 0 & \sigma_{14} \\ \sigma_{12} & \sigma_{22} & 0 & \sigma_{24} \\ 0 & 0 & \sigma_{33} & \sigma_{34} \\ 0 & 0 & 0 & \sigma_{44} \end{bmatrix} \tag{26}$$

Define

$$\mathbf{F}=\begin{bmatrix} f_{11} & f_{12} & f_{13} & f_{14} \\ f_{21} & f_{22} & f_{23} & f_{24} \\ f_{31} & f_{32} & f_{33} & f_{34} \\ f_{41} & f_{42} & f_{43} & f_{44} \end{bmatrix} \tag{27}$$

as an admissible transformation matrix if and only if the matrix $\mathbf{F}\boldsymbol{A}$ satisfies all a priori restrictions on $\boldsymbol{A}$, and $\mathbf{F}\boldsymbol{\Sigma}$ satisfies all a priori restrictions on $\boldsymbol{\Sigma}$. A necessary and sufficient condition for the identification of all four equations is that all possible admissible matrices **F** be diagonal (Johnston 1984).

Now,

$$\mathbf{F}\boldsymbol{A}=\begin{bmatrix} f_{11} & f_{12} & a_{1,3}f_{11}+a_{2,3}f_{12}+f_{13} & a_{1,4}f_{11}+a_{2,4}f_{12}+f_{14} \\ f_{21} & f_{22} & a_{1,3}f_{21}+a_{2,3}f_{22}+f_{23} & a_{1,4}f_{21}+a_{2,4}f_{22}+f_{24} \\ f_{31} & f_{32} & a_{1,3}f_{31}+a_{2,3}f_{32}+f_{33} & a_{1,4}f_{31}+a_{2,4}f_{32}+f_{34} \\ f_{41} & f_{42} & a_{1,3}f_{41}+a_{2,3}f_{42}+f_{43} & a_{1,4}f_{41}+a_{2,4}f_{42}+f_{44} \end{bmatrix}$$

$$\begin{matrix} -b_{1,1}f_{11}-b_{3,1}f_{13}-b_{4,1}f_{14} & -b_{2,2}f_{12}-b_{3,2}f_{13}-b_{4,2}f_{14} \\ -b_{1,1}f_{21}-b_{3,1}f_{23}-b_{4,1}f_{24} & -b_{2,2}f_{22}-b_{3,2}f_{23}-b_{4,2}f_{24} \\ -b_{1,1}f_{31}-b_{3,1}f_{33}-b_{4,1}f_{34} & -b_{2,2}f_{32}-b_{3,2}f_{33}-b_{4,2}f_{34} \\ -b_{1,1}f_{41}-b_{3,1}f_{43}-b_{4,1}f_{44} & -b_{2,2}f_{42}-b_{3,2}f_{43}-b_{4,2}f_{44} \end{matrix}$$

$$\left[\mathbf{F}\times\begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\\ \beta_4\end{bmatrix}\mathbf{F}\times\begin{bmatrix}\alpha_1\\ \alpha_2\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix}\mathbf{F}\times\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \gamma_3\\ \gamma_4\end{bmatrix}\right] \tag{28}$$

where a bold zero (**0**) indicates a row vector of zeros of an appropriate length to define the matrix.

With minimal algebra, it is clear that for **F** to be admissible implies that $f_{12}=f_{13}=f_{14}=f_{21}=f_{23}=f_{24}=f_{31}=f_{32}=f_{34}+f_{41}=f_{42}=f_{43}=0$; that is, **F** is diagonal. This implies that

$$\mathbf{F}\boldsymbol{\Sigma}=\begin{bmatrix} f_{11}\sigma_{11} & 0 & 0 & 0 \\ 0 & f_{22}\sigma_{22} & 0 & 0 \\ 0 & 0 & f_{33}\sigma_{33} & 0 \\ 0 & 0 & 0 & f_{44}\sigma_{44} \end{bmatrix} \tag{29}$$

and so $\mathbf{F}\boldsymbol{\Sigma}$ satisfies all a priori restrictions on $\boldsymbol{\Sigma}$. Therefore, all possible admissible transformation matrixes $\mathbf{F}$ are diagonal and the system is identified.

Appendix B: Stata Programs for Child Welfare Analyses

This Appendix contains the Stata programs used for the calculations in Chapter 2 of this dissertation.

I. Collation of Head of Household Data

```
#delim=;

use C:\NLSS03\DATA\Z11A1B.dta, clear;
collapse (sum) V11A1B_09, by(WWWHH WWW HH);
gen landvalue = ln(V11A1B_09+1);
save "C:\NLSS03\paper\landvalue.dta", replace;

use C:\NLSS03\DATA\Z01A.dta, clear;
collapse (count) IDC, by(WWWHH WWW HH);
rename IDC n;
save "C:\NLSS03\paper\HouseholdCount.dta", replace;

use C:\NLSS03\DATA\Z04.dta, clear;
drop if IDC != 1;
rename V04_01 HEADMIGRATED;
replace HEADMIGRATED = 0 if HEADMIGRATED == 2;
drop V*;
save "C:\NLSS03\paper\headmigrated", replace;

#delim=;
use "C:\NLSS03\DATA\Z15.dta", clear;
replace V15_02 = 0 if V15_02==.;
replace V15_03 = 0 if V15_03==.;
replace V15_04 = 0 if V15_04==.;
drop if V15_02 == 0 & V15_03 ==0 & V15_04 == 0;
collapse (count) V15_01, by(WWWHH WWW HH);
rename V15_01 finansoph;
save "C:\NLSS03\paper\finansoph", replace;

insheet using C:\NLSS03\RawWeights.csv, clear;
rename psuno WWW;
keep WWW hhweight;
save "C:\NLSS03\paper\hhweight.dta", replace;

use C:\NLSS03\DATA\Z01A.dta, clear;
keep if  V01A_03==1;
rename  V01A_01A ethcode;
rename  V01A_02 headfemale;
replace headfemale = 0 if headfemale ==1;
replace headfemale = 1 if headfemale ==2;
rename  V01A_06 unmarried;
replace  unmarried = 0 if unmarried == 1;
replace unmarried = 1 if unmarried != 0;
rename V01A_05 headage;
```

```
drop V0*;

/*one household had 2 heads; I drop the lesser educated*/
duplicates drop WWWHH, force;

joinby WWW HH WWWHH IDC using C:\NLSS03\DATA\Z07B.dta, unmatched(master);
rename V07B_02 headeduc;
drop V0*;
drop _merge;
joinby WWW HH WWWHH IDC using C:\NLSS03\DATA\Z07C.dta, unmatched(master);
rename V07C_02 headclass;
drop V0*;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\HouseholdCount.dta, sort;
drop _merge;

joinby WWWHH WWW HH using C:\NLSS03\paper\childincome.dta, unmatched(both);
replace childwages = 0 if childwages==.;
replace propsubschild = 0 if propsubschild == .;
drop _merge;

joinby WWWHH WWW HH using C:\NLSS03\paper\landvalue.dta, unmatched(both);
replace landvalue = 0 if landvalue == .;
drop _merge;

joinby WWWHH WWW HH IDC using C:\NLSS03\paper\jobdata.dta, unmatched(master);
drop monthsworked dayspermonth hoursperday daysworkedlastweek hoursdaylastweek hoursperyear;
drop _merge;

gen subsag = 1 if jobcode == 621;
replace subsag = 0 if subsag == .;

/*These steps lose data and are not appropriate for all purposes:*/
gsort -subsag;
duplicates drop WWWHH WWW HH IDC, force;
drop jobcode;

gen BAHUNCHHETRI = 1 if ethcode == 1 | ethcode == 2;
replace BAHUNCHHETRI = 0 if BAHUNCHHETRI != 1;
gen TAMAGURALI = 1 if ethcode == 3 | ethcode == 5 | ethcode == 10 | ethcode == 11 | ethcode == 13;
replace TAMAGURALI = 0 if TAMAGURALI != 1;
gen DAKASA = 1 if ethcode == 8 | ethcode == 12 | ethcode == 15;
replace DAKASA = 0 if DAKASA != 1;
gen TERAICASTE = 1 if ethcode == 4 | ethcode == 9 | ethcode == 27 | ethcode == 33;
replace TERAICASTE = 0 if TERAICASTE != 1;
gen NEWAR = 1 if ethcode ==6;
replace NEWAR = 0 if NEWAR !=1;
gen MUSLIM = 1 if ethcode == 7;
replace MUSLIM = 0 if MUSLIM != 1;
gen OTHERCASTE = 1 if BAHUNCHHETRI == 0 & TAMAGURALI == 0 & DAKASA == 0 &
TERAICASTE == 0 & NEWAR == 0 & MUSLIM == 0;
replace OTHERCASTE = 0 if OTHERCASTE != 1;

joinby WWW using C:\NLSS03\paper\hhweight.dta;
```

```
merge WWWHH WWW HH IDC using C:\NLSS03\paper\headmigrated.dta, sort;
replace HEADMIGRATED = 0 if HEADMIGRATED == .;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\finansoph.dta, sort;
drop _merge;
replace finansoph=0 if finansoph==.;

joinby WWW WWWHH using C:\NLSS03\DATA\pricedata.dta;
drop av03*;
drop district;
rename urbrural RURAL;
replace RURAL = 0 if RURAL ==1;
replace RURAL = 1 if RURAL ==2;
/* belt classification; terai unspecified */
gen MOUNTAIN = 1 if belt ==1;
replace MOUNTAIN = 0 if MOUNTAIN ==.;
gen HILL = 1 if belt == 2;
replace HILL = 0 if HILL ==.;
drop belt;
/*region classification; central unspecified*/
gen WESTERN = 1 if region == 3;
replace WESTERN = 0 if WESTERN ==.;
gen MIDWESTERN = 1 if region == 4;
replace MIDWESTERN = 0 if MIDWESTERN ==.;
gen EASTERN = 1 if region == 1;
replace EASTERN = 0 if EASTERN ==.;
gen FARWESTERN = 1 if region == 5;
replace FARWESTERN = 0 if FARWESTERN == .;
drop region;
rename ra_pindex pindex;

save "C:\NLSS03\paper\HeadData.dta", replace;
```

II. Calculation of Child Labor and Child Education

```
#delim =;
clear;

use C:\NLSS03\DATA\Z01C.dta, clear;
rename V01C_IDC IDC;
replace V01C_02 = 0 if V01C_02 ==.;
replace V01C_03 = 0 if V01C_03 ==.;
replace V01C_04 = 0 if V01C_04 ==.;
gen hoursperjobyear = V01C_02*V01C_03*V01C_04;
collapse (sum) hoursperjobyear, by(WWWHH WWW HH IDC);
rename hoursperjobyear hoursperyear;

merge WWWHH WWW HH IDC using C:\NLSS03\DATA\Z01A.dta, sort;
rename V01A_05 age;
rename V01A_02 female;
replace female = 0 if female == 1;
replace female = 1 if female ==2;
drop V* _merge;
```

```
replace hoursperyear = 0 if hoursperyear == .;

merge WWWHH WWW HH IDC using C:\NLSS03\DATA\Z07B.dta, sort;
rename V07B_01 formerschooltype;
rename V07B_02 classcompleted;
rename V07B_05 reasonleft;
drop _merge;
drop V*;

merge WWWHH WWW HH IDC using C:\NLSS03\DATA\Z07C.dta, sort;
drop _merge;

rename V07C_01 schooltype;
rename V07C_02 class;
rename V07C_05 transport;
rename V07C_06A transhours;
rename V07C_06B transmins;
rename V07C_07 edexp;
rename V07C_08 scholarship;
rename V07C_09 scholvalue;
replace scholarship = 0 if scholarship ==2;

merge WWWHH WWW HH IDC using C:\NLSS03\data\Z07A.dta, sort;
drop _merge;
rename  V07A_06 reasonnoschool;


save "C:\NLSS03\paper\EducationLabor.dta", replace;
```

III. Calculation of Proportion of Agricultural Income Due to Child Labor

```
#delim = ;
use c:\NLSS03\DATA\Z01C.dta, clear;
rename  V01C_IDC IDC;

joinby WWWHH WWW HH IDC using C:\NLSS03\DATA\Z01A.dta;
rename V01A_05 age;

keep if V01C_01C == 621;
replace V01C_02 = 0 if V01C_02 ==.;
replace V01C_03 = 0 if V01C_03 ==.;
replace V01C_04 = 0 if V01C_04 ==.;
gen hoursperjobyear = V01C_02*V01C_03*V01C_04;
collapse (sum) hoursperjobyear, by(WWWHH WWW HH IDC age);
rename hoursperjobyear hoursperyear;
gen CHILD = 1 if age<= 16;
replace CHILD = 0 if CHILD == .;
collapse (sum) hoursperyear, by(WWWHH WWW HH CHILD);
gen hpychild=hoursperyear;
replace hpychild = 0 if CHILD == 0;
rename hoursperyear hpyadult;
replace hpyadult = 0 if CHILD == 1;
```

```
collapse (sum) hpychild hpyadult, by(WWWHH WWW HH);
gen propsubschild = hpychild/(hpychild + hpyadult);
save "c:\NLSS03\paper\subsag.dta", replace;
```

