FACULTY OF SOCIAL SCIENCES
Department of Economics
University of Copenhagen

PhD Thesis

Ulrik Richardt Beck

Essays in Development Economics: Inequality Measurement and Household Factor Allocations

Academic advisor: Finn Tarp
Submitted: March 31, 2016

Contents

Acknowledgements

This dissertation is the result of my doctoral studies, which have taken place over the last four years at the Department of Economics at University of Copenhagen. Many people have been crucially important for me during this time and I wish to thank them for that.

First, I would like to express my sincere gratitude for the incomparable guidance and support that I have received from my advisor Professor Finn Tarp. I am also grateful for Finn's efforts to involve me in the Vietnam research program of the Development Economics Research Group at University of Copenhagen, as well as the Growth And Poverty Project, hosted at UNU-WIDER, which have made this dissertation possible. I must also thank Finn for facilitating my research visit at UC-Berkeley.

I owe many thanks to my co-author and friend Benedikte Bjerge. The collaborations between Benedikte and I have been both extremely enjoyable and productive. I also wish to thank my other co-authors, Marcel Fafchamps, Saurabh Singhal and Finn Tarp, for many thought-provoking discussions and great cooperation on the various papers. The environment and the people of the Development Economics Research Group have played a critical role in getting critiques, comments and encouragements over the years, which have been very important for me to make it to the finish line. In particular, I wish to thank Sam Jones and John Rand for many valuable comments and fruitful discussions.

I wish to thank the country teams of the Growth And Poverty Project who worked on Ethiopia, Malawi, Madagascar, Mozambique, Pakistan and Tanzania for letting me build upon their data work for Chapter 2. The empirics of chapter 3 and 4 are based on the *Gambia Networks Data 2009* dataset. I wish to thank Jean-Louis Arcand and Dany Jaimovich for graciously granting access to the data.

In the spring semester of 2014, I was a visiting student researcher at the Department of Agriculture and Resource Economics at University of California-Berkeley. I wish to thank Professors Alain de Janvry and Elisabeth Sadoulet for hosting me as well as providing stimulating input to my research.

Finally, I am thankful for the support I have received from my parents, my sister and my friends throughout the process. Last but not least, I am grateful to Mira for her unending love and support through this long and sometimes winding journey. Mira, this dissertation is only a small subset of the things I could not have achieved without you.

Ulrik Richardt Beck
Copenhagen, March 2016

Summary

The dissertation consists of an introductory chapter and four self-contained chapters.

Chapter 2, Keep It Real: Measuring Real Inequality Using Survey Data from Developing Countries, is concerned with the measurement of real inequality in developing countries. In particular, I investigate how two separate effects can drive wedges between inequality estimates based on nominal consumption aggregates and estimates based on real consumption aggregates. The first effect is caused by differences in the composition of consumption over the income distribution coupled with differential inflation of different consumption items. The second effect is caused by increasing quantity discounting as one moves up through the consumption distribution. I further argue that poverty estimation based on GDP data of national accounts and inequality estimates of consumption surveys should employ real, rather than nominal, inequality estimates. I estimate the magnitude of these effects using 15 nationally representative surveys from six countries (Ethiopia, Madagascar, Malawi, Mozambique, Pakistan and Tanzania) covering the period 1999–2011. In some of the studied countries, real inequality is higher than nominal inequality. This increases the level and reduces the decline of poverty over time, but the magnitude of the correction is country- and year-specific.

Chapter 3, Social Ties and the Efficiency of Factor Transfers, which is co-authored with Benedikte Bjerge and Marcel Fafchamps, introduces a test of whether social ties help or hinder efficiency-enhancing factor transfers. In particular, we test the impact of family connections, ethnic groups and geographical proximity. Our findings indicate that neighbors conduct more efficiency-enhancing transfers but that households that are members of the same ethnic group or kinship network conduct fewer efficiency-enhancing transfers. However, we find the latter effects to be driven by the presence of a few large landowners. When the presence of these large landowners is explicitly addressed, the finding is reversed and there are more efficiency-enhancing land transfers between kin-related households and between neighbors. Allocative efficiency in land is not achieved at the village level. This suggests that social ties do not reduce transfer barriers sufficiently to permit an efficient reallocation of land within villages.

In **Chapter 4, Insuring the Poor: Inter-household Land Transfers and the Importance of Land Abundance and Ethnicity in The Gambia**, which is co-authored with Benedikte Bjerge, we study whether patterns of land transactions in rural villages in The Gambia are consistent with the presence of norm-based land access rules, of which there is ample qualitative evidence. Our answer to this question is affirmative, as we find that land transactions are pro-poor on average. In particular, households in the lowest income quartile receive more land, and households in the richest income quartile donate more land. The result is stronger in less densely populated and less ethnically diverse villages where social norms are thought to be more important. The findings have implications for redistributive land reforms. We argue that while the impacts of

previous reforms may have been offset by a decrease in inter-household land transfers, this should not necessarily weigh against future reform efforts as local social security institutions break down due to increased population pressure.

In **Chapter 5, Coffee Price Volatility and Household Response: Evidence from Vietnam**, which is co-authored with Saurabh Singhal and Finn Tarp, we investigate responses among smallholder coffee farmers in the Central Highlands of Vietnam to fluctuations in the coffee price. We employ a panel dataset collected between 2006 and 2014. We find that coffee farmers are unable to perfectly smooth consumption when faced with changes in the coffee price. Instead, adults and adolescents of the household increase their off-farm wage labor supply as a coping strategy. We also find that children and adolescents increase their supply of on-farm labor. These findings are worrying for human capital formation of children and adolescents in these households.

Resumé (Danish summary)

Denne afhandling består af et introducerende kapitel samt fire kapitler, der kan læses hver for sig.

Kapitel 2, Keep It Real: Measuring Real Inequality Using Survey Data from Developing Countries, beskæftiger sig med måling af real ulighed i udviklingslande. Jeg undersøger hvordan to effekter kan medføre forskelle mellem estimater af ulighed, der baserer sig på henholdsvis nominelle og reale mål af forbrug. Den første effekt opstår på grund af forskelle i sammensætningen af forbrug over indkomstfordelingen i kombination med forskellig inflation for forskellige forbrugsgoder. Den anden effekt skyldes, at de relativt rige husholdninger opnår større mængderabatter. Jeg argumenterer for, at når fattigdom estimeres på baggrund af forbrugsdata fra nationalregnskabet samt estimater af ulighed, der stammer fra spørgeskemaundersøgelser om forbruget, bør estimater af real, snarere end af nominel, ulighed benyttes. Jeg estimerer størrelsen af de to effekter i 15 nationalt representative spørgeskemaundersøgelser fra seks forskellige lande (Etiopien, Madagaskar, Malawi, Mozambique, Pakistan og Tanzania). Alle undersøgelserne er indsamlet i perioden 1999 til 2011. I visse lande er den reale ulighed højere end den nominelle ulighed. Dette øger fattigdomsniveauet og reducerer faldet i fattigdom over tid, men størrelsen af effekterne afhænger både af hvilket land der studeres, og på hvilket tidspunkt.

Kapitel 3, Social Ties and the Efficiency of Factor Transfers, som er skrevet i samarbejde med Benedikte Bjerge og Marcel Fafchamps, introducerer en ny metode til at teste hvorvidt sociale bånd såsom familieforbindelser, etniske grupperinger og geografisk nærhed øger eller forringer muligheden for efficiensforbedrende overførsler af produktionsfaktorer. Vores resultater indikerer, at naboskab medfører flere efficiensforbedrende overførsler men at medlemsskab af den samme etniske gruppe eller det samme familienetværk er forbundet med færre efficiensforbedrende overførsler. Denne negative effekt kan tilskrives eksistensen en lille gruppe husholdninger, der ejer meget jord. Når vi kontrollerer for tilstedeværelsen af disse jordejere, finder vi et omvendt resultat, nemlig at der er flere efficiensforbedrende landoverførsler mellem husholdninger, der tilhører den samme udvidede familie samt mellem naboer. Allokativ efficiens i fordelingen af jord opnås ikke på landsbyniveau. Dette indikerer, at sociale bånd ikke er tilstrækkelige til at opnå en efficient reallokering af produktionsfaktorerne indenfor de enkelte landsbyer.

I **Kapitel 4, Insuring the Poor: Inter-household Land Transfers and the Importance of Land Abundance and Ethnicity in Gambia**, som er skrevet i samarbejde med Benedikte Bjerge, studerer vi hvorvidt jordtransaktioner i landsbyer i Gambia foregår i overensstemmelse med tilstedeværelsen af norm-baserede regler for adgang til jord. Sådanne regler er blevet beskrevet i talrige kvalitative studier. Vores svar på dette spørgsmål er bekræftende. Vores resultater indikerer, at den gennemsnitlige transak-

tion af jord er til fordel for de fattigste. I særdeleshed er det tilfældet for husholdninger, der tilhører den fattigste indkomstkvartil. Disse husholdninger modtager mere jord jo fattigere de er, og husholdninger i den rigeste indkomstkvartil donerer mere jord til dem, jo rigere de er. Resultatet skyldes stærke effekter i landsbyer med relativt lave befolkningstætheder og relativt lave niveauer af etnisk diversitet; netop hvor sociale normer menes at være vigtigst. Vi argumenterer for, at selvom effekten af tidligere jordreformer, der forsøgte at omfordele ejerskab af jord, muligvis er blevet kompenseret af en samtidig reduktion i landtransaktioner mellem husholdninger, kan fremtidige reformer blive nødvendige, når de lokale sikkerhedsnet bryder sammen i takt med, at befolkningstætheden stiger.

I **kapitel 5, Coffee Price Volatility and Household Response: Evidence from Vietnam**, som er skrevet i samarbejde med Saurabh Singhal og Finn Tarp, undersøger vi, hvordan kaffebønder i Det Centrale Højland i Vietnam reagerer på ændringer i kaffeprisen. Vi benytter et paneldatasæt indsamlet mellem 2006 og 2014. Vores resultater indikerer, at kaffebønder er ude af stand til at udglatte deres forbrug over tid, når kaffeprisen ændrer sig. I stedet finder vi, at unge og voksne i stigende grad tager lønarbejde andre steder end på deres egen gård. Vi finder også, at børn og unge i stigende grad arbejder på gården. Dette er bekymrende i forhold til deres muligheder for at akkumulere humankapital.

CHAPTER 1

INTRODUCTION

The term "developing countries" is at its core an optimistic one: Although it refers to countries that are currently underdeveloped, the term implies that they are "developing". To be sure, many developing countries have made tremendous gains over the last couple of decades (Chen and Ravallion, 2010). Perhaps most impressive is the growth of China, which has lifted hundreds of millions out of poverty in the last decades. However, other places in the world, in particular in Sub-Saharan Africa, things have not progressed as quickly.

This state of affairs begs questions along at least two dimensions. First, there are questions related to how to measure the extent of the progress that has taken place. Second, there are questions of behavior: Low income levels and the lack of well-functioning formal institutional frameworks twist incentives and restrict behavior. However, the ways in which this occurs differs in both space and time. These two broad questions have fascinated me since I started my PhD and it is within these areas that the four chapters of my dissertation lie.

Apart from this introduction, the dissertation contains four self-contained chapters. Chapter 2 adds to the discussion of how inequality has developed in some of these countries, with potential implications for the reduction of poverty. The next three chapters are concerned with the behavior of households living in countries that are still in the process of development. Chapter 3 and Chapter 4 focus on production factor transfers between households in rural Gambia. Chapter 3 asks to what extent transfers of production factors, namely land and labor, are efficiency improving and whether social ties help or hinder efficiency. Chapter 4 digs deeper into the motives behind land transfers in these villages. In Chapter 5, the focus shifts to Vietnam. The chapter investigates intra-household labor allocation decisions for coffee farmers when the price of coffee changes.

Another theme of the dissertation is an attempt to address the research questions at hand using the best available data. The research question of Chapter 2 called for a large and detailed consumption database at the household level, spanning multiple countries and multiple surveys within each country. A major part of the effort that went into this chapter was developing such a database. Chapters 3 and 4 are based on a highly detailed dataset of social and economic networks in rural villages in The Gambia. This dataset allowed us to cast new light on existing questions related to the properties of factor transfers in a rural developing economy. Finally, Chapter 5 is based on a panel survey of 12 Vietnamese provinces. The length of the panel, combined with low attrition rates and high-quality information, makes it possible to exploit temporal variation in the international coffee price to identify household responses following a change in the coffee price.

The dissertation contains a brief summary of the four self-contained chapters. In what follows below, I provide some additional context, which serves to further motivate each chapter.

1.1 Real inequality

Chapter 2 is concerned with the measurement of inequality. More specifically, it considers some of the issues related to using the same deflationary index for everyone. Price indices are necessary to compare expenditures separated by time or space in a meaningful way. Real gross domestic product (GDP) is a more useful construct than its nominal counterpart if one wants to estimate welfare changes over time: At a fundamental level, households do not get better off if all prices in the economy suddenly doubles. While price indices have traditionally been employed rigorously when estimating real GDP growth rates, price indices, or price deflators, are not normally used when inequality is measured. This is probably because inequality measures are unaffected by deflation procedures that apply a single price index to an entire cross-section. Temporal deflation, the most common form of deflation, belongs in this category.

However, this does not mean that prices and price changes do not matter for inequality. Chapter 2 considers two instances where they matter. First, when differences in consumption bundles over the consumption distribution are coupled with relative changes in prices over time, it leads to differential inflation across the income distribution. Second, differing prices over the consumption distribution due to quantity discounting effects can lead to differential real purchasing power across the income distribution. Chapter 2 investigates how inequality based on a real consumption aggregate, which accounts for these effects – called real inequality – and a more standard, nominal inequality measure, differs.

Why should one care about the measurement of real inequality? There are at least three reasons. First, many believe that the level of inequality is important in its own right. There is also empirical evidence of the detrimental effects of high levels of inequality on e.g., crime (Enamorado et al., 2016), institutions and schooling outcomes (Easterly, 2007) as well as GDP growth rates (Persson and Tabellini, 1994; Easterly, 2007; Ostry et al., 2014).

Second, while the pace of economic development is clearly important, the *quality* of economic development also matters a great deal. One such aspect is the pro-poorness of growth. While there is some debate regarding the exact definition of this term, the idea is clear: If the welfare gains from economic growth is captured by the upper end of the income distribution, leading to no or only slight improvements for the poorest part of the population, the growth that has occurred cannot be said to have been pro-poor. This intuition is formalized in the decomposition of poverty changes into the contribution of overall economic growth, and the contribution of distributional changes over time of (Datt and Ravallion, 1992). However, this approach does not take into account relative price shifts across the income distribution (Günther and Grimm, 2007). In other words, it fails to account for the difference between a nominal distributional change and a real distributional change.

Third, inequality estimates have more recently been used in an alternative approach to measuring poverty, pioneered in a series of papers by Sala-i-Martin and Pinkovskiy (Sala-i Martin, 2006; Pinkovskiy and Sala-i Martin, 2009, 2014; Sala-i Martin and Pinkovskiy, 2010). Relying on GDP estimates from national accounts and survey-based inequality estimates, the authors find rapid declines in poverty estimates. This stands in contrast to the more moderate findings of the standard approach that relies solely on household consumption data (Chen and Ravallion, 2010). I show how the use of real inequality estimates in some countries increases both the level of poverty and decrease the speed of poverty reduction, when applying the method of Sala-i-Martin and Pinkovskiy.

For the reasons mentioned above (and possibly others), inequality is a matter which is receiving increased attention. This is evident from the World Bank's focus since 2013 on shared prosperity, along with more traditional measures of economic growth. This has resulted in the World Bank's "shared prosperity indicator", which tracks the income growth of the bottom 40 percent (World Bank, 2013). It is also evident from the inclusion of a goal to "reduce inequality within and between countries", among the United Nations recently agreed set of Sustainable Development Goals (UN, 2016b); a goal that was absent from the previous iteration of UN development goals, the so-called Millennium Development Goals (UN, 2016a).

The data from Chapter 2 is based on country-specific databases that were constructed under the auspices of UNU-WIDER's Growth and Poverty Project (GAPP). I was part of the team that worked on the case of Malawi (Pauw et al., 2016; Beck et al., 2016). For Chapter 2, I have combined and standardized the databases for fifteen surveys from six different countries in order to create a unique dataset, which contains detailed information on the consumption patterns of over 220,000 households. This database allows comparisons across different countries without resorting to the more superficial summary statistics that are available from international databases, which analysts often rely on when moving beyond country case studies. The comparative approach, lost in single-country case studies, also turned out to be important, as the effects I find are highly country-dependent.

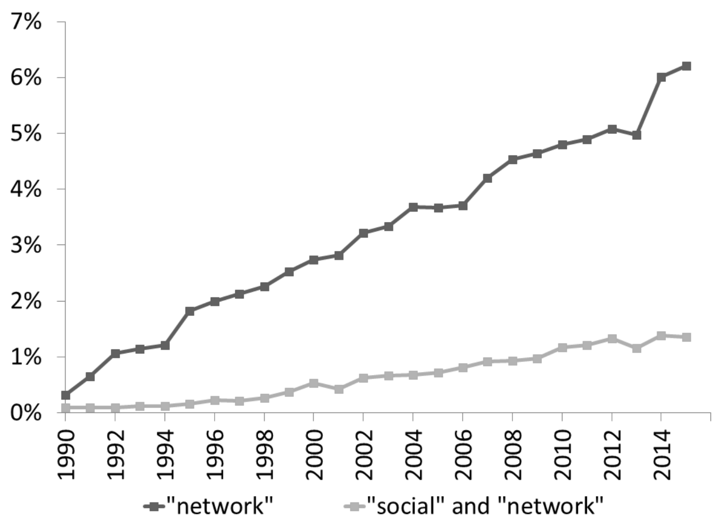
An abridged version of Chapter 2 is forthcoming as Beck (2016a).

1.2 Networks and land transfers in Gambia

Chapters 3 and 4 contribute to the literature on networks in development economics. It is far from all economic activity that takes place in the completely anonymous markets that are part of the assumption set of standard economic theory. In contrast, many economic transactions are conducted with a known counterpart. To a person who always buys the groceries from the same store, the grocer behind the counter stops being an unknown other. Instead, a personal relationship develops on this basis. Since they know each other, the grocer may extend credit or provide other benefits. The grocer has become part of the regular customer's *social network*. Economists are increasingly aware of the critical role that social networks play in economic exchange. This trend is illustrated by Figure 1, which shows a steady increase in the share of papers in the EconLit database that contain the word "network", and the combination of the words "social" and "network" since 1990.

Developing countries are often plagued by multiple market failures. Different strands of literature have showed how market failures give rise to compensating non-market responses such as alternative institutional frameworks and household coping behavior, e.g., how the absence of insurance markets make sharecropping arrangements optimal (Stiglitz, 1989); how a system of local judges can enforce cooperative trade in the absence of state-sanctioned punishment mechanisms (Milgrom et al., 1990); and how labor and food market failures increase farmers' reliance on own food production (de Janvry et al., 1991). In economic settings where property rights are weak, where contract enforcement is hard and where credit and factor markets are thin or non-existing, it

Figure 1: Papers in EconLit containing specific words



Note: Numbers are the share of the total number of papers published in scholarly journals in the database.
Source: The data was retrieved from the EconLit database on January 19, 2016.

is thought that the networks of a person or a household can help to offset the lack of formal institutional frameworks.

This is the starting point of Chapters 3 and 4. Both chapters use the so-called *Gambia Networks Data 2009* dataset, although the two chapters cut the raw sample in different ways. The data was collected in sixty Gambian villages in 2009 and contains information on several social and economic networks. The data is unusual as it covers the economic networks of *all* households in these villages. My co-author Benedikte Bjerge and I gained access to the data through Dany Jaimovich of Goethe University Frankfurt, who participated in the collection of the data.

We quickly zoomed in on the network of land transactions as a highly interesting part of the data as it showed what we found to be an unusual high frequency of land exchange. The high frequency was explained to us when we found out that these transfers are temporary in nature: In The Gambia, land is very rarely sold. Instead, usage rights are transferred. The transfer agreements last only a single or sometimes a few years before they must be renegotiated. It was this network of temporary transfers of usage rights that we saw in the data. Further, the literature suggested that these transfers of land were non-monetary in nature: The donor of land does not receive a reciprocating payment of commensurate value.

We found this mode of land exchange quite different compared to more standard modes of land allocation such as sales, rental agreements or sharecropping. This led to a desire to investigate the economic properties of the outcomes that arise under such arrangements. We also quickly realized that in this context, which is characterized exactly by multiple market failures as well as a lack of a well-functioning formal institutional framework, social networks could be highly influential on the patterns of land exchange. In Chapter 3, which Benedikte and I have co-authored with Marcel Fafchamps, we investigate how social networks affect the efficiency properties of land transfers. We find that, on average, only geographical proximity increases efficiency. On the other hand, kinship and ethnic group membership decreases efficiency. However, this latter finding is driven by a small group of large landowners who transfer land across the confines of the social networks. When the impact of the large landowners is addressed in the regressions, we find that social networks result in more efficiency-enhancing transfers.

While the results of Chapter 3 indicate that social networks matter for the *outcome* of land exchange, Benedikte and I also wished to dive deeper into the *motivations* behind the transfers of land usage rights. This led us down the path of studying traditional West African land market institutions. We were motivated by a wish to provide quantitative evidence consistent with informal livelihood insurance arrangements, the evidence of which is primarily based on qualitative studies. The output of this research led to Chapter 4, which focuses on one such arrangement, namely norm-based access rules to land. The chapter finds that land transfer behavior is consistent with motives of livelihood insurance in that the land-rich households transfer land rights to poorer households. However, we also find that these effects are only present in the villages with lower population densities and, to a lesser extent, in villages with low levels of ethnic heterogeneity. This is in accordance with literature on the impact of population increases and ethnic heterogeneity on social norms and community trust. We argue that, as population densities increase, redistributive land reforms may be needed as informal social insurance arrangements break down.

Chapter 3 is currently under review at The World Bank Economic Review. Chapter 4 is currently under review at World Development.

1.3 Coffee prices and household responses in Vietnam

Chapter 5 is also concerned with the allocation of production factors. While Chapters 3 and 4 study transfers of land in a network of households, Chapter 5 is concerned with intra-household allocations of labor between different activities. More specifically, the chapter studies the impact of coffee price fluctuations for smallholder coffee farmers in Vietnam. Vietnam has experienced impressive growth since a wide-ranging reform program was initiated in 1986. Part of the success story is the coffee boom of the Central Highlands region. Improvements in technology, coupled with state-subsidized coffee production starter kits and a beneficial international coffee price meant that coffee production in this region expanded rapidly in the 1990s: From 1.2 percent of the world's coffee in 1989 to 12.4 percent in 1999 (Luong and Tauer, 2006).

However, this switch has led to greater market dependence, and perhaps not all is so rosy in the Central Highlands. In Chapter 5, we investigate intra-household labor responses of smallholder coffee farmers when coffee prices change. The results indicate that household welfare decreases when coffee prices fall. The adults of the household increase their off-farm labor supply in order to cope with the loss of agricultural income. We also find increases in the labor supply of children (on the farm) and adolescents (both on the and off the farm). We argue that these effects of changes in the coffee price should be taken into account in formulating and implementing social protection schemes and inclusive growth policies.

The chapter is rooted in my involvement with the collection and analysis of the 2012 and 2014 rounds of the Vietnam Access to Resources Household Survey (VARHS), which the Development Economics Research Group at the Department of Economics, University of Copenhagen has been collecting since 2002. In 2013, I wrote a report chapter on crop production and commercialization with Do Huy Thiep, based on the 2012 wave of the survey (CIEM, 2013). The Central Highlands, with its high rates of commercialized agriculture, looked quite different compared to the other provinces of the survey, and this caught my interest. In the summer of 2014, I participated in the collection and cleaning of the 2014 wave of VARHS. A field trip with Saurabh Singhal to the Central Highlands province of Dak Nong really brought home to me the dynamism of the Vietnamese coffee economy and I was fascinated by the complex geographical and ethnic topographies of the region. This became even more apparent when I wrote an analysis of commune characteristics of the VARHS communes (Beck, 2016b) and, with Saurabh, an analysis of the role that ethnicity continues to play in rural Vietnam (Singhal and Beck, 2016). It was natural for me to choose the agricultural strategies of Vietnamese coffee farmers as the topic for my master's thesis (Beck, 2014). While the research question, identification strategy and data have changed fundamentally, the basic question of trying to understand how coffee farmers adapt to the volatile coffee prices on which their livelihood is based, remains the core issue of Chapter 5, which is co-authored with Saurabh Singhal and Finn Tarp.

Chapter 5 is currently under review at The Journal of International Economics.

References

Beck, U. (2014). The Price of Growing Coffee. Master's thesis, Department of Economics, University of Copenhagen.

Beck, U. (2016a). Keep It Real: Measuring Real Inequality Using Survey Data From Developing Countries. In Arndt, C. and Tarp, F., editors, *Measuring Poverty and Wellbeing in Developing Countries*. Oxford University Press, Oxford. Forthcoming.

Beck, U. (2016b). Local transformation in rural Vietnam: A commune level analysis. In Tarp, F., editor, *Growth, Structural Transformation and Rural Change in Viet Nam: A Rising Dragon on the Move*. Oxford University Press, Oxford. Forthcoming.

Beck, U., Mussa, R., and Pauw, K. (2016). Methods Matter: The sensitivity of Malawian poverty estimates to definitions, data, and assumptions. In Arndt, C. and Tarp, F., editors, *Measuring Poverty and Wellbeing in Developing Countries*. Oxford University Press, Oxford. Forthcoming.

Chen, S. and Ravallion, M. (2010). The Developing World Is Poorer Than We Thought, But No Less Successful in the Fight against Poverty. *The Quarterly Journal of Economics*, 125(4):1577–1625.

CIEM (2013). Characteristics of the Vietnamese Rural Economy: Evidence from a 2012 Rural Household Survey in 12 Provinces of Viet Nam. Research report, Central Institute for Economic Management, Hanoi.

Datt, G. and Ravallion, M. (1992). Growth and redistribution components of changes in poverty measures: A decomposition with applications to Brazil and India in the 1980s. *Journal of Development Economics*, 38(2):275–295.

de Janvry, A., Fafchamps, M., and Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. *Economic Journal*, 101(409):1400–1417.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, 84(2):755–776.

Enamorado, T., López-Calva, L. F., Rodriguez-Castelán, C., and Winkler, H. (2016). Income inequality and violent crime: Evidence from Mexico's drug war. *Journal of Development Economics*, 120:128–143.

Günther, I. and Grimm, M. (2007). Measuring pro-poor growth when relative prices shift. *Journal of Development Economics*, 82(1):245–256.

Luong, Q. V. and Tauer, L. W. (2006). A real options analysis of coffee planting in Vietnam. *Agricultural Economics*, 35(1):49–57.

Milgrom, P. R., North, D. C., and Weingast, B. R. (1990). The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs. *Economics & Politics*, 2(1):1–23.

Ostry, J. D., Berg, A., and Tsangarides, C. G. (2014). Redistribution, Inequality, and Growth. IMF Staff Discussion Note SDN/14/02, International Monetary Fund.

Pauw, K., Beck, U., and Mussa, R. (2016). Did Rapid Smallholder-led Agricultural Growth Fail to Reduce Rural Poverty? Making Sense of Malawi's Poverty Puzzle. In Arndt, C., Andy, M., and Tarp, F., editors, *Growth and Poverty in Sub-Saharan Africa*. Oxford University Press, Oxford. Forthcoming.

Persson, T. and Tabellini, G. (1994). Is Inequality Harmful for Growth? *The American Economic Review*, 84(3):600–621.

Pinkovskiy, M. and Sala-i Martin, X. (2009). Parametric Estimations of the World Distribution of Income. Working Paper 15433, National Bureau of Economic Research.

Pinkovskiy, M. and Sala-i Martin, X. (2014). Africa is on time. *Journal of Economic Growth*, 19(3):311–338.

Sala-i Martin, X. (2006). The World Distribution of Income: Falling Poverty and... Convergence, Period. *The Quarterly Journal of Economics*, 121(2):351–397.

Sala-i Martin, X. and Pinkovskiy, M. (2010). African Poverty is Falling...Much Faster than You Think! Working Paper 15775, National Bureau of Economic Research.

Singhal, S. and Beck, U. (2016). Ethnic Disadvantage in Vietnam: Evidence using Panel Data. In Tarp, F., editor, *Growth, Structural Transformation and Rural Change in Viet Nam: A Rising Dragon on the Move*. Oxford University Press, Oxford. Forthcoming.

Stiglitz, J. E. (1989). Markets, Market Failures, and Development. *The American Economic Review*, 79(2):197–203.

UN (2016a). United Nations Millennium Development Goals Webpage. http://www.un.org/millenniumgoals/. Accessed January 8, 2016.

UN (2016b). United Nations Sustainable Development Goals Webpage. https://sustainabledevelopment.un.org/. Accessed January 8, 2016.

World Bank (2013). Shared Prosperity: A New Goal for a Changing World. http://www.worldbank.org/en/news/feature/2013/05/08/shared-prosperity-goal-for-changing-world. Accessed January 8, 2016.

CHAPTER 2

KEEP IT REAL: MEASURING REAL INEQUALITY USING SURVEY DATA FROM DEVELOPING COUNTRIES

Ulrik Beck*

Abstract

This paper investigates how two effects drive wedges between nominal and real inequality estimates. The effects are caused by (i) differences in the composition of consumption over the income distribution coupled with differential inflation of consumption items, and (ii) larger quantity discounting effects for the non-poor. Household-specific deflators are estimated using 15 surveys collected in 6 countries in the period 1999–2011. In some countries (Mozambique, Tanzania, Malawi and Pakistan), nominal inequality is lower than real inequality. In other countries (Ethiopia and Madagascar), no differences are found. I argue that poverty estimation based on national account consumption means and estimates of inequality from consumption surveys should employ real, rather than nominal, inequality estimates. In some cases, this increases the level of poverty and reduces the decline of poverty over time.

I wish to thank Channing Arndt and Finn Tarp for valuable comments; seminar and workshop participants from the Department of Economics, University of Copenhagen for helpful inputs; and Azhar Hussein, Kristi Mahrt, Richard Mussa, Karl Pauw, Tiaray Razafimanantena, Vincenzo Salvucci, David Stifel, and Edward Whitney, who worked on various country teams for the Growth And Poverty Project initiated by UNU-WIDER, for helping me to understand the peculiarities of the consumption datasets. All interpretations and any remaining errors or omissions are my sole responsibility.

*Department of Economics, University of Copenhagen

2.1 Introduction

Measures of inequality are often used to direct and evaluate policy. In developing countries, estimates of inequality are typically based on consumption modules, which are included in nationally representative surveys. Based on this, a consumption aggregate is constructed. This aggregate is a measure of the value of consumption by a household. There are some technicalities involved with estimating the consumption aggregate; for instance, housing costs are often imputed, and the consumption of durable goods must be spread out over multiple years (Grosh and Deaton, 2000; Deaton and Zaidi, 2002). Nevertheless, at its core, the value of consumption is reached by multiplying current prices with quantities. As such, the standard consumption aggregate is a nominal concept. Inequality indices derived from nominal consumption aggregates are therefore also nominal in their nature.

There are at least two reasons why basing inequality estimates on a real consumption aggregate – and thereby estimating what I refer to as real inequality in the remainder of the paper – is relevant. First, the poorest households tend to dedicate a higher share of their spending towards basic food items, the prices of which have been rising faster than other prices in recent years. I refer to this as the composition effect. Second, if there is a systematic difference in the prices faced by households over the income distribution, nominal inequality will differ from real inequality. This can arise when the poor tend to purchase items in smaller quantities, which can lead to higher prices. I refer to this as the quantity discounting effect. This paper aims to empirically estimate household-specific deflators of these two effects, proceed to estimate real consumption aggregates, and use these to compute estimates of real inequality.

I overcome the substantial data requirements for this task by building on the set of country-specific databases that were constructed as part of the Growth And Poverty Project (GAPP) conducted under the auspices of the United Nations University – World Institute for Development Economic Research (UNU-WIDER). Using 15 nationally representative surveys from 6 different countries (Ethiopia, Madagascar, Malawi, Mozambique, Pakistan, and Tanzania), collected in the period 1999 to 2011 and covering over 220,000 households, I am able to construct household-specific indices of the composition and the quantity discounting effects.

The paper proceeds by investigating how real inequality estimates affect poverty estimates when poverty is estimated using a method developed in a series of papers by Sala-i-Martin and Pinkovskiy (hereafter SiMP). This method estimates poverty rates by fitting two-parameter consumption distributions using inequality estimates obtained from survey data and gross domestic product per capita estimates from the national accounts (Sala-i Martin, 2006; Pinkovskiy and Sala-i Martin, 2009; Sala-i Martin and Pinkovskiy, 2010; Pinkovskiy and Sala-i Martin, 2014). Using this approach, the authors find that poverty is falling much faster than what is observed by traditional methods of poverty estimation. Pinkovskiy and Sala-i Martin (2014) state that the discrepancy is mainly caused by differences in the growth rates of mean per capita consumption observed in the surveys and the mean per capita gross domestic product (GDP) observed in the national accounts. These differentials are not disputed; however, this paper shows that using the proper inequality estimates also matters.

The composition effect and the quantity discounting effect have both been studied before. This paper contributes to the existing literature by providing empirical evidence for six different developing countries. Furthermore, I develop a method to estimate the quantity discounting effect using existing nationally representative consumption data. Finally, the paper provides insights into the robustness of the SiMP methodology.

The strengths of the composition effect and the quantity discounting effect differ substantially across countries. In some countries, the poorest households were subject to a double penalty, which is the result of a combination of high food inflation rates for the consumption bundle of the poor and of the poor buying in smaller quantities. In continuation of this, the decline in poverty using the SiMP methodology may be overestimated, and the level of poverty underestimated. This paper therefore explains part of the gap between the very optimistic results of Pinkovskiy and Sala-i Martin (2014) and other, more mixed, findings.

The remainder of the paper is organized as follows. Section 2.2 describes the composition effect and the quantity discounting effect in more detail and how I plan to estimate household-specific deflators to correct for them. Section 2.3 describes the various data sources used and shows some descriptive statistics. Section 2.4 reports the estimated household-specific deflators and their impact on inequality and poverty. The section also contains a series of robustness checks of the main results. Section 2.5 concludes.

2.2 Conceptual framework

This section explains how the two effects described in the introduction arise and how the household-specific price indices are estimated. Following this, I outline the SiMP methodology and show how estimates of real inequality estimates affect it.

2.2.1 The composition effect

The composition effect arises as a consequence of differences in the consumption structure between rich and poor and relatively higher inflation in items that are disproportionately consumed by one of these groups. In developing countries, the relatively poor spend a larger part of their income on basic food items. If this is combined with disproportionate increases in the prices of basic food items, it leads to an increase in real inequality over nominal. The composition effect has been studied in some detail for developed countries (See, for instance, Muellbauer, 1974; Cage et al., 2002; Leicester et al., 2008). A higher consumption share of food items by the poor has also been found in developing countries (e.g., Pritchett et al., 2000; Deaton, 2003; Günther and Grimm, 2007; Aksoy and Isik-Dikmelik, 2008). However, the impact on inequality has not been the focus of the majority of this body of work although a few recent papers have explored the link between the composition effect and inequality in developing countries more directly (Goñi et al., 2006; Mohsin and Zaman, 2012; Arndt et al., 2015).

The hike in food price inflation after 2000, culminating in the price spike of the food price crisis of 2007–09 (Mitchell, 2008; Wiggins et al., 2010), provides a rationale for estimating the magnitude of this effect using recent data. A priori, poor households are expected to spend a larger share of their income on food items. This, coupled with high food price inflation, means that real inequality will be higher than nominal.