IV. Calculation of Household Income

```
#delim = ;

use c:\NLSS03\DATA\Z05A.dta, clear;
replace V05A_02 = 0 if V05A_02 ==.;
replace V05A_04 = 0 if V05A_04 == .;
gen ownfoodcons = V05A_04 * V05A_02;
collapse (sum) ownfoodcons, by(WWWHH WWW HH);
save "C:\NLSS03\paper\ownfoodcons.dta", replace;

use "C:\NLSS03\DATA\Z11A1C.dta", clear;
replace V11A1C_11C = 0 if V11A1C_11C == .;
replace V11A1C_11K = 0 if V11A1C_11K == .;
replace V11A1C_14C = 0 if V11A1C_14C == .;
replace V11A1C_14K = 0 if V11A1C_14K == .;
gen aglandrentrec =  V11A1C_11C+ V11A1C_11K+ V11A1C_14C+ V11A1C_14K;
collapse (sum) aglandrentrec, by(WWWHH WWW HH);
save "C:\NLSS03\paper\aglandrentrec.dta", replace;

use "C:\NLSS03\DATA\Z11A2B.dta", clear;
rename V11A2B_04 aglandrentpaid;
replace aglandrentpaid = 0 if aglandrentpaid ==.;
collapse (sum) aglandrentpaid, by(WWWHH WWW HH);
save "C:\NLSS03\paper\aglandrentpaid.dta", replace;

use c:\NLSS03\DATA\Z13C.dta, clear;
replace V13C_06 = 0 if V13C_06 == .;
replace V13C_12 = 0 if V13C_12 == .;
gen rentincome = V13C_06 + V13C_12;
collapse (sum) rentincome, by(WWWHH WWW HH);
save "C:\NLSS03\paper\rentincome.dta", replace;

#delim = ;
use c:\NLSS03\DATA\Z12B.dta, clear;
merge WWWHH WWW HH ENT using "C:\NLSS03\DATA\Z12A2.dta",sort;
replace V12A2_08 = 100 if V12A2_08==.;
gen enterprisenetrev = V12B_07 * V12A2_08/100;
drop V* _merge;
replace enterprisenetrev = 0 if enterprisenetrev ==.;
collapse (sum) enterprisenetrev, by(WWWHH WWW HH);
save "C:\NLSS03\paper\enterpriseincome.dta", replace;

#delim = ;
use c:\NLSS03\DATA\Z15.dta, clear;
replace V15_04 = 0 if V15_04 == .;
rename V15_04 otherincome;
```

```
collapse (sum) otherincome, by(WWWHH WWW HH);
save "C:\NLSS03\paper\otherincome.dta", replace;

#delim = ;
use c:\NLSS03\DATA\Z10A1.dta, clear;
merge WWWHH WWW HH ACT using c:\nlss03\data\z10A2.dta, sort;
rename V10A1_03 agwagescashperjobday;
rename V10A1_05A agwagesinkindperjobday;
rename V10A1_05B agwagesinkindtotal;
rename V10A2_06 agwagescashyear;
rename V10A2_08A agwagesinkindperjobday2;
rename V10A2_08B agwagesinkindtotal2;
replace agwagescashyear = 0 if agwagescashyear == .;
replace agwagesinkindperjobday2 = 0 if agwagesinkindperjobday2 ==.;
replace agwagesinkindtotal2 = 0 if agwagesinkindtotal2 == .;
replace agwagesinkindtotal = 0 if agwagesinkindtotal==.;
rename V10A2_13 agwagescontract;
replace agwagescontract = 0 if agwagescontract==.;
rename V10A1_IDC IDC;
drop V* _merge;
replace agwagescashperjobday = 0 if agwagescashperjobday ==.;
replace agwagesinkindperjobday = 0 if agwagesinkindperjobday ==.;
gen agwagesperjobday = agwagescashperjobday+agwagesinkindperjobday + agwagesinkindperjobday2;
gen agwagesyeartotal = agwagescashyear+agwagesinkindtotal+agwagesinkindtotal2+agwagescontract;
save "C:\NLSS03\paper\agwagestemp.dta", replace;
collapse (sum) agwagesyeartotal, by(WWWHH WWW HH IDC);
save "C:\NLSS03\paper\agwagestemp2.dta", replace;
use "C:\NLSS03\paper\agwagestemp.dta", replace;
rename IDC V01C_IDC;
merge WWWHH WWW HH ACT V01C_IDC using c:\nlss03\data\z01c.dta, nokeep sort;
rename V01C_IDC IDC;
replace V01C_02 = 0 if V01C_02 ==.;
replace V01C_03 = 0 if V01C_03 ==.;
gen agdaysperjobyear = V01C_02*V01C_03;
gen incomeperjobyear = agdaysperjobyear * agwagesperjobday;
collapse (sum) incomeperjobyear, by(WWWHH WWW HH IDC);
merge WWWHH WWW HH IDC using "C:\NLSS03\paper\agwagestemp2.dta", sort;
drop _merge;
gen agwages = incomeperjobyear + agwagesyeartotal;
save "C:\NLSS03\paper\agwagesindividual.dta", replace;
collapse (sum) agwages, by(WWWHH WWW HH);
save "C:\NLSS03\paper\agwages.dta", replace;

#delim = ;
use c:\NLSS03\DATA\Z10B1.dta, clear;
merge WWWHH WWW HH ACT using c:\nlss03\data\z10B2.dta, sort;
rename V10B1_04 otherwagescashperjobday;
rename V10B1_06A otherwagesinkindperjobday;
rename V10B1_06B otherwagesinkindtotal;
replace otherwagescashperjobday = 0 if otherwagescashperjobday ==.;
replace otherwagesinkindperjobday = 0 if otherwagesinkindperjobday ==.;
replace otherwagesinkindtotal = 0 if otherwagesinkindtotal==.;
replace V10B2_13 = 0 if V10B2_13==.;
replace V10B2_07C = 0 if V10B2_07C==.;
replace V10B2_07D = 0 if V10B2_07D==.;
replace V10B2_07E = 0 if V10B2_07E==.;
```

```
replace V10B2_07A = 0 if V10B2_07A==.;
replace V10B2_07B = 0 if V10B2_07B==.;
gen otherwagesyear =
V10B2_13+V10B2_07C+V10B2_07D+V10B2_07E+(V10B2_07A+V10B2_07B)*12+otherwagesinkindto
tal;
rename V10B1_IDC V01C_IDC;
drop V10* _merge;
gen otherwagesperjobday = otherwagescashperjobday + otherwagesinkindperjobday;
save "C:\NLSS03\paper\otherwagestemp.dta", replace;
collapse (sum) otherwagesyear, by(WWWHH WWW HH V01C_IDC);
rename V01C_IDC IDC;
save "C:\NLSS03\paper\otherwagestemp2.dta", replace;
use "C:\NLSS03\paper\otherwagestemp.dta", clear;
merge WWWHH WWW HH ACT V01C_IDC using c:\nlss03\data\z01c.dta, nokeep sort;
drop _merge;
rename V01C_IDC IDC;
gen otherdaysperjobyear = V01C_02*V01C_03;
gen incomeperjobyear = otherdaysperjobyear*otherwagesperjobday;
collapse (sum) incomeperjobyear, by(WWWHH WWW HH IDC);
merge WWWHH WWW HH IDC using "C:\NLSS03\paper\otherwagestemp2.dta", sort;
drop _merge;
gen otherwages = incomeperjobyear + otherwagesyear;
save "C:\NLSS03\paper\otherwagesindividual.dta",replace;
collapse (sum) otherwages, by (WWWHH WWW HH);
save "C:\NLSS03\paper\otherwages.dta", replace;

use c:\NLSS03\DATA\Z14B2.dta, clear;
rename V14B2_09A remittancecash;
rename V14B2_09B remittanceinkind;
rename V14B2_07A donordistrict;
drop V*;
replace remittancecash = 0 if remittancecash==.;
replace remittanceinkind = 0 if remittanceinkind ==.;
gen intrem = remittancecash + remittanceinkind if donordistrict>75;
gen domrem = remittancecash + remittanceinkind if donordistrict<=75;
replace intrem = 0 if intrem == .;
replace domrem = 0 if domrem == .;
collapse (sum) remittancecash remittanceinkind intrem domrem, by(WWWHH WWW HH);
save "C:\NLSS03\paper\remittance.dta", replace;

use c:\NLSS03\DATA\Z11D.dta, clear;
rename V11D_08 agrevenue;
rename V11D_23 agexpenditure;
replace agrevenue = 0 if agrevenue == .;
replace agexpenditure = 0 if agexpenditure ==.;

merge WWWHH WWW HH using c:\nlss03\data\Z11E2.dta, sort;
rename V11E2_08 lsrevenue;
rename V11E2_13 lsexpenditure;
drop _merge;
replace lsrevenue = 0 if lsrevenue ==.;
replace lsexpenditure = 0 if lsexpenditure ==.;

merge WWWHH WWW HH using c:\nlss03\data\Z02B.dta, sort;
gen homevalue = 12*V02B_03;
replace homevalue = 0 if V02B_01 != 1;
```

```
drop V* _merge;
replace homevalue = 0 if homevalue ==.;

merge WWWHH WWW HH using c:\nlss03\paper\remittance.dta, sort;
drop _merge;

merge WWWHH WWW HH using c:\nlss03\paper\enterpriseincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\agwages.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\otherwages.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\otherincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\rentincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\ownfoodcons.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\aglandrentrec.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\aglandrentpaid.dta, sort;
drop _merge;

replace agrevenue = 0 if agrevenue ==.;
replace agexpenditure = 0 if agexpenditure ==.;
replace lsrevenue = 0 if lsrevenue ==.;
replace lsexpenditure = 0 if lsexpenditure ==.;
replace rentincome = 0 if rentincome ==.;
replace remittancecash = 0 if remittancecash ==.;
replace remittanceinkind = 0 if remittanceinkind ==.;
replace intrem = 0 if intrem ==.;
replace domrem = 0 if domrem ==.;
replace enterprisenetrev = 0 if enterprisenetrev ==.;
replace agwages = 0 if agwages ==.;
replace otherwages = 0 if otherwages ==.;
replace otherincome = 0 if otherincome == .;
replace rentincome = 0 if rentincome == .;
replace ownfoodcons = 0 if ownfoodcons ==.;
replace aglandrentrec = 0 if aglandrentrec ==.;
replace aglandrentpaid = 0 if aglandrentpaid == .;

gen remittance = remittancecash + remittanceinkind;

gen income = agrevenue - agexpenditure + lsrevenue - lsexpenditure +rentincome+ownfoodcons
+remittance +enterprisenetrev +agwages + otherwages + otherincome + homevalue+aglandrentrec -
aglandrentpaid;

gen nonintremincome = income - intrem;
```

```
save "c:\nlss03\paper\income.dta", replace;
```

V. Calculation of Child Labor Income

```
#delim = ;

use c:\NLSS03\DATA\Z05A.dta, clear;
replace V05A_02 = 0 if V05A_02 ==.;
replace V05A_04 = 0 if V05A_04 == .;
gen ownfoodcons = V05A_04 * V05A_02;
collapse (sum) ownfoodcons, by(WWWHH WWW HH);
save "C:\NLSS03\paper\ownfoodcons.dta", replace;

use "C:\NLSS03\DATA\Z11A1C.dta", clear;
replace V11A1C_11C = 0 if V11A1C_11C == .;
replace V11A1C_11K = 0 if V11A1C_11K == .;
replace V11A1C_14C = 0 if V11A1C_14C == .;
replace V11A1C_14K = 0 if V11A1C_14K == .;
gen aglandrentrec =  V11A1C_11C+ V11A1C_11K+ V11A1C_14C+ V11A1C_14K;
collapse (sum) aglandrentrec, by(WWWHH WWW HH);
save "C:\NLSS03\paper\aglandrentrec.dta", replace;

use "C:\NLSS03\DATA\Z11A2B.dta", clear;
rename V11A2B_04 aglandrentpaid;
replace aglandrentpaid = 0 if aglandrentpaid ==.;
collapse (sum) aglandrentpaid, by(WWWHH WWW HH);
save "C:\NLSS03\paper\aglandrentpaid.dta", replace;

use c:\NLSS03\DATA\Z13C.dta, clear;
replace V13C_06 = 0 if V13C_06 == .;
replace V13C_12 = 0 if V13C_12 == .;
gen rentincome = V13C_06 + V13C_12;
collapse (sum) rentincome, by(WWWHH WWW HH);
save "C:\NLSS03\paper\rentincome.dta", replace;

#delim = ;
use c:\NLSS03\DATA\Z12B.dta, clear;
merge WWWHH WWW HH ENT using "C:\NLSS03\DATA\Z12A2.dta",sort;
replace V12A2_08 = 100 if V12A2_08==.;
gen enterprisenetrev = V12B_07 * V12A2_08/100;
drop V* _merge;
replace enterprisenetrev = 0 if enterprisenetrev ==.;
collapse (sum) enterprisenetrev, by(WWWHH WWW HH);
save "C:\NLSS03\paper\enterpriseincome.dta", replace;

#delim = ;
use c:\NLSS03\DATA\Z15.dta, clear;
replace V15_04 = 0 if V15_04 == .;
rename V15_04 otherincome;
collapse (sum) otherincome, by(WWWHH WWW HH);
save "C:\NLSS03\paper\otherincome.dta", replace;
```

```
#delim = ;
use c:\NLSS03\DATA\Z10A1.dta, clear;
merge WWWHH WWW HH ACT using c:\nlss03\data\z10A2.dta, sort;
rename V10A1_03 agwagescashperjobday;
rename V10A1_05A agwagesinkindperjobday;
rename V10A1_05B agwagesinkindtotal;
rename V10A2_06 agwagescashyear;
rename V10A2_08A agwagesinkindperjobday2;
rename V10A2_08B agwagesinkindtotal2;
replace agwagescashyear = 0 if agwagescashyear == .;
replace agwagesinkindperjobday2 = 0 if agwagesinkindperjobday2 ==.;
replace agwagesinkindtotal2 = 0 if agwagesinkindtotal2 == .;
replace agwagesinkindtotal = 0 if agwagesinkindtotal==.;
rename V10A2_13 agwagescontract;
replace agwagescontract = 0 if agwagescontract==.;
rename V10A1_IDC IDC;
drop V* _merge;
replace agwagescashperjobday = 0 if agwagescashperjobday ==.;
replace agwagesinkindperjobday = 0 if agwagesinkindperjobday ==.;
gen agwagesperjobday = agwagescashperjobday+agwagesinkindperjobday + agwagesinkindperjobday2;
gen agwagesyeartotal = agwagescashyear+agwagesinkindtotal+agwagesinkindtotal2+agwagescontract;
save "C:\NLSS03\paper\agwagestemp.dta", replace;
collapse (sum) agwagesyeartotal, by(WWWHH WWW HH IDC);
save "C:\NLSS03\paper\agwagestemp2.dta", replace;
use "C:\NLSS03\paper\agwagestemp.dta", replace;
rename IDC V01C_IDC;
merge WWWHH WWW HH ACT V01C_IDC using c:\nlss03\data\z01c.dta, nokeep sort;
rename V01C_IDC IDC;
replace V01C_02 = 0 if V01C_02 ==.;
replace V01C_03 = 0 if V01C_03 ==.;
gen agdaysperjobyear = V01C_02*V01C_03;
gen incomeperjobyear = agdaysperjobyear * agwagesperjobday;
collapse (sum) incomeperjobyear, by(WWWHH WWW HH IDC);
merge WWWHH WWW HH IDC using "C:\NLSS03\paper\agwagestemp2.dta", sort;
drop _merge;
gen agwages = incomeperjobyear + agwagesyeartotal;
save "C:\NLSS03\paper\agwagesindividual.dta", replace;
collapse (sum) agwages, by(WWWHH WWW HH);
save "C:\NLSS03\paper\agwages.dta", replace;

#delim = ;
use c:\NLSS03\DATA\Z10B1.dta, clear;
merge WWWHH WWW HH ACT using c:\nlss03\data\z10B2.dta, sort;
rename V10B1_04 otherwagescashperjobday;
rename V10B1_06A otherwagesinkindperjobday;
rename V10B1_06B otherwagesinkindtotal;
replace otherwagescashperjobday = 0 if otherwagescashperjobday ==.;
replace otherwagesinkindperjobday = 0 if otherwagesinkindperjobday ==.;
replace otherwagesinkindtotal = 0 if otherwagesinkindtotal==.;
replace V10B2_13 = 0 if V10B2_13==.;
replace V10B2_07C = 0 if V10B2_07C==.;
replace V10B2_07D = 0 if V10B2_07D==.;
replace V10B2_07E = 0 if V10B2_07E==.;
replace V10B2_07A = 0 if V10B2_07A==.;
replace V10B2_07B = 0 if V10B2_07B==.;
```

```
gen otherwagesyear =
V10B2_13+V10B2_07C+V10B2_07D+V10B2_07E+(V10B2_07A+V10B2_07B)*12+otherwagesinkindto
tal;
rename V10B1_IDC V01C_IDC;
drop V10* _merge;
gen otherwagesperjobday = otherwagescashperjobday + otherwagesinkindperjobday;
save "C:\NLSS03\paper\otherwagestemp.dta", replace;
collapse (sum) otherwagesyear, by(WWWHH WWW HH V01C_IDC);
rename V01C_IDC IDC;
save "C:\NLSS03\paper\otherwagestemp2.dta", replace;
use "C:\NLSS03\paper\otherwagestemp.dta", clear;
merge WWWHH WWW HH ACT V01C_IDC using c:\nlss03\data\z01c.dta, nokeep sort;
drop _merge;
rename V01C_IDC IDC;
gen otherdaysperjobyear = V01C_02*V01C_03;
gen incomeperjobyear = otherdaysperjobyear*otherwagesperjobday;
collapse (sum) incomeperjobyear, by(WWWHH WWW HH IDC);
merge WWWHH WWW HH IDC using "C:\NLSS03\paper\otherwagestemp2.dta", sort;
drop _merge;
gen otherwages = incomeperjobyear + otherwagesyear;
save "C:\NLSS03\paper\otherwagesindividual.dta",replace;
collapse (sum) otherwages, by (WWWHH WWW HH);
save "C:\NLSS03\paper\otherwages.dta", replace;

use c:\NLSS03\DATA\Z14B2.dta, clear;
rename V14B2_09A remittancecash;
rename V14B2_09B remittanceinkind;
rename V14B2_07A donordistrict;
drop V*;
replace remittancecash = 0 if remittancecash==.;
replace remittanceinkind = 0 if remittanceinkind ==.;
gen intrem = remittancecash + remittanceinkind if donordistrict>75;
gen domrem = remittancecash + remittanceinkind if donordistrict<=75;
replace intrem = 0 if intrem == .;
replace domrem = 0 if domrem == .;
collapse (sum) remittancecash remittanceinkind intrem domrem, by(WWWHH WWW HH);
save "C:\NLSS03\paper\remittance.dta", replace;

use c:\NLSS03\DATA\Z11D.dta, clear;
rename V11D_08 agrevenue;
rename V11D_23 agexpenditure;
replace agrevenue = 0 if agrevenue == .;
replace agexpenditure = 0 if agexpenditure ==.;

merge WWWHH WWW HH using c:\nlss03\data\Z11E2.dta, sort;
rename V11E2_08 lsrevenue;
rename V11E2_13 lsexpenditure;
drop _merge;
replace lsrevenue = 0 if lsrevenue ==.;
replace lsexpenditure = 0 if lsexpenditure ==.;

merge WWWHH WWW HH using c:\nlss03\data\Z02B.dta, sort;
gen homevalue = 12*V02B_03;
replace homevalue = 0 if V02B_01 != 1;
drop V* _merge;
replace homevalue = 0 if homevalue ==.;
```

```
merge WWWHH WWW HH using c:\nlss03\paper\remittance.dta, sort;
drop _merge;

merge WWWHH WWW HH using c:\nlss03\paper\enterpriseincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\agwages.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\otherwages.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\otherincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\rentincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\ownfoodcons.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\aglandrentrec.dta, sort;
drop _merge;

merge WWWHH WWW HH using C:\NLSS03\paper\aglandrentpaid.dta, sort;
drop _merge;

replace agrevenue = 0 if agrevenue ==.;
replace agexpenditure = 0 if agexpenditure ==.;
replace lsrevenue = 0 if lsrevenue ==.;
replace lsexpenditure = 0 if lsexpenditure ==.;
replace rentincome = 0 if rentincome ==.;
replace remittancecash = 0 if remittancecash ==.;
replace remittanceinkind = 0 if remittanceinkind ==.;
replace intrem = 0 if intrem ==.;
replace domrem = 0 if domrem ==.;
replace enterprisenetrev = 0 if enterprisenetrev ==.;
replace agwages = 0 if agwages ==.;
replace otherwages = 0 if otherwages ==.;
replace otherincome = 0 if otherincome == .;
replace rentincome = 0 if rentincome == .;
replace ownfoodcons = 0 if ownfoodcons ==.;
replace aglandrentrec = 0 if aglandrentrec ==.;
replace aglandrentpaid = 0 if aglandrentpaid == .;

gen remittance = remittancecash + remittanceinkind;

gen income = agrevenue - agexpenditure + lsrevenue - lsexpenditure +rentincome+ownfoodcons
+remittance +enterprisenetrev +agwages + otherwages + otherincome + homevalue+aglandrentrec -
aglandrentpaid;

gen nonintremincome = income - intrem;

save "c:\nlss03\paper\income.dta", replace;
```

VI. Calculation of Estimation Results

```
#delim = ;
use C:\NLSS03\paper\EducationLabor.dta, clear;
drop V0*;
drop if age<5 | age > 16;
joinby WWWHH WWW HH using C:\NLSS03\paper\HeadData.dta;
joinby WWWHH WWW HH using C:\NLSS03\paper\income;
drop agrevenue agexpenditure lsrevenue lsexpenditure rentincome remittancecash remittanceinkind
domrem enterprisenetrev agwages otherwages remittance;
drop scholarship scholvalue;
rename unmarried headunmarried;
replace headeduc = 0 if headeduc ==.;
gen transtime = transhours*60 + transmins;
drop transhours transmins;
drop if income<0;
gen realincome = income/pindex;
gen realintrem = intrem/pindex;
gen realnonintremincome = nonintremincome/pindex;
replace classcompleted = class - 1 if classcompleted ==.;
replace classcompleted = classcompleted+1;
gen edindex = (classcompleted + 5)/age;
replace edindex = 0 if edindex == .;
replace realnonintremincome = 0 if realnonintremincome==.;
gen lnincome = ln(realincome + 1);
gen lnintrem = ln(realintrem + 1);
gen lnnonintrem = ln(realnonintremincome + 1);
gen lnhours = ln(hoursperyear+1);
drop if realincome >=1000000;

gen laborpart = 1 if hoursperyear>0;
replace laborpart = 0 if laborpart==.;
gen currentlyinschool = 0 if class==.;
replace currentlyinschool = 1 if currentlyinschool==.;
gen schooling = 0 if classcompleted ==.;
replace schooling = 1 if schooling == .;

gen rchildincome = (childwages+propsubschild*ownfoodcons)/pindex;
gen lnonchildnonintrem = ln(realincome-realintrem-rchildincome+1);

/* tobit equivalent with weights and robust errors */
#delim=;
gen lnintrem1 = lnintrem;
replace lnintrem1 = . if lnintrem <=0;
gen lnintrem2 = lnintrem;
replace lnintrem2 = 0 if lnintrem<=0;

save "C:\NLSS03\paper\IncomeLaborEducation.dta", replace;

duplicates drop WWWHH, force;
```

```
intreg lnintrem1 lnintrem2 headunmarried headeduc headfemale headage subsag landvalue n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE MOUNTAIN HILL HEADMIGRATED finansoph
[pweight=hhweight], robust;
/* predicting expected values of fitted lnintrem */
predict flnintremxb, xb;
gen flnintremlambda = normalden(flnintremxb/12.74227)/normal(flnintremxb/12.74227);
gen flnintrem = normal(flnintremxb/12.74227)*(flnintremxb + 12.74227*flnintremlambda);

reg lnonchildnonintrem headunmarried headeduc headfemale headage subsag landvalue n TAMAGURALI
DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE MOUNTAIN HILL HEADMIGRATED finansoph
[pweight=hhweight], robust;
predict flnonchildnonintrem, xb;
save "C:\NLSS03\paper\FittedIncome.dta", replace;
#delim = ;
use C:\NLSS03\paper\IncomeLaborEducation.dta, clear;
joinby WWWHH using C:\NLSS03\paper\FittedIncome.dta;

#delim = ;

ivreg lnhours female headunmarried headeduc headfemale  n TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL (lnonchildnonintrem lnintrem =
headunmarried headeduc headfemale headage subsag landvalue n TAMAGURALI DAKASA
TERAICASTE
NEWAR MUSLIM OTHERCASTE MOUNTAIN HILL HEADMIGRATED finansoph)[pweight =
hhweight];
est store lnhoursivreg;
/*
ivreg edindex female headunmarried headeduc headfemale  n TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL (lnonchildnonintrem lnintrem =
headunmarried headeduc headfemale headage subsag landvalue n TAMAGURALI DAKASA
TERAICASTE
NEWAR MUSLIM OTHERCASTE MOUNTAIN HILL HEADMIGRATED finansoph)[pweight =
hhweight];

heckman lnhours flnonchildnonintrem flnintrem female headunmarried headeduc headfemale  n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], select(laborpart =
flnintrem flnonchildnonintrem female headunmarried headeduc headfemale n TAMAGURALI DAKASA
TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL) robust;

heckman edindex flnonchildnonintrem flnintrem female headunmarried headeduc headfemale  n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], select(schooling
= flnintrem flnonchildnonintrem female headunmarried headeduc headfemale n TAMAGURALI
DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL) robust;
*/

#delim = ;

reg lnhours lnonchildnonintrem lnintrem female headunmarried headeduc headfemale  n TAMAGURALI
DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight];
```

```
est store lnhoursols;
/*
logit laborpart flnonchildnonintrem flnintrem female headunmarried headeduc headfemale  n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], robust;

reg edindex flnonchildnonintrem flnintrem female headunmarried headeduc headfemale  n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], robust;
#delim = ;
logit schooling flnonchildnonintrem flnintrem female headunmarried headeduc headfemale  n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], robust;

reg flnhours flnonchildnonintrem lnintrem female headunmarried headeduc headfemale  n TAMAGURALI
DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], robust;

logit laborpart flnonchildnonintrem flnintrem female headunmarried headeduc headfemale  n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], robust;

reg edindex flnonchildnonintrem flnintrem female headunmarried headeduc headfemale  n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], robust;
#delim = ;
logit schooling flnonchildnonintrem flnintrem female headunmarried headeduc headfemale  n
TAMAGURALI DAKASA TERAICASTE
NEWAR MUSLIM OTHERCASTE RURAL MOUNTAIN HILL [pweight = hhweight], robust;
*/

suest lnhoursivreg lnhoursols;
```

Appendix C: Stata and R Programs for Engel Curve Analyses

This Appendix contains Stata and R programs used for calculations in Chapter 3 of this dissertation.

I. Stata programs

These programs were used to calculate income and consumption aggregates and to format household-level information.

A. Construction of Consumption Aggregates

```
#delim=;

/* food cons aggregation*/
/*aggregate currently includes tea, coffee, not alcohol (unlike offical aggregates) */
use "d:\NLSS03\DATA\Z05A.dta",clear;

destring (WWWHH),replace;

replace V05A_04 = 0 if V05A_04 ==.;
replace V05A_02 = 0 if V05A_02 ==.;
replace V05A_07 = 0 if V05A_07 ==.;
replace V05A_08 = 0 if V05A_08 ==.;
replace V05A_05 = 0 if V05A_05 ==.;

gen valuepergood = V05A_02*V05A_04 +V05A_05*V05A_07+ V05A_08;

drop if ITM>=111 & ITM <=124;
collapse(sum) valuepergood, by(WWW HH WWWHH);
rename valuepergood foodcons03;

save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\foodcons03.dta",replace;


#delim = ;
/*tobacco, alcohol, and other nonfood goods cons aggregation*/
use "d:\NLSS03\DATA\Z05A.dta",clear;

destring (WWWHH),replace;

replace V05A_04 = 0 if V05A_04 ==.;
```

```
replace V05A_02 = 0 if V05A_02 ==.;
replace V05A_07 = 0 if V05A_07 ==.;
replace V05A_08 = 0 if V05A_08 ==.;
replace V05A_05 = 0 if V05A_05 ==.;

gen valuepergood = V05A_02*V05A_04 +V05A_05*V05A_07+ V05A_08;

drop if ITM<111 | ITM >124;
collapse(sum) valuepergood, by(WWW HH WWWHH);
rename valuepergood tobalccons03;

save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\tobalccons03.dta",replace;

use "d:\NLSS03\DATA\Z06A.dta",clear;

destring WWWHH, replace;

/* dropping aggregates, wood, education, health care */
drop if ITM == 210 | ITM == 220 | ITM == 230 | ITM == 250 | ITM == 260 | ITM == 211
| ITM == 237 | ITM == 238;

replace V06A_02 = 0 if V06A_02 ==.;
replace V06A_03 = 0 if V06A_03 ==.;
gen freqnf03 = V06A_02*12;
replace freqnf03 = V06A_03 if freqnf03 == 0;
collapse (sum) freqnf03, by(WWWHH);
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\freqnf03.dta",replace;

use "d:\NLSS03\DATA\Z06B.dta",clear;
destring WWWHH, replace;
keep if ITM == 311 | ITM == 313 | ITM == 314 | ITM == 315 | ITM == 317 | ITM == 318 | ITM == 319 |
ITM == 321 | ITM == 322
| ITM == 324 | ITM == 325 | ITM ==326 | ITM == 327 | ITM == 328;
replace V06B_02 = 0 if V06B_02 == .;
collapse (sum) V06B_02, by(WWWHH);
rename V06B_02 infreqnf03;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\infreqnf03.dta",replace;

joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\tobalccons03.dta", unmatched (both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\freqnf03.dta",
unmatched (both);
drop _merge;
replace tobalccons03 = 0 if tobalccons03 ==.;
replace freqnf03 = 0 if freqnf03 ==.;
replace infreqnf03 = 0 if infreqnf03==.;
gen selectnfcons03 = tobalccons03+freqnf03+infreqnf03;

save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\selectnfcons03.dta",replace;


#delim = ;
/*education cons aggregation*/
```

```
use "d:\NLSS03\DATA\Z07C.dta",clear;
destring WWWHH, replace;
replace V07C_07 = 0 if V07C_07 ==.;
replace V07C_09 = 0 if V07C_09 ==.;
gen edcons03 = V07C_07  + V07C_09;

collapse (sum) edcons03, by(WWWHH WWW HH);
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\edcons03.dta",replace;

#delim = ;
/*durables cons aggregation*/

use "d:\NLSS03\DATA\Z06B.dta",clear;
destring WWWHH, replace;
drop if ITM <= 410;
gen durscons03 = V06B_02;
replace durscons03 = 0 if durscons03 == .;
collapse(sum) durscons03, by(WWWHH WWW HH);

save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\durscons03.dta",replace;

use "d:\NLSS03\DATA\Z06A.dta",clear;

destring WWWHH, replace;

/* dropping aggregates, wood, education, health care */
keep if ITM == 237 | ITM == 238;

replace V06A_02 = 0 if V06A_02 ==.;
replace V06A_03 = 0 if V06A_03 ==.;
gen healcons = V06A_02*12;
replace healcons = V06A_03 if healcons == 0;
collapse (sum) healcons, by(WWWHH);
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\healcons03.dta",replace;


/*housing*/
#delim = ;

/*calculating access to paved roads*/

use "d:\NLSS03\DATA\Z03.dta",clear;
destring WWWHH, replace;
drop if V03_01 != 104;
gen PAVEDROAD = 1 if V03_02==6;
replace PAVEDROAD = 0 if PAVEDROAD == .;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\paved03.dta",replace;

/*calculating HH assets*/
#delim = ;
use "d:\NLSS03\DATA\Z06C.dta",clear;
destring WWWHH, replace;
replace V06C_06 = 0 if V06C_06 == .;
collapse (sum) V06C_06, by(WWWHH);
rename V06C_06 valuedurs;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\valuedurs03.dta",replace;
```

```
use "d:\NLSS03\DATA\Z11A1B.dta", clear;
destring WWWHH, replace;
replace V11A1B_09 = 0 if V11A1B_09  ==.;
collapse (sum) V11A1B_09, by(WWWHH);
rename V11A1B_09 valueland;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\valueland03.dta",replace;

use "d:\NLSS03\DATA\Z11E1B.dta", clear;
destring WWWHH, replace;
replace V11E1B_03B = 0 if V11E1B_03B ==.;
collapse (sum) V11E1B_03B, by(WWWHH);
rename V11E1B_03B valuelivestock;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\valuelivestock03.dta",replace;

use "d:\NLSS03\DATA\Z11F2.dta", clear;
destring WWWHH, replace;
replace V11F2_04 = 0 if V11F2_04==.;
gen valueequip = V11F2_04;
collapse (sum) valueequip, by(WWWHH);
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\valueequip03.dta",replace;

use "d:\NLSS03\DATA\Z13C.dta", clear;
destring WWWHH, replace;
replace V13C_02 = 0 if V13C_02 == .;
replace V13C_08 = 0 if V13C_08 ==.;
gen valueother = V13C_02+V13C_08;
collapse (sum) valueother, by(WWWHH);
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\valueother03.dta",replace;

#delim = ;
use "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\valuedurs03.dta",clear;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\valueland03.dta",unmatched(both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\valuelivestock03.dta",unmatched(both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\valueequip03.dta",unmatched(both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\valueother03.dta",unmatched(both);
drop _merge;

replace valuedurs = 0 if valuedurs==.;
replace valueland = 0 if valueland==.;
replace valuelivestock = 0 if valuelivestock==.;
replace valueequip = 0 if valueequip==.;
replace valueother = 0 if valueother==.;
gen hhassets03 = valuedurs+valueland+valuelivestock+valueequip+valueother;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\hhassets03.dta",replace;

#delim = ;
use "d:\NLSS03\DATA\Z02B.dta",clear;
destring WWWHH, replace;
```

```
destring WWW, replace;
replace V02B_03 = 0 if V02B_03==.;
replace V02B_08 = 0 if V02B_08==.;
gen houscons03 = V02B_03*12+V02B_08*12;

/*calculation of imputed rental value*/

joinby WWWHH using "d:\NLSS03\DATA\Z02A.dta",unmatched(both);
drop _merge;
joinby WWWHH using "d:\NLSS03\DATA\Z02C1.dta", unmatched(both);
drop _merge;
joinby WWWHH using "d:\NLSS03\DATA\Z02C2.dta",unmatched(both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\paved03.dta",unmatched(both);
drop _merge;
replace PAVEDROAD = 0 if PAVEDROAD == .;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\hhassets03.dta",unmatched(both);
drop _merge;
replace hhassets03 = 0 if hhassets03==.;

joinby WWW using "d:\NLSS03\hhweight.dta",unmatched(master);
drop _merge;
gen RURAL = 1 if urbrur == 2;
replace RURAL = 0 if RURAL==.;
gen MOUNTAIN = 1 if belt == "M";
replace MOUNTAIN = 0 if MOUNTAIN==.;
gen HILL = 1 if belt == "H";
replace HILL = 0 if HILL == .;
gen EASTERN = 1 if devreg ==1;
replace EASTERN = 0 if EASTERN == .;
gen WESTERN = 1 if devreg == 3;
replace WESTERN = 0 if WESTERN ==.;
gen MIDWEST = 1 if devreg ==4;
replace MIDWEST =0 if MIDWEST == .;
gen FARWEST = 1 if devreg == 5;
replace FARWEST = 0 if FARWEST ==.;


gen lsize = ln(V02A_09);
replace lsize = 0 if lsize ==.;

gen rooms1 = 1 if V02A_02A==1;
replace rooms1 = 0 if rooms1==.;
gen rooms2 = 1 if V02A_02A==2;
replace rooms2 = 0 if rooms2==.;
gen rooms3 = 1 if V02A_02A==3;
replace rooms3 = 0 if rooms3==.;
gen rooms4 = 1 if V02A_02A==4;
replace rooms4 = 0 if rooms4==.;
gen rooms5 = 1 if V02A_02A==5;
replace rooms5 = 0 if rooms5==.;
gen rooms6 = 1 if V02A_02A==6;
replace rooms6 = 0 if rooms6==.;
gen rooms7 = 1 if V02A_02A==7;
```

```
replace rooms7 = 0 if rooms7==.;
gen rooms8 = 1 if V02A_02A==8;
replace rooms8 = 0 if rooms8==.;


gen KITCHEN = 1 if V02A_02B>0;
replace KITCHEN = 0 if KITCHEN==.;

gen CWALL = 1 if V02A_04==1 | V02A_04== 4;
replace CWALL = 0 if CWALL ==.;

gen CFLOOR = 1 if V02A_05==3 | V02A_05== 4;
replace CFLOOR = 0 if CFLOOR ==.;

gen WINDOWS = 1 if V02A_07!=1;
replace WINDOWS = 0 if WINDOWS==.;

gen PIPED = 1 if V02C1_01 == 1;
replace PIPED = 0 if PIPED ==.;

gen SEWAGE = 1 if V02C1_04==1;
replace SEWAGE = 0 if SEWAGE==.;

gen GARBAGE = 1 if V02C1_05==1 | V02C1_05==2;
replace GARBAGE = 0 if GARBAGE ==.;

gen MUNSEWAGE = 1 if V02C1_07==1;
replace MUNSEWAGE = 0 if MUNSEWAGE == .;

gen LIGHTING = 1 if V02C2_08 == 1;
replace LIGHTING = 0 if LIGHTING ==.;

gen PHONE = 1 if V02C2_11A == 1;
replace PHONE = 0 if PHONE==.;

gen lhhassets = ln(hhassets+1);


predlog houscons lsize  rooms2 rooms3 rooms4 rooms5 rooms6 rooms7 KITCHEN CWALL CFLOOR
WINDOWS PIPED SEWAGE GARBAGE
MUNSEWAGE LIGHTING PHONE lhhassets PAVEDROAD RURAL MOUNTAIN HILL EASTERN
WESTERN MIDWEST FARWEST
 if houscons03 > 25 & houscons03<25000 & V02A_02A<8;

replace houscons03 = YHTSMEAR if (houscons03 == 0 | houscons03 ==.) & V02B_01==1;
replace houscons03 = 0 if houscons03 ==.;


save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\houscons03.dta",replace;

joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\foodcons03.dta", unmatched(both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\selectnfcons03.dta", unmatched(both);
drop _merge;
```

```
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\edcons03.dta", unmatched(both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\durscons03.dta", unmatched(both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\healcons03.dta", unmatched(both);
drop _merge;

replace foodcons03 = 0 if foodcons03==.;
replace selectnfcons03 = 0 if selectnfcons03==.;
replace edcons03 = 0 if edcons03==.;
replace houscons03 = 0 if houscons03==.;
replace durscons03 = 0 if durscons03==.;

replace healcons = 0 if healcons ==.;

keep WWWHH WWW HH houscons03 foodcons03 selectnfcons03 edcons03 durscons03 healcons;

rename houscons03 houscons;
rename foodcons03 foodcons;
rename selectnfcons03 selectnfcons;
rename edcons03 edcons;
rename durscons03 durscons;
gen totcons = houscons+foodcons+selectnfcons+edcons+durscons+healcons;

save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata files\cons03.dta",replace;
```