In estimating the composition effect, I largely follow the method proposed by Arndt et al. (2015). The authors find that the structure of consumption bundles varies across the income distribution. Owing to more rapid inflation in the prices of basic goods, nominal inequality was found to underestimate real inequality by several Gini points for Mozambique in 2008. The authors divide consumption items into three groups: core food items, non-core food items, and non-food items. A household-specific Paasche price index that takes into account differential inflation rates of these three groups of items is given by

$$CPI_{COMP}^{i,t} = \left(\frac{p_c^1}{p_c^t} s_c^{i,t} + \frac{p_{nc}^1}{p_{nc}^t} s_{nc}^{i,t} + \frac{p_{nf}^1}{p_{nf}^t} s_{nf}^{i,t} \right)^{-1} \tag{1}$$

Here, p_a^t is the index price in year t of group a products, where a can be core (c), non-core (nc), or non-food (nf), and $s_a^{i,t}$ is the share of consumption used for group a products by household i in year t.[1]

There are two principal challenges associated with implementing this approach consistently across countries. The first is how to choose which food items should be included in the core and the non-core food groups, respectively. This choice should be country-specific since food consumption patterns vary substantially between countries. It should also be general since cross-country results can only be meaningfully compared if the decision rule is consistent across countries. An option that fits both of these criteria is to define the core food items as those included in the food poverty lines estimated by the GAPP country studies in year t of each country.[2] The poverty food

[1] Two differences to the methodology of Arndt et al. (2015) are worth noting. First, Arndt et al. (2015) also consider spatial differences in price levels. If richer households are overrepresented in spatial domains with higher price levels, failing to correct for this will overestimate inequality. I do not consider spatial differences in prices in the estimation of the composition effect; instead, a spatial price index is applied throughout where available. Thus, the 'nominal' inequality estimates of this paper contain spatial price corrections. Second, Arndt et al. (2015) do not use a Paasche index. This paper uses a true Paasche index as its properties are well known. Specifically, if there is substitution towards goods that become relatively cheaper, a Paasche index will underestimate the rate of inflation. This means that inflation estimates reported here are lower bounds on the true inflation rates in the presence of substitution. The Paasche index is written in share expenditure form to facilitate estimation.

[2] See Table 2 in the data section for references to the GAPP country studies.

basket is chosen consistently across countries, and across surveys within countries, in order to represent the most important food items for the poor. This makes this group of products an ideal candidate for the core food group. I choose not to use the inflation rates of the food poverty line as an estimate of the temporal change in p_c^t, since items are allowed to move in and out of the food poverty bundles between surveys, and since the prices used to estimate poverty lines are often estimated specifically for the poor. Instead, I re-estimate weights and price increases for the food items in the food poverty bundle directly from the survey data. Since the poverty lines vary at the sub-national level, and since this paper is concerned with estimating food inflation at the national level, a procedure to reconcile this difference is needed. I choose to keep only items that are present in the poverty lines of two or more spatial domains of year t, and also present in the first survey ($t = 1$), though not necessarily part of the poverty basket in the first survey. In order to increase precision of the estimated unit prices, I further restrict the group of food items to those items where each survey has at least 200 recorded purchases.

The second challenge is to estimate price changes of core foods, non-core foods, and non-food items separately. It is not feasible to estimate all price changes from survey information alone owing to missing prices and few purchases of some goods. Furthermore, detailed consumer price index (CPI) information at the product level is not always available, especially for rural areas. For the core food items, the surveys contain sufficient information to calculate price changes directly from the survey. However, this is not the case for the non-core food items and the non-food items. The non-core food items are not observed as frequently in the data, and using the survey prices is not an option. The non-food items are typically only reported as nominal values. Instead, I use external sources of CPI information that is available separately for food and non-food items. For the non-food group of items, the non-food CPI series is directly applicable.

I proxy the non-core food inflation by the total food CPI series. One can think of the total food CPI series as a weighted average of core food and non-core food CPI series. Therefore, estimation of the core food inflation from the household data means that the direction of the bias of the non-food inflation index is known. As will become clear, the bias tends to attenuate the magnitude of the composition effect; the estimates of this paper therefore represent lower bounds on the true effect sizes in most cases.

2.2.2 The quantity discounting effect

The quantity discounting effect arises when the poor purchase smaller amounts at a time, thereby missing potential quantity discounts. There are at least four explanations for why the poor would do so. First, the poor consume less. For perishable items, smaller purchases can be rational, especially if poor households lack the capacity to securely store food items. Second, the poor may be credit constrained, leading to smaller and more frequent purchases. Third, the poor may not have the means to transport large amounts at a time. Fourth, the state of being poor increases stress and takes up mental capacity, which can impede cognitive function and lead to sub-optimal decisions (Mani et al., 2013).

It is in theory possible to directly estimate the quantity discounting effect using unit prices, i.e. prices calculated from quantities and values reported by households. The main pitfall with this approach is that the quality of purchases is variable and unobserved. A specific item code in the consumption module of a questionnaire must, for practical purposes, cover a variety of qualities, even though higher quality items will tend to have higher unit prices. This is difficult to separate from a potential quantity discounting effect when high- and low-quality items share the same survey code. The problem of separating quality issues from true price variation has been referred to as the unit value problem (Deaton, 1988; Crawford et al., 2003; Chung et al., 2005; Beatty, 2010; McKelvey, 2011). One popular approach to deal with this was proposed by Deaton (1988). His study assumes that all differences in unit prices are caused by quality differences within sufficiently small geographical areas. By getting rid of the between-area price variation, remaining unit price variation can be used to estimate quality differences of purchases. However, as argued by Attanasio and Frayne (2006), another potential source of price variation within geographical areas is quantity discounting. Using a consumption survey from Colombia, the authors find that unit values are in fact negatively related to monthly spending and to sizes of purchases, conditional on monthly spending which is included to control for differences in demand for quality. Using a similar approach, Mussa (2015) find that the poor pay more for maize in Malawi in the 2010/11 Malawi Integrated Household Survey, which is also employed in this paper. However, for the purposes of this paper, this methodology is unsatisfactory, as it does not provide household-specific deflators, which are required for estimating real inequality.

An alternative way of reducing the confounding of quantity discounting with quality differences is to use a survey instrument that carefully separates different qualities of the same product into different product codes (Rao, 2000; Aguiar and Hurst, 2007). However, such specialized datasets are often not available, especially in developing countries. Alternatively, one could limit the study to reasonably homogeneous items, thereby reducing the bias introduced by quality differences (Attanasio and Frayne, 2006). However, when the topic of interest is national inequality, a method that works with all items of consumption and is feasible using existing nationally representative surveys is required.

In the following, I develop such a method by exploiting information in the survey about the size of the purchases. By exploiting this information, one can estimate a household-specific price index that at least partially controls for quality differences. As a point of departure, I take the expensiveness index of Aguiar and Hurst (2007). The authors construct an expensiveness index for each household in order to compare how expensively households buy their specific basket of goods. The index is given by

$$p^i_{AH} = \frac{\sum_m \left[p^i_m q^i_m\right]}{\sum_m \left[\bar{p}_m q^i_m\right]} \tag{2}$$

Here, p^i_m is the price paid for product m by household i, $\bar{p}_m$ is the average price paid for product m, and q^i_m is the quantity household i bought of product m.[3] This measure compares actual expenditures of household i with the cost of this bundle of food items, priced at the average prices. If the index is larger than one, the household is paying more for its bundle compared to what an average household would pay.[4] Next, I introduce product-specific quantity bins. Each bin contains an equally sized share of the quantity distribution of a single product. For instance, for maize, the first bin will contain the x percent smallest maize purchases. The last bin will contain the x percent largest maize purchases. Using such bins, a more specific version of the index can be calculated, where u indexes the quantity bins of each product:

$$p^i_{AH-u} = \frac{\sum_m \sum_u \left[p^i_{m,u} q^i_{m,u}\right]}{\sum_m \sum_u \left[\bar{p}_{m,u} q^i_{m,u}\right]} \tag{3}$$

This version of the index switches the unit of estimation from the product level to the level of the within-product quantity bin. It only compares prices of products with other purchases that fall into the same quantity bin. Index (2) and (3) are both biased estimates of quantity discounting due to the unit value problem. However, by taking the ratio between the two indices it is possible to get rid of all price variation that is not caused exactly by differences in the size of purchases. This gives the household-specific quantity discounting price index where I have exploited that the numerators of (2) and (3) are both equal to total household expenditure and therefore cancel out:[5]

$$\widetilde{CPI}^i_{QUANT} = \frac{p^i_{AH}}{p^i_{AH-u}} = \frac{\sum_m \sum_u \left[\bar{p}_{m,u} q^i_{m,u}\right]}{\sum_m \left[\bar{p}_m q^i_m\right]} \tag{4}$$

The necessary assumption for the quantity discounting index to exactly isolate the quantity discounting effect is that the size of purchase is uncorrelated with quality. If

[3]Some of the surveys employed for the empirical section of this paper do not allow for separation between what is purchased and what comes from other sources, such as barter, gifts, and own production. The first best solution is to use prices of purchased items only, so this is done wherever possible. This is possible to do for the surveys of Malawi, Mozambique, and Tanzania. However, not being able to do so does not invalidate the method as long as there is a correlation between the cost of total consumption and the cost of purchases, which is not unrealistic.

[4]The index of (2) is the Paasche version of the index used by Rao (2000), which is given by $\frac{\sum_m \left[p^i_m \bar{q}^i_m\right]}{\sum_m \left[\bar{p}^i_m \bar{q}^i_m\right]}$, where $\bar{q}^i_m$ is the average price paid for product m in the area in which i resides. This index is also used by Mussa (2014) to estimate the impact of price differences across the expenditure distribution in Malawi, using the same Malawian surveys as the current paper. However, upon closer inspection, the findings of Mussa (2014) were found to be driven by a coding error.

[5]The index is subsequently normalized to have a mean of one.

there is a correlation between quality and quantity of purchase, it will continue to affect Equation (4). Since one can expect richer households to buy higher-quality items, this effect will bias results in the opposite direction of quantity discounting. Therefore, if quantity discounting effects are found, the estimated effect can be seen a lower bound on the true effect size. As a baseline, I construct four bins separated at the 25th, 50th, and 75th percentile of the product-specific quantity distribution, but I check that results are robust to alternative numbers of bins.

Perhaps the best way to illustrate the mechanics of the quantity index is with an example, which can be found in Table 1. Consider the purchase decision of a single good (Good 1) with four cases determined by whether the good is bought in a high- or low-quality version and in either a small or a large amount. Alternatively, one can think of four different households buying four different versions of the same good. For simplicity, the total cost of all other goods is set to one.[6] Buying the high-quality version is more expensive and buying the small-amount version is also more expensive. Since the price of purchase varies between the four cases, the two indices of Equations (2) and (3) also vary between the four purchases. However, the ratio between the two, as found in the last column of Table 1, only varies between the small and big units since the quality effect has been divided out and only the quantity discounting effect remains.

The index of Equation (4) makes use of all variation in prices within the survey. However, if there is real price variation between geographical areas, the performance of the index can be improved by estimating average prices at a smaller geographical area than the national level (Deaton, 1988). This matters if the poor are disproportionately likely to live in either high- or low-price areas. The final index is shown in Equation (5). Here, $\bar{p}^{g}_{m,u}$ denotes the average price of unit size u of item m in geographical area g where household i lives. In this version, the household-specific deflator of household i is based only on variation within g.

$$CPI^{i}_{QUANT} = \frac{\sum_m \sum_u \left[\bar{p}^{g}_{m,u} * q^{i}_{m,u}\right]}{\sum_m \left[\bar{p}^{g}_{m} * q^{i}_{m}\right]} \tag{5}$$

The geographical area employed in the remainder of the paper is the survey stratum. This means that any differences in prices between strata are not included in the quantity discounting-adjusting CPI. The number of strata is survey-specific; the surveys used in this paper have between 8 and 31 strata.

Since quantities and therefore unit prices are only available for food items, the quantity discounting index is only estimable for the part of the consumption aggregate that is

[6] One can also think of a household that makes consumption decisions under a binding budget constraint. If we let other expenditure adjust such that total expenditure is the same in all four cases, the result is unchanged.

Table 1: A quality-adjusted index example

Good 1						Expenditure		p^i_{AH}	p^i_{AH-u}	$\widetilde{CPI}^i_{QUANT}$
Quality	Amount	$\bar{p}_1$	$\bar{p}_{1,u}$	q_1	p^i_1	Other	Total	(2)	(3)	(4)
High	Small	3.5	4	1	5	1	6	1.33	1.20	1.11
Low	Small	3.5	4	1	3	1	4	0.89	0.80	1.11
High	Big	3.5	3	1	4	1	5	1.11	1.25	0.89
Low	Big	3.5	3	1	2	1	3	0.67	0.75	0.89

Source: Author's calculations.

based on food expenditure. The non-food aggregate used in the current paper is not corrected for potential quantity discounting effects.

For a few surveys, the usage of quantity bins can be sidestepped by using units of purchase instead. In some surveys, households are asked to report the unit of consumption. For instance, for the surveys of Malawi, respondents have the option of choosing between more than 20 units for each item. These include cups and plates but also kilograms and liters. This variation can be exploited directly by using units instead of quantity bins in Equation (5).[7]

2.2.3 Estimating inequality

The deflated consumption aggregate for household *i* in year *t* is estimated as

$$Y^{i,t}_{real} = \frac{\left(y^{i,t}_{c} + y^{i,t}_{nc}\right) / CPI^{i,t}_{QUANT} + y^{i,t}_{nf}}{CPI^{i,t}_{COMP}} \tag{6}$$

where $y^{i,t}_{c}$, $y^{i,t}_{nc}$, and $y^{i,t}_{nf}$ are the nominal consumption aggregates of core, non-core, and non-food consumption of household i in year t. $Y^{i,t}_{real}$ is real consumption. All other notation is the same as described earlier. Using population weights, nationally representative real Gini coefficients are estimated.

2.2.4 Estimating poverty

The poverty rate is the share of people who consume less than a given poverty line. A widespread approach to estimating national poverty lines is to use information on consumption from nationally representative surveys in order to estimate the cost of consuming a predetermined amount of calories, given the actual consumption structure of the poor. Subsequently, non-food requirements are estimated. The sum of the food and non-food requirements equals the total poverty line. This is the so-called cost of basic needs (CBN) approach to poverty line estimation (Ravallion and Bidani, 1994; Tarp et al., 2002).

[7] This method requires a correspondence between the unit of purchase (i.e. the object of interest) and the unit of consumption (i.e. what is measured). The two are likely to be highly correlated.

The CBN methodology can be made robust to both the composition and the quantity discounting effects. The composition effect is implicitly handled since the poverty line is by definition the cost of a certain amount of the consumption bundle consumed by the poor. It is therefore only price changes experienced by the poor that influence the intertemporal change in the poverty line. The quantity discounting effect can be handled by pricing the consumption bundle using the prices paid by the poor, which is frequently done in practice.

Another common approach to estimating poverty rates is to impose an exogenously defined poverty line. The leading example of such a poverty line is USD 1.25 PPP (purchasing power parity)-adjusted to 2005 prices, as proposed by Ravallion et al. (2009).

Recently, Sala-i-Martin and Pinkovskiy (SiMP) have proposed a third approach (Pinkovskiy and Sala-i Martin, 2009; Sala-i Martin and Pinkovskiy, 2010; Pinkovskiy and Sala-i Martin, 2014). This approach uses survey-based inequality estimates and information from the national accounts on the gross domestic product (GDP) per capita to fit a two-parameter consumption distribution for each country. For their main estimates, Pinkovskiy and Sala-i Martin (2014) fit a log-normal distribution of consumption using mean GDP per capita from the World Bank's Development Indicators database to fix the mean of the distribution, and inequality estimates from Chen and Ravallion (2010) and the WIDER World Income Inequality Database (UNU-WIDER, 2014). For the countries considered in this paper, the inequality information stems from the same surveys as those used in the empirical analysis of this paper.

Using the fitted distribution and the USD 1.25-a-day poverty line, Pinkovskiy and Sala-i Martin (2014) estimate poverty based on the cumulative distribution function. The USD 1.25-a-day poverty line is measured in real 2005 international (PPP-adjusted) prices. For this reason, Pinkovskiy and Sala-i Martin (2014) use a measure of GDP in real 2005 PPP-adjusted international dollars to anchor the income distribution. If all households face the same prices, it is unnecessary to deflate inequality estimates, because the Gini coefficient is unaffected by scalar multiplications. However, as argued above, the deflator may not be constant over the income distribution. This possibility is ignored by the SiMP methodology. A similar point is made in reference to the SiMP methodology in relation to a composition-type effect in the context of Burkina Faso by Grimm et al. (2016). To conclude, if one wants to take seriously the notion of estimating poverty using a fitted two-parameter distribution, a real inequality estimate is necessary. The final section therefore investigates the impact of using real inequality poverty rates when substituting nominal inequality figures with their real counterparts.[8]

[8] Pinkovskiy and Sala-i Martin (2014) adjust estimates of consumption inequality to make them comparable with other surveys based on income. For the sake of simplicity, and since only consumption-based surveys are used in this paper, I do not consider such an adjustment.

In addition to the earlier discussion, there are at least three differences between the CBN and the SiMP methodologies (Guénard and Mesplé-Somps, 2010; Arndt et al., 2016c). First, consumption surveys often fail to sample sufficiently from the top of the distribution. This is not an issue under the CBN methodology since this part of the consumption distribution does not affect these poverty estimates. However, it can severely affect estimates of inequality, especially given that inequality estimates are sensitive to changes in the top end of the distribution. Second, the motivation behind using GDP instead of the survey-based consumption mean is that *for a meaningful analysis of the impact of growth on poverty, the income distribution used to calculate poverty must be consistent with observed growth rates* (Pinkovskiy and Sala-i Martin, 2014, p. 313). However, there are several valid explanations for why GDP growth rates should not equal the growth rate of household consumption, which include differences in the coverage of economic activities and in population. In terms of economic activities, GDP contains other components than household consumption. For example, an expansion of public services increases GDP but does not affect household expenditures directly. In terms of population coverage, household surveys usually cover only "ordinary" households whereas GDP also covers the part of the population that lives outside ordinary households such as prison populations and religious groups. Third, the USD 1.25-a-day poverty rate was defined keeping in mind the consumption definition of household surveys and cannot be directly transferred to the GDP measure of national income. To conclude, the definition of poverty used in the SiMP methodology differs fundamentally from the theoretical concept underlying the CBN estimates. Comparing the two directly is like comparing apples and oranges and I refrain from doing so in this paper. Instead, I compare SiMP-style measures of poverty to each other, using either nominal or real consumption aggregates to estimate them.

2.3 Data

The various data sources used for this paper, as well as some descriptive statistics, are detailed in Table 2. As mentioned previously, the results build upon work done in relation to GAPP. Building on this body of work, I have compiled a standardized database of consumption information that allows real inequality measures to be computed at the household level for the more than 220,000 household observations in the database. Nationally representative consumption questionnaires are often collected over an extended period of time, typically an entire year. Since prices can change within this time frame, all prices and consumption aggregates presented are deflated using a temporal (within-survey) price index.[9] Since prices also vary spatially, I deflate consumption aggregates by spatial indices. This gives the consumption aggregate

[9]The Madagascar surveys and the Ethiopia survey in 2000 and 2005 are exceptions where no such indices are used since those surveys were collected over a relatively short period of time (i.e. over a couple of months).

that "nominal" inequality estimates of this paper are based on. Units of purchases are available for the surveys of Madagascar 2001 and the two surveys of Malawi. For these surveys, I estimate the quantity discounting effect using both the approach that exploits the units of consumption as well as the approach that relies on binning of quantities.

The countries cover a range of different experiences. Consider the mean consumption per capita of the surveys, which have been converted into 2005 constant international dollars using the PPP-adjusted exchange rate. The mean per capita consumption in Pakistan in 2007/08 (2.54 dollars) was more than double that of Tanzania in 2007 (1.13) and three times that of Madagascar in 2005 (0.83). Trends also differ: At one end of the spectrum is Madagascar where the mean per capita consumption in 2001 was USD 0.91 a day; this fell slightly to USD 0.83 in 2005. At the other end of the spectrum are Ethiopia and Pakistan where mean per capita consumption increased annually by 4 percent annually from the first to the last survey (from USD 1.4 to 2.07 in Ethiopia and from USD 1.74 to 2.52 in Pakistan). The picture in terms of trends is generally consistent if one looks at the national accounts-based GDP per capita estimate instead; however, the level is generally substantially higher using the national accounts. This difference in levels is consistent with what is observed by Pinkovskiy and Sala-i Martin (2014).[10]

The level of inequality also varies across countries: Madagascar and Malawi are the most unequal; here, the 10th percentile of the population consume between 0.27 and 0.35 of mean income, whereas the 90th percentile consume between 1.72 and 2.06 of mean income. There are differences in inequality trends as well: While the consumption spread has decreased in Madagascar, it has increased in Malawi. Pakistan is the least unequal of the countries: The 10th and 90th percentiles consumed 0.53 and 1.57 of mean consumption in the latest survey round.

Information on nominal inequality in the form of Gini coefficients can be obtained from the World Income Inequality Database (WIID, UNU-WIDER, 2014). While I show the WIID estimates along with inequality estimates based on the semi-standardized database that I have compiled for this paper, the empirical analysis is based solely on nominal Gini coefficients estimated from the compiled database. This is necessary since only by using the micro-level datasets can the household specific deflators be applied. For the estimation of poverty using the SiMP methodology, I obtain time series of PPP-adjusted GDP per capita in constant 2005 US dollars from the 2012 version of the World Bank's world development indicators (World Bank, 2012). The same data sources were used by Pinkovskiy and Sala-i Martin (2014).

[10]Note that in this particular sample of countries, it is not the case that the growth rate of the national accounts-based GDP measure consistently outpaces the survey based consumption growth rate, even though this is the case for Sub-Saharan Africa as a whole (Pinkovskiy and Sala-i Martin, 2014). The national accounts-based GDP growth rate is higher than the survey based consumption growth rate for Ethiopia from 2004/05 to 2010/11 and Mozambique. The reverse is true for Malawi and Pakistan from 2001/02 to 2007/08. The two survey-on-survey growth rates are within a few percentage points of each other for Ethiopia from 1999/2000 to 2004/05, Madagascar and Tanzania.

Table 2: Data sources and descriptive statistics

Country and survey years	Consumption database reference (GAPP study)	CPI reference	No. of house-holds	No. of EAs	No. of strata	GDP per capita[1]		Consumption share at the x'th percentile		National poverty rate
						Survey	National accounts	10th	90th	
Ethiopia	Stifel and Woldehanna (2016)	NBE (2014); CSA (2015)								
HICES (1999/2000)			17,332	1264	20	1.40	1.44	0.48	1.60	46.8
HICES (2004/05)			21,595	1548	18	1.69	1.74	0.46	1.58	46
HICES (2010/11)			27,830	1966	20	2.07	2.56	0.42	1.64	23.8
Madagascar	Stifel et al. (2016)	INSTAT (2015)								
EPM (2001)			5080	303	12	0.91	2.54	0.27	2.06	57.8
EPM (2005)			11,781	561	12	0.83	2.38	0.31	1.83	59.1
Malawi	Pauw et al. (2016)	NSO (2015)								
IHS2 (2004/05)			11,280	564	30	1.33	1.77	0.35	1.72	47
IHS3 (2010/11)			12,271	768	31	1.89	2.17	0.29	1.78	38.8
Mozambique	Arndt et al. (2016b)	INE (2015)								
IAF (2002)			8700	857	11	1.29	1.60	0.31	1.78	54.1
IOF (2008)			10,832	1060	11	1.51	2.12	0.31	1.75	54.7
Pakistan	Nazli et al. (2015)	MoF (2015)								
HIES (2001/02)			14,649	1050	8	1.74	5.05	0.51	1.60	21.4
HIES (2005/06)			15,374	1109	8	2.21	5.87	0.51	1.61	23.0
HIES (2007/08)			15,441	1113	8	2.54	6.36	0.51	1.65	26.0
HIES (2010/11)			16,295	1180	8	2.52	6.60	0.53	1.57	27.0
Tanzania	Arndt et al. (2016a)	CountrySTAT (2015)								
HBS (2000)			22,176	1158	20	0.83	2.37	0.38	1.79	35.7
HBS (2007)			10,407	447	20	1.13	3.15	0.37	1.79	33.6

Note: EAs are enumeration areas. HICES is Ethiopia Household Income, Consumption and Expenditure Survey (CSA, 2000, 2005, 2011). EPM is Enquête Périodique auprès des Ménages (INSTAT, 2002, 2006). IHS is Integrated Household Survey (NSO, 2005, 2012). IAF is the Inquérito aos Agregados Familiares (MPF, 2004). IOF is Inquérito ao Orçamento Familiar (MPF and DNEAP, 2010). HIES is Household Integrated Economic Survey (FBS, 2003, 2007, 2009, 2013) and estimates exclude Azad Jammu and Kashmir, Federally Administered Tribal Areas and Northern Areas (PBS, 2006, 2007, 2009, 2013). HBS is Household Budget Survey (NBS, 2002, 2011), which covers mainland Tanzania. 1: 2005 PPP-adjusted international dollars. Source: Author's compilation based on the sources listed in the table except PPP conversion factors and national account information, which are from World Bank (2012).

2.4 Results

2.4.1 The composition effect

Table 3 shows the CPIs of core food, non-core food, and non-food inflation used for estimating the composition effect. Taking the first survey in each year as the baseline, prices of core food items rose faster than the prices of non-core food items, which are proxied by the total food CPI, in all countries except Ethiopia and Madagascar. Since total food inflation is a weighted average of core and non-core food inflation rates, in the four (two) countries where core inflation is higher (lower) than total food inflation, the use of total food inflation as a proxy measure of non-core food inflation overestimates (underestimates) the true rate of non-core inflation.

Why do core and non-core food prices in Ethiopia and Madagascar behave differently from the other countries and differently from what was expected a priori? A complete analysis of this is beyond the scope of this paper but two explanations are likely candidates. Between 2000 and 2005, Ethiopia experienced several good harvests that put downward pressure on food prices (Durevall et al., 2013). In particular, prices of domestically produced foods, which constitute the majority of core food items, were subjected to downward pressure. From 2004/05 to 2010/11, core food prices rose faster than non-core food prices, which is in line with expectations. A contributing factor in the case of Madagascar is that in 2004, because of a partially failed harvest of rice, the main staple of Madagascar, the Malagasy government intervened in the rice market by slashing import tariffs and importing state-bought rice (Dorosh and Minten, 2006). This, combined with a better domestic rice harvest in 2005, contributed to downward pressure on rice prices near the end of 2005, which is when the second Malagasy survey was conducted.

In all countries except for Ethiopia, the inflation of core foods outpaced the inflation of non-food items, compared to the first survey of each country. The magnitude of the price differentials vary between countries. For instance, core food prices in Mozambique rose 63 percent faster than non-food prices from 2002 to 2008. However, in Madagascar, the difference was only 2 percent from 2001 to 2005. To conclude, the data presented here shows that in many, but not all, of the studied countries, food price inflation has been higher than non-food price inflation in the period considered.

Figure 1 shows the mean consumption shares of the three groups of items for each percentile of the consumption distribution across countries and surveys. The first survey of each country is left out as it is used as to estimate baseline price levels of the consumption groups (p_a^1 of equation (1)). Composition effects are only estimable for the later surveys. The percentile-specific means are calculated for ease of illustration; deflators are household specific as indicated by Equation (1). A consistent picture, which matches what Arndt et al. (2015) found for Mozambique, emerges: As one moves up through the income distribution, the share of consumption expenditures

Table 3: Food and non-food CPI

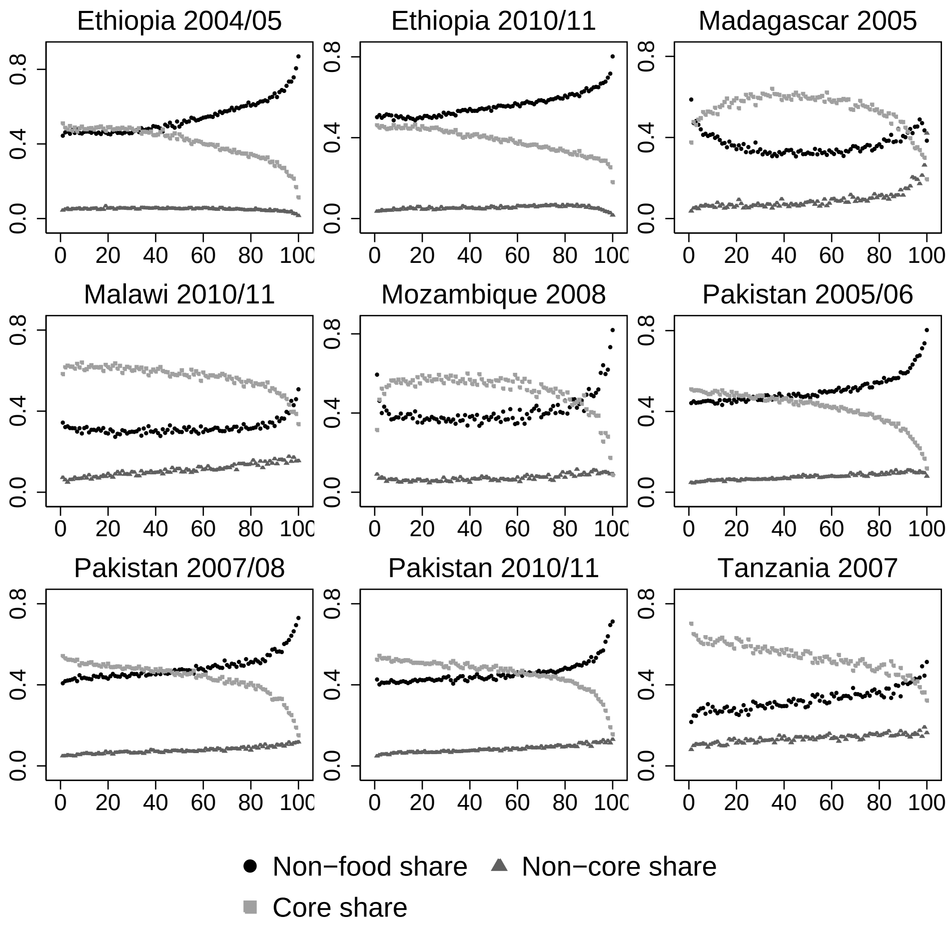
Country and year	Core food (CF)	Non-core food (NCF)	Non-food (NF)	CF/NCF	CF/NF
Ethiopia					
1999/2000	100.0	100.0	100.0	1.00	1.00
2004/05	98.9	145.7	112.8	0.68	0.88
2010/11	249.0	315.8	254.7	0.79	0.98
Madagascar					
2001	100.0	100.0	100.0	1.00	1.00
2005	152.4	176.3	149.9	0.86	1.02
Malawi					
2004/05	100.0	100.0	100.0	1.00	1.00
2010/11	248.5	177.6	188.2	1.40	1.32
Mozambique					
2002	100.0	100.0	100.0	1.00	1.00
2008	228.2	200.4	139.8	1.14	1.63
Pakistan					
2001/02	100.0	100.0	100.0	1.00	1.00
2005/06	132.0	131.1	124.4	1.01	1.06
2007/08	182.8	144.6	131.9	1.26	1.39
2010/11	290.3	279.2	207.1	1.04	1.40
Tanzania					
2000	100.0	100.0	100.0	1.00	1.00
2007	199.1	158.8	131.4	1.25	1.52

Note: Core CPIs are calculated based on survey data. Non-core and non-food inflation are calculated based on the sources listed in Table 1. All CPIs are normalized to 100 in the first survey year.
Source: Authors' calculations.

allocated to core foods decline. Instead, the non-food share, and in many cases also the non-core food share, increase. The core food consumption profiles of Madagascar and Mozambique have somewhat more U-shaped curves, where the very poorest spend less on food and more on non-food compared with those who consume a little more.

In all countries except Ethiopia, the non-core food share increases along the income distribution. This empirical regularity, combined with the use of the general food inflation index as the non-core food index, implies that the increase in inequality due to the composition effect is underestimated in all countries except Madagascar. Figure 2 shows the average composition CPI for each percentile of the consumption distribution. Results are as expected, given the inflation rates and the consumption shares reported here. In all countries except Ethiopia and Madagascar, the composition CPI is higher for the lower part of the distribution. This indicates that the consumption structure of the poor, combined with the observed price changes, resulted in higher price increases for the poor in this period. The magnitudes of the effects are country-specific. For instance, there is only a slight slope over the consumption distribution in Malawi. In Pakistan for the year 2005/06, only the top percentiles are notably different.

Figure 1: Consumption shares by consumption percentiles

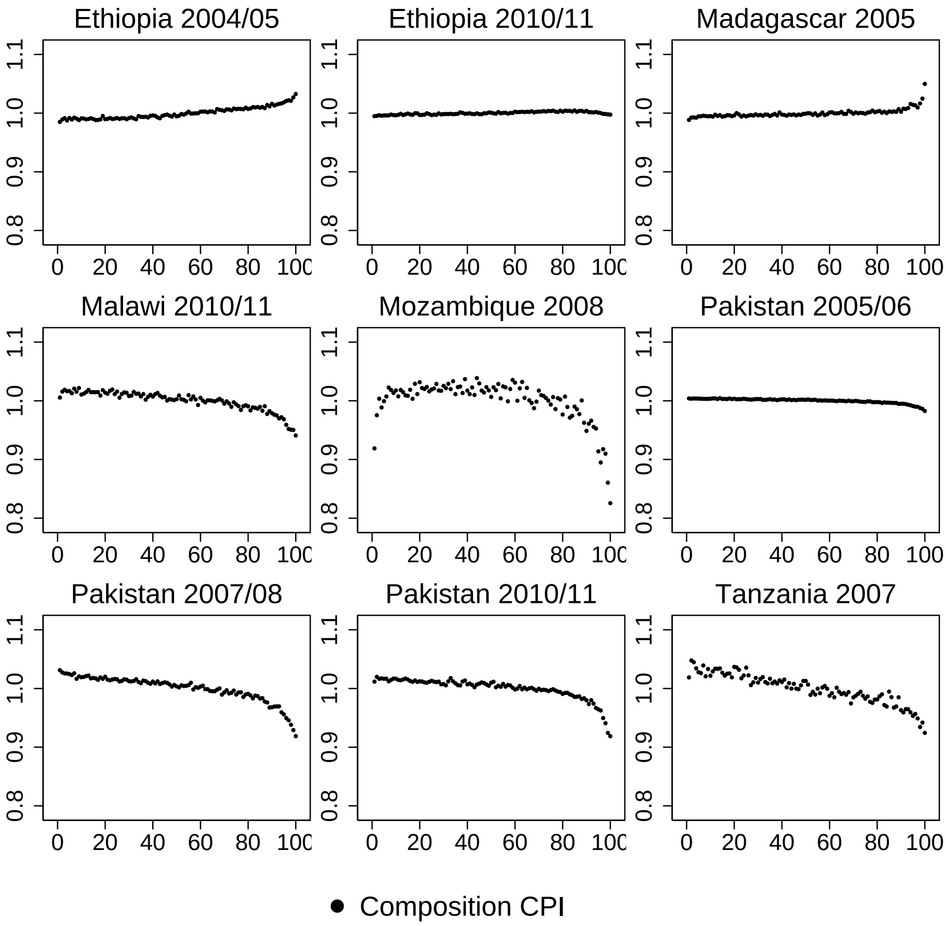


Note: In each scatter plot, each point represents the average for a percentile of the consumption distribution.
Source: Author's calculations.