B. Construction of Income Aggregates

```
#delim = ;

use d:\NLSS03\DATA\Z05B.dta, clear;
replace V05B_04 = 0 if V05B_04 == .;
collapse (sum) V05B_04, by(WWWHH WWW HH);
rename V05B_04 ownfoodcons;
save "d:\NLSS03\paper\ownfoodcons.dta", replace;

use "d:\NLSS03\DATA\Z11A1C.dta", clear;
replace V11A1C_11C = 0 if V11A1C_11C == .;
replace V11A1C_11K = 0 if V11A1C_11K == .;
replace V11A1C_14C = 0 if V11A1C_14C == .;
replace V11A1C_14K = 0 if V11A1C_14K == .;
gen aglandrentrec =  V11A1C_11C+ V11A1C_11K+ V11A1C_14C+ V11A1C_14K;
collapse (sum) aglandrentrec, by(WWWHH WWW HH);
save "d:\NLSS03\paper\aglandrentrec.dta", replace;

use "d:\NLSS03\DATA\Z11A2B.dta", clear;
rename V11A2B_04 aglandrentpaid;
replace aglandrentpaid = 0 if aglandrentpaid ==.;
collapse (sum) aglandrentpaid, by(WWWHH WWW HH);
```

```
save "d:\NLSS03\paper\aglandrentpaid.dta", replace;

use d:\NLSS03\DATA\Z13C.dta, clear;
replace V13C_06 = 0 if V13C_06 == .;
replace V13C_12 = 0 if V13C_12 == .;
gen rentincome = V13C_06 + V13C_12;
collapse (sum) rentincome, by(WWWHH WWW HH);
save "d:\NLSS03\paper\rentincome.dta", replace;

#delim = ;
use d:\NLSS03\DATA\Z12B.dta, clear;
merge WWWHH WWW HH ENT using "d:\NLSS03\DATA\Z12A2.dta",sort;
replace V12A2_08 = 100 if V12A2_08==.;
replace V12B_02 = 0 if V12B_02==.;
replace V12B_03 = 0 if V12B_03==.;
replace V12B_04 = 0 if V12B_04==.;
replace V12B_05A = 0 if V12B_05A==.;
replace V12B_06 = 0 if V12B_06==.;
replace V12B_08 = 0 if V12B_08 == .;
replace V12B_08 = 0 if V12B_08 == .;
replace V12B_10 = 0 if V12B_10 == .;
replace V12B_11 = 0 if V12B_11 == .;
gen enterprisenetrev = (V12B_02-V12B_03-V12B_04-V12B_05A-V12B_06
+V12B_10-V12B_11+V12B_09-V12B_08)* V12A2_08/100;
/* gen enterprisenetrev = (V12B_07) * V12A2_08/100; */
drop V* _merge;
replace enterprisenetrev = 0 if enterprisenetrev ==.;
collapse (sum) enterprisenetrev, by(WWWHH WWW HH);
save "d:\NLSS03\paper\enterpriseincome.dta", replace;

#delim = ;
use d:\NLSS03\DATA\Z15.dta, clear;
replace V15_04 = 0 if V15_04 == .;
rename V15_04 otherincome;
collapse (sum) otherincome, by(WWWHH WWW HH);
save "d:\NLSS03\paper\otherincome.dta", replace;

/*note - 2003 includes contract (V10A2_13); 96 does not*/
#delim = ;
use d:\NLSS03\DATA\Z10A1.dta, clear;
merge WWWHH WWW HH ACT using d:\nlss03\data\z10A2.dta, sort;
drop _merge;
joinby WWWHH WWW HH ACT using "d:\NLSS03\DATA\Z01c.dta", unmatched(both);
drop _merge;

destring WWWHH, replace;
destring WWW,replace;

replace V10A1_03 = 0 if V10A1_03 == .;
replace V10A1_05A  = 0 if V10A1_05A  ==.;
replace V10A1_05B = 0 if V10A1_05B == .;
replace V10A2_06 = 0 if V10A2_06==.;
replace V10A2_08A = 0 if V10A2_08A==.;
replace V10A2_08B = 0 if V10A2_08B==.;
replace V01C_02 = 0 if V01C_02 == .;
replace V01C_03 = 0 if V01C_03 ==.;
```

```
replace V10A2_13=0 if V10A2_13==.;

gen agwages =
(V10A1_03+V10A1_05A+V10A2_08A)*V01C_02*V01C_03+V10A1_05B+V10A2_06+V10A2_08B+V
10A2_13;
collapse (sum) agwages, by(WWWHH WWW HH);
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\agwages.dta",replace;


#delim = ;
use d:\NLSS03\DATA\Z10B1.dta, clear;
joinby WWWHH WWW HH ACT using "d:\NLSS03\DATA\Z10B2.dta", unmatched(both);
drop _merge;
joinby WWWHH WWW HH ACT using "d:\NLSS03\DATA\z01c.dta", unmatched(both);
drop _merge;

destring WWWHH, replace;
destring WWW,replace;

replace V10B1_04 = 0 if V10B1_04 == .;
replace V10B1_06A  = 0 if V10B1_06A  ==.;
replace V10B1_06B = 0 if V10B1_06B == .;
replace V10B2_07A = 0 if V10B2_07A==.;
replace V10B2_07B = 0 if V10B2_07B==.;
replace V10B2_07C = 0 if V10B2_07C==.;
replace V10B2_07D = 0 if V10B2_07D==.;
replace V10B2_07E = 0 if V10B2_07E==.;
replace V10B2_13  = 0 if V10B2_13 ==.;
replace V01C_02 = 0 if V01C_02 == .;
replace V01C_03 = 0 if V01C_03 ==.;

gen otherwages =
(V10B1_04+V10B1_06A)*V01C_02*V01C_03+(V10B2_07A+V10B2_07B)*V01C_02+V10B2_07C+V
10B2_07D
+V10B2_07E + V10B2_13 + V10B1_06B;
collapse (sum) otherwages, by(WWWHH WWW HH);
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\otherwages.dta",replace;
summ otherwages if otherwages>0;

#delim=;
use d:\NLSS03\DATA\Z14B2.dta, clear;
rename V14B2_09A remittancecash;
rename V14B2_09B remittanceinkind;
rename V14B2_07A donordistrict;
drop V*;
replace remittancecash = 0 if remittancecash==.;
replace remittanceinkind = 0 if remittanceinkind ==.;
gen intrem = remittancecash + remittanceinkind if donordistrict>75;
gen domrem = remittancecash + remittanceinkind if donordistrict<=75;
gen indiarem = remittancecas+remittanceinkind if donordistrict==81;
replace indiarem = 0 if indiarem==.;
replace intrem = 0 if intrem == .;
replace domrem = 0 if domrem == .;
collapse (sum) remittancecash remittanceinkind intrem domrem indiarem, by(WWWHH WWW HH);
gen indiadummy=1 if indiarem>intrem/2;
replace indiadummy = 0 if indiadumm == .;
```

```
save "d:\NLSS03\paper\remittance.dta", replace;

#delim = ;
use d:\nlss03\data\Z11E1B.dta, clear;
rename  V11E1B_05B lssales;
rename  V11E1B_06B lspurchases;
replace lssales = 0 if lssales == .;
replace lspurchases = 0 if lspurchases == .;
gen lssalesnet = lssales-lspurchases;
collapse (sum) lssalesnet, by (WWWHH WWW HH);
save "d:\NLSS03\paper\lssalesnet.dta", replace;

#delim = ;
use d:\NLSS03\DATA\Z11D.dta, clear;
replace V11D_01 = 0 if V11D_01==.;
replace V11D_02 = 0 if V11D_02==.;
replace V11D_03 = 0 if V11D_03==.;
replace V11D_04 = 0 if V11D_01==.;
replace V11D_05 = 0 if V11D_05==.;
replace V11D_06 = 0 if V11D_06==.;
replace V11D_07 = 0 if V11D_07==.;
replace V11D_09 = 0 if V11D_09==.;
replace V11D_10 = 0 if V11D_10==.;
replace V11D_11 = 0 if V11D_11==.;
replace V11D_12 = 0 if V11D_12==.;
replace V11D_13 = 0 if V11D_13==.;
replace V11D_14 = 0 if V11D_14==.;
replace V11D_15 = 0 if V11D_15==.;
replace V11D_16 = 0 if V11D_16==.;
replace V11D_17 = 0 if V11D_17==.;
replace V11D_18 = 0 if V11D_18==.;
replace V11D_19 = 0 if V11D_19==.;
replace V11D_20 = 0 if V11D_20==.;
replace V11D_21 = 0 if V11D_21==.;
replace V11D_22 = 0 if V11D_22==.;
rename V11D_08 agrevenue;
rename V11D_23 agexpenditure;
replace agrevenue = 0 if agrevenue == .;
replace agexpenditure = 0 if agexpenditure ==.;
gen agrevenue2 = V11D_01+V11D_02+V11D_07;
gen agexpenditure2 =
V11D_09+V11D_10+V11D_11+V11D_12+V11D_13+V11D_14+V11D_15+V11D_17+V11D_22;

#delim=;
merge WWWHH WWW HH using d:\nlss03\data\Z11E2.dta, sort;
rename V11E2_08 lsrevenue;
rename V11E2_13 lsexpenditure;
drop _merge;
replace lsrevenue = 0 if lsrevenue ==.;
replace lsexpenditure = 0 if lsexpenditure ==.;
#delim = ;


merge WWWHH WWW HH using d:\nlss03\data\Z02B.dta, sort;
gen homevalue = 12*V02B_03;
replace homevalue = 0 if V02B_01 != 1;
```

```
drop V* _merge;
replace homevalue = 0 if homevalue ==.;

merge WWWHH WWW HH using d:\nlss03\paper\remittance.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\nlss03\paper\lssalesnet.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\nlss03\paper\enterpriseincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\NLSS03\paper\agwages.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\NLSS03\paper\otherwages.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\NLSS03\paper\otherincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\NLSS03\paper\rentincome.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\NLSS03\paper\ownfoodcons.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\NLSS03\paper\aglandrentrec.dta, sort;
drop _merge;

merge WWWHH WWW HH using d:\NLSS03\paper\aglandrentpaid.dta, sort;
drop _merge;

replace agrevenue = 0 if agrevenue ==.;
replace agexpenditure = 0 if agexpenditure ==.;
replace lsrevenue = 0 if lsrevenue ==.;
replace lsexpenditure = 0 if lsexpenditure ==.;
replace rentincome = 0 if rentincome ==.;
replace remittancecash = 0 if remittancecash ==.;
replace remittanceinkind = 0 if remittanceinkind ==.;
replace intrem = 0 if intrem ==.;
replace domrem = 0 if domrem ==.;
replace enterprisenetrev = 0 if enterprisenetrev ==.;
replace agwages = 0 if agwages ==.;
replace otherwages = 0 if otherwages ==.;
replace otherincome = 0 if otherincome == .;
replace rentincome = 0 if rentincome == .;
replace ownfoodcons = 0 if ownfoodcons ==.;
replace aglandrentrec = 0 if aglandrentrec ==.;
replace aglandrentpaid = 0 if aglandrentpaid == .;
replace lssalesnet = 0 if lssalesnet==.;
replace indiadummy = 0 if indiadummy ==.;
replace agrevenue2 = 0 if agrevenue2 == .;
replace agexpenditure2 = 0 if agexpenditure2 == .;

gen remittance = remittancecash + remittanceinkind;
```

```
gen income = agrevenue - agexpenditure + lsrevenue - lsexpenditure +rentincome+ownfoodcons
+remittance +enterprisenetrev +agwages + otherwages + otherincome + homevalue+aglandrentrec -
aglandrentpaid
+lssalesnet+domrem;
/* check agprofits, lsstock profits(check), livestock sales , livestock being born (check rich poor distribution)
*/
gen nonintremincome = income - intrem;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata files\income03.dta",replace;

keep WWWHH WWW HH nonintremincome income intrem indiadummy;

save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata files\RemittanceIncome03.dta",replace;
```

C. Compilation of Household Characteristic Data

```
#delim=;

use D:\NLSS03\DATA\Z01A.dta, clear;
gen age0to4 = 1 if V01A_05 <=4;
replace age0to4 = 0 if age0to4 ==.;
gen age5to16 = 1 if V01A_05 > 4 & V01A_05 <= 16;
replace age5to16 = 0 if age5to16 ==.;
gen age17to49 = 1 if V01A_05 >16 & V01A_05<=49;
replace age17to49 = 0 if age17to49 == .;
gen age50plus = 1 if V01A_05 > 49;
replace age50plus = 0 if age50plus == .;
collapse (sum) age0to4 age5to16 age17to49 age50plus (count) IDC, by(WWWHH WWW HH);
rename IDC hhsize;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\householdcount03.dta", replace;


#delim=;
use "D:\NLSS03\DATA\Z03.dta", clear;
keep if V03_01 == 104;
gen TTRoad = V03_03A*24+V03_03B+V03_03C/60;
replace TTRoad = 0 if TTRoad == .;
drop V*;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\RoadAccess.dta", replace;

#delim=;
use "D:\NLSS03\DATA\Z01A.dta", clear;
rename V01A_02 female;
replace female = 0 if female == 1;
replace female = 1 if female == 2;
gen countdummy = 1;
collapse (sum) female (count) countdummy, by(WWWHH);
gen propfemale = female/countdummy;
keep propfemale WWWHH;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\propfemale.dta", replace;
```

```
#delim=;
use D:\NLSS03\DATA\Z08A.dta, clear;
joinby WWWHH WWW HH IDC using D:\NLSS03\DATA\Z08B1.dta, unmatched(both);
replace V08A_02 = 0 if V08A_02 == .;
replace V08A_02 = 0 if V08A_02 == 2;
rename V08A_02 chronicills;
gen acuteills = 1 if V08B1_01A==60;
gen ills = 1 if chronicills == 1 | acuteills == 1;
collapse (sum) ills, by(WWWHH WWW HH);
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\ills03.dta",replace;


use D:\NLSS03\DATA\Z04.dta, clear;
drop if IDC != 1;
rename V04_01 HEADMIGRATED;
replace HEADMIGRATED = 0 if HEADMIGRATED == 2;
drop V*;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\headmigrated03.dta", replace;

#delim=;
use "D:\NLSS03\DATA\Z15.dta", clear;
replace V15_02 = 0 if V15_02==.;
replace V15_03 = 0 if V15_03==.;
replace V15_04 = 0 if V15_04==.;
drop if V15_02 == 0 & V15_03 ==0 & V15_04 == 0;
collapse (count) V15_01, by(WWWHH WWW HH);
rename V15_01 finansoph;
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate files\finansoph03.dta", replace;


use D:\NLSS03\DATA\Z01A.dta, clear;
keep if  V01A_03==1;
rename  V01A_01A ethcode;
rename  V01A_02 headfemale;
replace headfemale = 0 if headfemale ==1;
replace headfemale = 1 if headfemale ==2;
rename  V01A_06 unmarried;
replace  unmarried = 0 if unmarried == 1;
replace unmarried = 1 if unmarried != 0;
rename V01A_05 headage;
drop V0*;

/*one household had 2 heads; I drop the lesser educated*/
duplicates drop WWWHH, force;

joinby WWW HH WWWHH IDC using D:\NLSS03\DATA\Z07B.dta, unmatched(master);
rename V07B_02 headeduc;
drop V0*;
drop _merge;

merge WWWHH WWW HH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\householdcount03.dta", sort;
drop _merge;

rename IDC V01C_IDC;
joinby WWWHH WWW HH V01C_IDC using "D:\NLSS03\DATA\Z01C.dta", unmatched(master);
```

```
rename V01C_IDC IDC;
gen subsag = 1 if V01C_01C == 621;
replace subsag = 0 if subsag == .;

/*These steps lose data and are not appropriate for all purposes:*/
gsort -subsag;
duplicates drop WWWHH WWW HH IDC, force;
drop V*;

gen BAHUNCHHETRI = 1 if ethcode == 1 | ethcode == 2;
replace BAHUNCHHETRI = 0 if BAHUNCHHETRI != 1;
gen TAMAGURALI = 1 if ethcode == 3 | ethcode == 5 | ethcode == 10 | ethcode == 11 | ethcode == 13;
replace TAMAGURALI = 0 if TAMAGURALI != 1;
gen DAKASA = 1 if ethcode == 8 | ethcode == 12 | ethcode == 15;
replace DAKASA = 0 if DAKASA != 1;
gen TERAICASTE = 1 if ethcode == 4 | ethcode == 9 | ethcode == 27 | ethcode == 33;
replace TERAICASTE = 0 if TERAICASTE != 1;
gen NEWAR = 1 if ethcode ==6;
replace NEWAR = 0 if NEWAR !=1;
gen MUSLIM = 1 if ethcode == 7;
replace MUSLIM = 0 if MUSLIM != 1;
gen OTHERCASTE = 1 if BAHUNCHHETRI == 0 & TAMAGURALI == 0 & DAKASA == 0 &
TERAICASTE == 0 & NEWAR == 0 & MUSLIM == 0;
replace OTHERCASTE = 0 if OTHERCASTE != 1;

joinby WWW using D:\NLSS03\DATA\hhweight.dta;
drop _merge;

merge WWWHH WWW HH IDC using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\headmigrated03.dta", sort;
replace HEADMIGRATED = 0 if HEADMIGRATED == .;
drop _merge;

merge WWWHH WWW HH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\finansoph03.dta", sort;
drop _merge;
replace finansoph=0 if finansoph==.;

merge WWWHH WWW HH using  "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\ills03.dta",sort;
drop _merge;
replace ills = 0 if ills == .;

joinby WWWHH using "C:\Users\Mike\Desktop\papers\Engel curves\Stata intermediate files\loans.dta",
unmatched(master);
drop _merge;

joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\RoadAccess.dta",unmatched(both);
drop _merge;

joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\propfemale.dta",unmatched(both);
drop _merge;

joinby WWW WWWHH using "D:\NLSS03\DATA\pricedata.dta";
```

```
drop av03*;
drop district;
rename urbrural RURAL;
replace RURAL = 0 if RURAL ==1;
replace RURAL = 1 if RURAL ==2;
/* belt classification; terai unspecified */
gen MOUNTAIN = 1 if belt ==1;
replace MOUNTAIN = 0 if MOUNTAIN ==.;
gen HILL = 1 if belt == 2;
replace HILL = 0 if HILL ==.;
drop belt;
/*region classification; central unspecified*/
gen WESTERN = 1 if region == 3;
replace WESTERN = 0 if WESTERN ==.;
gen MIDWESTERN = 1 if region == 4;
replace MIDWESTERN = 0 if MIDWESTERN ==.;
gen EASTERN = 1 if region == 1;
replace EASTERN = 0 if EASTERN ==.;
gen FARWESTERN = 1 if region == 5;
replace FARWESTERN = 0 if FARWESTERN == .;
drop region;
rename ra_pindex pindex;
replace loans=0 if loans==.;

replace headeduc=0 if headeduc==.;

drop ACT ethcode;

joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata intermediate
files\hhassets03.dta";
drop valueequip valuelivestock valueother valuedurs;
rename hhassets03 hhassets;

save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata files\HouseholdData03.dta",replace;
```

D. Compilation of Various Data

```
#delim = ;

use "C:\Users\Mike\Desktop\Papers\Engel curves\Stata files\HouseholdData03",clear;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata
files\RemittanceIncome03",unmatched(both);
drop _merge;
joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata files\cons03", unmatched(both);
drop _merge;

joinby WWWHH using "C:\Users\Mike\Desktop\Papers\Engel curves\Stata files\analysis03",
unmatched(both) _merge(merged);
drop merged;

gen nonindiadummy = 1 if indiadummy!=1 & intrem>0;
replace nonindiadummy=0 if nonindiadummy==.;
```

```
save "C:\Users\Mike\Desktop\Papers\Engel curves\Stata files\merged03",replace;
outsheet using "C:\Users\Mike\Desktop\Papers\Engel curves\R scripts\2003Data.txt", nolabel replace;
```