2.4.2 The quantity discounting effect

Figure 3 shows the simple expensiveness indices of Equations (2) and (3). The figure plots the average indices of each consumption percentile. The results shown are estimated using quantity bins rather than units, since the quantity bins approach is applicable to all surveys. Using either expensiveness index, less-poor households face higher unit prices in almost all cases. This is consistent with richer households buying higher quality. However, for Tanzania, Mozambique, and Malawi, there is a tendency for the index estimated without quantity bins to be higher than the index with quantity bins in the lower parts of the income distributions and for a reversal of this trend as one moves further along the distribution. This is precisely what one would expect in the presence of quantity discounting effects. It is also evident that the effects are noisily estimated. I return to this point in section 2.4.5.

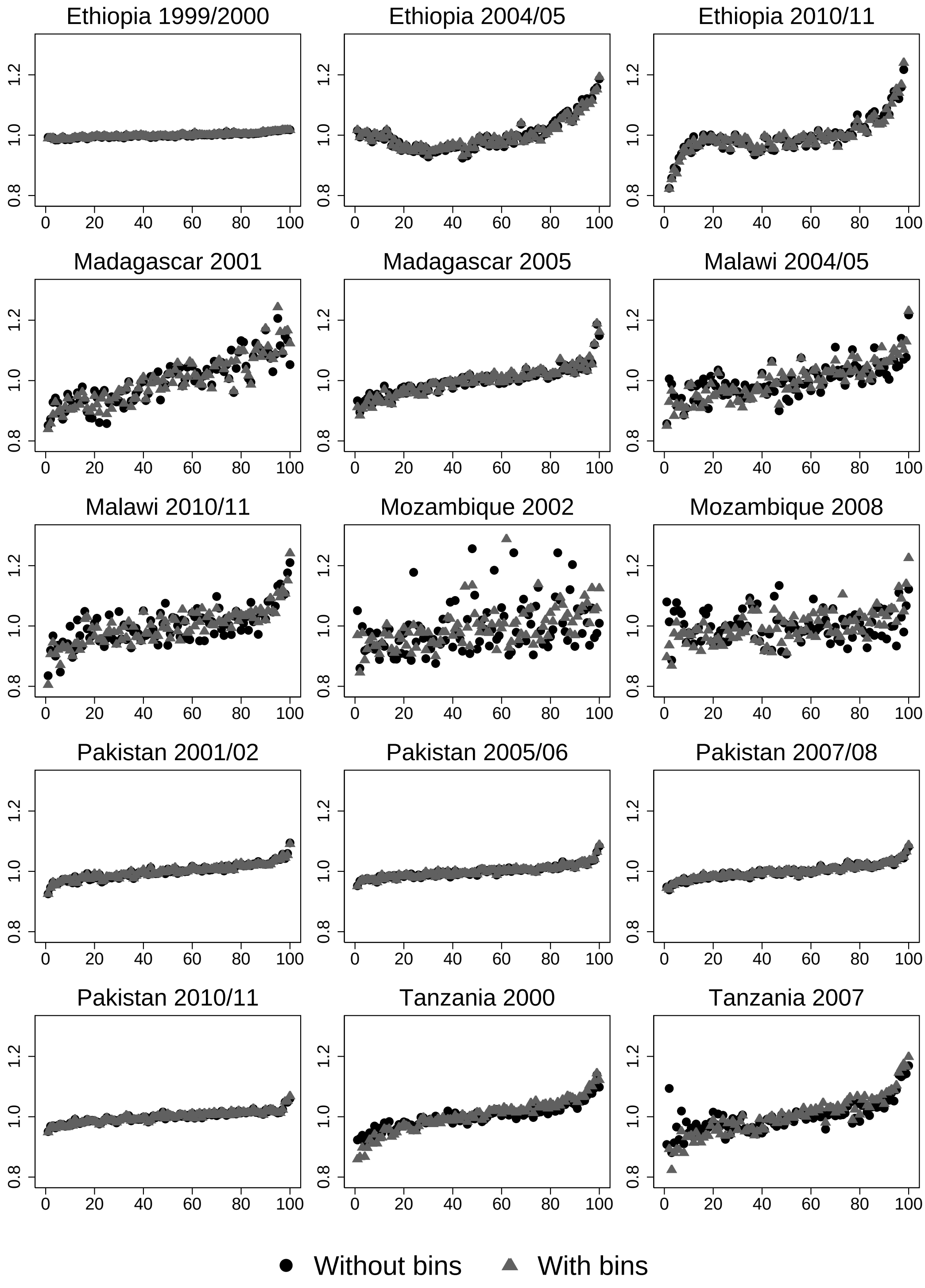
Figure 2: Composition CPI by country



Note: In each scatter plot, each point represents the average for a percentile of the consumption distribution. A few points at the extreme ends of the consumption distributions are outside the graph areas. The year of first survey for each country against which the effects are calculated are as follows: Ethiopia: 1999/2000; Madagascar: 2001; Malawi: 2004/05; Mozambique: 2002; Pakistan: 2001/02; Tanzania: 2000.
Source: Author's calculations.

For Ethiopia and Pakistan as well as Madagascar in 2005, there is almost no difference between the two expensiveness indices: The percentile-specific averages almost completely overlap. This complete lack of variation between the two expensiveness indices for some surveys is surprising. Even with just random variation in prices, one would expect at least some differences. A closer look at the price data reveals that the prices of the most common food items in the surveys of Ethiopia in 1999/2000 and 2004/05, Pakistan in all survey years, and Madagascar in 2005 have lower coefficients of variation than prices of other country surveys (Appendix Table A.1). For the surveys from Pakistan, most unit prices come out as integers when dividing values with quantities. This is unlikely to be the case if quantities and values were recorded separately, and is not the case in the other studied countries. This leads to some concerns regarding the

Figure 3: Expensiveness indices by country and survey using the bins-of-quantities formulation



Note: In each scatter plot, each point represents the average for a percentile of the consumption distribution. A few points at the extreme ends of the consumption distributions are outside the graph areas.
Source: Author's calculations.

amount of adjustments that the raw survey data of Ethiopia and Madagascar and, in particular, of Pakistan may have undergone before being made available.

Figure 4 shows the mean of the estimated quantity discounting CPIs by percentiles. The figure shows CPIs that are based on quantity binning for all surveys. For the surveys where it is possible (Madagascar in 2001, Malawi and Mozambique), the CPIs based on the approach that exploits information on units of purchase are also shown. The quantity binning-based CPIs exhibit a downward slope over the consumption distribution for the surveys of Mozambique, Tanzania, and Malawi. In these three countries, it appears that the quantity discounting effect is indeed at work in the sense that the poorest are paying higher unit prices solely because of the size of their purchase. On the other hand, Ethiopia, Pakistan, and Madagascar show no sign of a quantity discounting effect. The CPIs based on units of purchase are in general not very different from their counterparts based on quantity bins in the three surveys where both can be estimated. The results regarding inequality and poverty in what follows are therefore estimated using the quantity bin-based quantity discounting CPIs.

2.4.3 Inequality

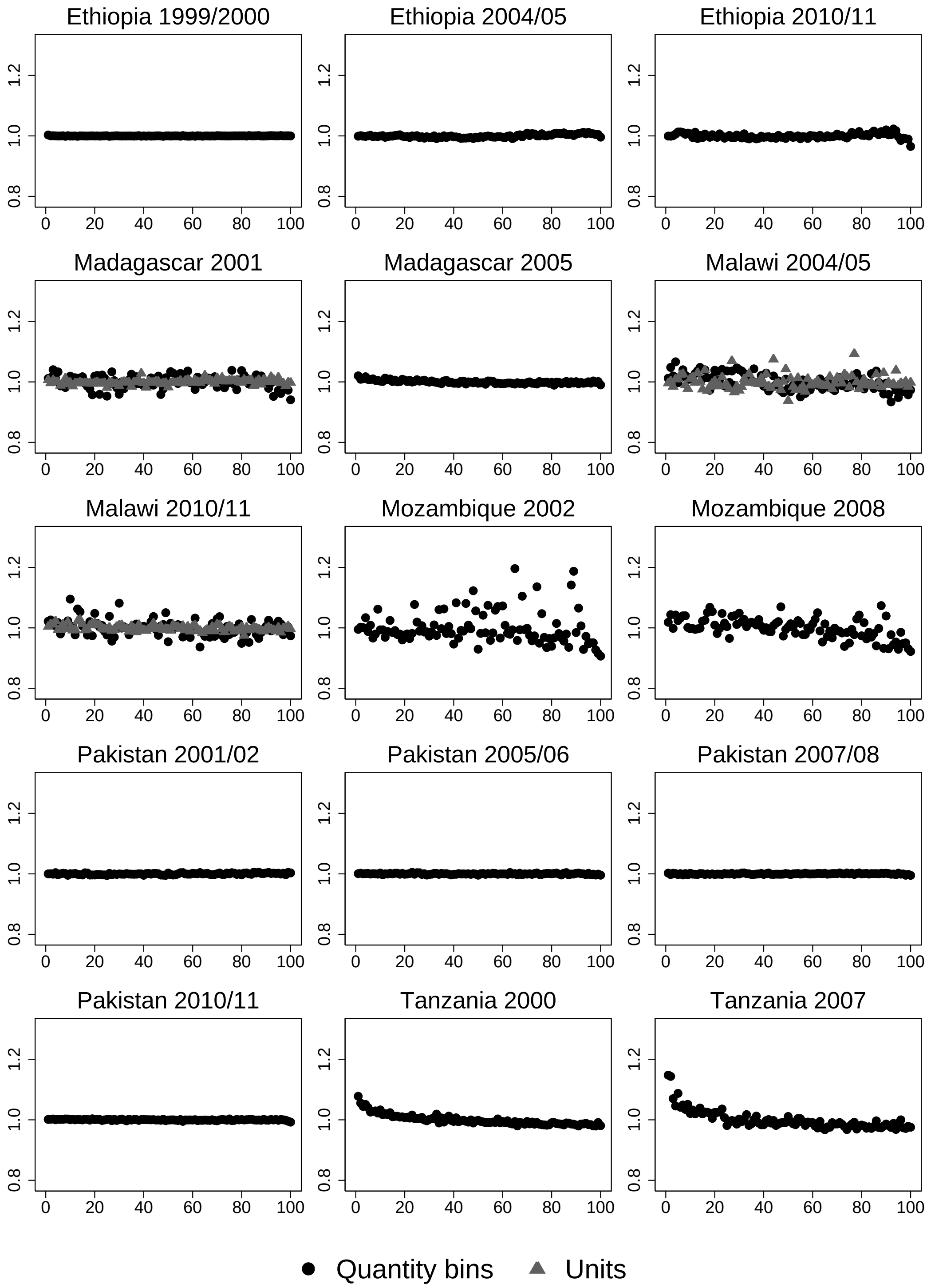
Table 4 shows the real Gini coefficients that result from applying the household-specific deflators of the previous sections. The first thing to note is that even the nominal Gini coefficients of WIID and those of the GAPP database differ. In some cases, such as Malawi, this is partly caused by the re-estimation of the consumption aggregate (Pauw et al., 2016). Another source of variation is the temporal and spatial deflation of the nominal consumption aggregates. However, these differences are not driving the results of what follows: The effects on the Gini coefficients would have been qualitatively similar if the household-specific deflators had been applied to consumption aggregates that exactly reproduce WIID Gini coefficients.

The composition effect

The composition effect results in higher real inequality than nominal inequality in all countries except Ethiopia and Madagascar where the effect is slightly negative. Effect sizes are economically meaningful. For example, although one would draw the conclusion from the nominal Gini coefficients that inequality in Mozambique was almost unchanged (or slightly decreasing, using the figures of the WIID database), the real Gini coefficients show an increase of 2.8 Gini points (i.e., from 41.5 to 44.3). In Tanzania, the nominal (GAPP) inequality measure increases by 1.1 Gini points from 2000 to 2007. However, applying the composition deflator more than doubles this to 2.5 Gini points (i.e., from 34.2 to 36.7).

The annual change in the composition-adjusted Gini coefficient compared to the annual change in the nominal (GAPP) Gini varies from 0.06 Gini points (for Pakistan from

Figure 4: Quantity CPI by country and survey



Note: In each scatter plot, each point represents the average for a percentile of the consumption distribution. A few points at the extreme ends of the consumption distributions are outside the graph areas.
Source: Author's calculations.

Table 4: Gini coefficients using alternative deflators

	WIID	GAPP	Quantity (Q)	Composition (C)	Both (B)	Q - GAPP	C - GAPP	B - GAPP
Ethiopia								
1999/2000	30.0	28.9	28.9			0.0		
2004/05	29.8	32.6	32.6	32.0	32.0	0.0	-0.6	-0.6
2010/11	29.8	32.1	32.3	32.0	32.2	0.1	-0.1	0.1
Madagascar								
2001	45.3	45.4	45.6			0.2		
2005	41.0	41.0	41.1	40.6	40.8	0.2	-0.4	-0.2
Malawi								
2004/05	41.0	41.9	42.7			0.8		
2010/11	39.3	44.5	45.3	45.4	46.2	0.8	1.0	1.7
Mozambique								
2002	47.1	41.5	42.1			0.6		
2008	41.4	41.4	42.7	44.3	45.4	1.3	2.8	4.0
Pakistan								
2001/02	30.4	26.8	26.8			0.0		
2005/06	32.7	28.5	28.5	28.7	28.7	0.0	0.2	0.3
2007/08	30.0	27.9	27.9	29.2	29.2	0.0	1.3	1.3
2010/11	30.6	26.0	26.1	27.2	27.2	0.1	1.1	1.2
Tanzania								
2000	34.6	34.2	34.8			0.6		
2007	35.0	35.3	36.2	36.7	37.6	0.9	1.4	2.3

Source: Author's calculations.

2007/08 to 2010/11) over 0.47 (for Mozambique from 2002 to 2008) to 1.08 (for Pakistan from 2005/06 to 2007/08.[11] To give an idea about magnitudes, these figures can be compared to the average annual absolute change in the nominal (GAPP) Gini coefficients of the countries included in this paper. The annual change in inequality is 0.5 Gini points. This means that composition adjustments of Gini coefficients are in some cases substantial, compared to the average change in the nominal Gini. Thus, the composition effect can severely alter the inequality track record for some, but not all, of the countries included in this paper.

The quantity discounting effect

The quantity discounting effect also increases the level of inequality substantially in some cases. For Mozambique, the level of inequality increases by between 0.6 and 1.3 Gini points, depending on the survey. For Tanzania, the increase is between 0.6 and 0.9 Gini points. For Malawi, the effect is 0.8 Gini points in both survey rounds. However, the effect is not found in all countries: Pakistan and Ethiopia show no signs of quantity discounting effects. Whereas the composition effect is one that – at least in the time period covered by the surveys included in this paper – builds up over time, the

[11] These figures can easily be calculated from Table 4.

quantity effect does not have this property. The quantity discounting effect is affecting the level, rather than the trend, of inequality.

The combined effect

The rightmost column in Table 4 shows results when both deflators are applied. In general, the combined effect is close (but not exactly equal) to the sum of the two effects. The combined effect of the quantity discounting effect and the composition effect implies that nominal inequality tends to underestimate the level of inequality and overestimate reductions in inequality in several countries. Since country growth performance and policy effectiveness are often evaluated in the context of such changes, it is important to consider the possibility that nominal inequality measures may be downwards biased.

2.4.4 Poverty

Table 5 shows the poverty rates calculated using the national accounts means and the Gini coefficients of Table 4. For countries such as Mozambique, Tanzania and Malawi where substantial differences in inequality were found, sizable differences in poverty are also found. For instance, a combination of the quantity discounting and the composition effect raises the poverty rate by 4.2 percentage points in Mozambique in 2008, by 2.7 percentage points in Tanzania in 2007, and by 1.8 percentage points in Malawi in 2010/11. The composition effect alone raises the poverty estimate by 3.0 percentage points in the 2008 Mozambique survey and by 1.7 percentage points in the 2007 Tanzania survey. For Pakistan, there are only very small effects on poverty from using real inequality in place of nominal. This is caused by the relatively high GDP per capita in this country, as measured from the national accounts, which gives poverty rates close to zero. The increase in inequality is not large enough to alter this result substantially. In Ethiopia and Madagascar, the estimated effect is smaller and sometimes even slightly negative. Since the composition effect builds up over time, the discrepancy in poverty estimates is bigger in later surveys. Given these results, the optimistic picture of very fast poverty reduction in Sub-Saharan African countries presented by Pinkovskiy and Sala-i Martin (2014) should be interpreted with caution: Although the technique still shows substantial poverty reductions when real inequality estimates are used, the level is higher and the pace of reduction is slower overall.

2.4.5 Robustness of results

It is likely that the household-specific deflators contain noise due to measurement error in the price and quantity data of the individual households. This is potentially a serious problem, since measurement error in the deflators will increase the gini coefficients. This is most easily seen by considering the connection between the gini coefficient and the variance of the consumption distribution. In the case where the consumption

Table 5: Poverty rates and changes using different inequality measures

	WIID	GAPP	Quantity (Q)	Composition (C)	Both (B)	Q - GAPP	C - GAPP	B - GAPP
Ethiopia								
1999/2000	50.3	49.5	49.5			0.0		
2004/05	36.6	39.8	39.8	39.1	39.1	0.0	-0.6	-0.7
2010/11	14.6	17.6	17.8	17.5	17.7	0.2	-0.1	0.1
Madagascar								
2001	34.2	34.3	34.5			0.2		
2005	32.2	32.1	32.3	31.7	31.9	0.2	-0.5	-0.3
Malawi								
2004/05	45.2	47.8	48.6			0.7		
2010/11	41.2	40.4	41.3	41.4	42.2	0.8	1.0	1.8
Mozambique								
2002	56.7	52.7	53.1			0.5		
2008	38.1	38.1	39.5	41.1	42.4	1.4	3.0	4.2
Pakistan								
2001/02	1.2	0.4	0.4			0.0		
2005/06	1.1	0.3	0.3	0.3	0.3	0.0	0.0	0.0
2007/08	0.3	0.1	0.1	0.3	0.3	0.0	0.1	0.1
2010/11	0.3	0.0	0.0	0.1	0.1	0.0	0.0	0.0
Tanzania								
2000	24.4	23.9	24.7			0.8		
2007	13.1	13.5	14.6	15.2	16.2	1.1	1.7	2.7

Note: Poverty rates are reported in percent.
Source: Author's calculations.

distribution is assumed to be log-normal with variance σ^2 as assumed in this paper as well as by Pinkovskiy and Sala-i Martin (2014), the Gini coefficient is given by $2\Phi\left(\sigma/\sqrt{2}\right) - 1) * 100$, where Φ is the cumulative standard normal distribution (Crow and Shimizu, 1988, p. 11). If one deflates all points in the consumption distribution by a deflator that is nothing but noise, it will increase the variance of the distribution and hence increase the Gini coefficient.

Visual inspection of Figures 2, 3 and 4 does give rise to concerns that deflators are noisily estimated. It is therefore interesting to check whether the results on inequality can be attributed to measurement error. In an extreme case, the estimated household-specific deflators are random draws from a distribution of noise terms. If this is the case, the estimated deflators form an empirical distribution function of the noise terms, which can be specific to each survey. This observation forms the basis of the following robustness check. For each household, I pick with replacement a deflator of a random household from the same survey. I then use these randomly allocated deflators to deflate consumption and re-estimate the Gini coefficient. For each survey, I conduct 1,000 iterations of this process. This provides a distribution of possible Gini coefficients that could have been estimated if the estimated deflators were random draws from the noise distribution. If the point estimates of the Gini coefficients of Table 4 are sufficiently large compared to this distribution, it is unlikely that the main results can be attributed to random measurement error of the type described above.

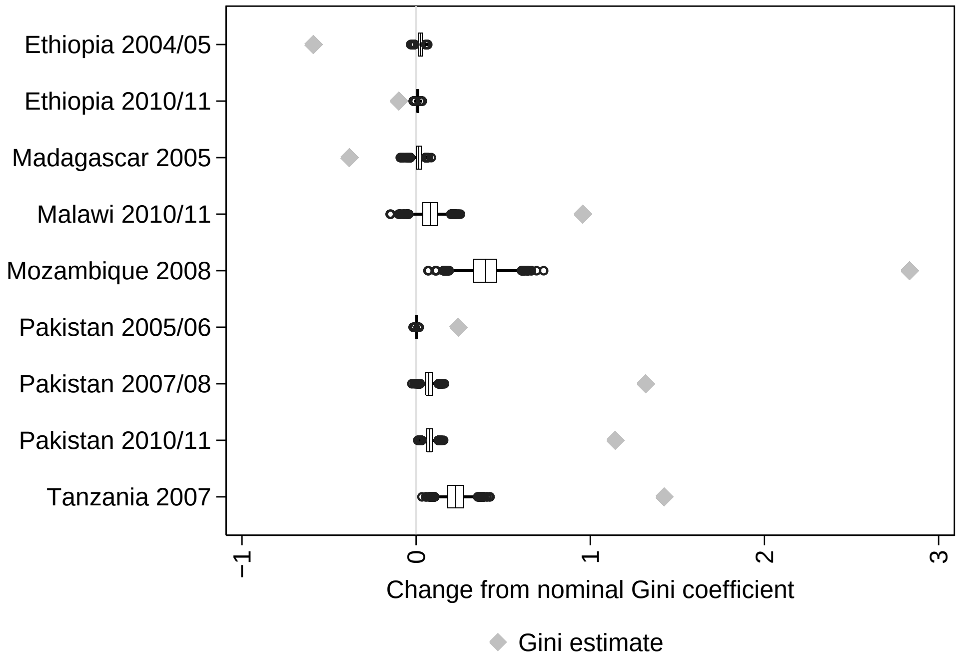
Figures 5 and 6 show the results of this robustness check, using the composition deflator and the quantity discounting deflator, respectively. For the composition deflator, results are encouraging. In Malawi, Mozambique, Pakistan and Tanzania where the composition effect led to an upwards adjustment of inequality, the estimated Gini coefficients are comfortably outside the distribution of counterfactual gini coefficients. In Ethiopia and Madagascar where the composition effect led to downward adjustments of inequality, the estimated Gini coefficients are also outside the distribution of counterfactual gini coefficients. This is not surprising, since the addition of random noise will on average increase the gini coefficient.

For the quantity deflator, I focus on the three countries where substantial effects were found, namely Malawi, Mozambique and Tanzania. The estimates of inequality for Mozambique and Tanzania are larger than any of the estimates of the 1,000 iterations. It is therefore highly unlikely that the quantity discounting effect can be attributed to measurement error, even though the median iteration does increase the gini coefficient, showing that the worry about measurement error is well founded. Malawi is an intermediate case. Even though the estimated Gini increases are of comparable magnitude to the increases found in Mozambique and Tanzania, the estimated gini increases fall within the span of estimates reached by the 1,000 iterations. For Malawi in 2004/05, the estimate is above the 81st percentile of iterations; for Malawi in 2010/11, the estimate is above the 93rd percentile of iterations. It can therefore not be ruled out that the particular noisiness of the Malawian surveys are driving the quantity discounting effect, even though the majority of iterations of the Malawi surveys result in lower estimates than the main results of this paper.

To conclude, while measurement error in the deflators appears to be a well-founded concern, the robustness check reveals a potential issue with the quantity discounting effect only in the case of Malawi. However, the point estimate is in the tail end of the simulated counterfactual distribution. For Mozambique and Tanzania, pure noise cannot explain the real Gini estimates of the main results. In none of the surveys in question can random measurement noise explain the estimated magnitudes of the composition effect.

I also conduct robustness checks of two choices made in the estimation of the quantity discounting effect, namely the number of bins chosen and the size of geographical area g within which prices are compared. First, I vary the number of quantity bins that are used in the estimation of the quantity discounting effect. The choice of four bins for the main results is not founded in theoretical considerations, so it is important that results are not driven by this choice. Appendix Table A.2 reports Gini coefficients using

Figure 5: Assigning random composition deflators

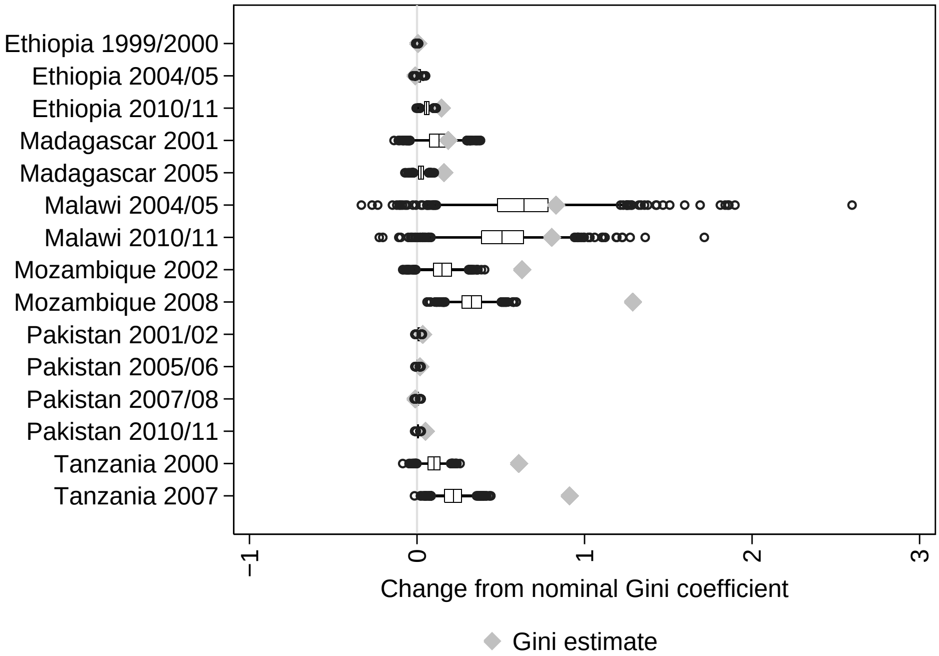


Note: The top and bottom of the box is the first and third quartile of 1,000 iterations of randomly drawn deflators; the top and bottom of the box represent the first and third quartiles; whiskers represent the 2.5th and 97.5th percentile; circles represent estimates outside the whiskers. Results are given in changes form the nominal (GAPP) gini coefficient for each survey. Grey diamonds show the difference from the GAPP gini coefficient when the composition deflator is applied.
Source: Author's calculations.

2, 4, 8, or 16 quantity bins. Varying the number of bins has only a limited effect on the estimated quantity discounting effect. The largest change in any of the alternative estimations, compared to the baseline option of four bins, is the case of Mozambique in 2008, where the 2-bin Gini estimate is 0.5 Gini points lower. Going from more to fewer bins has no systematic effect. For Malawi, Mozambique, and Tanzania, 2 bins produce lower equality estimates compared to 4 bins, but the reverse is true for Madagascar in 2001, and the estimate for Malawi in 2010/11 is also lower using 8 and 16 bins. In sum, the choice of number of bins does not appear to be driving the results of the quantity discounting effect.

Second, I vary the spatial unit within which quantity discounting effects are estimated. For the main results, the strata level was chosen as it smaller than the national level but retains a sizable number of households within each stratum in each survey. Appendix Table A.3 reports Gini coefficients using either no spatial adjustment, using the strata as the spatial unit (the baseline option) and using the enumeration area, of which there are many within each stratum. Compared to the within-strata estimation, the Gini coefficients increase in most instances when no spatial unit is used. The increase is substantial for some surveys such as the two Malawian surveys (1.4 and 1.9 Gini points)

Figure 6: Assigning random quantity discounting deflators



Note: The top and bottom of the box is the first and third quartile of 1,000 iterations of randomly drawn deflators; the top and bottom of the box represent the first and third quartiles; whiskers represent the 2.5th and 97.5th percentile; circles represent estimates outside the whiskers. Results are given in changes form the nominal (GAPP) gini coefficient for each survey. Grey diamonds show the difference from the GAPP gini coefficient when the quantity discounting deflator is applied.
Source: Author's calculations.

and the 2002 Mozambican survey (1.0 Gini points). This indicates a positive correlation between the average stratum price level and the average stratum consumption level, even though prices are deflated before estimation using regional spatial price indices. This correlation affects the estimates that do not account for spatial units, but does not affect estimates when the household-specific deflator is based on within-strata price variation only, as is the case of the main results of this paper. On the other hand, it is not clear that the between-strata price variation could not also be caused by quantity discounting effects. In this sense, employing the strata as the spatial unit gives a conservative estimate. The final column of Table A.3 uses the enumeration area as the spatial unit. This brings Gini estimates very close to the nominal Gini coefficients. It is possible that this means that there are no quantity discounting effects present once spatial price variation is controlled for at a sufficiently detailed level. However, given the high number of enumeration areas, the lack of effect is likely to be caused by a lack of a sufficient number of price observations for all the products of the survey to meaningfully estimate the quantity discounting effect within enumeration areas. To conclude, the level of the strata is a reasonable choice of spatial unit as it is conservative, compared to using no spatial unit, but retains sufficient price variation to estimate quantity discounting effects.

2.5 Conclusion

This paper shows how two different effects can drive wedges between estimates of nominal and real inequality. The first effect works through the combination of differential consumption structures across the consumption distribution and differential price increases of different product groups. The second effect works through quantity discounting: The poor pay more for their food consumption since they purchase food items in smaller quantities. Household-specific deflators are calculated for 15 surveys from 6 different countries which cover a range of varying experiences in terms of levels of consumption and inequality as well as their trends over time. A key advantage of the methodology of the current paper is that it relies only on information that is available in many existing nationally representative surveys of developing countries, and on widely available data on national food and non-food consumer price indices.

Composition effects were found in Malawi, Mozambique, Pakistan, and Tanzania but not in Ethiopia and Madagascar. Quantity discounting effects were found in Malawi, Mozambique and Tanzania; no effects were found in Pakistan, Madagascar, and Ethiopia. While measurement error is potentially a concern, a robustness check revealed that only for the quantity discounting effect and only in the case of Malawi was this effect large enough to sow some doubts about the existence of an effect.

In most cases, the estimated effects are lower bounds on the true effect sizes. Nonetheless, the impacts on inequality and on the derived poverty estimates are in some cases substantial. Estimated real Gini coefficients are between -0.6 and 4.0 Gini points higher than nominal Gini coefficients. In four countries (Malawi, Mozambique, Pakistan, and Tanzania), real inequality is higher than nominal inequality. Using real inequality indices can also affect inference on the speed of inequality changes. In extreme cases, it can change the direction of inequality trends so that a decrease in nominal inequality conceals an increase in real inequality (Mozambique). At the other end of the spectrum, real inequality was not found to be different from nominal inequality in two of the studied countries (Ethiopia and Madagascar).

Finally, inequality estimates matter for estimating poverty based on national account means and survey-based inequality estimates. In countries where the composition and quantity discounting effects affect the Gini coefficients, the poverty rates are also affected. While the quantity discounting effect potentially affects inequality indices in every year, the composition effect builds up over time as prices diverge. This means that in countries where later surveys are more heavily influenced by the composition effect, the use of nominal inequality indices does not only introduce a source of bias in the level of poverty, but may also overestimate the rate of poverty reduction.

The effects are highly country-specific. For the composition effect, this is caused by differences in consumption structures and in inflation rates. Inflation rates are affected by a complex interaction of domestic conditions such as harvest fortunes and government policies, as well as international changes in world market prices. Especially for the surveys conducted in the years of the food price crisis of 2007–09, world market prices of basic food items were very high. As new survey rounds become available, it will be interesting to see whether the composition effect shrinks again, or whether it is a longer-lasting phenomenon. For the quantity discounting effect, the cross-country differences are likely caused by a mix of real differences in the magnitude of quantity discounting present and of differences due to varying survey instruments and methodologies. The latter appears to be important for estimates of the quantity discounting effect in Pakistan and Ethiopia where there is little price variation for the most common food items, which helps to explain the lack of quantity effects in these countries.

The estimation of the composition and the quantity discounting effects requires only household consumption data and CPI information, which is available in many countries. Further, since the two effects are straight-forward to estimate, I suggest doing so for other countries, and when new surveys become available, in order to check whether keeping inequality in real terms matters in different country- and time-specific contexts.

References

Aguiar, M. and Hurst, E. (2007). Life-Cycle Prices and Production. *The American Economic Review*, 97(5):1533–1559.

Aksoy, M. A. and Isik-Dikmelik, A. (2008). Are low food prices pro-poor? Net food buyers and sellers in low-income countries. Policy Research Working Paper 4642, World Bank.

Arndt, C., Demery, L., McKay, A., and Tarp, F. (2016a). Growth and poverty reduction in Tanzania. In Arndt, C., Andy, M., and Tarp, F., editors, *Growth and Poverty in Sub-Saharan Africa*. Oxford University Press, Oxford. Forthcoming.

Arndt, C., Jones, S., and Salvucci, V. (2015). When do relative prices matter for measuring income inequality? The case of food prices in Mozambique. *The Journal of Economic Inequality*, 13(3):449–464.

Arndt, C., Jones, S. E., and Tarp, F. (2016b). Mozambique: Off-track or Temporarily Sidelined? In Arndt, C., Andy, M., and Tarp, F., editors, *Growth and Poverty in Sub-Saharan Africa*. Oxford University Press, Oxford. Forthcoming.

Arndt, C., Tarp, L., and McKay, A. (2016c). Two Cheers for the African Growth Renaissance (but not three). In Arndt, C., Andy, M., and Tarp, F., editors, *Growth and Poverty in Sub-Saharan Africa*. Oxford University Press, Oxford. Forthcoming.

Attanasio, O. P. and Frayne, C. (2006). Do the poor pay more? Unpublished, Institute for Fiscal Studies.

Beatty, T. K. M. (2010). Do the Poor Pay More for Food? Evidence from the United Kingdom. *American Journal of Agricultural Economics*, 92(3):608–621.

Cage, R. A., Garner, T. I., and Ruis-Castillo, J. (2002). Constructing Household Specific Consumer Price Indexes: An Analysis of Different Techniques and Methods. Technical Report Working Paper 354, U.S. Bureau of Labor Statistics.

Chen, S. and Ravallion, M. (2010). The Developing World is Poorer than We Thought, But No Less Successful in the Fight Against Poverty. *The Quarterly Journal of Economics*, 125(4):1577–1625.

Chung, C., Dong, D., Schmit, T. M., Kaiser, H. M., and Gould, B. W. (2005). Estimation of price elasticities from cross-sectional data. *Agribusiness*, 21(4):565–584.

CountrySTAT (2015). http://countrystat.org/. Accessed May 15 2015.

Crawford, I., Laisney, F., and Preston, I. (2003). Estimation of household demand systems with theoretically compatible Engel curves and unit value specifications. *Journal of Econometrics*, 114(2):221–241.

Crow, E. L. and Shimizu, K. (1988). *Log-normal Distributions: Theory and Application*. Marcel Dekker, New York.

CSA (2015). Country and regional level consumer price indices. Information 23, Central Statistical Agency, Addis Ababa.

CSA (2000). Household Income, Consumption and Expenditure Survey. Dataset 2000. http://microdata.worldbank.org/index.php/catalog. Central Statistical Agency. Accessed December 15 2015.

CSA (2005). Household Income, Consumption and Expenditure Survey. Dataset 2005. http://microdata.worldbank.org/index.php/catalog. Central Statistical Agency. Accessed December 15 2015.

CSA (2011). Household Income, Consumption and Expenditure Survey. Dataset 2011. http://microdata.worldbank.org/index.php/catalog. Central Statistical Agency. Accessed December 15 2015.

Deaton, A. (1988). Quality, Quantity, and Spatial Variation of Price. *The American Economic Review*, 78(3):418–430.