II. R program

This program was used for the calculations presented in Chapter 3 of this dissertation.

```
rm(list=ls(all=TRUE))
graphics.off()
library("sm")
library("scatterplot3d")
library("mgcv")
library("Rcmdr")
library("MASS")
library("survival")
library("micEcon")
library("rgl")
library("corpcor")

setwd("C:/Users/Workstation/Desktop/Papers/Engel curves/R scripts")
#setwd("C:/Users/Mike/Desktop/Papers/Engel curves/R scripts")
#data1996 <- read.table("1996Data.txt",header=TRUE)
data2003 <- read.table("2003Data.txt",header=TRUE)

#--------------------------------2003-----------------------------------------

A<-data2003

#drop negative incomes
A<-A[(A$nonintremincome)>=0,1:dim(A)[2]]
#A<-A[(A$totcons-A$intrem)>=0,1:dim(A)[2]]

#drop extreme nonintrem outliers
#A<-(A[((A$nonintremincome)/(A$hhsize))>(-5000),1:dim(A)[2]])
#A<-(A[((A$nonintremincome)/(A$hhsize))<(250000),1:dim(A)[2]])

pindex<-A$pindex
weight<-A$hhweight
incomerem<-A$intrem/A$hhsize
incomenonrem<-A$nonintrem/A$hhsize
income<-(incomerem+incomenonrem)/pindex
#lincrem<-log(incomerem+1)
#lincnonrem<-log(incomenonrem+1)
ycond<-A$ycond

food<-A$foodcons/A$hhsize
educ<-A$edcons/A$hhsize
durs<-A$durscons/A$hhsize
hous<-A$houscons/A$hhsize
totcons<-A$totcons/A$hhsize
```

```
nf<-A$selectnfcons/A$hhsize
heal<-A$healcons/A$hhsize

weighted.var <- function(x, w, na.rm = FALSE) {
  if (na.rm) {
    w <- w[i <- !is.na(x)]
    x <- x[i]
  }
  sum.w <- sum(w)
  sum.w2 <- sum(w^2)
  mean.w <- sum(x * w) / sum(w)
  (sum.w / (sum.w^2 - sum.w2)) * sum(w * (x - mean.w)^2, na.rm =
na.rm)
}


weighted.mean(log(totcons+1),weight)
sqrt(weighted.var(log(totcons+1),weight))
weighted.mean(log(food+1),weight)
sqrt(weighted.var(log(food+1),weight))
weighted.mean(log(educ+1,weight))
sqrt(weighted.var(log(educ+1),weight))
weighted.mean(log(durs+1),weight)
sqrt(weighted.var(log(durs+1),weight))
weighted.mean(log(hous+1),weight)
sqrt(weighted.var(log(hous+1),weight))
weighted.mean(log(nf+1),weight)
sqrt(weighted.var(log(nf+1),weight))
weighted.mean(log(heal+1),weight)
sqrt(weighted.var(log(heal+1),weight))



foodshare<-food/totcons
educshare<-educ/totcons
dursshare<-durs/totcons
housshare<-hous/totcons
nfshare<-nf/totcons
healshare<-heal/totcons

y<-log(totcons+1)

n<-length(totcons)

X<- (cbind(
AGE0TO4=A$age0to4,AGE5TO16=A$age5to16,AGE17TO49=A$age17to49,AGE50PLUS=A$age50plus
,
TAMAGURALI=A$TAMAGURALI,DAKASA=A$DAKASA,
TERAICASTE=A$TERAICASTE,NEWAR=A$NEWAR,MUSLIM=A$MUSLIM,OTHERCASTE=A$O
THERCASTE,
 SUBSAG=A$subsag,ills = (A$ills/A$hhsize),hhassets=log(A$hhassets/A$hhsize+1),
RURAL=A$RURAL, MOUNTAIN=A$MOUNTAIN, HILL=A$HILL,
#hhassets=A$hhassets/A$hhsize,
WESTERN=A$WESTERN,MIDWESTERN=A$MIDWESTERN,
FARWESTERN=A$FARWESTERN,EASTERN=A$EASTERN,indiarem=A$indiadummy
))
```

```
#x1<-totcons/pindex
x1<-incomenonrem/pindex
#x1<-(totcons-incomerem)/pindex
x1<-log(x1+1)
x2<-incomerem/pindex
x2<-log(x2+1)

weighted.mean(x1,weight)
sqrt(weighted.var(x1,weight))
weighted.mean(x2,weight)
sqrt(weighted.var(x2,weight))
weighted.mean(x2[x2>0],weight[x2>0])
sqrt(weighted.var(x2[x2>0],weight[x2>0]))
length(x2[x2>0])

Xcluded<-cbind(HEADUNMARRIED=A$unmarried,finansoph = A$finansoph,
HEADMIGRATED=A$HEADMIGRATED,HEADFEMALE=A$headfemale,HEADAGE=A$headage,
HEADEDUC=A$headeduc,TTRoad=A$TTRoad,propfemale=A$propfemale
)
        {
        print(paste("Xcluded.",i))
        print(weighted.mean(Xcluded[,i],weight))
        print(sqrt(weighted.var(Xcluded[,i],weight)))
        }


#endogenization of income

REMTEST <- as.logical(x2>0)
Xrempos <- rep(NA,length(REMTEST[REMTEST>0])*dim(X)[2])
dim(Xrempos) <- c(length(REMTEST[REMTEST>0]),dim(X)[2])
Xcludedrempos <- rep(NA,length(REMTEST[REMTEST>0])*dim(Xcluded)[2])
dim(Xcludedrempos) <- c(length(REMTEST[REMTEST>0]),dim(Xcluded)[2])
weightsrempos <- rep(NA,length(REMTEST[REMTEST>0]))


count <- 1
for (i in 1:n)      {
        if (REMTEST[i]==1)
                                {
                                Xrempos[count,]<-X[i,]
                                Xcludedrempos[count,]<-Xcluded[i,]
                                weightsrempos[count]<-weight[i]
                                count<-count+1
                                }
                        }



remincInst<-rlm(x2[x2>0]~Xrempos+Xcludedrempos, weights = weightsrempos, method = "MM")


x1unfitted<-x1
x2unfitted<-x2

remtobit<-survreg(Surv(x2,x2>0,type='left')~X+Xcluded,dist='gaussian')
```

```
prem<-predict(remtobit,type="lp")
sigma<-remtobit$scale
incremlambda<-dnorm(prem/sigma)/pnorm(prem/sigma)
incremtobit<-(pnorm(prem/sigma))*(prem+sigma*incremlambda)

X<- (cbind(X,nonindiarem=A$nonindiadummy))

for (i in 1:dim(X)[2])
        {
        print(paste("X.",i))
        print(weighted.mean(X[,i],weight))
        print(sqrt(weighted.var(X[,i],weight)))
        }


nonremincInst<-rlm(x1~X+Xcluded, weights = weight)
nonremincOLS<-lm(x1~X+Xcluded, weights = weight)


x1      <-fitted(nonremincInst)
x2      <-rep(0,n)
count   <-1
for (i in 1:n)     {
        if (REMTEST[i]==1)
                                {
                                x2[i]<-fitted(remincInst)[count]
                                count<-count+1
                                }
                        }
#to use tobit
x2<-incremtobit

x1sq<-x1unfitted*x1unfitted
x2sq<-x2unfitted*x2unfitted
x1x2<-x1unfitted*x2unfitted

x2sqtobit<-survreg(Surv(x2sq,x2>0,type='left')~X+Xcluded,dist='gaussian')
prem2<-predict(x2sqtobit,type="lp")
sigma2<-x2sqtobit$scale
incremlambda2<-dnorm(prem2/sigma2)/pnorm(prem2/sigma2)
incremtobit2<-(pnorm(prem2/sigma2))*(prem2+sigma2*incremlambda2)

#get rid of negative fitted incomes
#x1t<-x1[x1>=0 & x2>=0]
#x2t<-x2[x1>=0 & x2>=0]
#yt<-y[x1>=0 & x2>=0]
#x1unfittedt<-x1unfitted[x1>=0 & x2>=0]
#x2unfittedt<-x2unfitted[x1>=0 & x2>=0]
#weight<-weight[x1>=0 & x2>=0]
#dimX.2<-dim(X)[2]
#X<-X[x1>=0 & x2>=0]
#x1<-x1t
#x2<-x2t
#y<-yt
#x1unfitted<-x1unfittedt
#x2unfitted<-x1unfittedt
```

```
#dim(X)<-c(length(x1),dimX.2)

set.seed(2)

v1<-x1-x1unfitted
v2<-x2-x2unfitted


n<-length(x1)
x1<-x1unfitted
x2<-x2unfitted

#FOOD-------------------------------------------------------------------------
for (i in 1:dim(X)[2])
            {
            nam <- paste("X",i,sep=".")
            assign(nam, X[,i])
            }

xnam <- paste("X",1:dim(X)[2],sep = ".")

#parametric
EqLin <- gam(y~X+x1+x2+v1+v2, weights=weight)
#summary(EqLin)

EqQuad <- gam(y~X+x1+x1sq+x2+x2sq+x1x2+v1+v2+v1*v1+v2*v2+v1*v2, weights=weight)
#summary(EqQuad)

smoothpar<-1.4
EqGamAS <- gam(as.formula(paste("y ~
",paste(xnam,collapse="+"),"+s(x1)+s(x2)+s(v1)+s(v2)")),gamma=smoothpar,family=gaussian(link="ident
ity"))

#gam joint

EqGamNP          <-  gam(as.formula(paste("y ~
",paste(xnam,collapse="+"),"+s(x1,x2)+s(v1,v2)")),gamma=smoothpar,family=gaussian(link="identity"))
yhat       <-         predict(EqGamNP)
mse        <-         mean((yhat-y)^2)
rmse       <-         sqrt(mse)


#EqGamNP          <-  gam(as.formula(paste("y ~
",paste(xnam,collapse="+"),"+s(x1,x2)+s(v1,v2)")),family=Gamma(link="identity"))
FittedTerms<- predict.gam(EqGamNP,type="terms",se.fit = T)

#summary(EqGamNP)
gam.check(EqGamNP)
#variance increasing with mean, so use diff link ref. wood p. 234
#residNP <- residuals(EqGamNP)
#fittedNP <- fitted(EqGamNP)
#summary(log(fitted(EqGamNP))
#lm(log(residNP^2)~log(fitted(EqGamNP)))

Xmeanscalar<-colMeans(X)
for (i in 1:dim(X)[2])
```

```
		{
		namscalar <- paste("Xmeanscalar",i,sep=".")
		assign(namscalar, Xmeanscalar[i])
		}
xmeanScalarnam<-paste("Xmeanscalar",1:dim(X)[2],sep=".")

Xmean<-X
for (i in 1:dim(X)[2])
		{
		Xmean[,i]<-mean(X[,i])
		}

for (i in 1:dim(X)[2])
		{
		nam <- paste("Xmean",i,sep=".")
		assign(nam, Xmean[,i])
		}
xmeannam<-paste("Xmean",1:dim(X)[2],sep=".")


Xp<-predict(EqGamNP,type="lpmatrix")
coefNP<-coef(EqGamNP)

xboffset<-mean(rowSums(FittedTerms$fit[,1:dim(X)[2]]))+coefNP[1]

nsample<-1000
coefsample<-
mvrnorm(nsample,coef(EqGamNP)[(dim(X)[2]+2):(dim(X)[2]+30)],EqGamNP$Vp[(dim(X)[2]+2):(dim(X
)[2]+30),(dim(X)[2]+2):(dim(X)[2]+30)])

#-------------------------
plot3d(x1,x2,FittedTerms$fit[,dim(X)[2]+1]+xboffset)
plot3d(x1,x2,FittedTerms$fit[,dim(X)[2]+1]+1.96*FittedTerms$se.fit[,dim(X)[2]+1]+xboffset,add=T)
plot3d(x1,x2,FittedTerms$fit[,dim(X)[2]+1]-1.96*FittedTerms$se.fit[,dim(X)[2]+1]+xboffset,add=T)

lengthplot<-30
x1order	<-	seq(min(x1),max(x1),len=lengthplot)
x2order	<-	seq(min(x2),max(x2),len=lengthplot)

yorder	<-	matrix(NA,lengthplot,lengthplot)
yorderSEhi	<-	matrix(NA,lengthplot,lengthplot)
yorderSElo	<-	matrix(NA,lengthplot,lengthplot)
for (i in 1:lengthplot)
{
		for (j in 1:lengthplot)
		{
plotmcom	<-
		paste("data.frame(x1=x1order[i],x2=x2order[j],",(paste(xnam,"=",xmeanScalarnam,collapse=",")),
",v1=0,v2=0)")
plotmatrix	<-	eval(parse(text=plotmcom))
plotpredict	<-	predict.gam(EqGamNP,plotmatrix,type="terms",se.fit=T)
yorder[i,j]<-plotpredict$fit[dim(X)[2]+1]
yorderSEhi[i,j]<-plotpredict$fit[dim(X)[2]+1]+plotpredict$se.fit[dim(X)[2]+1]
yorderSElo[i,j]<-plotpredict$fit[dim(X)[2]+1]-plotpredict$se.fit[dim(X)[2]+1]
		}
}
```

```
persp3d(x1order,x2order,yorder,xlab = "ln x1",ylab="ln x2",zlab="f(ln x1,ln x2)",theta = -25,phi = -
8,col="slategray",ticktype="detailed",front="fill",back="lines",alpha=.8)
persp3d(x1order,x2order,yorderSEhi,add=T,col="red",alpha=.3)
persp3d(x1order,x2order,yorderSElo,add=T,col="green",alpha=.3)

persp(x1order,x2order,yorder,xlab = "ln x1",ylab="ln x2",zlab="f(ln x1,ln x2)",theta = -35,phi = -
3,col="slategray",ticktype="detailed",nticks=4)



#vis.gam(EqGamNP,se=0,view=c("x1","x2"),theta=-35,phi=-15,zlab="",xlab="ln x1",ylab="ln x2",col =
"slategray1")


#-------------------------
formulaboot<-as.formula(paste("yboot ~ ",paste(xnam,collapse="+"),"+s(x1,x2)+s(v1,v2)"))
formulaboot2<-as.formula(paste("y ~ ",paste(xnam,collapse="+"),"+s(x1,x2)+s(v1,v2)"))
br<-matrix(0,0,length(coef(EqGamNP)))
for (i in 1:19)
        {
        e<-rnorm(rep(1,n),yhat)-yhat
        yboot<-(yhat+e*EqGamNP$sig2^.5)
        yboot[yboot<0]<-0
        sp <- gam(formulaboot,gamma=smoothpar,family=gaussian(link="identity"))$sp
        b  <- gam(formulaboot2,gamma=smoothpar,family=gaussian(link="identity"),sp=sp)
        br <- rbind(br,mvrnorm(n=100,coef(b),b$Vp))
        }

br      <-      rbind(br,mvrnorm(n=100,coef(EqGamNP),EqGamNP$Vp))


x1difference<-  .01
x2standard      <-      median(x2[x2>0])

x1high  <-      x1+x1difference
x1low           <-      x1-x1difference

x1mathcom       <-
        paste("data.frame(x1=x1high,x2=rep(x2standard,n),",(paste(xnam,"=",xmeannam,collapse=",")),",
v1=rep(0,n),v2=rep(0,n))")
x1mathigh       <-      eval(parse(text=x1mathcom))
x1matlcom       <-
        paste("data.frame(x1=x1low,x2=rep(x2standard,n),",(paste(xnam,"=",xmeannam,collapse=",")),",
v1=rep(0,n),v2=rep(0,n))")
x1matlow        <-      eval(parse(text=x1matlcom))
x1matmcom       <-
        paste("data.frame(x1=x1,x2=rep(x2standard,n),",(paste(xnam,"=",xmeannam,collapse=",")),",v1=r
ep(0,n),v2=rep(0,n))")
x1matmiddle     <-      eval(parse(text=x1matmcom))
highpredict1<-  predict.gam(EqGamNP,x1mathigh,type="lpmatrix")
lowpredict1     <-      predict.gam(EqGamNP,x1matlow,type="lpmatrix")
midpredict1 <-  predict.gam(EqGamNP,x1matmiddle,type="lpmatrix")


highpredictvalue1 <- highpredict1 %*% coefNP
```

```
lowpredictvalue1 <- lowpredict1 %*% coefNP
elas1 <- (highpredictvalue1-lowpredictvalue1)/(2*x1difference)
plot(x1,elas1)
midpredictvalue1 <- midpredict1[,1:(dim(X)[2]+30)] %*% coefNP[1:(dim(X)[2]+30)]
yhatunlog1          <-          exp(midpredictvalue1+.5*rmse)
dydx1<-elas1*yhatunlog1/exp(x1)
plot(x1,dydx1,ylim=c(0,2))

highpredsample1 <- rep(NA,dim(br)[1]*n)
dim(highpredsample1) <- c(n,dim(br)[1])
lowpredsample1  <- rep(NA,dim(br)[1]*n)
dim(lowpredsample1) <- c(n,dim(br)[1])
dydx1sample <- rep(NA,dim(br)[1]*n)
dim(dydx1sample) <- c(n,dim(br)[1])
elassample1 <- rep(NA,dim(br)[1]*n)
dim(elassample1) <- c(n,dim(br)[1])
yhat1 <- rep(NA,n*dim(br)[1])
dim(yhat1) <- c(n,dim(br)[1])
yhatulsample1 <- rep(NA,n*dim(br)[1])
dim(yhatulsample1) <- c(n,dim(br)[1])

for (i in 1:dim(br)[1])
        {
        highpredsample1[,i] <- highpredict1[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]
        lowpredsample1[,i]  <- lowpredict1[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]
        elassample1[,i] <- (highpredsample1[,i]-lowpredsample1[,i])/(2*x1difference)
        yhatulsample1[,i] <- exp(midpredict1[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]+xboffset+.5*rmse^2)
        yhat1[,i] <- midpredict1[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]+xboffset
#       dydx1sample[,i] <- elassample1[,i]*yhatulsample1[,i]/exp(x1)
        }

elas1mean <- rep(NA,n)
CIelas1high <- rep(NA,n)
CIelas1low  <- rep(NA,n)
dydx1mean <- rep(NA,n)
CIdydx1high <- rep(NA,n)
CIdydx1low <- rep(NA,n)
yhat1mean <- rep(NA,n)
yhat1CIhigh <- rep(NA,n)
yhat1CIlow <- rep(NA,n)

for (i in 1:n)
        {
        elas1mean[i] <- mean(elassample1[i,])
        CIelas1high[i] <- quantile(elassample1[i,],prob=.975)
        CIelas1low[i] <- quantile(elassample1[i,],prob=.025)
#       CIelas1high[i] <- mean(elassample1[i,])+1.96*sd(elassample1[i,])
#       CIelas1low[i] <- mean(elassample1[i,])-1.96*sd(elassample1[i,])
#       dydx1mean[i] <- mean(dydx1sample[i,])
#       CIdydx1high[i] <- quantile(dydx1sample[i,],prob=.975)
#       CIdydx1low[i] <- quantile(dydx1sample[i,],prob=.025)
#       CIdydx1high[i] <- mean(dydx1sample[i,])+1.96*sd(dydx1sample[i,])
```

```
#	CIdydx1low[i] <- mean(dydx1sample[i,])-1.96*sd(dydx1sample[i,])
	yhat1mean[i] <- mean(yhat1[i,])
	yhat1CIhigh[i] <- quantile(yhat1[i,],prob=.975)
	yhat1CIlow[i] <- quantile(yhat1[i,],prob=.025)
	}

plot(x1,yhat1mean,xlim=c(3.6,13.6),title(ylab="Predicted Log Total
Consumption",xlab=expression(paste("ln ",x[1]))),ylab="",xlab="",cex=.5)
points(x1,yhat1CIhigh,pch=".")
points(x1,yhat1CIlow,pch=".")
points(x1[order(x1)][n/10],yhat1mean[order(x1)][n/10],pch="X",cex=1.5)
points(x1[order(x1)][n/4],yhat1mean[order(x1)][n/4],pch="X",cex=1.5)
points(x1[order(x1)][n/2],yhat1mean[order(x1)][n/2],pch="X",cex=1.5)
points(x1[order(x1)][n*3/4],yhat1mean[order(x1)][n*3/4],pch="X",cex=1.5)
points(x1[order(x1)][n*9/10],yhat1mean[order(x1)][n*9/10],pch="X",cex=1.5)
#textxy(x1[order(x1)][n/10],yhat1mean[order(x1)][n/10],"25",cx=1.5)

#plot(x1,dydx1mean,pch = ".",ylim=c(0,1.5))
#points(x1,CIdydx1high,pch=".")
#points(x1,CIdydx1low,pch=".")

#------

x2difference<-	.01
x2high	<-	rep(x2standard+x2difference,n)
x2low		<-	rep(x2standard-x2difference,n)
x2middle	<-	rep(x2standard,n)

x2mathcom	<-
	paste("data.frame(x1=x1,x2=x2high,",(paste(xnam,"=",xmeannam,collapse=",")),",v1=rep(0,n),v2
=rep(0,n))")
x2mathigh	<-	eval(parse(text=x2mathcom))
x2matlcom	<-
	paste("data.frame(x1=x1,x2=x2low,",(paste(xnam,"=",xmeannam,collapse=",")),",v1=rep(0,n),v2
=rep(0,n))")
x2matlow	<-	eval(parse(text=x2matlcom))
x2matmcom	<-
	paste("data.frame(x1=x1,x2=x2middle,",(paste(xnam,"=",xmeannam,collapse=",")),",v1=rep(0,n),
v2=rep(0,n))")
x2matmiddle	<-	eval(parse(text=x2matmcom))
highpredict2<-	predict.gam(EqGamNP,x2mathigh,type="lpmatrix")
lowpredict2	<-	predict.gam(EqGamNP,x2matlow,type="lpmatrix")
midpredict2 <-	predict.gam(EqGamNP,x2matmiddle,type="lpmatrix")

highpredictvalue2 <- highpredict2 %*% coefNP
lowpredictvalue2 <- lowpredict2 %*% coefNP
elas2 <- (highpredictvalue2-lowpredictvalue2)/(2*x2difference)
midpredictvalue2 <- midpredict2 %*% coefNP
plot(x1,elas2)
dydx2<-elas2*x2standard/midpredictvalue2
plot(x1,dydx2)

highpredsample2 <- rep(NA,dim(br)[1]*n)
dim(highpredsample2) <- c(n,dim(br)[1])
lowpredsample2  <- rep(NA,dim(br)[1]*n)
dim(lowpredsample2) <- c(n,dim(br)[1])
```

```
dydx2sample <- rep(NA,dim(br)[1]*n)
dim(dydx2sample) <- c(n,dim(br)[1])
elassample2 <- rep(NA,dim(br)[1]*n)
dim(elassample2) <- c(n,dim(br)[1])
yhat2 <- rep(NA,n*dim(br)[1])
dim(yhat2) <- c(n,dim(br)[1])
yhatulsample2 <- rep(NA,n*dim(br)[1])
dim(yhatulsample2) <- c(n,dim(br)[1])

for (i in 1:dim(br)[1])
        {
        highpredsample2[,i] <- highpredict2[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]
        lowpredsample2[,i]  <- lowpredict2[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]
        elassample2[,i] <- (highpredsample2[,i]-lowpredsample2[,i])/(2*x2difference)
#       yhatulsample2[,i] <- exp(midpredict2[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]+xboffset+.5*rmse^2)
        yhat2[,i] <-
midpredict2[,(dim(X)[2]+2):(dim(X)[2]+30)]%*%br[i,(dim(X)[2]+2):(dim(X)[2]+30)]+xboffset
#       dydx2sample[,i] <- elassample2[,i]*yhatulsample2[,i]/exp(x2standard)
        }



elas2mean <- rep(NA,n)
CIelas2high <- rep(NA,n)
CIelas2low  <- rep(NA,n)
dydx2mean <- rep(NA,n)
CIdydx2high <- rep(NA,n)
CIdydx2low <- rep(NA,n)

for (i in 1:n)
        {
        elas2mean[i] <- mean(elassample2[i,])
        CIelas2high[i] <- quantile(elassample2[i,],probs=.975)
        CIelas2low[i] <- quantile(elassample2[i,],probs=.025)
#       CIelas2high[i] <- mean(elassample2[i,])+1.96*sd(elassample2[i,])
#       CIelas2low[i] <- mean(elassample2[i,])-1.96*sd(elassample2[i,])
#       dydx2mean[i] <- mean(dydx2sample[i,])
#       CIdydx2high[i] <- quantile(dydx2sample[i,],probs=.975)
#       CIdydx2low[i] <- quantile(dydx2sample[i,],probs=.025)
#       CIdydx2high[i] <- mean(dydx2sample[i,])+1.96*sd(dydx2sample[i,])
#       CIdydx2low[i] <- mean(dydx2sample[i,])-1.96*sd(dydx2sample[i,])
        }

#engel curves for x2
x2mat.x1.10     <-
        paste("data.frame(x1=rep(quantile(x1,prob=.1),n),x2=x2,",(paste(xnam,"=",xmeannam,collapse=",
")),",v1=rep(0,n),v2=rep(0,n))")
x2mat.x1.10     <-      eval(parse(text=x2mat.x1.10))
x2mat.x1.50     <-
        paste("data.frame(x1=rep(quantile(x1,prob=.5),n),x2=x2,",(paste(xnam,"=",xmeannam,collapse=",
")),",v1=rep(0,n),v2=rep(0,n))")
x2mat.x1.50     <-      eval(parse(text=x2mat.x1.50))
```

```
x2mat.x1.90      <-
        paste("data.frame(x1=rep(quantile(x1,prob=.9),n),x2=x2,",(paste(xnam,"=",xmeannam,collapse=",
")),",v1=rep(0,n),v2=rep(0,n))")
x2mat.x1.90      <-      eval(parse(text=x2mat.x1.90))

x2matpredict.10 <-       predict.gam(EqGamNP,x2mat.x1.10,type="lpmatrix")
x2matpredict.50 <-       predict.gam(EqGamNP,x2mat.x1.50,type="lpmatrix")
x2matpredict.90 <-       predict.gam(EqGamNP,x2mat.x1.90,type="lpmatrix")


x2mat.x1.10sample        <- x2mat.x1.50sample <- x2mat.x1.90sample <- matrix(NA,n,dim(br)[1])

for (i in 1:dim(br)[1])
        {
        x2mat.x1.10sample[,i] <- x2matpredict.10[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]+xboffset
        x2mat.x1.50sample[,i] <- x2matpredict.50[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]+xboffset
        x2mat.x1.90sample[,i] <- x2matpredict.90[,(dim(X)[2]+2):(dim(X)[2]+30)] %*%
br[i,(dim(X)[2]+2):(dim(X)[2]+30)]+xboffset
        }

x2mat.10mean <- rep(NA,n)
x2mat.10hi       <-      rep(NA,n)
x2mat.10lo       <-      rep(NA,n)
x2mat.50mean <- rep(NA,n)
x2mat.50hi       <-      rep(NA,n)
x2mat.50lo       <-      rep(NA,n)
x2mat.90mean <- rep(NA,n)
x2mat.90hi       <-      rep(NA,n)
x2mat.90lo       <-      rep(NA,n)


for (i in 1:n)
        {
        x2mat.10mean[i] <-      mean(x2mat.x1.10sample[i,])
        x2mat.10hi[i]   <-      quantile(x2mat.x1.10sample[i,],prob=.975)
        x2mat.10lo[i]   <-      quantile(x2mat.x1.10sample[i,],prob=.025)
        x2mat.50mean[i] <-      mean(x2mat.x1.50sample[i,])
        x2mat.50hi[i]   <-      quantile(x2mat.x1.50sample[i,],prob=.975)
        x2mat.50lo[i]   <-      quantile(x2mat.x1.50sample[i,],prob=.025)
        x2mat.90mean[i] <-      mean(x2mat.x1.90sample[i,])
        x2mat.90hi[i]   <-      quantile(x2mat.x1.90sample[i,],prob=.975)
        x2mat.90lo[i]   <-      quantile(x2mat.x1.90sample[i,],prob=.025)
        }

plot(x2,x2mat.10mean,ylim = c(8.5,12.5),title(ylab="Predicted Log Total
Consumption",xlab=expression(paste("ln ",x[2]))),ylab="",xlab="",cex=.5)
points(x2,x2mat.10hi,pch=".")
points(x2,x2mat.10lo,pch=".")
points(x2,x2mat.90mean,cex=.5)
points(x2,x2mat.90hi,pch=".")
points(x2,x2mat.90lo,pch=".")
textxy(5.7,10,"10th non-remittance income percentile")
textxy(6,11.26,"90th non-remittance income percentile")
```

```
nrempos<-length(x2[x2>0])
points(x2[x2>0][order(x2[x2>0])][nrempos/4],x2mat.10mean[x2>0][order(x2[x2>0])][nrempos/4],pch="X
",cex=1.5)
points(x2[x2>0][order(x2[x2>0])][nrempos/4],x2mat.90mean[x2>0][order(x2[x2>0])][nrempos/4],pch="X
",cex=1.5)
points(x2[x2>0][order(x2[x2>0])][nrempos/2],x2mat.10mean[x2>0][order(x2[x2>0])][nrempos/2],pch="X
",cex=1.5)
points(x2[x2>0][order(x2[x2>0])][nrempos/2],x2mat.90mean[x2>0][order(x2[x2>0])][nrempos/2],pch="X
",cex=1.5)
points(x2[x2>0][order(x2[x2>0])][nrempos*3/4],x2mat.10mean[x2>0][order(x2[x2>0])][nrempos*3/4],pc
h="X",cex=1.5)
points(x2[x2>0][order(x2[x2>0])][nrempos*3/4],x2mat.90mean[x2>0][order(x2[x2>0])][nrempos*3/4],pc
h="X",cex=1.5)

#----------------------------

plot(x1,elas1mean,xlim=c(3.6,13.6),ylim=c(0,1),title(ylab="Elasticity of Total
Consumption",xlab=expression(paste("ln ",x[1]))),ylab="",xlab="",col="blue",cex=.5)
points(x1,CIelas1high,pch=".",col="blue")
points(x1,CIelas1low,pch=".",col="blue")
points(x1[order(x1)][n/10],elas1mean[order(x1)][n/10],pch="X",cex=1.5, col = "blue")
points(x1[order(x1)][n/4],elas1mean[order(x1)][n/4],pch="X",cex=1.5, col = "blue")
points(x1[order(x1)][n/2],elas1mean[order(x1)][n/2],pch="X",cex=1.5, col = "blue")
points(x1[order(x1)][n*3/4],elas1mean[order(x1)][n*3/4],pch="X",cex=1.5, col = "blue")
points(x1[order(x1)][n*9/10],elas1mean[order(x1)][n*9/10],pch="X",cex=1.5, col = "blue")

points(x1,elas2mean,col="red",cex=.5)
points(x1,CIelas2high,pch=".",col="red")
points(x1,CIelas2low,pch=".",col="red")
points(x1[order(x1)][n/10],elas2mean[order(x1)][n/10],pch="X",cex=1.5, col = "red")
points(x1[order(x1)][n/4],elas2mean[order(x1)][n/4],pch="X",cex=1.5, col = "red")
points(x1[order(x1)][n/2],elas2mean[order(x1)][n/2],pch="X",cex=1.5, col = "red")
points(x1[order(x1)][n*3/4],elas2mean[order(x1)][n*3/4],pch="X",cex=1.5, col = "red")
points(x1[order(x1)][n*9/10],elas2mean[order(x1)][n*9/10],pch="X",cex=1.5, col = "red")

c("quantile","elas1","se")
c(10,elas1mean[order(x1)][n/10],(CIelas1high[order(x1)][n/10]-elas1mean[order(x1)][n/10])/1.96)
c(25,elas1mean[order(x1)][n/4],(CIelas1high[order(x1)][n/4]-elas1mean[order(x1)][n/4])/1.96)
c(50,elas1mean[order(x1)][n/2],(CIelas1high[order(x1)][n/2]-elas1mean[order(x1)][n/2])/1.96)
c(75,elas1mean[order(x1)][n*3/4],(CIelas1high[order(x1)][n*3/4]-elas1mean[order(x1)][n*3/4])/1.96)
c(90,elas1mean[order(x1)][n*9/10],(CIelas1high[order(x1)][n*9/10]-elas1mean[order(x1)][n*9/10])/1.96)

c("quantile","elas2","se")
c(10,elas2mean[order(x1)][n/10],(CIelas2high[order(x1)][n/10]-elas2mean[order(x1)][n/10])/1.96)
c(25,elas2mean[order(x1)][n/4],(CIelas2high[order(x1)][n/4]-elas2mean[order(x1)][n/4])/1.96)
c(50,elas2mean[order(x1)][n/2],(CIelas2high[order(x1)][n/2]-elas2mean[order(x1)][n/2])/1.96)
c(75,elas2mean[order(x1)][n*3/4],(CIelas2high[order(x1)][n*3/4]-elas2mean[order(x1)][n*3/4])/1.96)
c(90,elas2mean[order(x1)][n*9/10],(CIelas2high[order(x1)][n*9/10]-elas2mean[order(x1)][n*9/10])/1.96)

orderelassample1<-elassample1
orderelassample2<-elassample2
for (i in 1:dim(elassample1)[2])
        {
        orderelassample1[,i]<-elassample1[,i][order(x1)]
        orderelassample2[,i]<-elassample2[,i][order(x1)]
        }
```

```
#t.10<-sqrt(n)*(elas2mean[order(x1)][n/10]-
elas1mean[order(x1)][n/10])/sqrt((CIelas2high[order(x1)][n/10]-
#elas2mean[order(x1)][n/10])/1.96+(CIelas1high[order(x1)][n/10]-elas1mean[order(x1)][n/10])/1.96)

ttest.10<-t.test(orderelassample1[(n/10),],orderelassample2[(n/10),])
ttest.25<-t.test(orderelassample1[(n/4),],orderelassample2[(n/4),])
ttest.50<-t.test(orderelassample1[(n/2),],orderelassample2[(n/2),])
ttest.75<-t.test(orderelassample1[(n*3/4),],orderelassample2[(n*3/4),])
ttest.90<-t.test(orderelassample1[(n*9/10),],orderelassample2[(n*9/10),])

ttest.10
ttest.25
ttest.50
ttest.75
ttest.90

#plot(x1,dydx2mean,ylim=c(0,5))
#points(x1,CIdydx2high,pch=".")
#points(x1,CIdydx2low,pch=".")
"
#quantile remittance engel curves
x2ten		<-	quantile(x2[x2>0],prob=.1)
x2twentyfive<-	quantile(x2[x2>0],prob=.25)
x2fifty	<-	quantile(x2[x2>0],prob=.5)
x2sevfive	<-	quantile(x2[x2>0],prob=.75)
x2ninety<-	quantile(x2[x2>0],prob=.9)

x2mcten<-
	paste("data.frame(x1=x1,x2=rep(x2ten,n),",(paste(xnam,"=",xmeannam,collapse=",")),",v1=rep(0,
n),v2=rep(0,n))")
x2mctwfive	<-
	paste("data.frame(x1=x1,x2=rep(x2twentyfive,n),",(paste(xnam,"=",xmeannam,collapse=",")),",v1
=rep(0,n),v2=rep(0,n))")
x2mcfifty	<-
	paste("data.frame(x1=x1,x2=rep(x2fifty,n),",(paste(xnam,"=",xmeannam,collapse=",")),",v1=rep(
0,n),v2=rep(0,n))")
x2mcsevfive <-
	paste("data.frame(x1=x1,x2=rep(x2sevfive,n),",(paste(xnam,"=",xmeannam,collapse=",")),",v1=re
p(0,n),v2=rep(0,n))")
x2mcninety	<-
	paste("data.frame(x1=x1,x2=rep(x2ninety,n),",(paste(xnam,"=",xmeannam,collapse=",")),",v1=re
p(0,n),v2=rep(0,n))")

x2matten	<-	eval(parse(text=x2mcten))
x2mattwefv	<-	eval(parse(text=x2mctwfive))
x2matfifty	<-	eval(parse(text=x2mcfifty))
x2matsevfive<-	eval(parse(text=x2mcsevfive))
x2matninety	<-	eval(parse(text=x2mcninety))

predict10	<-	predict.gam(EqGamNP,x2matten,type="lpmatrix")
predict25	<-	predict.gam(EqGamNP,x2mattwefv,type="lpmatrix")
predict50	<-	predict.gam(EqGamNP,x2matfifty,type="lpmatrix")
predict75	<-	predict.gam(EqGamNP,x2matsevfive,type="lpmatrix")
predict90	<-	predict.gam(EqGamNP,x2matninety,type="lpmatrix")
```

```
yhat2.10sample          <-      rep(NA,n*nsample)
dim(yhat2.10sample)     <-      c(n,nsample)
yhat2.25sample          <-      rep(NA,n*nsample)
dim(yhat2.25sample)     <-      c(n,nsample)
yhat2.50sample          <-      rep(NA,n*nsample)
dim(yhat2.50sample)     <-      c(n,nsample)
yhat2.75sample          <-      rep(NA,n*nsample)
dim(yhat2.75sample)     <-      c(n,nsample)
yhat2.90sample          <-      rep(NA,n*nsample)
dim(yhat2.90sample)     <-      c(n,nsample)

for (i in 1:nsample)
{
        yhat2.10sample[,i]<-predict10[,(dim(X)[2]+2):(dim(X)[2]+30)] %*% coefsample[i,]
        yhat2.25sample[,i]<-predict25[,(dim(X)[2]+2):(dim(X)[2]+30)] %*% coefsample[i,]
        yhat2.50sample[,i]<-predict50[,(dim(X)[2]+2):(dim(X)[2]+30)] %*% coefsample[i,]
        yhat2.75sample[,i]<-predict75[,(dim(X)[2]+2):(dim(X)[2]+30)] %*% coefsample[i,]
        yhat2.90sample[,i]<-predict90[,(dim(X)[2]+2):(dim(X)[2]+30)] %*% coefsample[i,]
}

yhat2.10<-      rep(NA,n)
yhat2.25<-      rep(NA,n)
yhat2.50<-      rep(NA,n)
yhat2.75<-      rep(NA,n)
yhat2.90<-      rep(NA,n)
yhat2.10CIhi<-  rep(NA,n)
yhat2.25CIhi<-  rep(NA,n)
yhat2.50CIhi<-  rep(NA,n)
yhat2.75CIhi<-  rep(NA,n)
yhat2.90CIhi<-  rep(NA,n)
yhat2.10CIlo<-  rep(NA,n)
yhat2.25CIlo<-  rep(NA,n)
yhat2.50CIlo<-  rep(NA,n)
yhat2.75CIlo<-  rep(NA,n)
yhat2.90CIlo<-  rep(NA,n)

for (i in 1:n)
{
        yhat2.10[i]             <-      mean(yhat2.10sample[i,])+xboffset
        yhat2.10CIhi[i]   <-    quantile(yhat2.10sample[i,]+xboffset,prob=.975)
        yhat2.10CIlo[i]   <-    quantile(yhat2.10sample[i,]+xboffset,prob=.025)
        yhat2.25[i]             <-      mean(yhat2.25sample[i,])+xboffset
        yhat2.25CIhi[i]   <-    quantile(yhat2.25sample[i,]+xboffset,prob=.975)
        yhat2.25CIlo[i]   <-    quantile(yhat2.25sample[i,]+xboffset,prob=.025)
        yhat2.50[i]             <-      mean(yhat2.50sample[i,])+xboffset
        yhat2.50CIhi[i]   <-    quantile(yhat2.50sample[i,]+xboffset,prob=.975)
        yhat2.50CIlo[i]   <-    quantile(yhat2.50sample[i,]+xboffset,prob=.025)
        yhat2.75[i]             <-      mean(yhat2.75sample[i,])+xboffset
        yhat2.75CIhi[i]   <-    quantile(yhat2.75sample[i,]+xboffset,prob=.975)
        yhat2.75CIlo[i]   <-    quantile(yhat2.75sample[i,]+xboffset,prob=.025)
        yhat2.90[i]             <-      mean(yhat2.90sample[i,])+xboffset
        yhat2.90CIhi[i]   <-    quantile(yhat2.90sample[i,]+xboffset,prob=.975)
        yhat2.90CIlo[i]   <-    quantile(yhat2.90sample[i,]+xboffset,prob=.025)
}

plot(x1,yhat2.10)
```

```
points(x1,yhat2.10CIhi,pch=".")
points(x1,yhat2.10CIlo,pch=".")
points(x1,yhat2.25,pch=".")
points(x1,yhat2.50,pch=".")
points(x1,yhat2.75,pch=".")
points(x1,yhat2.90,pch=".")
"


anova(EqLin,EqGamNP,test="F")
anova(EqQuad,EqGamNP,test="F")
anova(EqGamAS,EqGamNP,test="F")

#F<-.26/5.81/(774.62/3777.73)
#1-pf(F,5.81,3771.92)
```