Deaton, A. (2003). Prices and Poverty in India, 1987-2000. *Economic and Political Weekly*, 38(4):362–368.

Deaton, A. and Zaidi, S. (2002). *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*. Number 14101 in World Bank Publications. The World Bank.

Dorosh, P. A. and Minten, B. (2006). Rice Price Stabilization in Madagascar: Price and Welfare Implications of Variable Tariffs. 2006 Annual Meeting, August 12-18, International Association of Agricultural Economists.

Durevall, D., Loening, J. L., and Ayalew Birru, Y. (2013). Inflation dynamics and food prices in Ethiopia. *Journal of Development Economics*, 104:89–106.

FBS (2003). Household Integrated Economic Survey. Round 4: 2001-02. Technical report, Federal Bureau of Statistics Pakistan, Islamabad, Pakistan.

FBS (2007). Household Integrated Economic Survey. 2005-06. Technical report, Federal Bureau of Statistics Pakistan, Islamabad, Pakistan.

FBS (2009). Household Integrated Economic Survey. 2007-08. Technical report, Federal Bureau of Statistics Pakistan, Islamabad, Pakistan.

FBS (2013). Household Integrated Economic Survey. 2011-12. Technical report, Federal Bureau of Statistics Pakistan, Islamabad, Pakistan.

Günther, I. and Grimm, M. (2007). Measuring pro-poor growth when relative prices shift. *Journal of Development Economics*, 82(1):245–256.

Goñi, E., López, H., and Servén, L. (2006). Getting Real about Inequality. Evidence from Brazil, Colombia, Mexico, and Peru. Technical Report 3815, World Bank.

Grimm, M., Wetta, C., and Nikiema, A. (2016). Burkina Faso: Shipping Around the Malthusian Trap. In Arndt, C., Andy, M., and Tarp, F., editors, *Growth and Poverty in Sub-Saharan Africa*. Oxford University Press, Oxford. Forthcoming.

Grosh, M. and Deaton, A. (2000). ch. 17: Consumption. In Glewwe, P., Deaton, A., and Grosh, M., editors, *Designing household survey questionnaires for developing countries: Lessons from ten years of LSMS experience*. World bank.

Guénard, C. and Mesplé-Somps, S. (2010). Measuring Inequalities: Do Household Surveys Paint a Realistic Picture? *Review of Income and Wealth*, 56(3):519–538.

INE (2015). Índices de preços no Consumidor Moçambique. National Institute of Statistics Mozambique. Data obtained as supplementary material from the National Directorate for Studies and Policy Analysis (DNEAP) in Mozambique.

INSTAT (2002). Enquete aupres des menages 2001 - rapport principal. Technical report, Institut National de la Statistique.

INSTAT (2006). Enquete aupres des menages 2005 - rapport principal. Technical report, Institut National de la Statistique.

INSTAT (2015). Institut National de la Statistique de Madagascar. www.instat.mg, accessed May 15 2015.

Leicester, A., O'Dea, C., and Oldfield, Z. (2008). The Inflation Experience of Older Households. IFS Commentary 106, The Insititute for Fiscal Studies.

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341(6149):976–980.

McKelvey, C. (2011). Price, unit value, and quality demanded. *Journal of Development Economics*, 95(2):157–169.

Mitchell, D. (2008). A Note on Rising Food Prices. SSRN Scholarly Paper ID 1233058, Social Science Research Network, Rochester, NY.

MoF (2015). Pakistan Economic Survey. Technical report, Ministry of Finance Pakistan.

Mohsin, A. and Zaman, K. (2012). Distributional effects of rising food prices in Pakistan: Evidence from HIES 2001-02 and 2005-06 survey. *Economic Modelling*, 29(5):1986–1995.

MPF (2004). Poverty and well-being in Mozambique: the second national assessment. Technical report, Ministry of Planning and Finance and International Food Policy Research Institute and Purdue University.

MPF and DNEAP (2010). Poverty and well-being in Mozambique: third national assessment. Technical report, Ministry of Planning and Finance and National Directorate of Studies and Policy Analysis.

Muellbauer, J. (1974). Prices and Inequality: The United Kingdom Experience. *The Economic Journal*, 84(333):32–55.

Mussa, R. (2014). Food Price Heterogeneity and Income Inequality in Malawi: Is Inequality Underestimated? MPRA Paper 56080, Munich Personal RePEc Archive.

Mussa, R. (2015). Do the Poor Pay More for Maize in Malawi? *Journal of International Development*, 27(4):546–563.

Nazli, H., Whitney, E., and Marht, K. (2015). Poverty Trends in Pakistan. WIDER Working Paper 2015/136, UNU-WIDER.

NBE (2014). Annual Report: National Bank annual report 2013-2014. Technical report, National Bank of Ethiopia, Dar es Salaam, Tanzania. Accessed May 1st 2015.

NBS (2002). Household Budget Survey 2000/01. Technical report, National Bureau of Statistics Tanzania, Dar es Salaam, Tanzania.

NBS (2011). National Household Budget Survey 2007. Technical report, National Bureau of Statistics Tanzania.

NSO (2005). Integrated Household Survey 2004-2005. Volume I: Household Socio-Economic Characteristics. Technical report, National Statistics Office Malawi, Zomba, Malawi.

NSO (2012). Integrated Household Survey 2010-2011. Household Socio-economic Characteristics Report. Technical report, National Statistics Office Malawi, Zomba, Malawi.

NSO (2015). National Statistics Office of Malawi webpage. www.nsomalawi.mw. Accessed March 6, 2015.

Pauw, K., Beck, U., and Mussa, R. (2016). Did Rapid Smallholder-led Agricultural Growth Fail to Reduce Rural Poverty? Making Sense of Malawi's Poverty Puzzle. In

Arndt, C., Andy, M., and Tarp, F., editors, *Growth and Poverty in Sub-Saharan Africa*. Oxford University Press, Oxford. Forthcoming.

PBS (2006). Household Integrated Economic Survey. 2004-05. Technical report, Pakistan Bureau of Statistics, Islamabad, Pakistan.

PBS (2007). Household Integrated Economic Survey. 2005-06. Technical report, Pakistan Bureau of Statistics, Islamabad, Pakistan.

PBS (2009). Household Integrated Economic Survey. 2007-08. Technical report, Pakistan Bureau of Statistics, Islamabad, Pakistan.

PBS (2013). Household Integrated Economic Survey (HIES). 2011-12. Technical report, Pakistan Bureau of Statistics, Islamabad, Pakistan.

Pinkovskiy, M. and Sala-i Martin, X. (2009). Parametric Estimations of the World Distribution of Income. Working Paper 15433, National Bureau of Economic Research.

Pinkovskiy, M. and Sala-i Martin, X. (2014). Africa is on time. *Journal of Economic Growth*, 19(3):311–338.

Pritchett, L., Suryahadi, A., Sumarto, S., and Suharso, Y. (2000). The evolution of poverty during the crisis in Indonesia, 1996-99. Policy Research Working Paper Series 2435, The World Bank.

Rao, V. (2000). Price Heterogeneity and "Real" Inequality: A Case Study of Prices and Poverty in Rural South India. *Review of Income and Wealth*, 46(2):201–211.

Ravallion, M. and Bidani, B. (1994). How Robust Is a Poverty Profile? *The World Bank Economic Review*, 8(1):75–102.

Ravallion, M., Chen, S., and Sangraula, P. (2009). Dollar a Day Revisited. *The World Bank Economic Review*, 23(2):163–184.

Sala-i Martin, X. (2006). The World Distribution of Income: Falling Poverty and... Convergence, Period. *The Quarterly Journal of Economics*, 121(2):351–397.

Sala-i Martin, X. and Pinkovskiy, M. (2010). African Poverty is Falling...Much Faster than You Think! Working Paper 15775, National Bureau of Economic Research.

Stifel, D., Razafimanantena, T., and Rakotomanana, F. (2016). Utility-Consistent Poverty in Madagascar, 2001-10: Snapshots in the Presence of Multiple Economy-Wide Shocks. In Arndt, C., Andy, M., and Tarp, F., editors, *Growth and Poverty in Sub-Saharan Africa*. Oxford University Press, Oxford. Forthcoming.

Stifel, D. and Woldehanna, T. (2016). Poverty in Ethiopia, 2000-11: Welfare Improvements in a Changing Economic Landscape. In Tarp, F., Arndt, C., and Andy, M., editors, *Growth and Poverty in Sub-Saharan Africa*. Oxford University Press.

Tarp, F., Simler, K., Matusse, C., Heltberg, R., and Dava, G. (2002). The Robustness of Poverty Profiles Reconsidered. *Economic Development and Cultural Change*, 51(1):77–108.

UNU-WIDER (2014). World Income Inequality Database (WIID 3.0b). wider.unu.edu. Accessed December 15, 2015.

Wiggins, S., Keats, S., and Compton, J. (2010). What Caused the Food Price Spike of 2007/08? Lessons for World Cereals Markets. Food prices project report, Overseas Development Institute.

World Bank (2012). World Development Indicators 2012. data.worldbank.org. Accessed December 15 2015.

Appendix

Table A.1: Coefficient of variation of prices of the three most consumed food items in each survey

	Most consumed	2nd most consumed	3rd most consumed	Simple average
Ethiopia				
1999/2000	0.16	0.18	0.26	0.20
2004/05	0.25	0.29	0.22	0.25
2010/11	0.41	0.39	0.39	0.40
Madagascar				
2001	0.59	0.31	0.26	0.39
2005	0.19	0.17	0.19	0.18
Malawi				
2004/05	0.92	0.64	0.65	0.74
2010/11	0.79	0.54	0.40	0.58
Mozambique				
2002	0.56	0.21	0.66	0.48
2008	0.59	0.56	0.60	0.58
Pakistan				
2001/02	0.33	0.15	0.08	0.19
2005/06	0.23	0.13	0.22	0.19
2007/08	0.30	0.24	0.09	0.21
2010/11	0.48	0.31	0.12	0.30
Tanzania				
2000	0.64	0.64	0.16	0.48
2007	0.54	0.15	0.62	0.44

Note: Most consumed is defined in terms of frequency of consumption among households. In order to remove any effect of extreme outliers, prices below the 1st (and above the 99th) percentile are replaced by the value of the 1st (99th) percentile.
Source: Author's calculations.

Table A.2: Quantity discounting-deflated Gini's with different numbers of bins

	Number of bins			
	2	4	8	16
Ethiopia				
1999/2000	28.9	28.9	28.9	28.9
2004/05	32.6	32.6	32.6	32.6
2010/11	32.2	32.3	32.3	32.4
Madagascar				
2001	45.7	45.6	45.2	45.1
2005	41.1	41.1	41.1	41.1
Malawi				
2004/05	42.6	42.7	43.0	42.9
2010/11	45.1	45.3	45.2	45.2
Mozambique				
2002	41.9	42.1	42.1	42.1
2008	42.2	42.7	42.9	42.9
Pakistan				
2001/02	26.8	26.8	26.8	26.8
2005/06	28.5	28.5	28.5	28.5
2007/08	27.9	27.9	27.9	27.9
2010/11	26.1	26.1	26.1	26.1
Tanzania				
2000	34.6	34.8	34.9	34.8
2007	36.0	36.2	36.3	36.2

Note: The number of bins used in the main text is 4.
Source: Author's calculations.

Table A.3: Gini coefficients deflated by quantity discounting index using alternative spatial domains

	Nominal (GAPP)	Alternative spatial domains		
		No spatial unit	Strata	Enumeration area
Ethiopia				
1999/2000	28.9	28.9	28.9	28.9
2004/05	32.6	32.6	32.6	32.6
2010/11	32.1	32.3	32.3	32.1
Madagascar				
2001	45.4	45.9	45.6	45.4
2005	41.0	41.3	41.1	41.0
Malawi				
2004/05	41.9	44.1	42.7	42.1
2010/11	44.5	47.1	45.3	44.5
Mozambique				
2002	41.5	43.1	42.1	41.5
2008	41.4	42.9	42.7	41.4
Pakistan				
2001/02	26.8	26.9	26.8	26.8
2005/06	28.5	28.5	28.5	28.5
2007/08	27.9	28.0	27.9	27.9
2010/11	26.0	26.2	26.1	26.0
Tanzania				
2000	34.2	34.9	34.8	34.3
2007	35.3	36.6	36.2	35.5

Note: The spatial unit used in the main text is the strata.
Source: Author's calculations.

CHAPTER 3

SOCIAL TIES AND THE EFFICIENCY OF FACTOR TRANSFERS

Ulrik Beck* Benedikte Bjerge† Marcel Fafchamps‡

Abstract

We introduce a novel way of testing whether social and geographical proximity help or hinder efficiency-enhancing factor transfers. The approach is implemented empirically using unusually rich data collected in The Gambia. We find that neighbors conduct more efficiency-enhancing transfers but that membership of the same ethnic group or kinship network is associated with fewer efficient transfers. The latter effect is primarily driven by land transfers from a few large landowners. If the presence of large landowners is controlled for, the finding is reversed and we find more efficiency-enhancing land transfers between kin-related households and between neighbors. Labor transfers are not found to equilibrate factor ratios across households. While most land transfers improve allocative efficiency, an efficient allocation is not achieved at the village level, which suggests that social ties are not sufficiently fluid to permit a fully efficient reallocation of factors of production within villages.

We benefited from comments from seminar participants at the Universities of Oxford, UC-Berkeley, Paris 1, Gothenburg and Goethe University Frankfurt, as well as conference participants at the CSAE conference 2014, the International Economic Association's World Congress 2014 and the Nordic Conference in Development Economics 2015. We thank Dany Jaimovich, Thomas Markussen, Carol Newman, John Rand, Måns Söderbom and Finn Tarp for very valuable comments. Finally, we thank Jean-Louis Arcand and Dany Jaimovich for making the Gambia Networks Data 2009 available to us. All interpretations and any remaining errors are the sole responsibility of the authors.

*Department of Economics, University of Copenhagen

†Department of Food and Resource Economics, University of Copenhagen

‡Freeman Spogli Institute for International Studies, Stanford University

3.1 Introduction

Since Coase (1937), economists have known that transaction costs potentially matter for the efficiency of exchange. Many avenues have been discussed by which economic agents seek to reduce transaction costs of exchange. One such avenue is reliance on pre-existing social networks. The economics literature has successfully documented the role that pre-existing social ties play in informal insurance schemes in diverse contexts (Fafchamps and Lund, 2003; Fafchamps and Gubert, 2007; Attanasio et al., 2012; Mazzocco and Saini, 2012) as well as the importance of such ties in the selection of a trading partner – e.g., in labor markets (Granovetter, 1995; Topa, 2001) and in international trade (Rauch, 2001; Rauch and Casella, 2003; Chaney, 2014). The finding that pre-existing ties matter in bilateral exchange is usually interpreted as suggesting that efficiency of exchange is probably not achieved. But the truth is that the link between pre-existing social networks and the efficiency properties of exchange has seldom been studied directly (see Sadoulet et al., 1997; Holden and Ghebru, 2005 and Macours et al., 2010 for exceptions).[1]

In this paper, we use a unique social network dataset that covers 51 rural Gambian villages to cast new light on this research question. We examine a closed factor economy in which efficient factor transactions are easily identified based on initial endowments. We observe all pairwise factor transactions between all agents in this economy, and have data on pre-existing social ties for each pair of agents. To understand the nature of the limits imposed by social networks, we investigate whether allocative efficiency is achieved by frictionless trading over the pre-existing social network.

Factor markets in developing countries are plagued by incomplete information and unclear property rights, which lead to high transaction costs and hinder the efficient allocation of production factors (de Janvry et al., 1991; Pender and Fafchamps, 2005; Otsuka, 2007). Pre-existing social networks can play an important role in alleviating these issues by increasing trust and decreasing information asymmetries, thereby reducing transaction costs. Land usage rights in rural Gambia are vested in the hands of the descendants of the ancestral settlers, and thus the distribution of land ownership rights is highly unequal. This results in an active network of land exchange where temporary land usage rights are often transferred on an annual basis. This makes The Gambia a suitable candidate for testing how pre-existing networks in the form of ethnicity, kinship and geographical proximity affect the efficiency properties of land exchange.

[1] Sadoulet et al. (1997) compare contracts between kin related households and non-kin related households and find that the former are more efficient. Holden and Ghebru (2005) show that tenants in communities with larger shares of contracts between kin related households face lower access constraints to the land market. Macours et al. (2010) find that insecurity of property rights in the Dominican Republic reduces rental activity, and causes the tenancy market to match up along socio-economic lines.

While there are several potential motivations for conducting transfers that improve efficiency, we sidestep this issue by focusing only on the properties of the *outcome* of such transfers. In the penultimate section, we briefly discuss possible motivations. But the main results of the paper do not depend on the underlying mechanism behind the transfers.

We test if pairwise factor transactions are efficiency improving and if the pattern of exchange is shaped by pre-existing ties. We also test whether economic efficiency is achieved at the village level. Our response to the first question is positive: Factor transactions on average serve to improve allocative efficiency. Our response to the second question is mixed. On average over all transactions, we find that, if anything, factor transfers are more efficiency enhancing across ethnic groups and outside the extended family network. This result is likely due to historical patterns of village settlements by which large holdings of village land are in the hands of members of a specific ethnic or kin group, who then share it with relative newcomers. If we ignore large landowners, we find more efficiency-enhancing land transfers between kin and neighbors. Regarding the final question, we find no evidence that allocative efficiency is achieved at the village level. Taken together, the evidence is consistent with the idea that trade over the social network is not frictionless: Agents located between surplus and deficit households in the social network are unable to channel enough of excess factors to equalize land–labor ratios between surplus and deficit households.

This paper contributes to the literature in several ways. First, we introduce a novel way of testing how pre-existing networks affect the efficiency of exchange. Most evidence of network effects in markets does little more than document that exchange is more likely among agents with social ties. While this kind of evidence suggests that social networks have equity implications, it does not demonstrate that trade on networks leads to an inefficient outcome: Goods may have to travel between agents along social networks but still end up in the right hands. Providing evidence that trade flows follow pre-existing social ties is not, by itself, sufficient to conclude that allocative efficiency is not achieved. The same can be said of value chains: That consumers cannot purchase directly from producers or wholesalers does not imply that consumption goods do not end in the hands of consumers. Whether or not efficiency is achieved depends on the "bandwidth" and the architecture of the network or value chain: If enough goods can travel from surplus nodes to deficit nodes, then efficiency will be achieved. Second, this paper contributes to the literature on the importance of pre-existing social networks for offsetting the negative effects of market imperfections. We find some evidence to this effect. Third, this paper provides empirical evidence on the efficiency properties of social networks in The Gambia, a country that is fairly representative of West African rural institutions. If pre-existing social networks are important for the selection of trading partners, this has direct implications for the prospects of reducing poverty and inequality as well as economic growth (Holden et al., 2009). This issue is particularly important in West Africa where access to land is an increasing concern as land scarcity increases due to population growth (World Bank, 2005).

The paper is organized as follows. In section 3.2, we develop a simple model to test allocative efficiency in a market with one factor and subsequently in a market with multiple factors of production. Section 3.3 describes the tenure system in rural Gambia and the data used in the empirical analysis, while section 3.4 presents the results of the analysis. Section 3.5 discusses alternative interpretations of the empirical findings and provides a number of robustness checks. Section 3.6 concludes this paper.

3.2 Conceptual framework

We wish to test whether transfers of factors of production across households serve to increase allocative efficiency and, if so, whether the extent of efficiency gains is limited by the structure of social networks. We begin by deriving the efficient allocation of factors across farmers for different types of production functions, with and without transaction costs. Some configurations of production functions and transaction costs are rejected trivially by the data, so we do not consider them further. Other configurations make factor allocation predictions that cannot be tested by the data at hand. An important and empirically relevant subset of configurations makes testable predictions regarding the reallocation of factors between farmers. This is the set of configurations that we focus on. For these configurations, we derive a method for testing whether efficiency gains are shaped by the structure of social networks.

We consider a village economy with N farming households observed during a single cropping season. Each household starts the season with an endowment of factors of production, i.e., land and labor. We characterize the efficient allocation of factors for different production technologies. Initial factor endowments need not be allocated efficiently across households, so that there are possible efficiency gains from trade. The object of this section is to derive testable predictions regarding the transfers of factors across households that are necessary to achieve, or at least improve, efficiency in factor allocation.

3.2.1 One factor

We start by considering an economy with a single factor of production a. Each household starts with a_i units of the single factor. The total supply of a in the economy is denoted $A \equiv \sum_{i=1}^{N} a_i$. Each agent shares the same production function $q(x_i)$ where x_i denotes the amount of a used by agent i.

If the production function exhibits constant returns to scale (CRS) in x_i, any allocation of factors across households is efficient. Given this, we should not observe any factor transfers if, as is likely, there is any transaction cost in transferring factors across households. Since we do observe factor transfers in our data, this particular configuration is rejected trivially by the data and we do not consider it further. Similarly, increasing returns to scale implies the concentration of all factors into a single production unit, which is not what we observe either.

A more interesting case is if the production function exhibits decreasing returns to scale (DRS) in x_i. In this case, the efficient factor allocation is to set $x_i = \overline{a} = \frac{A}{N}$, i.e., to equalize the allocation of factors across all households. The reallocation of factor a across households can be organized in a variety of ways – e.g., through market transactions or gift exchange. Since our focus is purely on allocative efficiency, we focus on actual factor transfers, not on institutional mechanisms by which these transfers are achieved (e.g., market transactions or gift exchange). Conditional on allocative efficiency, these institutional mechanisms only affect the distribution of welfare among agents.

Let w_i be the net amount of a that is transferred out by household i. Allocative efficiency requires that households have equal amounts of land after transfers have taken place, i.e., $a_i - w_i = \overline{a} \Leftrightarrow w_i = a_i - \overline{a}$ for all i. A simple regression of the following form can be employed to investigate whether the average factor transfer is efficiency enhancing:

$$w_i = \alpha_0 + \alpha_1(a_i - \overline{a}) + u_i \tag{1}$$

Allocative efficiency requires $\alpha = 0$ and $\alpha_1 = 1$ since that means that $w_i = a_i - \bar{a}$. However, transfers can be said to be efficiency enhancing on average if $0 < \alpha_1 < 1$. This means that relatively land-rich households transfer land out and relatively land-poor households transfer land in. This approach is easily implementable but it does not allow for testing the effect of pre-existing networks on the efficiency properties of factor transfers. In order to do so, we move the analysis to the dyadic level.

Any pattern of factor transfers can be represented as a weighted directed graph or network in which agents are nodes and commodity flows are links. Let w_{ij} denote a transfer of factor a from agent i to agent j and let the gross transfer matrix $W \equiv [w_{ij}]$ denote the set of all transfers. Simple accounting implies that transfers must satisfy:

$$x_i = a_i + \sum_j (w_{ji} - w_{ij}) \text{ and}$$

$$\sum_i x_i = A$$

This yields the following feasibility condition:

$$\sum_j \sum_i (w_{ji} - w_{ij}) = 0 \tag{2}$$

which should always be satisfied in practice if the data are internally consistent.

Let W^* denote a transfer matrix that satisfies the feasibility condition and achieves efficiency, i.e., such that $x_i = A/N$. We call such matrices efficient. There is a very large number of efficient matrices W^* for a given endowment structure of factor a across agents: The specific factor transfers that are needed to achieve efficiency are undetermined. The reason is simple. Say a net factor transfer from i to j is needed for efficiency. This transfer can be achieved via a net flow directly from i to j – or through a series of transfers from i to k to m etc... until j.[2]

While it is not possible to determine which specific trades take place without introducing additional assumptions, it is possible to provide some characterization of *average* trade flows if they serve to improve allocative efficiency. The logic is simple and is easily illustrated with an example. Starting from an unequal factor distribution, allocative efficiency requires that those with above average endowments transfer out $a_i - \overline{a}$ and those with below average endowments transfer in $\overline{a} - a_i$, i.e., transfer out $-(\overline{a} - a_i)$.[3]

Although there are many gross out-transfers w_{ij} that can achieve this result, allocative efficiency requires that, in aggregate, each agent exactly transfers these net amounts. It follows that, for each i, the *sum* of out-transfers is $\sum_j w_{ij} = a_i - \overline{a}$. Similarly, the sum of in-transfers is $\sum_i w_{ij} = \overline{a} - a_j$. Thus if we observe all $N(N-1)$ possible gross transfers in the economy, we can estimate a regression of the form:

$$w_{ij} = \alpha_0 + \alpha_1(a_i - \overline{a}) + \alpha_2(\overline{a} - a_j) + u_{ij} \tag{3}$$

over all $N(N-1)$ dyads. If transfers achieve an allocative efficient outcome, we should find $\alpha_0 = 0$ and $\alpha_1 = \alpha_2 = \frac{1}{N}$.[4] Alternatively, we could estimate a model of the form:

$$w_{ij} = \alpha_0 + \alpha_1 \frac{a_i - \overline{a}}{N} + \alpha_2 \frac{a_j - \overline{a}}{N} + u_{ij} \tag{4}$$

and test whether $\alpha_1 = \alpha_2 = 1$.

[2]One could ask which transfer matrix W^* minimizes the number of separate transfers needed to achieve efficiency, but this does not guarantee uniqueness of w^* and it is unclear how these trades could be implemented in a decentralized manner, so we do not discuss this further.

[3]Let N_1 and N_2 denote the number of agents with an above and below average endowment, i.e., with $a_i - \overline{a} > 0$ and $a_j - \overline{a} < 0$, respectively. The $N - N_1 - N_2$ others have an allocatively efficient endowment such that $a_k = \overline{a}$. Allocative efficiency requires that N_1 agents collectively transfer out $\sum_{i \in N_1} a_i - \overline{a}$ and N_2 agents collectively transfer in $\sum_{j \in N_2} \overline{a} - a_j$.

[4]To verify, we have:

$$\begin{aligned}
\sum_i w_{ij} &= \alpha_1 \sum_i (a_i - \overline{a}) + \alpha_2 \sum_i (\overline{a} - a_j) \\
&= \alpha_i (\sum_i a_i - N\overline{a}) + \alpha_2 N(\overline{a} - a_j) \\
&= \alpha_2 N(\overline{a} - a_j)
\end{aligned}$$

which implies that $\alpha_2 = 1/N$. Similarly, we have:

$$\begin{aligned}
\sum_j w_{ij} &= \alpha_1 \sum_j (a_i - \overline{a}) + \alpha_2 \sum_j (\overline{a} - a_j) \\
&= \alpha_1 N(a_i - \overline{a}) \\
\alpha_1 &= 1/N
\end{aligned}$$

Since regressions (3) and (4) are dyadic regressions, standard errors must be corrected to allow for non-independence across observations. One major concern is the possible negative correlation across observations driven by the fact that if i transfers his excess factor to j, he cannot transfer it also to another agent k. Hence, u_{ij} is negatively correlated with u_{ik}. We deal with this problem by clustering standard errors at the village level, so that no assumptions about the correlation structure inside the village are made.

Now suppose that transfers can only take place between certain pairs of nodes – e.g., because social ties are needed to mitigate information asymmetries and enforcement problems. To fix ideas, let $S \equiv [s_{ij}]$ denote the set of maximum socially feasible flows such that:

$$w_{ij} \leq s_{ij} \text{ for all } ij \tag{5}$$

Can efficiency be achieved? This depends on whether there exists an efficient transfer matrix W^* that satisfies the set of constraints (5). If such a matrix exists, allocative efficiency can be achieved even though transfers are restricted to take place between existing social ties.

Restrictions imposed by the socially feasible flow matrix S may prevent allocative efficiency from being reached. S is unobservable. However, we expect pre-existing social ties between i and j to affect s_{ij}. In essence, we use the pre-existing social networks as proxy variables for S and test whether s_{ij}'s between socially connected households are larger than s_{ij}'s between unconnected households, and whether s_{ij}'s are sufficiently large to achieve efficiency.

One possibility is that the economy is segmented into distinct network components. Formally, this means that the set of nodes N can be partitioned into mutually exclusive components $\{N_1, ...N_p\}$ such that (1) there is a path through connected nodes that leads from any node in N_i to any other node in N_i; and (2) there is no path through connected nodes that leads from a node in N_i to a node in $N_{j\neq i}$.[5] If s_{ij}'s are higher between households that belong to the same component, it is possible that allocative efficiency is possible within each component, but not between components. By extension of the reasoning behind (3), allocative efficiency within components can be tested by estimating a regression model of the form:

$$\begin{aligned} w_{ij} = \alpha_0 + \alpha_1(a_i - \overline{a}) + \alpha_2(\overline{a} - a_j) \\ + \alpha_3(a_i - \overline{a}_k)D_{ij} + \alpha_4(\overline{a}_k - a_j)D_{ij} + u_{ij} \end{aligned} \tag{6}$$

where $\overline{a}_k \equiv \sum_{i \in N_k} a_i$ is the average endowment in component k, $D_{ij} = 1$ if i and j belong to the same component, and 0 otherwise. If components are of equal size n and allocative efficiency is achieved only within socially distinct components, then

[5] This arises when the columns and rows of matrix S can be simultaneously reordered to make it block diagonal.

$\alpha_1 = \alpha_2 = 0$ and $\alpha_3 = \alpha_4 = 1/n$. If components differ in size, we can emulate (4) and estimate a model of the form:

$$\begin{aligned} w_{ij} &= \alpha_0 + \alpha_1 \frac{a_i - \overline{a}}{N} + \alpha_2 \frac{\overline{a} - a_j}{N} \\ &+ \alpha_3 (a_i - \overline{a}_k) \widetilde{D}_{ij} + \alpha_4 (\overline{a}_k - a_j) \widetilde{D}_{ij} + u_{ij} \end{aligned} \tag{7}$$

where $\widetilde{D}_{ij} = 1/N_k$ if i and j belong to the same component k and 0 otherwise. N_k is the size of the component k. Allocative efficiency within components implies $\alpha_0 = 0$ and $\alpha_1 + \alpha_3 = \alpha_2 + \alpha_4 = 1$.

An extreme case arises when transfers are only possible across pairs of agents. In this case, the regression model simplifies to:

$$\begin{aligned} w_{ij} &= \alpha_0 + \alpha_1 \frac{a_i - \overline{a}}{N} + \alpha_2 \frac{\overline{a} - a_j}{N} \\ &+ \alpha_5 (a_i - a_j) L_{ij} + u_{ij} \end{aligned} \tag{8}$$

where $L_{ij} = 1$ if there is a direct link between i and j in S, and 0 otherwise.[6] If allocative efficiency is only achieved across linked pairs of agents, then $\alpha_1 = \alpha_2 = 0$ and $\alpha_5 = 1$.

So far we have discussed situations in which allocative efficiency is achieved either at the level of the economy, or within each component. It is also conceivable that factor transfers go in the direction of allocative efficiency but fail to achieve full efficiency. This would arise if there is "friction", i.e., if transfer paths in S do not have the capacity or "bandwidth" required to transfer all the necessary factors from surplus agents to deficit agents. In this case, some transfers occur but allocative efficiency is not achieved either in the economy as a whole, or within components. To test friction in the economy as a whole, we estimate model (4) and test whether $0 < \alpha_1 < 1$ or $0 < \alpha_2 < 1$.[7] To test friction within components, we estimate (7) and test whether $0 < \alpha_3 < 1$ and $0 < \alpha_4 < 1$.

We can also combine all three ideas into a single estimation model of the form:

$$\begin{aligned} w_{ij} &= \alpha_0 + \alpha_1 \frac{a_i - \overline{a}}{N} + \alpha_2 \frac{\overline{a} - a_j}{N} \\ &+ \alpha_3 (a_i - \overline{a}_k) \widetilde{D}_{ij} + \alpha_4 (\overline{a}_k - a_j) \widetilde{D}_{ij} \\ &+ \alpha_5 (a_i - a_j) L_{ij} + u_{ij} \end{aligned} \tag{9}$$

and apply the same logic to test for improvements of allocative efficiency along direct links (α_5), indirect links (α_3 and α_4), and without links (α_1 and α_2).

[6] Note that the social networks we consider are undirected such that $L_{ij} = L_{ji}$. If the social network is directed, $(a_i - a_j)L_{ji}$ could be included as an additional term in the regression.

[7] We note that while efficiency requires $(\alpha_1 = 1 \wedge \alpha_2 = 1)$, efficiency is enhanced as long as $(\alpha_1 > 0 \vee \alpha_2 > 0)$ holds. An example is the case where $0 < \alpha_1 < 1$ and $\alpha_2 = 0$. This means that while the average household who conducts out-transfers has a higher endowment of a than the village average, the receiving households are on average not different from the village average. However, even if it is random who receives such a transfer, the transfer will on average reduce the inequality of the distribution of a, thus increasing allocative efficiency.

The estimation procedure of this paper bears some resemblance with the gravity models of the international trade literature, which also investigate determinants of gross trade flows between dyads (for a review, see Bergstrand and Egger, 2011). A gravity model estimates to what extent the gross flow x_{st} of goods from country s to country t is determined by their gross domestic products (GDP) and the distance between them, d_{st}. The canonical version of the gravity model is of the form $\ln(x_{ij}) = \beta_1 \ln(GDP_s) + \beta_2 \ln(GDP_t) + \beta_3 \ln(d_{st})$, plus additional country-specific and dyad-specific regressors as needed. The literature typically finds $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_3 < 0$, meaning that larger countries trade more and that distance between countries reduces the amount of trade due to transport costs. As with physical gravity, large "masses" of GDP in close proximity attract each other, resulting in more trade in both directions. This is in contrast with the process studied in the current paper, where we expect opposites to attract: Households with different factor endowments are predicted to transact more. Furthermore, while the gravity model posits that trade falls in both directions with distance, our modeling framework predicts directed flows from factor-rich to factor-poor households.

The gravity model has been used to study the effect of trade blocks (e.g., Frankel et al., 1995; Baier and Bergstrand, 2007). This is done by adding a dummy that equals one if both countries belong to the same trade block. If countries within the same trade block trade more than what can be explained by the gravity equation parameters, i.e., if the coefficient of the dummy is positive and significant, it is interpreted as evidence that trade blocks lower the barriers to trade. The assumption is that trade takes place whenever it is profitable, and that more trade is profitable when barriers are lower; the benchmark case of gravity models is a world with no costs to trade where the share of spending in country s on goods from country t should be equal to the world share of spending on goods from country t. Trade blocks have the same formal structure as the components of the model of this paper. However, our setup is somewhat different since more land transfers between households do not necessarily improve upon allocative efficiency. Rather, it depends on who transfers land to whom. For this reason, component dummies D_{ij} enter our estimating equation not as separate dummies but as interaction terms with $a_i - \bar{a}_k$ and $\bar{a}_k - a_j$, as explained above.

3.2.2 Two factors

We now consider what happens if there are two factors of production, say, x and y. The production function is now of the form $q(x_i, y_i)$ where y_i is a second factor of production. Each agent is endowed with a factor vector $\{a_i, l_i\}$ while the agent's factor usage is denoted $\{x_i, y_i\}$. The total supply of factors in the economy is denoted $A \equiv \sum_{i=1}^{N} a_i$ and $L \equiv \sum_{i=1}^{N} l_i$.

Allocative efficiency again depends on whether production is characterized by decreasing, constant, or increasing returns to scale. If the function $q(x_i, y_i)$ has increasing returns to scale in x and y, the efficient allocation of factors is to have a single production unit concentrating all available factors – e.g., a plantation. Since this is not at all what we observe in our data, we do consider this case further. With constant returns to scale, there is no specific efficient scale of production. But efficiency requires that marginal rates of substitution between factors be equalized across households:

$$\frac{\frac{\partial q(x_i,y_i)}{\partial x_i}}{\frac{\partial q(x_i,y_i)}{\partial y_i}} = \frac{\frac{\partial q(x_j,y_j)}{\partial x_j}}{\frac{\partial q(x_j,y_j)}{\partial y_j}} \text{ for all } i \text{ and } j \tag{10}$$

If we further assume that the production function is homothetic in x and y, an efficient factor allocation must satisfy:

$$\frac{x_i}{y_i} = \frac{x_j}{y_j} = \bar{r} \text{ for all } i \text{ and } j \tag{11}$$

where $\bar{r} \equiv \frac{A}{L}$ is the average factor ratio in the economy.[8] Hence, for transfers of factors across households to improve production efficiency, they must flow so as to equalize factor ratios across all households in a given village.[9]

This observation provides a partial characterization of factor transfers that can be used to test the role of social networks. Agents that have a factor ratio endowment $r_i \equiv \frac{a_i}{l_i} > \bar{r}$ should on average transfer out part of their a_i endowment such that their factor usage ratio x_i/y_i is equal to $\bar{r}$. This can be tested by estimating a regression model of the form:

$$w_{ij} = \alpha_0 + \alpha_1(r_i - \bar{r}) + \alpha_2(\bar{r} - r_j) + u_{ij} \tag{12}$$

where w_{ij} as before denotes out-transfers of a. If factor transfers improve allocative efficiency, we should find that $\alpha_1 > 0$ and $\alpha_2 > 0$: On average, factor a flows from agents relatively rich in that factor to agents poor in that factor.[10]

We are now ready to introduce limits to transfers imposed by social matrix S. The testing logic remains the same as before. If transfers are not feasible across components, we expect convergence in factor ratios within components only. The estimating equation

[8]The first part of the equality follows from the fact that homothetic production functions have linear expansion paths. The second part follows from clearing conditions of the factor markets. Homotheticity is likely to hold in this study, at least approximately, given the limited range of farm sizes observed in our data.

[9]Equalization of factor ratios does not require transfers of both factors across households; trade in one factor is sufficient. It follows that, in the CRS case, efficiency can be achieved even with one non-traded factor.

[10]In an efficient factor allocation, we should have:

$$\frac{r_i + \sum_j [w_{ij} - w_{ji}]}{l_i} = \bar{r} \text{ and}$$
$$\frac{r_j + \sum_i [w_{ij} - w_{ji}]}{l_j} = \bar{r}$$

becomes:

$$\begin{aligned} w_{ij} &= \alpha_0 + \alpha_1(r_i - \bar{r}) + \alpha_2(\bar{r} - r_j) \\ &+ \alpha_3(r_i - \bar{r}_k)D_{ij} + \alpha_4(\bar{r}_k - r_j)D_{ij} + u_{ij} \end{aligned} \tag{13}$$

where $\bar{r}_k \equiv \frac{\sum_{i \in N_k} a_i}{\sum_{i \in N_k} l_i}$. The testing strategy is similar as before: If factors only flow within components of the social network, we should find $\alpha_1 = \alpha_2 = 0$ and $\alpha_3 > 0$ and $\alpha_4 > 0$. Other regression models presented in the previous sub-section also extend to this case, and need not be discussed further.

If returns to scale are decreasing, allocative efficiency requires that:

$$\frac{\partial q(x_i, y_i)}{\partial x_i} = \frac{\partial q(x_j, y_j)}{\partial x_j} \text{ for all } i \text{ and } j \tag{14}$$

$$\frac{\partial q(x_i, y_i)}{\partial y_i} = \frac{\partial q(x_j, y_j)}{\partial y_j} \text{ for all } i \text{ and } j \tag{15}$$

which implies that each household should use the same quantity of x and y. The above also implies that:

$$\frac{\frac{\partial q(x_i, y_i)}{\partial x_i}}{\frac{\partial q(x_i, y_i)}{\partial y_i}} = \frac{\frac{\partial q(x_j, y_j)}{\partial x_j}}{\frac{\partial q(x_j, y_j)}{\partial y_j}} \text{ for all } i \text{ and } j \tag{16}$$

and hence efficiency requires an equalization of factor ratios across farms, as in the CRS case. We thus have, for the DRS case, two available testing strategies: To apply model (6) to an individual factor; or to apply model (12) to factor ratios. Both strategies should yield similar results in the DRS case. In the CRS case, only model (12) can be used; model (6) is not identified since farm size is undetermined.

3.2.3 Unobserved heterogeneity

Identifying efficiency-enhancing factor flows becomes more difficult if agents differ in their unobserved productivity – or, equivalently, in an unobserved fixed factor (e.g., management capacity). In an efficient allocation, agents that are more productive should attract more factors of production for the marginal productivity of factors to be equalized across agents. If we do not know who is more productive, we cannot test whether the *absolute* allocation of factors is efficient.

To illustrate the difficulty, let the production of agent i be written $\theta_i q(x_i)$ where θ_i is the total factor productivity of agent i and x_i is a single factor of production.[11] With

[11] Unobserved individual productivity θ_i can be thought of as a total factor productivity shifter. If we posit a Cobb-Douglas production function, θ_i can also represent a fixed unobserved factor (e.g., management capacity) or factor-biased productivity differentials: In a Cobb-Douglas production function, both factor out as a multiplicative constant.

decreasing returns to scale,[12] allocative efficiency requires that:

$$\theta_i \frac{\partial q(x_i)}{\partial x_i} = \theta_j \frac{\partial q(x_j)}{\partial x_j} \text{ for all } i \text{ and } j \tag{17}$$

The above implies that $x_i > x_j$ if $\theta_i > \theta_j$: More productive agents use more x. Since we do not observe θ_i, identification fails in all the regression models discussed so far: Without knowing who is most productive, we cannot test whether the allocation of the single factor x_i across farmers is efficient. By extension, the same conclusion applies in the CRS case if there are two factors of production but only one is traded.

Some identification can be recovered if we observe not one but two traded factors, x and y. The production function is now of the form $\theta_i q(x_i, y_i)$. We continue to assume DRS in x and y.[13] Allocative efficiency now requires that:

$$\theta_i \frac{\partial q(x_i, y_i)}{\partial x_i} = \theta_j \frac{\partial q(x_j, y_j)}{\partial x_j} \text{ for all } i \text{ and } j \tag{18}$$

$$\theta_i \frac{\partial q(x_i, y_i)}{\partial y_i} = \theta_j \frac{\partial q(x_j, y_j)}{\partial y_j} \text{ for all } i \text{ and } j \tag{19}$$

which yields the same expression as (10). If we continue to assume that the production function is homothetic in x and y, an efficient factor allocation still must satisfy (11): factors should flow so as to equilibrate factor ratios among agents.

It follows that, unless factor endowments are perfectly correlated with productivity differentials, *on average* agents with $r_i \equiv \frac{a_i}{l_i} > \bar{r}$ should transfer out some of their excess factors to others, and vice versa. In the earlier section, this property held exactly for each farmer; here it only holds on average, and requires that factor endowments and productivity not be strongly positively correlated. With these assumptions, it is possible to apply the same testing strategy as that outlined above. The strategy requires a strong maintained assumption about the distribution of factors and productivity. If θ_i is strongly correlated with r_i, then it is possible that the efficient allocation is to give even more land to land-rich farmers. We rule this out by assumption. However, this assumption is reasonable in the context of our data, where land is mostly inherited from the household's lineage and manpower depends on where the household is in its life cycle. There is no compelling reason to expect these exogenous factors to be strongly positively correlated with θ_i, even if some correlation may be present. Moreover, given the low technological level of the study population, it is unlikely that productivity differentials differ massively among them – or at least not as much as the very large differentials in land endowment.

[12] If returns to scale are constant in x_i, it is efficient to allocate all the supply of x to the most productive producer. Since we do not observe such cases in our data – far from it – we do not consider this case any further.

[13] If $q(x, y)$ is CRS, the optimal allocation is for all factors to go to the single most productive farmer. Again, this is not what we observe.

3.2.4 Summary

To summarize, we have devised a method for testing the role of social network structure in improving the efficiency of factor allocation in farming villages. This method does not work if production exhibits increasing returns to scale, or if constant returns to scale are combined with unobserved heterogeneity in productivity: In these cases, all factors should go to a single producer. Since this is not what we observe in the data, we can probably dismiss this possibility for our study. The method also fails when unobserved heterogeneity is combined with decreasing returns to scale and a single traded factor. In this case, we are unable to empirically characterize the efficient allocation of factors, and hence we cannot test whether factor transfers serve to improve efficiency.

The good news is that there are several empirically relevant cases in which we can characterize the efficient allocation of factors, either in absolute terms or in terms of factor ratios. In these cases, this characterization can be used as basis for a test of the role of social network structure in permitting or curtailing efficiency gains in factor allocation. Table A.1 lists all the various cases and their predictions. The testable cases include decreasing returns to scale with one factor that is traded and no unobserved heterogeneity; and decreasing returns to scale with two traded factors, with or without unobserved heterogeneity. In the first case, we can characterize efficiency in absolute terms and use model (6) as starting point for our testing strategy; in the second case, we can characterize efficiency in terms of factor ratios and construct a testing strategy based on regression model (12). The latter method also applies to the CRS case with two factors but no unobserved heterogeneity, even if one factor is de facto not traded.

3.3 Data

In order to implement the testing strategy outlined in the previous section, we rely on a unique dataset from The Gambia. The *Gambia Networks Data 2009* was collected in six out of eight Local Government Areas between February and May 2009. A sample of 60 villages was randomly selected among villages with between 300 and 1000 inhabitants in the latest (2003) census. Restricting the sample to small villages was motivated by the desire to obtain complete dyadic data for each entire village. Collecting such data for much larger villages would be impractical. The sample is representative of the 20 percent smallest villages of The Gambia (Arcand and Jaimovich, 2014). A household survey was administered to all households in the villages, using a structured group interview approach. Information was collected on the land a_i and labor l_i endowments of each household. The dataset also contains information on all pre-existing social links as well as on all land and labor transactions within each village over an entire agricultural year. For more detailed information on the data collection strategy, see Jaimovich (2015).

Six villages are dropped from the analysis due to substantial amounts of missing household-level information. We also drop three semi-urban villages because our focus is on factor transactions in agriculture.[14] For the same reason, we restrict the sample to households for whom the main activity of the household head is farming.[15] The sample we use for the empirical analysis consists of 1,625 households across 51 villages, corresponding to 57,060 within-village dyads.

Rural villages in The Gambia are organized around compounds. A compound is a group of buildings used for housing, storage and other purposes, which are located in relative proximity to each other, typically around a central courtyard. Most often, a compound has a single household head that makes decisions regarding production and other daily activities. However, sometimes a single physical compound has more than a single decision maker. In this case, independent production units (*dabadas*) can exist within the compound and independent consumption and cooking units (*sinkiros)* can exist within a single compound or within a single *dabada* (von Braun and Webb, 1989; Webb, 1989). Since we are interested in production decisions, the *dabada* should be used as the unit of observation. If several *dabadas* reside within the same compound, the networks between these are present in the dataset. Fourteen percent of household heads in our sample are not heads of the compound in which they live.

Descriptive statistics for our variables of interest are presented in Table 1. They paint a picture typical of rural Gambia. The largest households in the sample have more than 50 members, but they only account for 0.01 percent of the sample. These large households arise as the result of polygamous marriages. Fifty percent of household heads in the sample have more than one wife. The households are predominately headed by poorly educated men: Only 9 percent have any formal education although an additional 42 percent have received Koranic education, which provides basic literacy in Arabic. Household heads are quite old with an average age of 54 years.[16] The share of income which comes from agricultural outputs is 16 percent. This is low, considering that The Gambia is a big exporter of groundnuts, cashews and other cash crops. This may partly be caused by the sampling of the villages, which rules out larger villages. Another contributing factor is that, at the household level, cash crops are occasionally

[14]In terms of network activity, only 2 percent of the households in the semi-urban villages participate in the land market, while 10 percent participate in the labor market. The main reason for the absence of land sharing is probably the very small landholdings in these areas (0.24 hectares per household compared to 10.28 in the rural villages), and the much higher availability of employment opportunities outside the village.

[15]The village level variables such as $\bar{a}$ and $\bar{r}$ are calculated as averages for farmers in each village only.

[16]The high rates of polygamy, the very low education rates and the high age of household heads are unusual compared to other parts of Africa, but the information lines up well with what is known about The Gambia. According to the nationally representative Multiple Indicator Cluster Survey (MICS) from 2010, 41 percent of women aged 14-49 are in a polygamous marriage. More than two-thirds of the women aged 14-49 are currently married (The Gambia Bureau of Statistics (GBOS), 2011). MICS also reports that 68.4 percent of household heads have no education. This share is expectedly higher in our dataset due to the rural nature of the sample. The report of the 2003 Gambian census which constituted the sampling frame for the survey used in the present paper reports the mean age of household heads to be 46 (The Gambia Bureau of Statistics (GBOS), 2008).

Table 1: Household-level descriptive statistics

	Mean	Std. Dev.
Household size	14.100	15.002
Age of head	54.319	16.229
Female headed household	0.057	0.232
Illiterate	0.489	0.500
Formal education	0.087	0.282
Ethnicity: Mandinka	0.541	0.498
Ethnicity: Fula	0.186	0.390
Ethnicity: Wollof	0.101	0.301
Ethnicity: Jola	0.065	0.247
Ethnicity: Sererr	0.056	0.230
Number of working adults	5.371	10.39
Have primary usage rights to land	0.819	0.385
Land owned with official rights (hectares)	10.502	24.478
land–labor ratio (hectares land per active worker)	2.762	7.386
Income per capita (1,000 GMD/PPP)	0.282	0.285
Agricultural sales (share of income)	0.161	0.271
Receive remittances	0.455	0.498
Relative wealth level (=1 (low), 2, 3 or 4 (high))	1.755	0.806
Participates in land transfers	0.452	0.498
Land transfer (hectares)	1.040	2.827
Land transfer (% of donors' land endowment)[1]	0.376	0.307
Participates in labor transfers (within the village)	0.544	0.498
Labor transfer (days)	5.263	15.382
Household has family links in the village	0.958	0.202
Household head has family links in the village	0.850	0.357
Wife of household head has family links in the village	0.508	0.500
Household have marriage links in the village	0.642	0.480
Observations	1,625	

1: Conditional on transferring out land. Excludes observations where the donor is also a receiver
Source: Authors' calculations.

bartered and not sold. Finally, the structured group interview approach may lead to under-reporting of agricultural income.

The descriptive statistics also show that more than 45 percent of households receive remittances. This high rate is evidence of a substantial exodus of manpower from these villages. This is partly explained by the high rate of urbanization Gambia has experienced in the last decades. 65 percent of the households receiving remittances are receiving them from people within The Gambia. However, The Gambia also has a high rate of international remittances. According to the World Development Indicators, Gambia ranked second-highest in international remittances as a share of GDP in Sub-Saharan Africa in 2010. Out-migration of persons from labor-rich households can reduce discrepancies in factor ratios. This effect does not invalidate the analysis of the current paper, but it does make the results conditional on the manpower available in the village at the beginning of the cropping season.

The sample is representative of the ethnic diversity in The Gambia. The largest ethnic group (Mandinka) accounts for 54 percent of households in the sample. Four other ethnic groups each account for five percent or more of the households.[17] The average monetary income per capita is 2,750 Gambian Dalasis a year, which is equivalent to 282 USD a year in PPP terms.[18]

3.3.1 Endowments and transfers of land and labor

In The Gambia, all land is nominally owned by the state, but de facto usage rights to land are determined by a complex indigenous land tenure system (Freudenberger, 2000). Two principal types of usage rights exist, which are referred to as *primary* and *secondary* rights. A household with primary rights over a plot of land can decide which crops to grow and whether to lend all or some of the land to another farmer – in which case this farmer gains a secondary usage right on the land. A household with secondary rights over a plot has full control over the agricultural management of plot while their secondary rights persist.

Landless or land-poor households obtain secondary usage rights to land in two ways (Freudenberger, 2000). First, households who possess surplus land have a moral obligation to transfer usage rights over some land to those in need, and the village chief can allocate land to households in need of land. Second, landless or land-poor households can access land through market-based transactions. The questionnaire was designed to capture both monetary and non-monetary transfers of secondary usage rights. In practice, most transfers of secondary usage are non-monetary, although symbolic payments of kola nuts, labor services or cash do sometimes take place (Eastman, 1990; Freudenberger, 2000; Jaimovich, 2011). Field observations from the collection of the *Gambia Networks data 2009* support this interpretation.[19] Sharecropping is not a common practice in Gambia. In fact, Dey (1982) describes how donor-supported sharecropping schemes were unsuccessful and quickly abandoned. Further supporting evidence of this comes from the dataset of the current paper where only 10 percent of all land transfers were reciprocated in the labor market.

Secondary usage rights are temporary in nature. Plots are most commonly borrowed on an annual or seasonal basis (Chavas et al., 2005). This means that households with

[17]Using the Herfindahl index to measure ethnic diversity, we find that ethnic fragmentation inside villages ranges from 0 (completely homogeneous) to 0.84, with a mean of 0.28.

[18]Using Penn World Tables 7.1 PPP-adjusted exchange rate for 2009 (Heston et al., 2012). Note that consumption and bartering of own production is not included in this figure. The Gambia Integrated Household Survey of 2010 found that mean consumption of own production and gifts in The Gambia in 2010 amounted to 6,283 dalasis per household per year, or 776 dalasis per person, using the mean household size of 8.1 from that survey (GBS, 2011).

[19]We thank Dany Jaimovich, who participated in the data collection, for this observation.

few or no primary usage rights to land must secure a transfer of secondary usage rights every year. A cross-sectional survey such as the one we employ will therefore capture almost all transfers of secondary usage rights.

According to Table 1, 82 percent of the household have *primary* usage rights to land. The distribution of primary usage rights is highly unequal, however, with some households having less than 0.5 hectares and others more than 100 hectares. The average land holding among rural farmers is 10.5 hectares per household. The unequal land endowment distribution is explained by the fact that primary usage rights were originally obtained by the founders of the village that first cleared the land. Since primary rights are in most cases inherited, primary usage rights are concentrated in the hands of descendants of the village founders. The unequal ownership structure of land is also reflected in the social structure: Households of the founder lineages have intermarried for generations and tend to belong to the same ethnic group. Therefore, these households are more likely to be part of a single ethnic and family-related component. There is therefore more scope for increasing allocative efficiency by transferring secondary land usage rights from the descendants of the village founders to relative newcomers. Indeed, in many villages one ethnic group has much larger land endowments than others. Similarly, the land endowment of households outside the largest kinship group is less than a third of that of households that belong to this component. Due to the moral obligation for households who possess surplus land to transfer usage rights to land-poor households, we would expect to see many efficiency-enhancing transfers of land from founder lineages to other ethnic or kinship components.

In the empirical analysis, we use the primary usage rights of a household as measure of their land endowment. The data on land endowments is likely to contain measurement error because many households had difficulties reporting the size of their land endowment into standard measurement units. This will undoubtedly generate some attenuation bias in our results. We return to this issue in the discussion section.

Labor endowments are determined by household size and composition. Households have an average of five adult working members. This corresponds to an average land–labor ratio of 2.8 hectares per working adult among the farming households in the sample. While this indicates that land is relatively abundant on average, land endowments are inequitably distributed: 18 percent of households have no land and the median land–labor ratio of households is 1 hectare per working adult.

Table 2 provides descriptive statistics on participation in land and labor transfers, and separately by initial land–labor ratios. Around 45 percent of households engage in land transfers, and households that export land typically transfer land to more than one household. Land transfers account for 38 percent of the initial land endowments of the land-exporting households. At the receiving end, 36 percent of landless households in the sample receive land from other households.[20]

A large proportion of households participate in labor transfers (54 percent). However, the number of labor days transferred is small: The households on average devote five days annually to working on other farms in the village. Compared to the land transfers where each household on average sends more than one hectare of land, or around 10 percent of the average land endowment, this number is small. The limited transfers of labor suggest that transaction costs are higher in labor transfers than in land transfers. This is not surprising for this part of the developing world: Labor supervision is known to be problematic in African agriculture, and labor markets in subsistence agriculture are very thin throughout most of West Africa (Otsuka, 2007; Holden et al., 2009). Since land transfers by definition take place before transfers of agricultural labor, it is also conceivable that land transfers equalize land–labor ratios across households ex ante, thereby reducing the need for ex post adjustment through labor transactions – in which case the role of labor transfers is primarily to deal with shocks in household manpower availability (e.g., illness).

The relatively small geographical size of the villages means that soil quality is fairly uniform within each village. This suggests that information asymmetries and moral hazard are probably less problematic in land transactions than in labor transactions. Furthermore, we note that the dataset does not include seasonal migrant workers typically referred to as strange farmers (Swindell, 1987). Strange farmers used to be a relatively common phenomenon in rural Gambia, but were not commonly found during the collection of the *Gambia Networks Data 2009*.[21]

Table 2 indicates that households with higher initial land–labor ratios are more likely to transfer out land and less likely to transfer in land. This descriptive evidence is consistent with the idea that land transfers serve to improve efficiency. The same pattern is not evident for labor: Land abundant households are not more likely to transfer in or transfer out labor. This constitutes prima facie evidence that labor transfers do not serve the same efficiency-enhancing function. Thus, there are both theoretical and empirical reasons for why we focus our analysis primarily on land transactions.

The dataset does not contain information on households that reside outside the village but it does contain information about land transfers between households in the village and households outside the village. When calculating $\bar{a}$, transfers of land outside the network must be netted out. To see why, consider the simple model in (1). Allocative efficiency is achieved if land use is equal to the average amount of land available to surveyed farming households. Therefore, $\bar{a}$ is calculated as the net land endowment in the village, i.e., including land transferred in from outside the village and from

[20] At first glance, the 2 percent of landless households who transfer out land seem like a paradox. However, these households also transfer in land and thus end up with a nonnegative amount of land.

[21] We thank Dany Jaimovich for clarifying this to us.

Table 2: Land and labor market participation rates by initial land–labor ratios

	All	Split by hectares per adult				
		No land	]0-0.6]	]0.6-1.5]	]1.5-3.0]	>3.0
Land transfers:						
Participates	45.2	36.1	42.4	41.1	53.8	56.5
Land sender	21.0	2.0	14.9	19.0	30.1	40.5
Land receiver	28.9	35.7	33.3	26.2	30.5	21.4
Labor transfers:						
Participates	54.4	42.5	49.8	56.6	64.3	57.5
Labor sender	36.9	30.6	34.1	39.5	43.6	34.7
Labor receiver	32.3	23.8	27.5	30.6	41.0	39.8
Observations	1,625	294	255	516	266	294

Source: Authors' calculations.

non-farming households as well as excluding land transferred to households outside the village and to non-farming households.

3.3.2 Social networks

With the available data, we construct three types of social networks, each with its own unique characteristics: ethnicity, kinship, and geographical proximity. The ethnicity network is defined by positing a link between all pairs of households whose head belongs to the same ethnic group. This implies that, when we use ethnicity to define social networks, $L_{ij} = D_{ij}$ always holds. This means that, in this case, α_5 is not separately identified from α_3 and α_4 in model (9).[22] In this case, we simply investigate whether there are more efficiency-enhancing land transfers between members of the same ethnic components than across the village as a whole.

The kinship network is a symmetric network in which a link between two households exists if the households are kin related, either through the household head, the wife(s) of the head, or through marriage. Unlike in other African contexts (e.g., Barr et al., 2012), kinship networks in our sample are relatively dense. This is driven by the fact that descendants of the village founders and original settlers often reside in the village. Around 85 percent of household heads have relatives living in the village. Similarly, in 51 percent of households, the wife (or wives) has relatives living in the village, and 64 percent of households have marriage ties to other households in the village. Around 96 percent of households have at least one of these three kinds of links. This means that the components in the kinship network are often large: In the majority of villages, most

[22] Alternatively, one could allow for multiple ethnicities for each household. This would allow for separate identification of α_3, α_4 and α_5. However, combining the information on marriage linkages and ethnicity, we find that only nine percent of marriage links are between households where the head belongs to different ethnic groups. This amounts to an average of 2.2 marriage links per village between households whose heads have different ethnicities. Due to the low likelihood of inter-ethnic marriages, we do not pursue this line of investigation further.

(though not all) households are connected, directly or indirectly, to each other through consanguinity and marriage – i.e., they belong to the same network component. In this case, we investigate whether there are more efficiency-enhancing land transfers between kinship components and between directly kin-related households than across the village as a whole.

The geographical network is measured by a dummy variable taking the value 1 if a pair of households live in neighboring compounds, and 0 otherwise. In the geographical network, there is no natural partition of households into components, given that households are living relatively close to each other. Coefficients α_3 and α_4 are thus not estimated when we use geographical proximity as a measure of social networks. Rather, we exclusively consider the direct ties between pairs of households. The data on geographical proximity were only collected in a sub-sample of 19 villages, containing 624 households. In the reduced sample with geographical proximity, almost 10 percent of households are compound neighbors.

Kinship, ethnicity, and geographical distance are the dimensions of social proximity for which we have data. They offer the advantage of being immutable or pre-determined and are thus reasonably exogenous to factor transfers taking place within an agricultural season. We nonetheless expect that factor transfers between two households are often embedded in long-term relationships of favor exchange that we do not observe. From the available literature, it is likely that these favor exchange relationships are shaped, as least partially, by kinship, ethnicity, and geographical distance. If gift exchange transcends pre-determined social categories such that they do not matter, this will show up in our analysis as zero α_3, α_4 and α_5 coefficients on network regressors. Even in this case, gift exchange may still not achieve efficiency, e.g., if people choose their network links in order to maximize gains from factor exchange, but can only sustain a number of links that is insufficient to ensure equalization of factor ratios in all circumstances. This situation will manifest itself in the form of α_1 and α_2 being less than their efficiency-maximizing value (e.g., 1 or $\frac{1}{N}$, depending on the estimated model). These observations form the basis for our testing strategy.

3.4 Empirical results

This section presents the empirical results. We have noted in the previous section that although the frequency of labor transfers is higher than that of land transfers in our data, land transfers last for the entire season and are much larger in terms of the proportion of traded factors. A fair approximation of our data is that land transfers serve to achieve a desired farm size *ex ante*, while labor transfers are used to deal with temporary surpluses and deficits in labor *ex post* – probably in response to unforeseen shocks in the production process (Fafchamps, 1993). For these reasons, we primarily focus on land transfers.

3.4.1 One factor models

We first estimate the household-level model (1). Results are presented in Table 3. We find that $0 < \alpha_1 < 1$. This means that land transfers do improve efficiency, i.e., land on average flows from land-rich households towards land-poor households. However, efficiency is not achieved, i.e., $\alpha_1 \neq 1$. This result also holds when including village fixed effects as well as a battery of control variables, specifically household size; age of household head; indicator variables for literacy of household head; for whether the household head has any formal education; for whether the household head is female; for whether household receives remittances; and a relative wealth-level indicator which takes a value between one and four.[23] While the household-level regressions provide a first indication that land flows are on average efficiency-enhancing direction, this approach does not enable us to test the importance of pre-existing networks. We therefore proceed to the dyadic regressions.

Table 3: Household level model results

		(1)	(2)
Regressor	Coefficient		
$a_i - \overline{a}$	α_1	0.022**	0.021**
		(0.009)	(0.009)
Constant	α_0	0.005	-0.500
		(0.003)	(0.366)
Observations		1,625	1,625
Village fixed effects		NO	YES
Control variables		NO	YES

Note: The dependent variable is the net out transfer of land measured in hectares. Standard errors are clustered at the village level. *** p<0.01, ** p0<.05, * p<0.1.
Source: Authors' calculations.

We proceed by estimating the various models outlined in section 3.2.1. We start by reporting estimates of our simplest model (4) in which regressors are normalized, reproduced here:

$$w_{ij} = \alpha_0 + \alpha_1 \frac{a_i - \overline{a}}{N} + \alpha_2 \frac{a_j - \overline{a}}{N} + u_{ij}$$

The average endowment $\overline{a}$ is calculated as the village endowment of land net of the external transfers of land with non-village members. Results are presented in the first column of Table 4. Standard errors are clustered at the village level throughout our analysis.

We find that land transfers flow in a direction consistent with efficiency: Households that are relatively land rich tend to transfer land out, while households that are rela-

[23] We note that control variables are not required for the results to hold if the interest is purely in understanding whether efficiency-enhancing land transfers coincide with social networks. However, it is still of interest to see if other factors that potentially correlate with land endowments explain the results.

Table 4: One factor model: Do land transfers equalize farm size?

Regressor	Coefficient	(1)	(2) Ethnicity	(3) Kinship	(4) Kinship	(5) Kinship	(6) Neighbor
$(a_i - \overline{a})/N$	α_1	0.020**	0.029***	0.038***	0.019**	0.038***	0.034**
		(0.008)	(0.009)	(0.012)	(0.008)	(0.012)	(0.013)
$(\overline{a} - a_j)/N$	α_2	0.004	0.003	0.010**	0.002	0.009**	-0.010
		(0.003)	(0.004)	(0.004)	(0.002)	(0.004)	(0.007)
$(a_i - \overline{a}_k)\tilde{D}_{ij}$	α_3		-0.009**	-0.018***		-0.020***	
			(0.004)	(0.007)		(0.007)	
$(\overline{a}_k - a_j)\tilde{D}_{ij}$	α_4		0.001	-0.006**		-0.007**	
			(0.002)	(0.003)		(0.003)	
$(a_i - a_j)L_{ij}$	α_5				0.000	0.000**	0.001***
					(0.000)	(0.000)	(0.000)
Constant	α_0	0.019***	0.019***	0.019***	0.019***	0.019***	0.014***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
N		57,060	57,060	57,060	57,060	57,060	23,917

Note: Note: The dependent variable is the net out transfer of land measured in hectares. *** p<0.01, ** p0<.05, * p<0.1.
Source: Authors' calculations.

tively land-poor tend to import land from other households. However, we reject the hypothesis that land transfers result in all households having the same cultivated land area: The test ($\alpha_1 = \alpha_2 = 1$) is strongly rejected ($p < 0.0001$). The statistically positive coefficient α_1 imply that land abundant farmers transfer out more land, but not enough to equalize land across households within each village.

Next, we examine whether there are more equalizing land transfers between members of the same ethnic or kin group. Estimation results for the normalized model (7) are presented in the columns 2 and 3 of Table 4. Transfers that take place inside ethnically defined components are less efficiency enhancing than transfers across different social components ($\alpha_3 < 0$ and $\alpha_4 < 0$). This is most likely explained by the fact that a few households in a single ethnic and family-related component hold primary rights over a large fraction of the land in each village as described in the preceding section. In this setting, social distance does not appear to be an obstacle to land transfers. This is consistent with the result of Arcand and Jaimovich (2014). Also using the *Gambia Networks Data 2009*, the authors find that ethnic divisions do not reduce land transfers.

Next, we introduce direct kinship links between households L_{ij} and estimate the normalized model (8). Results are reported in column 4 of Table 4. The coefficient α_5 is positive but not statistically significant, suggesting that partial allocative efficiency is not improved between directly linked households. In column 5, we test for partial allocative efficiency along all three dimensions of kinship ties: components (α_3 and α_4), direct links (α_5) and without links (α_1 and α_2). By including the component variables $(a_i - \overline{a}_k)\tilde{D}_{ij}$ and $(\overline{a}_k - a_j)\tilde{D}_{ij}$ we obtain a positive coefficient estimate for α_5, suggesting that land flows more easily between closely related households conditional on the negative component effect. Other coefficient estimates remain similar to those reported in column 3.

To investigate whether the negative result on ethnic and kinship components is driven by transfers made by the largest landowners, we conduct two robustness checks. First, $(a_i - \bar{a}_k)/N_k$ and $(\bar{a}_k - a_j)/N_k$ are replaced with dummies equal to one when $(a_i - \bar{a}_k)/N_k$ or $(\bar{a}_k - a_j)/N_k$ are greater than zero, respectively. Second, we winsorize outliers by replacing values of $(a_i - \bar{a}_k)/N_k$ and $(\bar{a}_k - a_j)/N_k$ below the first percentile and above the 99th percentile with the value of the first and the 99th percentile, respectively. Results are shown in Table 5. The negative estimates of α_3 and α_4 appear to be driven by such outliers. When we replace the original regressors $(a_i - \bar{a}_k)/N_k$ and $(\bar{a}_k - a_j)/N_k$ with dummy variables in columns 1 and 3, the coefficient α_3 becomes positive and significant for both the ethnicity and kinship components. This suggests that a larger number of transfers within components are improving efficiency than between components, although some transfers from very land-rich households mean that this effect does not hold when magnitudes are considered. Supporting this, we also find a positive and significant α_3 when outliers are winsorized.

In the last column of Table 4 we consider geographical proximity. As explained earlier, the data does not contain information on geographical components so the focus here is solely on compound neighbors. Transfers between compound neighbors are found improve efficiency as expected: $\alpha_5 > 0$.

To conclude, in the simple model with only one factor, land transfers improve efficiency. Not all pre-existing social networks appear to improve efficiency further: We only

Table 5: Robustness of one factor results: outliers

Regressor	Coefficient	(1) Ethnicity	(2) Ethnicity	(3) Kinship	(4) Kinship
$(a_i - \bar{a})/N$	α_1	0.015**	0.009	0.014**	0.007
		(0.007)	(0.008)	(0.007)	(0.009)
$(\bar{a} - a_j)/N$	α_2	0.005*	-0.002	0.004	0.001
		(0.003)	(0.003)	(0.003)	(0.003)
$(a_i - \bar{a}_k)\tilde{D}_{ij}$	α_3	0.021***	0.021*	0.021***	0.024*
		(0.006)	(0.011)	(0.006)	(0.013)
$(\bar{a}_k - a_j)\tilde{D}_{ij}$	α_4	-0.004	0.010**	-0.001	0.006
		(0.002)	(0.005)	(0.003)	(0.005)
Constant	α_0	0.017***	0.020***	0.014***	0.020***
		(0.002)	(0.003)	(0.003)	(0.003)
Outlier method		Dummies	Winsorize	Dummies	Winsorize
Observations		57,060	57,060	57,060	57,060

Note: The dependent variable is the net out transfer of land measured in hectares. *** p<0.01, ** p0<.05, * p<0.1.
Source: Authors' calculations.

observe a positive effect on efficiency for compound neighbors. When outliers are controlled for, however, kinship and ethnicity components are seen to improve efficiency. The picture that emerges is thus one in which a few households with large land endowments distribute some of this land to newcomer lineages and ethnic minorities. But equalization of farm sizes is not achieved, and once the impact of the large landowners is reduced, we find evidence that social distance is an obstacle to efficiency-enhancing land transfers.

3.4.2 Two factor model

We now move on to the models with two factors of production outlined in section 3.2.2. Results of the simple model, which does not consider social networks, are presented in column 1 of Table 6. We find that, as anticipated, land is transferred so as to equilibrate factor ratios among households in each village: Land flows from households with a relatively high land–labor ratio ($\alpha_1 > 0$) to households with a relatively low land–labor ratio ($\alpha_2 > 0$).

Obstacles to land transfers imposed by pre-existing social ties are estimated using model (13) and reported in the remaining columns of Table 6. As in Table 3, ethnicity and kinship components have a negative impact on the efficiency properties of the land transfer network. This is only statistically significant for the kinship network, however. The same caveats regarding large landowners apply to these results as in the previous subsection. Table 7 report results where outliers are handled through either dummies or winsorization. As before, either approach leads to positive and significant

Table 6: Two factor model: Do land transfers equalize farm size?