Appendix D: Stata Program for Poverty Analyses

This Appendix contains the Stata program used for calculations in Chapter 4 of this dissertation.

```
insheet using "D:\WDI Data\Rem Data.csv",clear
save "C:\Users\Mike\Desktop\papers\Cross-national study\Stata files\Remittances received percent GDPEE2.dta",replace
/*gdp data from wdi "gdp per capita"*/


insheet using "C:\Users\Mike\Desktop\papers\Cross-national study\DistanceMatrix.csv",clear
save "C:\Users\Mike\Desktop\papers\Cross-national study\Stata files\DistanceMatrix.dta",replace

insheet using "C:\Users\Mike\Desktop\papers\Cross-national study\InstGDP.csv",clear
save "C:\Users\Mike\Desktop\papers\Cross-national study\Stata files\InstGDP.dta",replace

insheet using "D:\WDI Data\PovertyEE.csv", clear
joinby countrycode year using "C:\Users\Mike\Desktop\papers\Cross-national study\Stata files\Remittances received percent GDPEE2.dta",unmatched(both)
drop _merge

joinby countrycode using "C:\Users\Mike\Desktop\papers\Cross-national study\Stata files\DistanceMatrix.dta",unmatched(both)
drop _merge

joinby year using "C:\Users\Mike\Desktop\papers\Cross-national study\Stata files\InstGDP.dta",unmatched(both)
drop _merge

/*dropping poverty measure from eca report*/
/* drop gdppercapit2000ppp */

tsset countrycode year

drop if countrycode>=24 & countrycode <=26
gen poverty = p0215/100
gen lpov = ln(poverty+1)
gen lpov1 = ln(p1215/100+1)
gen lpov2 = ln(p2215/100+1)
gen lpov4=ln(p0430/100+1)
gen lpov14=ln(p1430/100+1)
gen lpov24=ln(p2430/100+1)


gen lgdp = ln(gdppcppp+1)
gen lgin = ln(ginicoefficient+1)
gen lpovrural = ln(rural215+1)
gen lpovurban = ln(urban215+1)
gen lenroll = ln(ssenroll+1)

gen InstGDPWeight = germanygdp/gerdist+russiagdp/russdist+unitedstatesgdp/usdist
```

```
gen InstGDPWeight2 = germanygdp/gerdist^2+russiagdp/russdist^2+unitedstatesgdp/usdist^2
gen lInstGDPWeight = ln(germanygdp)/gerdist+ln(russiagdp)/russdist+ln(unitedstatesgdp)/usdist
gen lInstGDPWeight2 = ln(germanygdp)/gerdist^2+ln(russiagdp)/russdist^2+ln(unitedstatesgdp)/usdist^2
gen lInstGDPWeights =
ln(germanygdp)/ln(gerdist)+ln(russiagdp)/ln(russdist)+ln(unitedstatesgdp)/ln(usdist)
gen lInstGDPWeights2 =
ln(germanygdp)/ln(gerdist^2)+ln(russiagdp)/ln(russdist^2)+ln(unitedstatesgdp)/ln(usdist^2)
gen InstGer = (germanygdp)/(gerdist^2)
gen InstRuss = russiagdp/russdist^2
gen InstUS = unitedstatesgdp/usdist^2


gen pcremppp = rempct*gdppcppp*.01
gen lrem = ln(pcremppp+1)
gen tradepc = ln(imports+exports+1)/population
gen tradeopen = tradepc/gdppercapit2000ppp

/* dropping russia */
drop if countrycode == 18
/*
xtabond poverty lgdp ginicoefficient
xtreg poverty lgdp ginicoefficient,fe
/* xtreg poverty lgdp ginicoefficient,be */
xtreg poverty lgdp ginicoefficient

mean gdppcppp if lpov!=.
mean ginicoefficient if lpov!=.
mean pcremppp if lpov!=.

mean lpov
mean lpov1
mean lpov2

xtabond poverty lgdp ginicoefficient lrem
xtreg poverty lgdp ginicoefficient lrem,fe
/* xtreg poverty lgdp ginicoefficient lrem,be */
xtreg poverty lgdp ginicoefficient lrem
reg poverty lgdp ginicoefficient lrem
*/

xtabond lpov lgdp lgin
xtreg lpov lgdp lgin, fe
xtreg lpov lgdp lgin, be
xtreg lpov lgdp lgin

gen dlrem = D.lrem
gen d2lrem= D2.lrem
reg lrem lgdp ginicoefficient dlrem d2lrem lInstGDPWeight2
predict lremfitted, xb

/*
drop lremfitted
reg lrem lgdp lgin dlrem d2lrem InstGer InstRuss InstUS
predict lremfitted, xb
*/
```

```
reg lpov lgdp lgin lrem
xtabond lpov lgdp lgin lrem
xtreg lpov lgdp lgin lrem, fe
xtreg lpov lgdp lgin lrem, be
xtreg lpov lgdp lgin lrem

/*these are the results I use*/

ivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re
xtoverid

ivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re
xtoverid

ivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re
xtoverid


ivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re
xtoverid

ivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re
xtoverid

ivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re
xtoverid

/* Robustness checks: */
/* Different instruments*/
/*
ivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen)
xtivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), fe
xtivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), re

ivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen)
xtivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), fe
xtivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), re

ivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen)
xtivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), fe
xtivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), re

ivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen)
xtivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), fe
```

```
xtivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), re

ivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen)
xtivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), fe
xtivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), re

ivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen)
xtivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), fe
xtivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 tradeopen), re
*/

ivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll)
xtivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), fe
xtivreg lpov lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), re

ivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll)
xtivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), fe
xtivreg lpov1 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), re

ivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll)
xtivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), fe
xtivreg lpov2 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), re

ivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll)
xtivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), fe
xtivreg lpov4 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), re

ivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll)
xtivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), fe
xtivreg lpov14 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), re

ivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll)
xtivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), fe
xtivreg lpov24 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2 lenroll), re

/*urban, rural analysis*/

xtivreg lpovrural lgdp lgin (lrem = lgdp ginicoefficient dlrem d2lrem lInstGDPWeight2), fe
xtivreg lpovrural lgdp lgin (lrem = lgdp ginicoefficient dlrem d2lrem lInstGDPWeight2), re
xtivreg lpovurban lgdp lgin (lrem = lgdp ginicoefficient dlrem d2lrem lInstGDPWeight2), fe
xtivreg lpovurban lgdp lgin (lrem = lgdp ginicoefficient dlrem d2lrem lInstGDPWeight2), re

/* no log of poverty rate */

ivreg p0215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg p0215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg p0215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re

ivreg p1215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg p1215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg p1215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re

ivreg p2215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg p2215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg p2215 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re
```

```
ivreg p0430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg p0430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg p0430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re

ivreg p1430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg p1430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg p1430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re

ivreg p2430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2)
xtivreg p2430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), fe
xtivreg p2430 lgdp lgin (lrem = lgdp lgin dlrem d2lrem lInstGDPWeight2), re

/* no endogenization */

reg lpov lgdp lgin lrem
xtreg lpov lgdp lgin lrem, fe
xtreg lpov lgdp lgin lrem, re

reg lpov1 lgdp lgin lrem
xtreg lpov1 lgdp lgin lrem, fe
xtreg lpov1 lgdp lgin lrem, re

reg lpov2 lgdp lgin lrem
xtreg lpov2 lgdp lgin lrem, fe
xtreg lpov2 lgdp lgin lrem, re

reg lpov4 lgdp lgin lrem
xtreg lpov4 lgdp lgin lrem, fe
xtreg lpov4 lgdp lgin lrem, re

reg lpov14 lgdp lgin lrem
xtreg lpov14 lgdp lgin lrem, fe
xtreg lpov14 lgdp lgin lrem, re

reg lpov24 lgdp lgin lrem
xtreg lpov24 lgdp lgin lrem, fe
xtreg lpov24 lgdp lgin lrem, re

/*Aggregated by country:*/
keep if lpov!=.
collapse (mean) lpov lpov1 lpov2 lpov4 lpov14 lpov24 lgdp lgin lInstGDPWeight2 lrem, by(countrycode)
ivreg lpov lgdp lgin (lrem = lgdp lgin lInstGDPWeight2)
ivreg lpov lgdp lgin lrem
```

References

Adams, Jr., Richard H. 2005. "Remittances, Household Expenditure, and Investment in Guatemala." World Bank Policy Research Working Paper no. 3532. Washington, DC: World Bank.

Alam, Asad, Mamta Murthi, Ruslan Yemtsov, Edmundo Murrugarra, Nora Dudwick, Ellen Hamilton, and Erwin Tiongsin. 2005. *Growth, Poverty, and Inequality: Eastern Europe and the Former Soviet Union.* World Bank: Washington, DC.

Amin, Shahina, M. Shakil Quayes, and Janet M. Rives. 2004. "Poverty and Other Determinants of Child Labor in Bangladesh." *Southern Economic Journal* 70: 2, 876-892.

Arrehag, Lisa, Örjan Sjöberg, and Mirja Sjöblom. 2005. "Cross-Border Migration and Remittances in a Post Communist Society: Return Flows of Money and Goods in the Korçë District, Albania." *South Eastern Europe Journal of Economics* 1, 9-40.

Baker, Rachel and Rachel Hinton. 2001. "Approaches to Children's Work and Rights in Nepal." *Annals of the American Academy of Political and Social Science* 575, 176-193.

Basu, Kaushik and Pham Hoang Van. 1998. "The Economics of Child Labor." *The American Economic Review* 88:3, 412-427.

Beneito, P. 2003. "A complete system of Engel curves in the Spanish economy." *Applied Economics* 35, 803-816.

Bhalotra, Sonia and Clif Attfield. 1998. "Intrahousehold Resource Allocation in Rural Pakistan: A Semiparametric Analysis." *Journal of Applied Econometrics* 13:5, 463-480.

Blundell, Richard, and Alan Duncan. 1998. "Kernel Regression in Empirical Microeconomics." *Journal of Human Resources* 33:1, 62-87.

Blundell, Richard, Alan Duncan, and Krishna Pendakur. 1998. "Semiparametric Estimation and Consumer Demand." *Journal of Applied Econometrics* 13:5, 435-461.

Blundell, Richard and James L. Powell. 2001. "Endogeneity in Nonparametric and Semiparametric Regression Models." CEMMAP working paper CWP09/01. London: Center for Microdata Methods and Practice.

Blundell, Richard W., Martin Browning, and Ian A. Crawford. 2003. "Nonparametric Engel Curves and Revealed Preference." *Econometrica* 71:1, 205-240.

Blundell, Richard. 2005. "How Revealing Is Revealed Preference?" *Journal of the European Economic Association* 3:2-3, 211-235.

CARANA Corporation. 2004. "Diasporas, Émigrés and Development : Economic Linkages and Programmatic Respons." Special Study under USAID Trade Enhancement for the Services Sector Project.

Central Bureau of Statistics (CBS). 2004. Nepal Living Standards Survey Statistical Report. Kathmandu: CBS.

Chami, Ralph, Connel Fullenkamp and Samir Jahjah. 2003. "Are Immigrant Remittance Flows a Source of Capital for Development?" IMF Working Paper WP/03/189. Washington, DC: International Monetary Fund.

Cox, Alejandra and Manuelita Ureta. 2003. "International Migration, Remittances, and Schooling: Evidence from El Salvador." NBER Working Paper 9766. Cambridge, Massachusetts: National Bureau of Economic Research.

Edmonds, Eric V. and Nina Pavcnik. 2005. "Child Labor in the Global Economy." *The Journal of Economic Perspectives* 19:1, 199-220.

Federal Deposit Insurance Corporation (FDIC). 2004. "Linking International Remittance Flows to Financial Services: Tapping the Latino Immigrant Market." *Supervisory Insights* 4:2, pp. 17-24.

Ferrari, Aurora, Guillemette Jaffrin, and Sabin Shrestha. 2007. *Access to Financial Services in Nepal.* Washington, DC: World Bank Publications.

Friedman, Milton. 1957. *A Theory of the Consumption Function.* Princeton University Press.

Gong, Xiaodong, Arthur van Soest, and Ping Zhang. 2005. "The Effects of the Gender of Children on Expenditure Patterns in Rural China: A Semiparametric Analysis." *Journal of Applied Econometrics* 20, 509-527.

Graner, Elvira and Ganesh Gurung. 2003. "Arab Ko Lahure: Looking at Nepali Labour Migrants to Arabian Countries." *Contributions to Nepalese Studies* 30:2, 295-325.

Hamilton, Eve and Manuel Orozco. 2006. "Remittance, Diasporas, and Economic Development: Issues, Lessons Learned, and Recommendations for Donor Interventions. Chemonics International, Inc.

Huber, P.J. 1981. *Robust Statistics*. New York: John Wiley and Sons.

Inter-American Development Bank (IADB). 2009. "The Multi-Lateral Investment Fund: Promoting Inclusive Growth through Private Sector Investment." http://www.iadb.org/mif/projects.cfm?language=English. Accessed March 1, 2009.

Jensen, Peter and Helena Skyt Nielson. 1997. "Child Labor or School Attendance? Evidence from Zambia." *Journal of Population Economics* 10, 407-424.

Johnston, J. *Econometric Methods*. McGraw-Hill.

Keely, Charles B. and Bao Nga Tran. 1989. "Remittances from Labor Migration: Evaluations, Performances, and Implications." *International Migration Review* 23:3, 500-525.

Kireyev, Alexei. 2006. "The Macroeconomics of Remittances: The Case of Tajikistan." International Monetary Fund Working Paper 602.

Krinsky, Itzhak and A. Leslie Robb. 1986. "On Approximating the Statistical Properties of Elasticities." *Review of Economics and Statistics* 68:4, 715-719.

Lopez Cordoba, Ernesto. 2004. "Globalization, Migration, and Development: The Role of Mexican Migrant Remittances." New York: Inter-American Development Bank.

Lyssiotou, Panayiota, Panos Pashardes, and Thanasis Stengos. 2002. "Age effects on consumer demand: an additive partially linear regression model." *Canadian Journal of Economics* 35:1, 153-165.

McKenzie, David J. 2005. "Beyond Remittances: The Effects of Migration on Mexican Households" in Maurice Shiff and Çaḡlar Özden (eds.), *International Remittances, Migration, and the Brain Drain*. Washington, DC: World Bank, 123-147.

Modigliani, Franco. 1986. "Life Cycle, Individual Thrift, and the Wealth of Nations." *American Economic Review* 76(June): 297-313.

Newey, Whitney K., James L. Powell, and Francis Vella. 1999. "Nonparametric Estimation of Triangular Simultaneous Equations Models." *Econometrica* 67:3, 565-603.

Orozco, Manuel. 2004. "International Financial Flows and Worker Remittances." United Nations Population and Mortality division.

Orozco, Manuel. 2007. "Looking Forward and Including Migration in Development: Remittance Leveraging Opportunities for Moldova." International Organization for Migration.

Pinger, Pia R. 2007. "Come Back or Stay? Spend Here or There? Temporary versus Permanent Migration and Remittance Patterns in the Republic of Moldova." Kiel Institute for the World Economy working paper 438.

R Development Core Team. 2009. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ratha, Dilip, Sanket Mohapatra and Zhimei Xu. 2008. "Outlook for Remittance Flows 2008-2010: Growth expected to moderate significantly, but flows to remain resilient." Migration and Development Brief 8. Washington, DC: World Bank.

Ruan, Xiaomin, Kristina Grimsrud, and Alok K. Bohara. 2009. "The Evidence of the Impact of Intra Household Balance of Power on Children's Welfare: Evidence from Nepal." Unpublished manuscript, Department of Economics, University of New Mexico.

Stahl, Charles W. and Fred Arnold. 1986. "Overseas Workers' Remittances in Asian Development." *International Migration Review* 20: 4, 899-925.

State Secretariat for Economic Affairs (SECO). 2007. *Development Financing and the Remittance Market in Serbia and Switzerland***.** Geneva: SECO, 2007.

Stahl CW and Arnold F. 1986. "Overseas Workers' Remittances in Asian Development." *International Migration Review* 20:4, 899-925.

Thieme, Susan. 2003. "Savings and Credit Associations and Remittances: The Case of Far West Nepalese Labour Migrants in Delhi, India." Social Finance Programme Working Paper 39. Geneva, Switzerland: International Labour Office.

Thieme, Susan and Simone Wyss. 2005. "Migration Patterns and Remittance Transfer in Nepal: A Case Study of Sainik Basti in Western Nepal." *International Migration* 43:5, 59-98.

Wood, Simon N. 2003. "Thin plate regression splines." *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 65:1, 95-114.

Wood, Simon N. 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall.

Wood, Simon N. 2009. mgcv: GAMs with GCV smoothness estimation and GAMMs by REML/PQL. Software package for use with R statistical program.

World Bank. 2006. *Global Economic Prospects 2006: Economic Impacts of Remittances and Migration*. Washington, DC: World Bank Publications.

World Bank. 2008. "World Development Indicators." Washington, DC: World Bank Group.

World Economic Outlook (2005). *Globalization and External Imbalances*. International Monetary Fund: Washington, DC.

Yang, Dean. 2006. "International Migration, Remittances, and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks." NBER Working Paper 12325, National Bureau of Economic Research.