Regressor	Coefficient	(1)	(2) Ethnicity	(3) Kinship	(4) Kinship	(5) Kinship	(6) Neighbor
$(r - \bar{r})$	α_1	0.001***	0.002***	0.003***	0.001**	0.003***	0.001
		(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$(\bar{r} - r_j)$	α_2	0.000**	0.001***	0.001***	0.000*	0.001***	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(r_i - \bar{r}_k)\tilde{D}_{ij}$	α_3		-0.001	-0.002***		-0.003***	
			(0.001)	(0.001)		(0.001)	
$(\bar{r}_k - r_j)\tilde{D}_{ij}$	α_4		-0.000	-0.001**		-0.001**	
			(0.000)	(0.000)		(0.000)	
$(r_i - r_j)L_{ij}$	α_5				0.000	0.001	0.001**
					(0.000)	(0.000)	(0.001)
Constant	α_0	0.019***	0.019***	0.019***	0.019***	0.019***	0.014***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations		57,060	57,060	57,060	57,060	57,060	23,917

Note: The dependent variable is the net out transfer of land measured in hectares. *** p<0.01, ** p0<.05, * p<0.1.
Source: Authors' calculations.

Table 7: Robustness of two factor results: outliers

Regressor	Coefficient	(1) Ethnicity	(2) Ethnicity	(3) Kinship	(4) Kinship
$(r - \bar{r})$	α_1	0.001**	0.000	0.001*	-0.000
		(0.000)	(0.001)	(0.001)	(0.001)
$(\bar{r} - r_j)$	α_2	0.000**	0.000	0.000**	0.000**
		(0.000)	(0.000)	(0.000)	(0.000)
$(r_i - \bar{r}_k)\tilde{D}_{ij}$	α_3	0.017***	0.003**	0.016***	0.003**
		(0.005)	(0.002)	(0.004)	(0.001)
$(\bar{r}_k - r_j)\tilde{D}_{ij}$	α_4	-0.001	0.001	-0.002	-0.000
		(0.002)	(0.001)	(0.002)	(0.001)
Constant	α_0	0.015***	0.019***	0.015***	0.019***
		(0.002)	(0.002)	(0.003)	(0.002)
Outlier method		Dummies	Winsorize	Dummies	Winsorize
Observations		57,060	57,060	57,060	57,060

Note: The dependent variable is the net out transfer of land measured in hectares. *** p<0.01, ** p0<.05, * p<0.1.
Source: Authors' calculations.

component effects ($\alpha_3 > 0$) for both the kinship and the ethnicity components.

To conclude, results of the two factor model are quite similar to those of the one factor model: Land transfers increases efficiency but do not achieve equalization in farm size or land–labor ratios. When considering the raw data, the only pre-existing network that improves efficiency further is being a compound neighbor. However, when we estimate models that are robust to the impact of a few large landowners, we find that both kinship and ethnicity also increase efficiency within components.

3.5 Discussion

We have shown that land transfers on average increase allocative efficiency. Some pre-existing social networks can increase the amount of transfers that improve efficiency. This does not imply that the perceived motive for land transfers is to increase efficiency. An alternative interpretation is that transfers of land are a way of helping poor households (see Chapter 4 of this dissertation for an in-depth discussion of this possibility). Using the same data, Jaimovich (2011) finds that differences in income inequality across villages drive transfers of inputs and credit – but not exchange of land. No matter whether the intent behind land transfers is equity or efficiency, the end result is the same: Transferring land to a land-poor household improves the allocation of factors – and is a more efficient way of assisting needy households than keeping them with insufficient land and helping them out after the harvest (Fafchamps, 1992). Thus, based on our model and results, we would expect more connected villages to have less unequal distributions of land–labor ratios.

Table A.2 shows regressions of village-level inequality in land–labor ratios on social network density. Here we define efficiency as equal land–labor ratios; hence, lower inequality in land–labor ratios implies higher efficiency in the two factor model. To ease interpretation, the network density measures have been normalized by dividing with their standard deviation. We estimate regressions using Gini coefficients on the land–labor ratios before and after land transfers have taken place. The results show a negative association between village network densities and factor ratio inequality, suggesting that more socially connected villages are less unequal both before and after land transfers. Regarding the absolute change in factor inequality, more than 75 percent of villages have a less unequal factor distribution after the reallocation of land and the mean village gini decreases from 55.8 to 51.2. Social network density is not only correlated with ex-post production factor inequality, but also strongly correlated with ex-ante inequality. This is perhaps not surprising: The same underlying factors that hinder or facilitate the short-term land transfers we study here may also, over time, hinder or facilitate permanent land transfers between households, for instance, through inheritance or transfers inter vivo at the time of marriage. The estimated effects are small but economically meaningful: A one standard deviation increase in the density measure decreases ex-ante land–labor inequality by between 0.4 and 0.7 Gini points, depending on which network is used to calculate density measures.

How vulnerable are our results to productivity differentials across farmers? We have argued that, provided some assumptions are made about productivity (i.e., multiplicative θ_i's and homotheticity) and about the correlation between factor endowments and productivity, the interpretation of the two-factor results presented in Table 4 is qualitatively the same as long as transfers of land and labor are both possible. This interpretation carries over to the case of unobservable factor-specific productivity differences, as long as the production function is (approximately) Cobb-Douglas. Indeed, in that case, factor-specific productivity shifters factor out of the land and labor aggregates and are observationally equivalent to θ_i.

Things are different if the production function is not Cobb-Douglas. In this case, if labor or land is of varying quality, allocative efficiency need not imply factor ratio equalization. How serious a threat is this to our identification approach? We suspect that unobservable productivity differences matter more in modern agriculture, and less in traditional agriculture, which is fairly undifferentiated and follows simple heuristic rules of behavior. Chavas et al. (2005) find evidence that technical inefficiency among Gambian farmers is modest and that differences in factor mix are not driven by technological differences. The authors nonetheless find substantial allocative inefficiency in farm input allocation, which motivates the focus of the present paper.

To ascertain how sensitive our findings are to factor productivity differentials across and within villages, we re-estimate our regressions with village fixed effects and household-specific controls in model (13). Village fixed effects control for differences in land quality across villages. The territory of each individual village is fairly small so that within-village differences in land quality are likely to be small. The household-specific control variables are the same as those employed in Table 3, i.e., household size; age of household head; indicators for the literacy of the household head; for whether the household head has any formal education; for whether the household head is female; for whether the household receives remittances; and a relative wealth-level indicator. These controls, which are added for both the sending and receiving household, are intended to proxy for unobserved labor productivity differences. Results of the two-factor model with village fixed effects and household characteristics are reported in Table A.3. Our main findings do not change, suggesting that productivity differences are unlikely to be driving the variation in land endowment across households in our study area.

As pointed out in the data section, another possible threat to identification is measurement error in the reported land endowment and land transfers. Measurement error in the land endowment variable a_i is expected to generate attenuation bias, in which case estimated coefficients should be seen as a lower bound of the true coefficients. Measurement error in the dependent variable, in contrast, should normally result in a higher variance of the error term and thus weaker inference and a loss of power. It is also conceivable that this higher variance resulted in a fluke, i.e., a false rejection. We therefore wish to investigate how sensitive our results are to measurement error in land transfers.

To this effect, we start from the reasonable assumption that even when households are unable to accurately report the actual size of the land they transferred, they are much less likely to misreport the *fact* of transferring land to or from a particular household. For this reason, information on whether land was transferred is probably more reliable than the size of the land transfer. To investigate whether our results depend on the magnitude of the observed land transfers, we re-estimate our model replacing as dependent variable the size of reported transfers (i.e. w_{ij}) with a binary indicator taking the value of 1 if household i transfers land to j, and zero otherwise. Formally, it is possible to have an increase in efficiency even though most transactions go from land-poor to land-rich households if these transactions are outweighed by few but larger transactions in the opposite direction. Therefore, estimation on the binary transfer indicator variable is not a valid test of the efficiency hypothesis. However, if results are qualitatively similar to our main specification, it supports that the main results are not driven by misreporting in the size of transferred land.

Estimation results for the two factor model are reported in Table A.4. Results are similar to the baseline results reported in Table 6, except that the coefficient of the kinship component variable is less statistically significant. We take this as reassuring evidence that our main findings are not simply an artifact of measurement error in the size of land transfers.

Land transfers are not the only margin that can be used to equalize factor ratios across households. Another option is to adjust the amount of labor during the cropping season. To formally examine this we estimate model (12) using labor transfers as dependent variable w_{ij}. The estimated model is of the form:

$$w_{ij} = \beta_0 + \beta_1(r_i - \bar{r}) + \beta_2(\bar{r} - r_j) + u_{ij}$$

Labor transfers that improve allocative efficiency require $\beta_1 < 0$ and $\beta_2 < 0$: labor l should on average to flow from agents with a low land–labor endowment ratios r_i to agents with high land–labor ratios relative to the village average $\bar{r}$.

Estimation results are reported in Table A.5. We find only weak evidence that labor transfers improve allocative efficiency in factors of production: β_2 is negative and statistically significant at the 10 percent level in two out of seven specifications. In the kinship regressions, β_5 is significant and negative at the 10 percent level. This is weak evidence that that kinship may improve efficiency of labor transactions. All other coefficients are statistically insignificant. In particular, and in contrast to the previous findings for land transfers, geographical proximity does not improve allocative efficiency of labor transactions. These findings suggest that labor transfers do not help equalize ex-ante factor ratios in the studied villages. We suspect that labor transfers primarily serve as an ex-post adjustment to labor shocks, and labor shocks are not correlated with factor ratios.

If labor transfers primarily help farmers adjust to labor shocks or during spike periods of demand during the growing season, these transfers will typically take place after land transactions have taken place, i.e., after cultivation rights have been allocated and planting has begun. In this case, it would be correct to compare labor transfers to a land–labor ratio that uses the amount of cultivated land, as opposed to the amount of owned land, as the numerator. To investigate this possibility, we redefine r_i in model (12) and (13) as $\tilde{r}_i \equiv \frac{x_i}{l_i}$, i.e., the ratio of cultivated land after all land transactions have taken place to the endowment of labor. If we take land transfers as given, labor should on average flow from agents with low $\tilde{r}_i$'s to agents with high $\tilde{r}_i$'s relative to the village average $\bar{r}$ in order to achieve ex-post allocative efficiency. Results are reported in Table A.6. Similar to the previous results using the households' land endowment, we find that β_2 is significant and negative at the 10 percent level in two out of seven specifications. β_5 is also found to be negative and significant at the 10 percent level in the two kinship regressions. In sum, there is no strong evidence that labor transfers play an important role in equilibrating factor ratios across households, either ex ante or ex post. The presence of social ties does not influence this conclusion.

3.6 Conclusion

In this paper, we have examined whether pre-existing social networks affect the allocative efficiency of factor transfers. It is widely believed that frictions caused by lack of trust or observability explain why many economic transactions take place along ethnic and social networks, or between neighbors. More efficiency gains are therefore expected to be achieved through exchange within ethnic groups or between kin and between neighbors – unless networks are sufficiently dense and fluid to allow resources to flow to their optimal user.

To our knowledge, this paper constitutes the first attempt to test this prediction directly. We introduce a novel way of testing whether social and geographical distances reduce the likelihood of efficiency-improving factor transfers. The approach can be used to test whether social structure inhibits allocative efficiency. The data requirements of this approach are quite daunting, however: We need exhaustive information about all factor transfers and all social links. Using such data for a large number of villages in The Gambia, we implement the approach empirically.

The results support the hypothesis that land transfers improve allocative efficiency. However, we strongly reject that allocative efficiency is achieved. We find evidence that efficiency-enhancing transfers of land are more likely between households that are geographically or socially proximate. Contrary to expectations, we do not find that efficiency-enhancing transfers of land are more likely within ethnic groups or kinship groups – if anything, we find the opposite result. This finding, however, is largely driven by land transfers originating from a small number of households with large landholdings. Once we mitigate their impact on the regression, we find the anticipated result that land flows more freely to equalize endowments within ethnic groups. We find no evidence that labor transfers serve to increase efficiency of factor allocations. But we do not rule out that they serve a beneficial role ex post, e.g., to deal with unanticipated health shocks – on which we have no data.

References

Arcand, J.-L. and Jaimovich, D. (2014). Does ethnic diversity decrease economic interactions? Evidence from exchange networks in rural Gambia. MPRA Paper No. 60497.

Attanasio, O., Barr, A., Cardenas, J. C., Genicot, G., and Meghir, C. (2012). Risk Pooling, Risk Preferences, and Social Networks. *American Economic Journal: Applied Economics*, 4(2):134–167.

Baier, S. L. and Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics*, 71(1):72–95.

Bank, A. D. (2005). *Appraisal report: farmer managed rice project*. Agricultural and Rural Development Department, Central and West Regions, African Development Bank Group. Tunis, Tunisia.

Barr, A., Dekker, M., and Fafchamps, M. (2012). Who Shares Risk with Whom under Different Enforcement Mechanisms? *Economic Development and Cultural Change*, 10(4):677–706.

Bergstrand, J. H. and Egger, P. (2011). Gravity equations and economic frictions in the world economy. In Greenway, D., Falvey, R., Kreickemeier, U., and Bernhofen, D., editors, *Palgrave Handbook of International Trade*, pages 532–570. Palgrave Macmillan Press, New York.

Chaney, T. (2014). The Network Structure of International Trade. *American Economic Review*, 104(11):3600–3634.

Chavas, J.-P., Petrie, R., and Roth, M. (2005). Farm Household Production Efficiency: Evidence from The Gambia. *American Journal of Agricultural Economics*, 87:160–179.

Coase, R. H. (1937). The nature of the firm. *economica*, 4(16):386–405.

de Janvry, A., Fafchamps, M., and Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. *The Economic Journal*, 101:1400–1417.

Dey, J. (1982). Development planning in The Gambia: The gap between planners' and farmers' perceptions, expectations and objectives. *World Development*, 10(5):377–396.

Eastman, C. (1990). Gambian usufruct tenure. Help or hindrance to agricultural production? *Land Use Policy*, 7(1):27–40.

Fafchamps, M. (1992). Solidarity Networks in Preindustrial Societies: Rational Peasants with Moral Economy. *Economic Development and Cultural Change*, 41(1):147–174.

Fafchamps, M. (1993). Sequential Labor Decisions Under Uncertainty: An Estimable Household Model of West-African Farmers. *Econometrica*, 61(5):pp. 1173–1197.

Fafchamps, M. and Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics*, 83:326–350.

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

Frankel, J., Stein, E., and Wei, S.-j. (1995). Trading blocs and the Americas: The natural, the unnatural, and the super-natural. *Journal of Development Economics*, 47(1):61–95.

Freudenberger, M. S. (2000). Tenure and Natural Resources in The Gambia: Summary of Research Findings and Policy Options. Land Tenure Center. University of Wisconsin-Madison. Working paper no. 40.

GBS (2011). Integrated Household Survey. Income and Expenditure. Poverty Assessment - 2010. Technical report, Gambia Bureau of Statistics.

Granovetter, M. (1995). *Getting a job: A study of contacts and careers*. University of Chicago Press.

Heston, A., Summers, R., and Aten, B. (2012). Penn World Table Version 7.1.

Holden, S. and Ghebru, H. (2005). Kinship transaction costs and land rental market participation. Department of Economics and Resource Management. Norwegian University of Life Sciences.

Holden, S. T., Otsuka, K., and Place, F., editors (2009). *The Emergence of Land Markets in Africa: Impacts on Poverty, Equity, and Efficiency.* Washington, DC: Resources for the Future.

Jaimovich, D. (2011). Macrostructure and microstructure: Evidence from overlapping village networks in The Gambia. MPRA paper 38932, University of Munich.

Jaimovich, D. (2015). Missing Links, Missing Markets: Evidence of the Transformation Process in the Economic Networks of Gambian Villages. *World Development*, 66:645–664.

Macours, K., Janvry, A. d., and Sadoulet, E. (2010). Insecurity of property rights and social matching in the tenancy market. *European Economic Review*, 54(7):880–899.

Mazzocco, M. and Saini, S. (2012). Testing Efficiency Risk Sharing with Heterogeneous Risk Preferences. *American Economic Review*, 1:428–468.

Otsuka, K. (2007). Efficiency and Equity Effects of Land Markets. *Handbook of Agricultural Economics*, 3:2671–2703.

Pender, J. and Fafchamps, M. (2005). Land Lease Markets and Agricultural Efficiency in Ethiopia. *Journal of African Economies*, 15(2):251–284.

Rauch, J. E. (2001). Business and Social Networks in International Trade. *Journal of Economic Literature*, 39(4):1177–1203.

Rauch, J. E. and Casella, A. (2003). Overcoming Informational Barriers to International Resource Allocation: Prices and Ties. *The Economic Journal*, 113(484):21–42.

Sadoulet, E., de Janvry, A., and Fukui, S. (1997). The meaning of kinship in sharecropping contracts. *American Journal of Agricultural Economics*, 79(2):394–407.

Swindell, K. (1987). Family Farms and Migrant Labour: The Strange Farmers of the Gambia. *Canadian Journal of African Studies*, 12(1):3–17.

The Gambia Bureau of Statistics (GBOS) (2008). Population and housing census 2003 - volume 6: Housing and household characteristics. Technical report, The Gambia Bureau of Statistics (GBOS).

The Gambia Bureau of Statistics (GBOS) (2011). The Gambia Multiple Indicator Cluster Survey 2010, Final Report. Technical report, The Gambia Bureau of Statistics.

Topa, G. (2001). Social Interactions, Local Spillovers and Unemployment. *The Review of Economic Studies*, 68(2):261–295.

von Braun, J. and Webb, P. J. R. (1989). The Impact of New Crop Technology on the Agricultural Division of Labor in a West African Setting. *Economic Development and Cultural Change*, 37:513–534.

Webb, P. (1989). Intrahousehold Decisionmaking and Resource Control: The Effects of Rice Commercialization in West Africa. Technical report, International Food Policy Research Institute.

Appendix

Table A.1: Summary table of model predictions

Case	Hetero-geneity?	No. of factors	No. of traded factors	Returns to scale	Prediction of factor structure across farms	Model(s)
1	NO	1	1	DRS	Equality of factors.	(4), (6)-(9)
2	NO	1	1	CRS	Undetermined. If any TC, no trade.	Ruled out
3	NO	2	1	DRS	Unidentified.	-
4	NO	2	1	CRS	Equality of factor ratios. Cultivated land is multiple of ex-ante labor.	(12), (13)
5	NO	2	2	DRS	Equality of both factors.	(6)-(9), (12), (13)
6	NO	2	2	CRS	Equality of factor ratios. If any TC, trade only in factor with lowest TC.	(12), (13)
7	YES	1	1	DRS	Unidentified.	-
8	YES	1	1	CRS	One plantation.	Ruled out
9	YES	2	1	DRS	Unidentified.	-
10	YES	2	1	CRS	Unidentified.	-
11	YES	2	2	DRS	If θ is not strongly correlated with r: Equality of factor ratios.	(12), (13)
12	YES	2	2	CRS	One plantation.	Ruled out

Note: DRS/CRS are decreasing and constant returns to scale, respectively. TC stands for transaction costs. In the cases with two factors and one transfer (case 9 and 10), we assume that trade is in land and that labor is untraded.
Source: Authors' compilation.

Table A.2: Inequality in land–labor ratios before and after land transactions have taken place.

Density measure	(1) Before	(2) After	(3) Before	(4) After	(5) Before	(6) After
Ethnicity/SD	-0.037*	-0.007				
	(0.022)	(0.020)				
Kinship/SD			-0.069***	-0.037*		
			(0.020)	(0.020)		
Neighbor/SD					-0.060***	-0.037*
					(0.021)	(0.020)
Constant	0.660***	0.531***	0.687***	0.580***	0.610***	0.544***
	(0.064)	(0.059)	(0.042)	(0.041)	(0.027)	(0.026)
Observations	51	51	51	51	51	51

Note: The dependent variable is within village Gini-inequality in land–labor ratios. Network density is the is the ratio of all social links of the given type divided by the total number of possible within-village links. The density measures are then normalized by dividing with their village-level standard deviation. *** p<0.01, ** p<0.05, * p<0.1.
Source: Authors' calculations.

Table A.3: Robustness of two factor results: adding controls

Regressor	Coefficient	(1)	(2) Ethnicity	(3) Kinship	(4) Kinship	(5) Kinship	(6) Neighbor
$(r-\bar{r})$	α_1	0.001**	0.002***	0.003***	0.001**	0.003***	0.001
		(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
$(\bar{r}-r_j)$	α_2	0.000**	0.001***	0.001***	0.000*	0.001***	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(r_i-\bar{r}_k)\tilde{D}_{ij}$	α_3		-0.001	-0.002***		-0.002***	
			(0.001)	(0.001)		(0.001)	
$(\bar{r}_k-r_j)\tilde{D}_{ij}$	α_4		-0.000	-0.001**		-0.001***	
			(0.000)	(0.000)		(0.000)	
$(r_i-r_j)L_{ij}$	α_5				0.000	0.001	0.001**
					(0.000)	(0.000)	(0.001)
Constant	α_0	-0.019*	-0.019*	-0.018*	-0.019*	-0.018*	0.049***
		(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
Observations		57,060	57,060	57,060	57,060	57,060	23,917
Village fixed effects		YES	YES	YES	YES	YES	YES
Control variables		YES	YES	YES	YES	YES	YES

Note: The dependent variable is the net out transfer of land measured in hectares. Control variables include the following receiver and sender characteristics: household size; household age; dummies for whether the household head is literate, whether household head has any formal education, whether household head is female, whether household receives remittances as well as for self-reported wealth quartile. *** p<0.01, ** p<0.05, * p<0.1.
Source: Authors' calculations.

Table A.4: Robustness: Two factor model using binary dependent variable for land transfers

Regressor	Coefficient	(1)	(2) Ethnicity	(3) Kinship	(4) Kinship	(5) Kinship	(6) Compound neighbor
$r_i - \bar{r}$	α_1	0.001***	0.001***	0.001***	0.001***	0.001***	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\bar{r} - r_j$	α_2	0.000**	0.000***	0.000***	0.000	0.000***	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$(r_i - \bar{r}_k)\tilde{D}_{ij}$	α_3		-0.000	-0.000**		-0.001***	
			(0.000)	(0.000)		(0.000)	
$(\bar{r}_k - r_j)\tilde{D}_{ij}$	α_4		-0.000	-0.000		-0.000*	
			(0.000)	(0.000)		(0.000)	
$(r_i - r_j)L_{ij}$	α_5				0.000	0.000*	0.001***
					(0.000)	(0.000)	(0.000)
Constant	α_0	0.010***	0.010***	0.010***	0.010***	0.010***	0.008***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations		57,060	57,060	57,060	57,060	57,060	23,917

Note: The dependent variable is equal 1 if there is a net out-transfer of land.
Source: Authors' calculations.

Table A.5: Two factor model: Do labor transfers equalize farm size?

Regressor	Coefficient	(1)	(2) Ethnicity	(3) Kinship	(4) Kinship	(5) Kinship	(6) Neighbor
$(r - \bar{r})$	β_1	0.001	0.000	0.000	0.003	0.000	-0.000
		(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
$(\bar{r} - r_j)$	β_2	-0.003	-0.001	-0.003*	-0.002	-0.003*	0.002
		(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
$(r_i - \bar{r}_k)\tilde{D}_{ij}$	β_3		0.001	0.001		0.003	
			(0.002)	(0.002)		(0.002)	
$(\bar{r}_k - r_j)\tilde{D}_{ij}$	β_4		-0.003	-0.001		0.001	
			(0.003)	(0.002)		(0.002)	
$(r_i - r_j)L_{ij}$	β_5				-0.009*	-0.009*	0.001
					(0.005)	(0.005)	(0.001)
Constant	β_0	0.105***	0.105***	0.105***	0.105***	0.105***	0.086***
		(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.015)
Observations		57,060	57,060	57,060	57,060	57,060	23,917

Note: The dependent variable is the amount of labor transfered from j to i measured in working days. land–labor ratios (r) are calculated using ex-ante measures of both land and labor. *** p<0.01, ** p<0.05, * p<0.1.
Source: Authors' calculations.

Table A.6: Two factor model: Do labor transfers equalize farm size? Using cultivated land.

Regressor	Coefficient	(1)	(2) Ethnicity	(3) Kinship	(4) Kinship	(5) Kinship	(6) Neighbor
$(\tilde{r}-\bar{r})$	β_1	0.001	-0.000	0.000	0.002	0.000	-0.000
		(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
$(\bar{r}-\tilde{r}_j)$	β_2	-0.003	-0.000	-0.002*	-0.002	-0.002*	0.002
		(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
$(\tilde{r}_i-\bar{r}_k)\tilde{D}_{ij}$	β_3		0.001	0.001		0.002	
			(0.002)	(0.001)		(0.002)	
$(\bar{r}_k-\tilde{r}_j)\tilde{D}_{ij}$	β_4		-0.004	-0.001		0.001	
			(0.003)	(0.002)		(0.002)	
$(\tilde{r}_i-\tilde{r}_j)L_{ij}$	β_5				-0.009*	-0.009*	0.001
					(0.005)	(0.005)	(0.001)
Constant	β_0	0.106***	0.106***	0.105***	0.106***	0.105***	0.086***
		(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.015)
Observations		57,060	57,060	57,060	57,060	57,060	23,917

Note: The dependent variable is the amount of labor transfered from j to i measured in working days. land–labor ratios (r) are calculated using ex-ante measures of both land and labor. *** p<0.01, ** p<0.05, * p<0.1.
Source: Authors' calculations.

CHAPTER 4

INSURING THE POOR: INTER-HOUSEHOLD LAND TRANSFERS AND THE IMPORTANCE OF LAND ABUNDANCE AND ETHNICITY IN THE GAMBIA

Ulrik Beck* Benedikte Bjerge†

Abstract

We use a network-level dataset to study patterns of inter-household land transactions in rural villages in The Gambia. We find that land transactions are pro-poor: Poorer households receive more land, and richer households donate more land. This is consistent with the presence of norm-based land access rules, of which there is ample qualitative evidence. The result is driven by less densely populated and less ethnically diverse villages where social norms tend to be more important. This has implications for redistributive land reforms: While previous reform efforts may have been crowded out by a decline in inter-household land transfers, increases in population densities may require future reforms as local social security institutions break down.

We benefitted from comments from seminar participants at University of Copenhagen and the Goethe University in Frankfurt, as well as conference participants at the Nordic Conference in Development Economics 2015, the Annual World Bank Conference on Land and Poverty 2015, and the International Conference of Agricultural Economists 2015. The authors also thank Dany Jaimovich, Thomas Markussen, Tseday Mekasha, Carol Newman, John Rand, Måns Söderbom and Finn Tarp for their very valuable comments. Finally, we thank Jean-Louis Arcand and Dany Jaimovich for making the Gambia Networks Data 2009 available to us. All interpretations and any remaining errors are the sole responsibility of the authors.

*Department of Economics, University of Copenhagen

†Department of Food and Resource Economics, University of Copenhagen

4.1 Introduction

Material exchange between households is an important insurance mechanism for households in many developing countries where formal insurance opportunities are often non-existent, prohibitively expensive or ineffective (Morduch, 1995; Cox and Fafchamps, 2007). When households are hit by idiosyncratic shocks, they often rely on help from their network of relatives and close friends (Fafchamps and Lund, 2003; Fafchamps and Gubert, 2007; Mazzocco and Saini, 2012).

A related issue is insurance of livelihoods in the absence of shocks, or before shocks have occurred. Households who do not own land and have no other sources of income may have to rely on help from other members of the community to ensure their livelihood. Many village-based communities in developing countries have traditionally relied on a set of norm-based access rules to grant land usage rights to those in need (Wolf, 1957; Cohen, 1980; Noronha, 1985; Eastman, 1990; Platteau, 1991; Freudenberger, 2000; Pamela, 2010). Under such a set of rules, membership in the rural community provides short-term usage rights to land. This ensures the livelihoods of households who lack formal land rights. In this way, norm-based access rules to land can function as a local social security institution.

The topic of norm-based access rules to land has received relatively little attention in the development economics literature, even though customary tenure systems still exist in many sub-Saharan rural areas (Deininger, 2003; Pamela, 2010). One reason for this scarce attention is that traditional tenure systems and norm-based access rules are thought to be disappearing (Devereux, 2001; Platteau, 2006). A long-standing hypothesis for explaining this disappearance is that norm-based access rules can only exist when land is sufficiently abundant. Increases in population density make land more valuable and markets emerge as a way to allocate the increasingly scarce land (Boserup, 1965; Platteau, 2002; Fenske, 2013). A complementary explanation is that well-functioning norm-based access rules depend on a high level of intra-village trust and solidarity, which may be difficult to maintain in the presence of ethnic heterogeneity.

In this paper, we ask whether it is possible to find empirical evidence that supports the existence of norm-based access rules in the present day. We also ask whether the strength of such rules depend upon land abundance and ethnic fractionalization at the local level. The empirical evidence on the transformation at the microeconomic level of such indigenous systems is quite scarce. This paper aims to fill this gap.

We do so by investigating the effect of household income on inter-household land transfers using a unique social network dataset collected in 2009 in rural villages in The Gambia. We proceed by investigating whether village-level differences in population density and ethnic heterogeneity can explain differences in the strength of the norm-based access rule. The use of a network dataset with information at the dyadic level allows us to address several identification issues that plague the closely related literature of inter-household gift-giving behavior (e.g., Cox et al., 2004; Kazianga, 2006; Mitrut and Nordblom, 2010): we correct for potential omitted variable bias by including both recipient and donor attributes; we check for potential endogeneity of monetary income using information about households' pre-transfer income; and we account for unobserved household heterogeneity.

Why does the existence of norm-based access rules to land matter? There are at least two reasons. First, the existence of such rules has implications for our understanding of welfare and inequality in these societies: Where short-term transfers of land usage rights are frequent and goes to the poor, outright land ownership is a poor indicator of welfare. Second, informal social security schemes can affect the impact of redistributive land reforms. The presence of norm-based social security systems may explain the relatively poor effects on growth and equity of such reforms in the past (for a review of West African land reforms, see Fenske, 2011) since such reforms can crowd out private voluntary transfers of land to those in need.[1] This means that the impact of the reform is likely to be smaller than anticipated. In fact, it is not certain that the displacement of informal safety nets with public or market-based alternatives is socially and economically preferable (Devereux, 2001). However, as the strength of social security norms wanes as land becomes scarce, or where ethnic fractionalization lowers trust, reforms may become needed to safeguard the livelihoods of the poorest part of the rural population.

Rural Gambia is an illustrative case for studying traditional social security systems. While the land system in urban areas has been subject to land reform, the traditional system of land rights is still in effect in rural areas (Freudenberger, 2000; Chavas et al., 2005; Pamela, 2010). This implies that those who possess surplus land as well as the village chief have a moral obligation to allocate land to those in need. Allocations of land usage rights are temporary and often non-monetary in nature: Donors of land rarely receive monetary payment for the land that they lend to other farmers. A World Bank report states: "In the absence of any state-supported welfare programs, social safety nets in The Gambia are based on social and religious traditions. [...] anyone with above average earnings is expected to support near relatives and friends with lower income levels" (World Bank, 1993). The characteristics of the traditional land rights system of The Gambia is not unusual compared to that of other West African countries. Common characteristics include the unequal land ownership structure and the central role of village chiefs in the reallocation of land (Eastman, 1990; Platteau, 2002; Otsuka, 2007; Fenske, 2011; Holden et al., 2009). Therefore, the results of this paper concern not only The Gambia, but also other developing countries where informal land markets still exist.

[1] For a theoretical model which supports this argument, see Bourlès and Bramoullé (2013) who show how a Pigou–Dalton transfer from rich to poor, e.g., from a redistributive reform, can increase post-transfer inequality in a network setting which allows for altruism.

Our findings support that inter-household transfers of land are motivated by social security considerations as we find that inter-household land transfers are pro-poor. The poorest households with no or little monetary income per capita receive more land and more land is transferred to landless households. This pattern stands in contrast to gift-giving behavior in other contexts where "transfers occur through social networks and poor individuals are excluded from these networks" (Kazianga, 2006). This result is driven by exchange in villages with low population densities, which supports the hypothesis that population pressure affects the functioning of the norm-based access rule. The result is robust to inclusion of household and village fixed effects as well as an alternative income measure and alternative definitions of high- and low-density villages. We also find relatively weaker evidence that norm-based access rules function less well in more ethnically diverse areas. We explain the relative weakness of this finding by the long history of peaceful co-existence of several different ethnicities in The Gambia. Taken together, the evidence suggests that future land reforms should take into account that the extent to which norm-based transfers are crowded out may vary at the local level depending on the scarcity of land and the ethnic composition.

The paper is organized as follows. Section 4.2 discusses the traditional land access norms in more detail and why traditional land access norms are thought to be disappearing. In Section 4.3, we outline the empirical method. Section 4.4 describes the dataset and provides descriptive information. Section 4.5 presents the results and a series of robustness checks. Section 4.6 concludes.

4.2 Traditional land-access norms

One can distinguish between two types of traditional social institutions aimed at the reduction of food shortages of individual households, namely informal mutual insurance arrangements and norm-based access rules for vital resources (Platteau, 1991, 2002). While the existing empirical literature has been successful at documenting the existence and effects of informal mutual insurance arrangements (e.g., Fafchamps and Lund, 2003; Kazianga and Udry, 2006; Dercon and Weerdt, 2006; Fafchamps and Gubert, 2007; Mazzocco and Saini, 2012; Caudell et al., 2015), this paper focuses instead on the mechanics of norm-based access rules for land. Informal mutual insurance arrangements normally kick in after a shock has occurred, and can therefore be regarded as an ex-post insurance mechanism. On the other hand, norm-based access rules for land is an ex-ante insurance mechanism used to secure livelihoods on an annual basis, independent of whether a shock occurs. While ex-post arrangements compensate for a shortfall in income or consumption, ex-ante arrangements attempt to prevent the occurrence of a shortfall.

There are several examples of such ex-ante insurance arrangements (e.g., Platteau and Abrahamb, 1987; Fafchamps, 1992; Platteau, 1997; Devereux, 2001; McGuire, 2008; Krishnan and Sciubba, 2009). Examples include informal farmer seed distribution systems to ensure supply and access to affordable seeds (McGuire, 2008) and labor sharing arrangements during the cropping and harvesting season to warrant completion of farm operations in time (Krishnan and Sciubba, 2009). This paper studies another such arrangement, namely temporary and non-monetary land transfers to poor or landless households before the planting season to ensure that these households can support their own livelihood.

Ex-ante arrangements that reduce food shortages can be preferable to ex-post relief since it reduces the potential for moral hazard, thereby avoiding the waste of community resources (Fafchamps, 1992). To fix ideas, consider an economy of agricultural households who produce agricultural output using identical production functions that do not have increasing returns to scale. The production function takes land and labor as inputs. Total consumption is the sum of agricultural production and income from non-agricultural activities. Some households are land-abundant in the sense that in the absence of shocks, they can maintain more than the subsistence level of consumption from agricultural production and non-agricultural income. Other households are land-poor or landless and do not have compensating non-agricultural income. Suppose further that households care about the consumption level of other households in the sense that they do not want them to fall below subsistence level. This goal may be achieved in a variety of ways: One possibility is that the land-abundant households employ the land-poor households to work on their surplus land, through either a wage contract or a sharecropping agreement (Eswaran and Kotwal, 1985). Another option is that land-abundant households compensate land-poor households with output after production has taken place. However, these options are not as attractive as temporary transfers of land that take place ex ante, since such transfers plausibly minimize costs related to labor supervision, shirking and moral hazard by allowing recipients to keep all the gains from their efforts (Fafchamps, 1992). Thus, ex-ante land transfers are a way to avoid the need for future assistance, while making better use of the total endowment of labor and land resources.

On the other hand, land transfers do not work well as ex-post insurance, i.e., to smooth consumption after the occurrence of a shock, due to the substantial delay between planting and harvesting. Ex-post insurance instead takes place through transfers of labor, cash or goods. A network of land transfers is therefore well suited to investigate the existence of an ex-ante insurance mechanism.

4.2.1 The Gambia

As mentioned in the introduction, this paper uses data from villages in The Gambia. This section outlines the traditional system of norm-based access rules to land in The Gambia.

Traditional ex-ante insurance schemes, such as land access norms, are rooted in a shared belief of how to behave towards other community members. The definition of "member" varies over time and space but it typically includes those who can claim descent from the founding lineages of the village as well as former migrants who over time have been accepted as members of the village community (Platteau, 2002). In The Gambia, feelings of duty, cooperation, trust, and obligation towards kin and friends form the basis of the local moral economy (Platteau, 2006; Pamela, 2010). Such feelings are central to the concept referred to as *badingya* by the largest ethnic group in The Gambia, the Mandinkas. *Badingya* represents cooperation, obligation, harmony, and productivity (von Braun and Webb, 1989; Freudenberger, 1993). In contrast, the concept of *fadingya* refers to the negative traits of individual selfish ambitions and competition. These twin concepts bind relatives and communities together by shaping notions of social justice. They also impose a limit on accumulation of private productive assets, thereby ensuring social stability by avoiding a delinking of individual actions and the best interest of the social group as a whole (von Braun and Webb, 1989; Platteau, 2006).

Two principal types of land usage rights exist in The Gambia. They are referred to as primary and secondary rights. Primary usage rights are obtained by clearing bush land. Primary rights are similar, although not equivalent, to the Western concept of land ownership: The household with primary rights can decide which crops to grow and whether to lend some of the land to other farmers. Under the indigenous tenure system, outright sale of primary rights, as well as leasing of land, is prohibited (Freudenberger, 1993). Instead, primary usage rights are inherited and descendants of those who first settled and cultivated the land, as well as the village chief, called the *Alkalo*, retain considerable shares of the primary usage rights in most villages (Dey, 1982; Freudenberger, 2000; Pamela, 2010). This creates a highly unequal distribution of primary land rights between households, where relative newcomers have very little land or none at all. As a result of the inequalities of primary usage rights, allocation of secondary usage rights is an important mode of access to land for land-poor households. Secondary usage rights are temporary rights to cultivate plots. The rights to the land revert to the donor after one or sometimes a few cropping seasons.[2]

Descendants of the first settlers who possess surplus land have a moral obligation to give secondary usage rights to some of their land to those in need.[3] The *Alkalo* can also decide that one household must transfer usage rights to another household (Freudenberger, 2000).[4] If the *Alkalo* is among the major landholders in the community,

[2]Previously, land was often lent for several seasons at a time. To prevent people from borrowing land over long periods of time to claim land ownership rights, land owners tend to insist on seasonal loans in order to maintain control (Dey, 1982, p.388). It should also be noted that borrowers are not necessarily allocated the same fields every year (Freudenberger, 2000, p.82).

[3]In a field study in the Gambian village Dumbutu, it was found that the principle that all residents should have access to land if they have the means to either cultivate or to build compounds is reinforced by the fact that many land-rich compounds do not have enough labor to use all their land, partly caused by the considerable urban migration of youth (Freudenberger, 2000).

[4]However, if the land is in use for agricultural purposes, the *Alkalo* needs the consent of the landholder.

he will often allocate some land of his own to poorer village members. Access to secondary usage rights is contingent on community membership status.

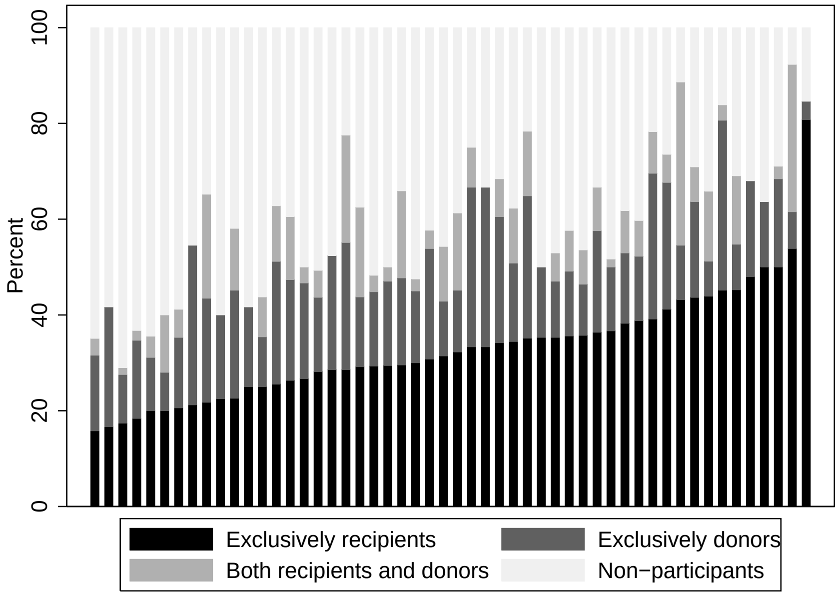
Transfers of temporary usage rights are often non-monetary in nature. This means that donors of land rarely receive monetary payment for the land they lend to other farmers (Eastman, 1990). Thus, transfers of secondary usage rights to land are a way to ensure the livelihoods of poor community members. In this way, the transfers fulfill a social security function. However, the lending of land itself may create ties that can be called upon when needed (Cashdan, 1985). Thus, in periods of labor shortages, often before and during the rainy season in relation to weeding and harvesting, donors of land may receive some labor or other inputs from land-receiving households. But much less labor is transferred than land, compared to the total endowments. While district-level authorities in Gambia administer some land for the benefit of district inhabitants, they do not have the legal capacity to interfere with the allocation of land owned by individual families (Freudenberger, 1993). Access to vital resources is generally not governed by explicit agreements or external enforcement but instead by social norms, such as those embedded in the concepts of *fadingya* and *badingya*.

Figure 1 confirms that land transfers are frequent in the surveyed sample of rural villages in The Gambia. Each bar in the figure represents a village; the top block represents the fraction of households that do not participate in the land market, while the bottom two blocks represent households that are exclusively recipients or donors of land. More than 55 percent of the households participate actively in the land market as either donors or recipients. A highly unequal distribution of primary land rights means that some households transfer land to more than one household. Around 40 percent of households are land recipients.

4.2.2 Threats to traditional norm based transfer systems

The traditional land institutions described above were developed under conditions of land abundance. However, it has been recognized, at least since the seminal work of Boserup (1965), that population increase can put pressure on existing land access institutions. A prerequisite for the effectiveness of transfers of land as a social security device is that productive resources are sufficiently plentiful compared to the size of the population. As land becomes increasingly scarce, land use patterns shift from extensive to intensive production, and the traditional land institutions are transformed (Noronha, 1985; Platteau, 2002; Goetghebuer and Platteau, 2005; Platteau, 2006; Peters, 2009). Population growth and increased commercialization result in the emergence of markets a means to allocate scarce production resources. Moreover, shorter fallow periods are possible, which means that less land is available for redistribution to the poorest households. In such settings of increasing scarceness of land and market-based allocation of resources, traditional social norms have less influence over actions than they do in the conventional village societies in which they emerged. The erosion of

Figure 1: Household participation in the land market by village



Note: Each bar represents a village.
Source: Authors' calculations.

norm-based access rules has the potential to reduce food security for individuals and social groups who previously relied on these norms for access to land. The Gambia has experienced high rates of population growth in the last few decades, resulting in a substantially higher population density than other regions in Sub-Saharan Africa.[5] It is therefore possible that Gambian villages – and especially those with high population densities – are experiencing the deteriorating effects of land scarcity on the strength of norm-based access rules, and it is an open question whether norm-based access rules still exist at all in this setting.

There is some evidence that increasing land pressure due to population increases has been leading to a breakdown of the prohibition on monetary land rentals over the last few decades in The Gambia (Dey, 1982; Freudenberger, 2000). However, an anthropological study based on key informants in a countrywide sample of 52 villages found that "no respondent reported a case of land being rented for money or a fixed share of the crop" (Eastman, 1990). This was also the experience during the collection of the dataset used in this paper, which took place in 2009.[6] Even if land transfers are still primarily moneyless assignments, it does not rule out that the basis of transfers could be changing from norm-based transfers towards a more market-based approach. Instead

[5]According to the World Development Indicators (World Bank, 2015), the total population in The Gambia was last recorded at 1.8 million people in 2012, up from 0.4 million in 1960, an increase of 384 percent during the last 50 years. This means that the population density has increased from 40 people per sq. km in 1960 to 171 in 2011. This should be compared to an average population density in Sub-Saharan Africa of 38 people per sq. km in 2011.

[6]We thank Dany Jaimovich, one of the collectors of the dataset, for this insight.

of money, such transfers can be paid for in other media such as labor or agricultural output.

Population growth and subsequent market formalization are not the only factors that can hinder traditional practices of norm-based land access. The traditional social security scheme is enforced through social norms which themselves depend on feelings of solidarity and trust within the community. The majority of interactions between households take place along lines of pre-existing social ties such as between family members, neighbors or members of the same ethnic group. People who share, for example, the same ethnic background face a reduced social distance on which a relationship of trust is facilitated (Zak and Knack, 2001). Corroborating this, evidence from US cities show that ethnic heterogeneity leads to poorer provision of public goods (Alesina et al., 1999). We therefore investigate whether high levels of ethnic heterogeneity affect the level of social security provided by village institutions through the level of trust between community members.

In respect to ethnicity, The Gambia is an interesting case since many Gambian villages have a long history of mixed ethnicities. These ethnic groups have lived side by side for many decades and grown the same crops under similar environmental conditions without any large-scale conflict (Dey, 1982; Eastman, 1990; Arcand and Jaimovich, 2014). Using the same dataset as this paper, Jaimovich (2011) found that that high ethnic diversity was related to a higher density of links and more clustering in the land network at the village level. Furthermore, members of minority ethnic groups were not found to receive less land through the network of transfers of secondary usage rights and shared ethnic group membership was not found to predict land transfers. However, these findings do not rule out that ethnicity impacts the strength of the social security mechanism. Instead, the results could simply be caused by higher levels of market-based transfers in ethnically diverse villages.

4.3 Method

This section outlines our empirical strategy. We begin by testing whether the observed patterns of land transfers are consistent with norm-based access rules to land. Next, we subject this baseline estimation to a series of robustness checks. Finally, we investigate whether the land transfer structure is consistent with a stronger effect of norm-based access rules in villages with low population densities and low levels of ethnic diversity.

4.3.1 Testing pro-poorness of land transfers

The empirical strategy is inspired by the method used in a number of previous studies to identify the effect of recipient income on gift-giving behavior (e.g., Cox, 1987; Cox and Jakubson, 1995; Kazianga, 2006; Mitrut and Nordblom, 2010). In that literature, a negative relationship between recipient income and gift giving is interpreted as evidence that lower income induces gift-giving behavior.

The current paper takes temporary transfers of land usage rights as the dependent variable. The land transfer is observed at the level of the dyad, i.e., the level of the link between two households within a single village. The dyad-level data allow us to extend the method of the gift literature in three ways. First, in contrast to many previous studies (Cox, 1987; Kazianga, 2006; Mitrut and Nordblom, 2010), we explicitly control for both donor and recipient characteristics in order to address potential omitted variable bias. Second, since all households have many potential partners that are observed, we estimate models that include either donor or recipient fixed effects. Third, we use an alternative method to correct for endogeneity of the income variable.

The baseline regression takes the following form:

$$A_{ij} = \alpha + \gamma_1 y_j + \gamma_2 y_i + \gamma_3 w_{ij} + \beta_1 z_j + \beta_2 z_i + \sum_k a_k + \epsilon_{ij} \tag{1}$$

A_{ij} is the amount of land that household i transfers to household j, measured in hectares. Thus, i denotes the donor (i.e., the sender of land), while j denotes the recipient. y_i and y_j are i and j's log-transformed monetary income per capita. The income variable measures cash-in-hand earned from off-farm employment and agricultural sales. Subsequent regressions split income into agricultural and non-agricultural income per capita. The vectors z_i and z_j are household-specific attributes. w_{ij} are link-specific characteristics, and a_k is a village fixed effect for village k.

If recipients with lower income are allocated more land, we expect to find $\gamma_1 < 0$, which would be consistent with the presence of norm-based access rules as it indicates that poorer households are allocated more land. If $\gamma_2 > 0$, land is more likely to be transferred from households with higher income.

To the extent that other variables are correlated with the actual amount of transacted land and with household income, it is essential to control for these. The baseline specification therefore includes a range of control variables. At the level of the household, we control for the share of households who share ethnicity with the household; an indicator variable equal to one if the household has no primary usage rights to land; the logarithm of the age of the household head; the logarithm of the number of working-age adults in the household; indicator variables for whether the household head is illiterate, has any formal education and whether it is female; and an indicator equal to one if the household receives remittances. We include these controls for both donors and recipients. At the level of the link, we include a control for whether households i and j are kin related, through either the household head, the wife(s) of the household head, or marriage ties. Finally, we expect that the value of land and the amount of arable land vary across villages and are likely to affect which transfers take place as well as the size of the inter-household transfer. Hence, village fixed effects are included in all estimations.

Due to a large number of zero observations, equation (1) is estimated using a tobit regression. We check that our results are not driven by the specific assumptions of the tobit model in Appendix A. Residuals from dyadic observations involving the same individual i are likely to be correlated. Specifically, we need to account for the possibility that when i donates land to j, i is less likely to donate land to some other household s. Likewise, if j receives land from i, j may be less likely to also receive land from s. To allow for this sort of interdependence at the dyadic level, all standard errors are clustered at the village level. This approach is conservative in the sense that we do not assume anything about the dependency of dyadic observations inside the villages (Fafchamps and Söderbom, 2014).

The effect of income on land transactions need not be linear. It is possible that the underlying motive that guides land transfers depends on the level of income per capita of the recipient (Cox et al., 2004). For instance, the poorest households were found to be excluded from inter-household gift exchange in Burkina Faso (Kazianga, 2006). In our setting, we expect that norm-based land access rules, if they exist, insure only the poorest households. If this is the case, only recipients with the lowest incomes should be allocated land based on their income. Moreover, if villagers care about the general level of equity in the village, one would expect the richest households to transfer more land. To allow for this, we let the marginal effects of donor and recipient income vary between quartiles of the income distribution. We implement this with a spline regression approach (Greene, 2002, ch. 7). The regression takes the following form:

$$\begin{aligned} A_{ij} &= \alpha + \sum_{k=1}^{4} \left[(\gamma_1^k * \min(y_j, q^k) - q^{k-1}) * I(y_j \geq q^{k-1}) \right] \\ &+ \sum_{k=1}^{4} \left[(\gamma_2^k * \min(y_i, q^k) - q^{k-1}) * I(y_i \geq q^{k-1}) \right] \\ &+ \gamma_3 w_{ij} + \beta_1 z_j + \beta_2 z_i + \sum_k a_k + \epsilon_{ij} \end{aligned} \tag{2}$$

where k denotes the four bins, which are partitioned by the values q^1, q^2, and q^3, and we define $q^0 \equiv 0$ and $q^4 \equiv \infty$. $I(y_j \geq q^{k-1})$ is an indicator variable equal to one if $y_j \geq q^{k-1}$. In this specification, the marginal effect of e.g., receiver income for households whose income belongs to the lowest quarter is given by γ_1^1 whereas the marginal effect of receiver income for potential receivers whose income are above q^1 but below q^2 is given by γ_1^2. [7]

[7] The quartiles are constructed using the entire sample. One concern is that it is instead the village-specific quartiles that are relevant. Unfortunately, there is no straightforward way of estimating the model with village-level splines. Out of the 52 villages, there are only 6 or fewer villages that do not have any households in all four quartiles. Therefore, this is not likely to be a substantial problem.

4.3.2 Effects of population density and ethnic diversity

To test whether village characteristics influence inter-household transaction motives, we introduce interaction terms between village-specific characteristics v_k and the variables of interest. Two village-specific measures are considered. First, to investigate the impact of population density, we construct an indicator variable based on village density, which is equal to one if the village has a population density above that of the median village and zero otherwise.[8] Second, to investigate the impact of ethnic heterogeneity, we construct an indicator for ethnic diversity, which is equal to one if the level of ethnic heterogeneity, as measured by a Herfindahl fractionalization index, is above that of the median village, and zero otherwise.

In order to investigate how access rules for land differ depending on village-specific characteristics, demeaned donor and recipient incomes are interacted with the village-specific characteristics. The model can be written as

$$A_{ij} = \alpha + \gamma_1 y_j + \gamma_2 y_i + \gamma_3 w_{ij} + \gamma_4 \bar{y}_j * v_k + \gamma_5 \bar{y}_i * v_k + \beta_1 z_j + \beta_2 z_i + \sum_k a_k + \epsilon_{ij} \quad (3)$$

where $\bar{y}_i$ and $\bar{y}_j$ denote the demeaned incomes of households i and j.[9] The rest of the notation is the same as in Section 4.3.1. The uninteracted village characteristic is not included in the estimation as it is swept away by the village fixed effects. A joint test of $\gamma_1 + \gamma_4 < 0$ is a test of whether the norm-based access rule guides transfer motives in high-density (ethnically diverse) villages, i.e., poorer households receive more land, while a test of $\gamma_1 < 0$ is a test of whether the norm-based access rule guides transfer motives in low-density (ethnically homogeneous) villages.

4.4 Data and descriptive statistics

We use a unique dataset from The Gambia. The *Gambia Networks Data 2009* was collected in six out of eight Local Government areas between February and May 2009. A sample of 60 villages was randomly selected among villages with between 300 and 1000 inhabitants in the latest 2003 census. The data includes a standard household survey, which covers all households in the villages. The data also has information on pre-existing social links as well as land transactions between all households in the villages. In particular, the land network contains information about the size and direction of land transfers between all households residing in the village over an

[8] We opt to use a binary measure in order to avoid that results are driven by outliers that arise from interacting two continuous variables.

[9] The income variables are demeaned in order to keep the interpretation of γ_1 and γ_2 unchanged between estimation of (1) and (3) (Ozer-Balli and Sørensen, 2013).

agricultural year.[10] These transfers are temporary transfers of secondary usage rights as described in section 4.2.1.

Five villages were dropped due to substantial amounts of missing household-level information in these villages. Second, as we are concerned with land transactions in rural areas, three semi-urban villages were dropped.[11] This brings the final number of villages to 52. The average village has 521 inhabitants. This corresponds to a village population density of 235 inhabitants per square kilometer.[12] The economic conditions resemble that of other rural communities in West Africa: Very few households have access to electricity, and the majority of households do not have access to improved water sources. The sample also represents ethnic diversity of The Gambia: The largest ethnic group, the Mandinkas, comprises 55 percent of households in the sample and four other ethnic groups each comprise more than 5 percent each.[13] The villages span a range of different diversity levels. Using the Herfindahl fragmentation index, ethnic fragmentation inside villages ranges from 0 (completely homogeneous) to 0.84 (the maximum fragmentation index value is 1), with a mean of 0.28. The distribution of self-reported income measured in terms of the Gini coefficient is on average 0.31, with a maximum of 0.60.

Rural villages in The Gambia are organized in compounds, which correspond to a group of people who work jointly on common fields, eat together, and organize daily activities (von Braun and Webb, 1989; Pamela, 2010). Depending on the size of the compound, independent production units (*dababas*) can within a single compound. The *dababa* is used as the unit of analysis. If several households exist within one compound, links between these are observed. Around 16 percent of household heads are not the head of the compound in which they live.

Household-level descriptive statistics for all households, and separately for donors and recipients of land, can be found in Table 1.

The data is consistent with the description of rural Gambia. Of note is that 0.5 percent of households have more than 50 members which is explained by the polygamous nature of the rural Gambian society (49 percent of household heads have more than one wife). Households have an average of five adult working members. The households are predominately led by men with low levels of education; only 13 percent have any formal education. For the majority of households the main economic activity is related to agriculture (79 percent), through many households also engage in other income-generating activities. The average monetary income per capita is 2,931 Gambian Dalasis

[10] The data was collected using a structured group approach with a median household coverage rate in the villages of 94 percent. For detailed information on the sampling methodology and data description, see Jaimovich (2015).

[11] Table B.1 in the Appendix shows village-level descriptive statistics.

[12] The denominator is the sum of the cultivated area of the village and the village itself. As some areas are uncultivated, the population density of the country as a whole is lower.

[13] Compared to the 2003 Census for The Gambia, the Mandinka ethnicity is slightly overrepresented and the Serahule ethnicity is underrepresented (Arcand and Jaimovich, 2014).

Table 1: Descriptive statistics: Household level

	All households		By participation type (Mean)		
	Mean	Std. Dev.	Recipients	Donors	Non-participants
	N=2,029		N=818	N=516	N=860
Total income per cap.[1]	2.931	3.330	2.751	2.915	2.996
Agricultural income per cap.[1]	0.257	0.818	0.292	0.269	0.224
Other income per cap.[1]	2.674	3.268	2.460	2.646	2.772
Agricultural share of income	0.146	0.255	0.159	0.173	0.127
Receive remittances	0.458	0.498	0.455	0.529	0.437
Land owned (hec.)	10.034	22.907	8.348	17.118	8.285
Household size	13.358	8.952	14.49	15.434	11.750
Number of working adults	4.893	3.894	5.279	5.535	4.390
Work on other farms (days)	3.631	11.018	3.592	4.083	3.295
Emigrated household member	0.526	0.499	0.517	0.599	0.503
Age of head	52.679	16.13	53.243	55.122	51.274
Agricultural work	0.785	0.411	0.793	0.797	0.766
Non-agricultural	0.178	0.383	0.167	0.157	0.203
HH has family the village	0.946	0.226	0.932	0.977	0.944
Female headed household	0.047	0.211	0.040	0.035	0.057
Illiterate	0.444	0.497	0.458	0.409	0.442
Formal education	0.125	0.331	0.104	0.132	0.136
Monogamous	0.458	0.498	0.438	0.428	0.483
Polygamous	0.489	0.500	0.521	0.527	0.452
Ethnicity: Mandinka	0.548	0.498	0.553	0.541	0.564
Ethnicity: Fula	0.171	0.377	0.181	0.140	0.177
Ethnicity: Wollof	0.097	0.296	0.084	0.130	0.088
Ethnicity: Jola	0.079	0.270	0.075	0.085	0.074
Ethnicity: Sererr	0.060	0.238	0.048	0.079	0.056

Note: 165 households participate on both sides of the market.
1: 1.000 GMD's
Source: Authors' calculations.

a year, PPP-equivalent to 301 USD a year, of which around 15 percent stems from agricultural activities.[14] Considering that The Gambia is a big exporter of groundnuts, cashews and other cash crops, the agricultural income share is low. This may partly be caused by the sampling of the villages, which rules out larger villages. Another contributing factor is that households occasionally use cash crops for barter.[15]

On average, donors have higher incomes than recipients. This serves to motivate the search for econometric evidence of traditional land-access norms. Recipient households tend to be smaller and more likely to be female headed. While land recipients and non-recipients have similar amounts of land with primary rights (around 8 hectares per

[14] Using Penn World Tables PPP-adjusted exchange rate. Note that consumption and bartering of own production is not included in this figure.

[15] We thank Dany Jaimovich who participated in the collection of the Gambia Networks Data 2009 for this insight.

household on average), land donors have substantially larger primary rights landholdings (around 17 hectares per household on average). Perhaps surprisingly, household heads of households that donate land are more likely to work on other farms than land recipients and non-participants in the land market. One might expect that the poorer recipient households would be more likely to work on other households' farms in order to earn some additional income. However, most labor transactions in this setting are not proper jobs but rather short-term arrangements made in order to cope with the seasonal spikes in labor demand, mostly before or during the rainy season or in response to idiosyncratic shocks to for example health. This is reflected in the low levels of work carried out on farms that belong to other households. Households who work for other households can expect to receive some token goods or labor in return, but rarely a monetary payment (Swindell, 1987; Jaimovich, 2015). It is therefore not obvious that lower-income households would participate more in this kind of work. Nearly half of the households receive remittances, which are an important source of cash (just over 20 percent of the sample report that over half of their monetary income comes from remittances). This reflects the rapid rate of urbanization of The Gambia where many rural households have family members living in the urban areas; some households also have family members living abroad.

4.4.1 Inter-household land transfers

The data contains information on all temporary transfers of secondary usage rights that took within the last year. We note two limitations. First, the data does not include information about eventual contracts that form the basis of land transactions. Qualitative field observations carried out in relation to data collection of the *Gambia Networks Data 2009*, as well as existing studies conducted in The Gambia, support that most transactions are not market transactions, as the donor receives no payment for the land he transfers out. Land transactions are however sometimes reciprocated through small symbolic payments in kola nuts, labor services, or cash (Eastman, 1990; Freudenberger, 2000; Jaimovich, 2011). The observed land transfers are not a result of sharecropping arrangements, which are not common in The Gambia. Dey (1982) outlines how early sharecropping arrangements imposed by donor-implemented programs were unsuccessful and subsequently abandoned. If sharecropping was widespread, one would expect to find large labor transactions in the labor network, which is not the case. On average, households spend 3.6 days per year working on other farms.

Secondly, the data does not contain information on households located outside the village. It is therefore necessary to assume that the village is the natural domain for norm-based land transfers of secondary usage rights. This is a natural assumption, as the village community is the unit of social security. This assumption is supported by the fact that external actors do not play a major role in the land network: Around 6 percent of the households in the sample either receive land from or transfer land to households outside the village. The relatively low level of land transactions involving households residing outside the village is likely to be explained by the immobility of land.

Table 2 reports descriptive statistics of the land network. In each village, an average of one fifth of households transfers out land. The average donor allocates land to more than one household. The average amount of land transferred per land transaction is almost 2 hectares. These sizable transfers stand in contrast to the sparse transfers of labor described earlier. In total, donors allocate an average of 33 percent of their initial land holdings (disregarding households that participate on both sides of the market).

The descriptive statistics of Table 1 indicated that land is transacted from land-abundant households to land-poor households. This observation is confirmed by Figure 2, which shows changes in land distribution before and after transfers have taken place. The level of land inequality is estimated using a Theil index, which takes values between zero (total equality) and one (total inequality). Each bar corresponds to a village, and negative values indicate a decrease in land inequality due to land transfers. In most villages, land transfers serve to make land usage rights less unequal than land endowments. This is consistent with norm-based access rules and serves to further motivate the regression-based analysis of the following section.

4.5 Results

4.5.1 Baseline results

Table 3 reports the results of (1), i.e., the effect of recipient and donor income per capita on the land network. Column 1 shows results without control variables. The income of land recipients affects land transfers negatively and significantly, as expected in the presence of norm-based access rules. When including control variables, the income

Table 2: Descriptive statistics: transfers

	Mean	Std.	Min	Max	Obs.
No. of households[1]	39.02	14.45	12	71	52
No. of donors[1]	9.29	4.48	0	24	52
No. of donors, pct. of households[1]	24.18%	9.42%	0	48.98%	52
No. of recipients per donor	1.78	1.40	1	14	483
No. of recipients, pct. of households[1]	32.98	13.57%	0	80.77%	52
Size of the land transactions (hec.)	1.83	1.50	0.25	20	859
Size of the land transactions (% of donors' initial landholding)[2]	33.40%	26.48%	0.001%	100%	353

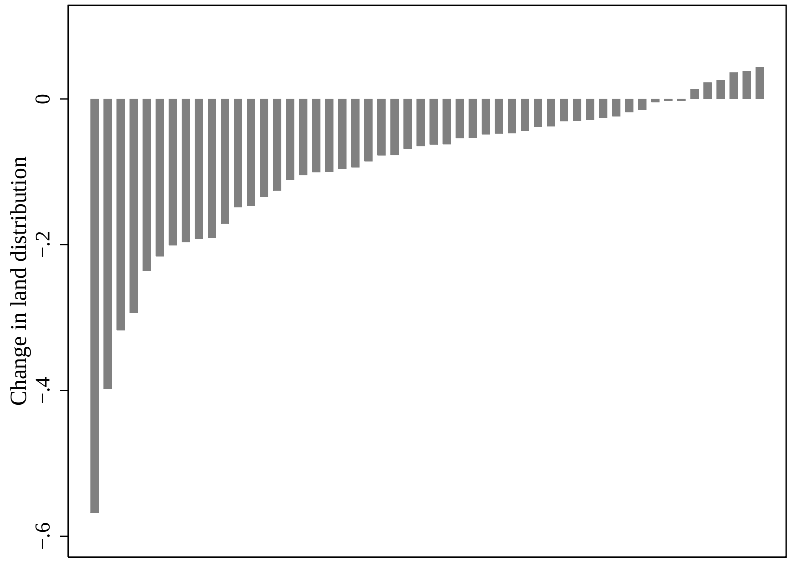
Note: Total number of land transactions is 859. Total number of recipients and donors is 657 and 483, respectively.
1: Simple mean of village levels.
2: Excludes observations where the donor participates on both sides of the market.
Source: Authors' calculations.

Table 3: Pro-poorness of land transactions

	(1)		(2)		(3)		(4)	
j's total income	-0.379**	(0.162)	-0.351*	(0.198)				
i's total income	0.051	(0.317)	0.441*	(0.257)				
j's non-agric. income					-0.439**	(0.196)		
- 1st quartile							-0.748*	(0.387)
- 2nd quartile							0.233	(0.860)
- 3rd quartile							-0.772	(1.021)
- 4th quartile							-0.442	(0.488)
i's non-agric. income					0.478*	(0.264)		
- 1st quartile							0.662	(0.851)
- 2nd quartile							0.067	(1.137)
- 3rd quartile							-0.675	(1.536)
- 4th quartile							1.024**	(0.455)
j's agric. income					0.683**	(0.278)		
- 3rd quartile							1.539	(0.976)
- 4th quartile							0.378	(0.392)
i's agric. income					0.267	(0.338)		
- 3rd quartile							0.052	(1.114)
- 4th quartile							0.280	(0.491)
Link controls:								
Kinship tie			1.550***	(0.263)	1.549***	(0.261)	1.547***	(0.260)
Recipient controls (j):								
Share of same ethnicity			-1.228***	(0.295)	-1.242***	(0.299)	-1.224***	(0.293)
No land (=1)			1.438***	(0.312)	1.496***	(0.313)	1.511***	(0.314)
Age of head (log)			-0.004	(0.005)	-0.004	(0.005)	-0.003	(0.005)
Adult labor (log)			0.033	(0.020)	0.028	(0.019)	0.027	(0.019)
Illiterate (=1)			-0.085	(0.134)	-0.084	(0.134)	-0.078	(0.135)
Formal school (=1)			-0.528**	(0.229)	-0.510**	(0.227)	-0.497**	(0.230)
Female head (=1)			-0.584	(0.362)	-0.559	(0.360)	-0.542	(0.360)
Receives remittances (=1)			0.206	(0.189)	0.216	(0.184)	0.213	(0.183)
Donor controls (i):								
Share of same ethnicity			1.838**	(0.788)	1.841**	(0.786)	1.834**	(0.796)
No land (=1)			-4.219***	(0.510)	-4.239***	(0.506)	-4.217***	(0.507)
Age of head (log)			0.021***	(0.006)	0.021***	(0.006)	0.020***	(0.006)
Adult labor (log)			0.067**	(0.027)	0.069**	(0.028)	0.064**	(0.028)
Illiterate (=1)			-0.583**	(0.243)	-0.575**	(0.241)	-0.579**	(0.245)
Formal school (=1)			0.040	(0.327)	0.035	(0.334)	0.033	(0.339)
Female head (=1)			-0.176	(0.726)	-0.193	(0.715)	-0.195	(0.722)
Receives remittances (=1)			0.107	(0.231)	0.105	(0.231)	0.110	(0.232)
Observations			87,788		87,788		87,788	

Note: Estimation is done using a tobit model. Coefficients reported. Dependent variable: Amount of land j is allocated from i in hectares. All regressions includes village fixed effects. Standard errors reported in parentheses are clustered at the village level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.
Source: Authors' calculations.

Figure 2: Change in land equity by village



Note: Each bar denotes the change in the village-level Theil index of land usage rights due to land transfers. Negative values represent a fall in inequality
Source: Authors' calculations.

of land recipients continue to affect land transfers negatively while the income of senders of land affect land transfers positively (column 2). This effect is significant at the 10 percent level. Conditional on receiving a transfer, the average partial effect of an increase in recipient income per capita of 100 percent of household *j* (corresponding to 0.85 of a standard deviation at the mean) results in a decrease in the amount of land transferred by around 0.03 hectares, or 300 square meters of land. This effect is calculated conditional on the existence of a transfer. The effect is economically meaningful, especially considering that recipients are those households with little income. The average partial effect without conditioning on the existence of a transfer is a decrease in the latent variable of a land transfer from *i* to *j* of around 37 square meters, when per capita income of *j* increases by 100 percent. This means that household *j* loses 37 square meters of the latent transfer from each other household in the village. The average number of households in each village is 39. Donor income positively affects the probability of sending land, which is also consistent with the presence of a norm-based access rule. This shows that recipient characteristics, which have been the focus of most of the inter-household gift-giving literature due to data limitations, are not the only important characteristics.

When splitting the income variable of the donor and recipient into agricultural and non-agricultural income (column 3), it becomes clear that the negative effect of recipient income is driven by variation in non-agricultural income, which is significant and negative at the 5 percent level. Recipient agricultural income, on the other hand, is positive and significant. This is likely caused by the fact that agricultural income picks up market-based transactions of cash-crop farmers: Households that acquire income from cash crops conduct more commercialized agriculture, and it makes sense that in order to conduct this type of farming, these households must acquire access to land, although they are not eligible for norm-based transfers. Non-agricultural donor income is still positive and significant at the 10 pct. level.

Of particular interest is whether there is a stronger effect for the poorest households. The estimation results using a spline regression are shown in column 4.[16] On the recipient side, only the lowest non-agricultural income quartile is negative and statistically significant (at the 10 percent level). On the donor side, only the highest non-agricultural income quartile is positive and statistically significant. These results are consistent with the existence of a norms-based access rule: Land is redistributed from the richest to the poorest households. In fact, only within the poorest quarter of households does the income level affect in-transfers of land and only for the richest quarter of households does the income level affect out-transfers. There are no significant effects for any quartile of agricultural income. This may simply be caused by many of the poorest households having zero agricultural income.

Turning to the control variables reported in Table 3, coefficient estimates are consistent across specifications. In line with the results of Jaimovich (2011), kin-related households are more likely to transact land. Recipients with formal schooling are less likely to receive land. It is possible that educated household heads have more alternative income options and are therefore less likely to be allocated land. The share of households in the village with the same ethnicity as the recipient is negatively correlated with receiving land. This suggests that village minorities are allocated more secondary usage rights, perhaps because they are less likely to possess primary land rights. Moreover, donor households are typically older, larger in terms of the number of working adults, and better educated.

A control variable of particular interest is the indicator for whether the household has no primary usage rights. Landless households are more likely to be allocated land and less likely to donate land. This result has potential policy implications, as reform-based redistribution of land to landless households will likely crowd out voluntary land transactions. To investigate whether this result is applicable to households with little land, we include the initial landholding of both the donor and recipient. The estimation results are shown in Table 4.

The estimate on donors' initial landholding is positive and statistically significant, suggesting that households with more primary usage rights donate more land. However, the coefficient on the recipients' amount of initial land is insignificant. What seems to be important is whether the household has any land rights at all.

[16] A substantial share of households have zero agricultural income. Therefore, only the third and fourth income quartiles are included.

Table 4: Baseline results: Initial land holdings

	(1)		(2)		(3)	
j's non-agricultural income	-0.439**	(0.196)	-0.301*	(0.177)	-0.435**	(0.197)
i's non-agricultural income	0.478*	(0.264)	0.295	(0.294)	0.404	(0.278)
j's agricultural income	0.683**	(0.278)	0.569**	(0.281)	0.661**	(0.279)
i's agricultural income	0.267	(0.338)	0.284	(0.359)	0.259	(0.343)
No land (=1): j	1.496***	(0.313)			1.410***	(0.299)
No land (=1): i	-4.239***	(0.506)			-3.764***	(0.452)
Land endowment: j			-0.014	(0.012)	-0.008	(0.009)
Land endowment: i			0.026***	(0.006)	0.022***	(0.006)
Observations	87,788		87,788		87,788	

Note: Tobit. Dependent variable: Amount of land j is allocated from i. All regressions include the same control variables as Table 3 and village fixed effects. Standard errors are clustered on the village level. Land endowment is measured in hectares. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

In combination with the evidence of norm-based land transfers on the income measure, the results suggest that future land-redistribution reforms in West Africa should take into account potential crowding-out effects of the social security mechanism: Land redistribution reforms may not have as big an impact as one may expect if welfare effects are offset by informal norm-based land allocations.

4.5.2 Robustness of main findings

This section investigates the robustness of our results to concerns relating to the main specification, namely the potential confounding influence of unobserved household-specific factors; the possibility that our income measure is endogenous; the use of income as a measure of welfare; and the validity of the assumptions that the tobit model relies on.

Unobserved household heterogeneity

Even when including control variables, there may still be unobserved household-level characteristics affecting behavior. In the case of land transfers, a salient issue is that some households may be better farmers. Since unobserved household characteristics such as farming ability are likely to be both positively correlated with observed income and the size of land transactions, omitting farmers' ability is likely to bias our estimates upwards. This would result in a less negative coefficient estimate. Finding an effect in support of the norm-based access rule ($\gamma_1 < 0$) therefore limits the concern related to unobserved heterogeneity.

Nevertheless, we address the issue of unobserved heterogeneity by estimating a version of model (1) with household fixed effects. Using this approach, the effect of recipient (donor) income can be identified when donor (recipient) fixed effects are included. The models are:

$$A_{ij} = \alpha + \delta_1 y_i + \delta_2 w_{ij} + \delta_3 z_j + a_t + \epsilon_{ij} \quad (4)$$

$$A_{ij} = \alpha + \delta_1 y_j + \delta_2 w_{ij} + \delta_3 z_i + a_t + \epsilon_{ij} \quad (5)$$

Model (4) estimates the effect of donor income with recipient fixed effects and model (5) estimates the effect of recipient income with donor fixed effects.

Table 5 reports the estimation of the baseline model with either donor or recipient fixed effects, estimated by OLS. The results on donor and recipient income are consistent in sign with the results reported in Table 3. Moreover, the results are at least as significant as, and often more significant than, the main results. In sum, it does not appear that the presence of unobservable characteristics is driving our findings.

Endogeneity of income

It is possible that households strategically lower their monetary income, e.g., by consuming more of the household's own production instead of selling it, or through migration decisions, in order to increase the likelihood of receiving land. Furthermore, it is possible that there is reverse causality if the allocated land is used to produce marketable output, a land transfer can in itself affect income.

Several approaches to handle these endogeneity issues have been suggested in the literature on gift giving, which face similar issues. Kazianga (2006) use a pre-transfer income variable and use long-run rainfall data under the assumption that rainfall affects farm outcomes. Mitrut and Nordblom (2010) do not explicitly correct for endogeneity of income in their study of Romanian gift transfers, but instead check the quality of the income data using household consumption instead. We address both issues of strategic

Table 5: Land transaction results including fixed effects

	h = Donor		h = Recipient	
	(1)	(2)	(3)	(4)
Total income$_h$	0.005		-0.006***	
	(0.003)		(0.002)	
Non-agri. income$_h$		0.006*		-0.007***
		(0.004)		(0.002)
Agri. income$_h$		-0.002		0.005
		(0.004)		(0.003)
Fixed effects	Recipient FE's	Recipient FE's	Donor FE's	Donor FE's
Controls	Link + donor char's	Link + donor char's	Link + recipient char's	Link + recipient char's
Observations	87,788	87,788	87,788	87,788

Note: OLS. Standard errors reported in parentheses are clustered at the village level. Dependent variable: Amount of land j is allocated from i in hectares. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively. Link-varying controls are included in all regressions.
Source: Authors' calculations.

behavior and reverse causality by estimating a measure of the pre-transfer income for all households. This is possible due to the completeness of the transfer network data for all households in the sample.

Using household-level data (2,029 households), a prediction of observed income per capita as a function of land available to the household, i.e., the combined amount of land that the household holds primary and secondary usage rights over, is estimated:

$$y_i = \alpha + \lambda_1 \log(A_i) + \beta z_i + \epsilon_i \tag{6}$$

where y_i is the natural logarithm of observed income per capita for household i, and $\log(A_i)$ is the natural logarithm of land available to household i. Also included is a set of control variables collected in vector z_i.[17] Denote the predicted realized log of income by $\hat{y}_i$. Using the estimated parameters and the amount of land available to the household before transfers take place, A_i^{exante}, it is possible to compute the counterfactual income level in the absence of any transfers as $\hat{y}_i^{exante} = \hat{\alpha} + \hat{\lambda}_1 ln(A_i^{exante}) + \hat{\beta} z_i$. Therefore, $\Delta = (e^{\hat{y}_i} - e^{\hat{y}_i^{exante}})$ is a measure of how much income changes due to inter-household land transfers. This measure can now be used to obtain an estimate of potential income in the absence of land transfers: $\hat{y}_i^{potential} = log(e^{y_i} - \Delta)$. This approach hinges on the assumption that transferred and non-transferred land are of the same quality. If there is a quality difference, it is likely that households choose to transfer out their lowest quality land, keeping the best land for their own production. In this case, Δ will overestimate the change in income due to the land exchanged. Therefore, if there is no substantial difference between observed and potential income even in the presence of this overestimation, it severely limits the concern regarding endogeneity of income.

However, the effect of land usage rights on income can go both ways. On the one hand, access to land can increase cash crop production. This can increase agricultural income. This is what leads to concerns over reverse causality. On the other hand, access to more land can mean that farmers substitute away from off-farm labor, reducing non-agricultural income in the process. In order to capture both effects, we estimate equation (6) separately for non-agricultural and agricultural income.

Results of (6) are reported in Table 6. We find that land endowments affect agricultural income positively. However, the economic impacts are limited: Figure 3 shows that there is little difference in the distribution of potential and observed income per capita for households participating in the land market. This implies that observed monetary income is not highly endogenous to land transfers. This likely reflects that while transfers of land can affect the *welfare* of recipient households, the impact on *monetary*

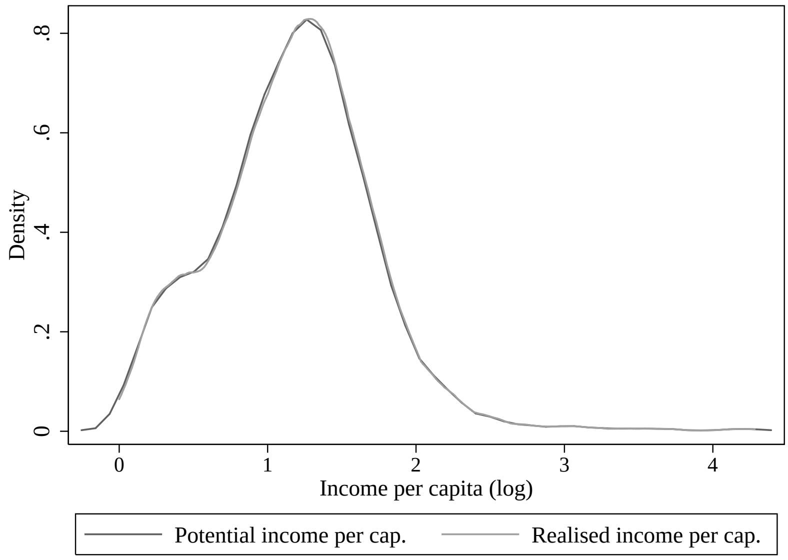
[17] The control variables include all the household controls included in estimation of model (1) as well as household size; an indicator variable for whether the household head is married to a single wife, or is married to multiple wives; an indicator variable for whether the household owns any cattle; an indicator variable for whether the household owns land of relatively high fertility (self-reported); and indicator variables for different ethnicities.

Table 6: Household-level income regression

	(1) Non-agricultural per cap. income (log)		(2) Agricultural per cap. income (log)	
Cultivated land (log)	0.019	(0.012)	0.017**	(0.007)
Observations	2,029		2,029	
Control variables	YES		YES	
R-squared	0.358		0.243	

Note: All regressions include household-level control variables similar to the subsequent dyad-regression as well as village fixed effects. Standard errors reported in parentheses are clustered at the village level. ***, **, and * indicate significance at the 1 percent, 5 percent and 10 percent level.
Source: Authors' calculations.

Figure 3: Realized and potential income



Note: The distribution is shown for the 1,140 households that actively participate in the land market and therefore experience a change in the amount of land cultivated due to land transactions.
Source: Authors' calculations.

income is limited. We are therefore not concerned with the use of observed income per capita instead of the measure of potential income in the main specifications of the paper. An advantage to using observed income is that this does not make any assumptions regarding the average quality of exchanged versus non-exchanged land. Unsurprisingly, given the small differences between realized and potential income, all results are robust to using the potential non-agricultural and agricultural income measures instead (results not shown).

Alternative welfare measure

One may worry that monetary income is a poor measure of welfare. Specifically, monetary income does not take account of consumption of own production and barter as well as asset ownership and wealth. This is potentially worrying, as welfare is the underlying concept that is believed to drive norm-based transfers of land usage rights. In addition, there are two other issues with the monetary income measure. First, self-reported income may be measured with error. In addition, monetary income is likely misreported when households have several income sources (World Bank, 2007). If measurement errors are random, this leads to attenuation bias, and results are lower bounds on the true absolute effect size. Another issue is that respondents may be unwilling to reveal their true income, especially considering that the data we are using is collected using a structured group approach. If this is the case, it is likely that households in the village do not know each other's true income. In this case, the income measure of the dataset may be a better approximation to the relevant decision variable than the true income.

To test the sensitivity of our results to the choice of welfare measure, we construct a wealth index and test the robustness of our results using this instead of the monetary income measure. The index is constructed using principle component analysis and is based on three variables: the relative self-reported wealth quartile of the household; the number of corrugated steel huts owned by the household; and the number of cattle for which the household pays taxes. This index has the additional benefit that, unlike monetary income, the number of huts and cattle of each household are visible to everyone in the small and geographically compact villages we consider.

The correlation coefficient between the income variable and our wealth index is 0.60, which is quite high. This may in itself alleviate some of the concerns regarding the use of monetary income in the main specifications. Table 7 reports estimation results using the wealth index rather than the monetary income variable. Results are consistent with the main results, and significance levels are even higher using the wealth index. In accordance with the results of Table 3, we find that less wealthy households receive more land in line with presence of norm-based transfers. The estimate is now statistically significant at the 1 percent level. Including controls in column 2 confirm these results: More wealthy households donate more land and less wealthy household receive more land. Based on this, we conclude that our main results are not driven by the use of the monetary income as the main independent variable.

Assumptions of the tobit model

We check that the main findings are not driven by the assumptions of the tobit model, namely the assumptions of normality and heteroscedasticity of the latent variable as well as the assumption that the partial effect of an explanatory variable x on $p(A_{ij}|x)$

Table 7: Pro-poorness of land transfers using a wealth index

	(1)		(2)	
j's wealth	-0.601***	(0.170)	-0.261**	(0.127)
i's wealth	0.073	(0.215)	0.525**	(0.225)
Link characteristics:				
Kinship tie			1.545***	(0.267)
Recipient characteristics (j):				
Share of same ethnicity			-1.278***	(0.292)
No land			1.406***	(0.307)
Age of head (log)			-0.003	(0.005)
Adult labor (log)			0.032	(0.021)
Illiterate			-0.074	(0.137)
Formal school dummy			-0.509**	(0.234)
Female head			-0.532	(0.363)
Receive remittances			0.238	(0.185)
Donor characteristics (i):				
Share of same ethnicity			1.845**	(0.799)
No land			-4.164***	(0.513)
Age of head (log)			0.019***	(0.005)
Adult labor (log)			0.079***	(0.029)
Illiterate			-0.580**	(0.241)
Formal school dummy			0.064	(0.330)
Female head			-0.183	(0.735)
Receive remittances			0.037	(0.240)
Observations	87,294	87,294	87,294	87,294

Note: Tobit. Coefficients reported. Dependent variable: Amount of land j is allocated from i. All regressions includes village fixed effects. Standard errors reported in parentheses are clustered at the village level. The number of observations is sligthly lower than in Table 3 due to missing wealth indicators for a few households. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.
Source: Authors' calculations.

and on $p(A_{ij}|x, A_{ij} > 0)$ are of the same sign. The results can be found in Appendix A. Results are encouraging: We find that tests of normality and heteroscedasticity of the latent errors cannot be rejected; we find qualitatively similar results using an OLS regression which does not require normality and heteroscedasticity for consistency; and we find that two-step hurdle-models give qualitatively similar results for both $p(A_{ij}|x)$ and $p(A_{ij}|x, A_{ij} > 0)$.

4.5.3 The strength of the social security norm

In order to investigate whether some villages exhibit stronger signs of the norm-based access rule we estimate Model (3). From the description in Section 4.2.2, we expect land transfers in villages with high population densities and villages with high levels of ethnic diversity to be less affected by norm-based access rules. We investigate these two possibilities one by one. It is possible that ethnically diverse and densely populated villages share the common characteristic of being more welcoming towards newcomers, and that this could drive findings on both population density and ethnic diversity. This does not appear to be the case: A cross-tabulation of the two village-level indicators for having above-median population density and above-median ethnic diversity puts 13 to 15 villages in each of the four resulting cells.

Table 8 reports the estimation results of equation (3), which includes interactions between the income variables and village-level characteristics of population density and ethnic diversity. Allowing for varying effects between villages with high and low population densities, an interesting finding emerges (column 1a). The interaction terms of both donor and recipient income are insignificant, and point estimates are of opposite signs compared to the non-interacted income variables. The joint test of the sum of the interaction and the income variable yields an insignificant estimate. This indicates that transactions are not driven by norm-based access rules in villages with high population densities. In fact, there is no evidence of norm-based access rules in these villages. The leftover main effects of recipient and donor incomes account for the effects of income on land transfers in villages with relatively low population densities. These have the expected signs in the presence of norm-based access rules to land. These results are in line with what we expect under the hypothesis that norm-based access rules to land are stronger in low-density villages.

When splitting villages by the level of ethnic diversity, the same result is found for recipient income, which is negative for low-diversity villages and insignificant for high-diversity villages. This is indicative of norm-based access rules being more important in less diverse villages. The positive effect of donor income is now entirely captured by the interaction term and is not significant for low-diversity villages. While this is not necessarily evidence against the functioning of norm-based access rules in low-diversity villages – evidence of this effect is primarily based on the recipients being poor rather than on the donors being rich – it is somewhat puzzling and stands in contrast to the results on population density. A closer investigation reveals that the effect is driven by a few households with very high incomes that are also located in high-ethnicity villages. When we restrict the sample to links where the donor has an income equal to or less than the 99th percentile of income, the result disappears.

When the income variable is split into agricultural and non-agricultural components, another interesting finding emerges (columns 2a and 2b). The effect of recipient non-agricultural income is again only significant and negative in the low-density villages. In the comparable baseline estimation, there was a significant and positive effect from the recipient agricultural income. We suggested that this could be caused by these households' need for land to conduct commercialized agriculture, even though they are not eligible for norm-based land donations. This effect is entirely picked up by the high-density villages, which is where we would expect commercialized agriculture to be present. Turning to ethnic diversity, the negative effect on non-agricultural recipient income is again only present in low-diversity villages, consistent with the theory that

Table 8: Village characteristics and pro-poorness of land transfers

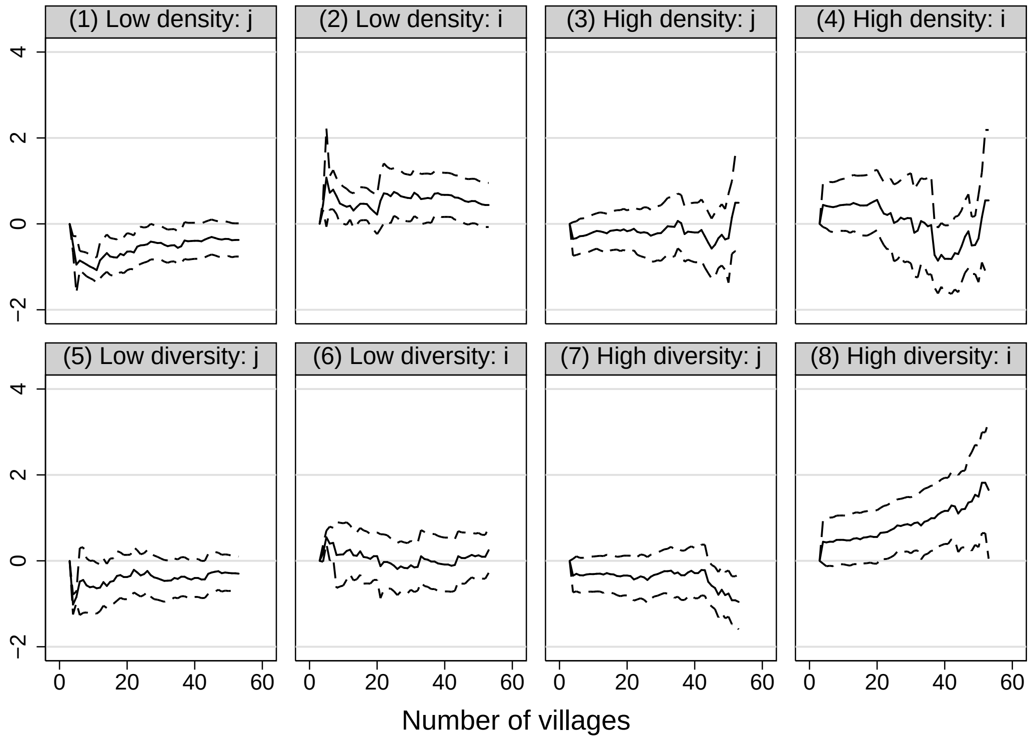
Village characteristic	None (1)	Density (1a)	Diversity (1b)	None (2)	Density (2a)	Diversity (2b)
j's total income	-0.351*	-0.520***	-0.466*			
	(0.198)	(0.197)	(0.248)			
i's total income	0.441*	0.673**	-0.151			
	(0.257)	(0.276)	(0.293)			
Characteristic is high * j's income		0.465	0.226			
		(0.363)	(0.323)			
Characteristic is high* i's income		-0.829	**1.030****			
		(0.648)	(0.442)			
j's non-agricultural income				-0.439**	-0.580***	-0.486*
				(0.196)	(0.205)	(0.250)
i's non-agricultural income				0.478*	0.656**	-0.086
				(0.264)	(0.281)	(0.330)
j's agricultural income				0.683**	0.244	-0.101
				(0.278)	(0.347)	(0.524)
i's agricultural income				0.267	0.521	0.033
				(0.338)	(0.392)	(0.578)
Characteristic is high* j's non-agri. income					0.375	0.074
					(0.340)	(0.317)
Characteristic is high* i's non-agri. income					-0.685	**0.934****
					(0.648)	(0.454)
Characteristic is high* j's agri. income					**1.131****	**1.131***
					(0.525)	(0.598)
Characteristic is high* i's agri. income					-0.819	0.338
					(0.757)	(0.701)
Observations	87,788	87,788	87,788	87,788	87,788	87,788

Note: Tobit regression. Dependent variable is amount of land j is allocated from i. All regressions include the same control variables as Table 3 and village fixed effects. Standard errors are clustered at the village level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively. Coefficients in bold refer to cases where the interaction and main effect are jointly significant at the 5 percent level. Source: Authors' calculations.

low-diversity villages adhere more to norm-based access rules. The effect is, however, only significant at the 10 percent level. The positive and significant effect of donor income is driven by differences in non-agricultural income. Again, restricting the sample to donors whose incomes are equal to or less than the 99th percentile removes this result. As with population density, the positive effect on recipient agricultural income also found in the baseline specification is picked up by high-diversity villages.

The indicator variables marking high population density and ethnic diversity use the median density and diversity of villages as the cutoff point for being categorized as a village with either high or low density/diversity. Figure 4 shows parameter estimates using alternative cutoffs for the indicator variables. The topmost row shows results on population density while the lower row uses ethnic diversity. The two first columns shows the main effects for donors and recipients, while columns 3 and 4

Figure 4: Alternative cutoffs on village characteristics: Land transactions



Note: *i* and *j* denote donors and receivers of land, respectively. Dotted lines give 95 percent confidence intervals. The x-axis denotes the number of villages not included in the interaction, ranked by either density or ethnic diversity. In order to avoid spurious results, results are only shown for cut-offs where there are at least four villages in both main and interaction term. Source: Authors' calculations.

report the joint test of interaction plus main effect, i.e., the estimate of the effect in more population-dense and more ethnically diverse villages. Moving to the right in each diagram corresponds to increasing the number of villages included in the main effect. The effects of donor and recipient income in low-density villages are significant and of the same sign as in the baseline model for most alternative values of the cutoff (subfigure 1 and 2). There is no effect for donors and recipients across most cutoffs in more population-dense village, which is consistent with Table 8 (subfigure 3 and 4). Turning to the results on ethnic diversity, the main result of a significant estimate on recipient income (subfigure 5) is not consistently negative or significant across cutoffs. It is, however, borderline significant regardless of the cutoff. Low ethnic diversity does not appear to motivate donors of land further (subfigure 6). The effect is significantly negative when only few villages are covered by the main effect. However, this appears to be driven by a single outlying village, corresponding to the large negative spike in the figure. The highly significant and positive estimate of donor income in high diversity villages is robust – and the effect becomes increasingly large when only the most ethnic diverse villages are included (subfigure 8).

It is also possible to include recipient or donor fixed effects and estimate the effects if population density and ethnic diversity using OLS. Results are reported in Table B.2 in the appendix. Our main conclusions from Table 8 are unchanged. When we include recipient fixed effects, we still find a positive effect of donor income in low-density villages. Splitting income into agricultural and non-agricultural income, we again find a positive effect of non-agricultural income in low-density villages. Turning to the results for land-receiving households, we again find that recipients with lower incomes are allocated more land in less population-dense villages and in villages characterized by lower ethnic diversity.

To conclude, the results on population density are highly consistent across different cutoffs while the results on ethnic diversity are not as strong as they are only borderline significant for some alternative cutoffs. Inclusion of household fixed effects did not change the conclusion of the main findings.

4.6 Conclusion

It is widely believed that traditional tenure systems and local social security systems are disappearing as solidarity feelings and norm-based insurance mechanisms are undermined by increasing population pressure and market integration. This paper looks for evidence of still-existing norm-based access rules to land in rural areas of The Gambia. We find evidence that supports the hypothesis that inter-household land transfers are at least partially driven by social security considerations as poor and landless households are allocated more land through inter-household transfers of temporary usage rights to land. We further find that the norm-based access rule only functions for the poorest households, more specifically those in the lowest income quartile. Donors are more likely to be rich households, i.e., those in the uppermost income quartile. These findings are robust to a series of robustness checks and alternative specifications.

Taking advantage of a complete network-level dataset, we are able to deal with three issues which are often not addressed in studies on inter-household transfers: (i) we deal with omitted variable bias arising from omission of donor characteristics by including both donor and recipient characteristics at the same time, and find that donor characteristics are also important; (ii) we check the extent of endogeneity of income to land transactions; and (iii) we correct for unobserved household heterogeneity by including donor and recipient fixed effects.

We further find that the effect of income on transfers of land depend on village characteristics. Land does not flow towards relatively poorer households in villages characterized by higher population densities. We also find relatively weaker evidence that higher levels of ethnic diversity have similar effects on land transfers. The former result is consistent with what we would expect to see under the hypothesis that land scarcity undermines traditional access rules to land (Platteau, 2002). As to the latter result, despite the history of The Gambia, which has experienced a long period of relatively peaceful co-existence of different ethnic groups, it seems that ethnic heterogeneity still has the potential to undermine community-level social security systems, although we note that this result is not as consistent across different specifications as results of the effect of population density.

Other explanations can rationalize the observed patterns in the land network. One possibility is that donors are altruistic, i.e., they care about the well-being of others, which induces them to donate land usage rights to the poor members of the village. This is virtually impossible to distinguish empirically from the presence of a norm-based access rule. Indeed, a norms-based access rule is likely to be at least partly upheld by altruism. It is also very possible that altruistic norms, like norm-based access rules, can be threatened by increases in population densities and ethnic diversity. We do therefore not attempt to differentiate between these two tightly linked explanations. Another possibility is that households donate land usage rights as an insurance policy. Donating land may allow households to call upon favors in times of need, and the data does reveal that donors of land are more likely to be receivers of labor. This is another case that is hard to distinguish from a system of norm-based access rules, in part because such informal insurance mechanisms may be part of the framework that supports the traditional system of norm-based access rules.

Finally, land transfers can provide efficiency gains. If a donor household has too much land to cultivate using its own labor, it can temporarily transfer land usage rights to landless households in order to increase allocative efficiency (Freudenberger, 2000). Using the same dataset, the effect of land transactions on allocative efficiency was explored in Chapter 3. The results of Chapter 3 indicate that land transactions increase efficiency. If households with low land–labor ratios are also poorer, this could explain the present finding. This, however, is not a strong finding in the data: The household land–labor ratio and income are not significantly correlated at the ten percent level when running a simple regression of income on the land–labor ratio. We have also attempted to explicitly control for this possibility by including a dummy variable for whether potential donors and receivers have no land in the set of control variables. In this way, the findings of this paper are consistent with those of the experimental literature where both efficiency and concerns for those who are less well-off have been found to drive decisions by subjects in the lab (Charness and Rabin, 2002).

The results line up well with the existing literature on norm-based access to land. For example, Devereux (1999, 2001) argues that traditional practices of redistribution from wealthier to poorer households are rapidly disappearing due to commercialization. This is exactly what we find in more population-dense villages, where commercialization of the land market is likely further progressed. However, the results of this paper also imply that community security systems in terms of access to land are still in effect in less densely populated villages in rural Gambia.

In terms of policy, the findings suggest that the displacement of informal safety nets should be taken into account when land redistribution reforms are planned and implemented. First, standard measures of land ownership may overestimate land inequality if allocations of temporary usage rights are not taken into account. Second, redistribution of land towards landless households is likely to crowd out norm-based inter-household transfers of land to the rural poor, but the extent to which such transfers exist may vary at the local level, depending on the level of population density and ethnic diversity. However, this does not mean that land reforms are always unnecessary or even harmful. This paper does not show that the norm-based system is better than what a redistributive land reform can achieve. If land transfers follow social ties, which the results of this paper as well as the paper of Chapter 3 indicate, the status quo may not be efficient at the aggregate level. And the existence of even symbolic reciprocating payments mean that the poor receivers of land are not quite as well off as they would be if the owned the land outright. Finally, the results of the paper suggest that as population densities rise and land access becomes increasingly market based, the need for land redistribution reforms is likely to increase as the strength of norm-based access rules decline.

References

Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114:1243–1284.

Arcand, J.-L. and Jaimovich, D. (2014). Does ethnic diversity decrease economic interactions? Evidence from exchange networks in rural Gambia. MPRA Paper 60497, University Library of Munich.

Boserup, E. (1965). *The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure*. London: George Allen & Unwin LTD.

Bourlès, R. and Bramoullé, Y. (2013). Altruism in Networks. Working Paper 56, Marseille School of Economics.

Cashdan, E. A. (1985). Coping with Risk: Reciprocity Among the Basarwa of Northern Botswana. *Man (N.S.)*, 20:454–4745.

Caudell, M., Rotolo, T., and Grima, M. (2015). Informal lending networks in rural Ethiopia. *Social Networks*, 40:34–42.

Charness, G. and Rabin, M. (2002). Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics*, 117(3):817–869.

Chavas, J.-P., Petrie, R., and Roth, M. (2005). Farm Household Production Efficiency: Evidence from The Gambia. *American Journal of Agricultural Economics*, 87:160–179.

Cohen, J. M. (1980). *Agricultural Development in Africa - Issue of Public Policy*, chapter Land Tenure and Rural Development in Africa, pages 349–400. Prager.

Cox, D. (1987). Motives for Private Income Transfers. *Journal of Political Economy*, 95:508–546.

Cox, D. and Fafchamps, M. (2007). Extended Family and Kinship Networks: Economic Insights and Evolutionary Directions. *Handbook of Development Economics, Chapter 58*, 4:3711–3784.

Cox, D., Hansen, B. E., and Jimenez, E. (2004). How responsive are private transfers to income? Evidence from a laissez-faire economy. *Journal of Public Economics*, 88:2193–2219.

Cox, D. and Jakubson, G. (1995). The connection between public transfers and private interfamily transfers. *Journal of Public Economics*, 57:129–167.

Deininger, K. (2003). *Land Policies for Growth and Poverty Reduction*. The International Bank for Reconstruction and Development and The World Bank.

Dercon, S. and Weerdt, J. D. (2006). Risk-sharing networks and insurance against illness. *Journal of Development Economics*, 81(2):337–356.

Devereux, S. (1999). 'Making less last longer': Informal safety nets in Malawi. ICS Discussion Paper 373.

Devereux, S. (2001). Livelihood Insecurity and Social Protection: A Re-emerging Issue in Rural Development. *Development Policy Review*, 19(4):507–519.

Dey, J. (1982). Development Planning in The Gambia: The Gap between Planners' and Farmers' Perceptions, Expectations and Objectives. *World Development*, 10(5):377–396.

Eastman, C. (1990). Gambian usufruct tenure. Help or hindrance to agricultural production? *Land Use Policy*, 7(1):27–40.

Eswaran, M. and Kotwal, A. (1985). A Theory of Contractual Structure in Agriculture. *The American Economic Review*, 75(3):352–367.

Fafchamps, M. (1992). Solidarity Networks in Preindustrial Societies: Rational Peasants with Moral Economy. *Economic Development and Cultural Change*, 41(1):147–174.

Fafchamps, M. and Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics*, 83:326–350.

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

Fafchamps, M. and Söderbom, M. (2014). Network Proximity and Business Practices in African Manufacturing. *World Bank Economic Review*, 28:99–129.

Fenske, J. (2011). Land tenure and investment incentives: Evidence from West Africa. *Journal of Development Economics*, 95:137–156.

Fenske, J. (2013). Does Land Abundance Explain African Institutions? *The Economic Journal*, 123(573):1363–1390.

Freudenberger, M. S. (1993). Institutions and natural resource management in The Gambia: A case study of the Foni Jarrol district. LTC Research Paper 114.

Freudenberger, M. S. (2000). Tenure and Natural Resources in The Gambia: Summary of Research Findings and Policy Options. Land Tenure Center. University of Wisconsin-Madison. Working paper no. 40.

Goetghebuer, T. and Platteau, J.-P. (2005). *The social economics of poverty*, chapter 15. Community ties and land inheritance in the context of rising outside opportunities. Evidence from the Peruvian Highlands, pages 361–401. Routledge.

Greene, W. H. (2002). *Econometric Analysis*. Prentice Hall, New Jersey, USA, 5th edition.

Holden, S. T., Otsuka, K., and Place, F., editors (2009). *The Emergence of Land Markets in Africa: Impacts on Poverty, Equity, and Efficiency.* Resources for the Future, Washington, DC.

Jaimovich, D. (2011). Macrostructure and microstructure: Evidence from overlapping village networks in The Gambia. MPRA paper 38932, University of Munich.

Jaimovich, D. (2015). Missing Links, Missing Markets: Evidence of the Transformation Process in the Economic Networks of Gambian Villages. *World Development*, 66:645–664.

Kazianga, H. (2006). Motives for household proivate transfers in Burkina Faso. *Journal of Development Economics*, 79:73–117.

Kazianga, H. and Udry, C. (2006). Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso. *Journal of Development Economics*, 79(2):413–446.

Krishnan, P. and Sciubba, E. (2009). Links and Architecture in Village Networks. *The Economic Journal*, 119:917–949.

Mazzocco, M. and Saini, S. (2012). Testing Efficiency Risk Sharing witrh Heterogenous Risk Preferences. *American Economic Review*, 1:428–468.

McGuire, S. J. (2008). Securing Access to Seed: Social Relations and Sorghum Seed Exchange in Eastern Ethiopia. *Human Ecology*, 36:217–229.

Mitrut, A. and Nordblom, K. (2010). Social norms and gift behavior: Theory and evidence from Romania. *European Economic Review*, 54:998–1015.

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

Noronha, R. (1985). A Review of the Literature on Land Tenure Systems in sub-Saharan Africa. Report ARU 43, World Bank, Washington, DC.

Otsuka, K. (2007). Efficiency and Equity Effects of Land Markets. *Handbook of Agricultural Economics*, 3:2671–2703.

Ozer-Balli, H. and Sørensen, B. E. (2013). Interaction effects in econometrics. *Empirical Economics*, 45(1):583–603.

Pamela, K. (2010). *Land, labour and entrustment: West African female farmers and the politics of difference*. Leiden, Boston.

Peters, P. E. (2009). Challenges in Land Tenure and Land Reform in Africa: Anthropological Contributions. *World Development*, 37(8):1317–1325.

Platteau, J.-P. (1991). *Traditional Systems of Social Security and Hunger Insurance: Past Achievements and Modern Challenges*, pages 112–70. Clarendon Press, Oxford.

Platteau, J.-P. (1997). Mutual insurance as an elusive concept in traditional rural communities. *The Journal of Development Studies*, 33(6):764–796.

Platteau, J.-P. (2002). The gradual erosion of the social security function of customary land tenure arrangements in lineage-based societies. Discussion Paper No. 2002/26.

Platteau, J.-P. (2006). *Solidarity norms and institutions in village societies: static and dynamic considerations*. In: Kolm, S.-C., Mercier Ythier, J. (Eds.), Handbook of Economics of Giving, Altruism and Reciprocity, vol. 1. Elsevier.

Platteau, J.-P. and Abrahamb, A. (1987). An inquiry into quasi-credit contracts: The role of reciprocal credit and interlinked deals in small-scale fishing communities. *The Journal of Development Studies*, 23(4):461–490.

Swindell, K. (1987). Family Farms and Migrant Labour: The Strange Farmers of the Gambia. *Canadian Journal of African Studies*, 12(1):3–17.

von Braun, J. and Webb, P. J. R. (1989). The Impact of New Crop Technology on the Agricultural Division of Labor in a West African Setting. *Economic Development and Cultural Change*, 37:513–534.

Wolf, E. R. (1957). Closed Corporate Peasant Communities in Mesoamerica and Central Java. *Southwestern Journal of Anthropology*, 13(1):1–18.

World Bank (1993). The Gambia - An assessment of Poverty. Report No. 11941-GM.

World Bank (2007). World Development Report 2008: Agriculture for Development. Technical report, World Bank.

World Bank (2015). World Development Indicators 2015. Accessed May 11, 2015.

Zak, P. J. and Knack, S. (2001). Trust and Growth. *The Economic Journal*, 111:295–321.

Appendix A: Tobit specification checks and robustness

In the main text, equation (1) is estimated using a tobit model. We do so due to the large number of zeroes observed in the dyadic land transfer network. The zeroes can be interpreted as the outcome of a corner solution of the land transfer decision between two households. A transfer is observed only when the benefits of making a transfer exceed the costs of participating in the land market. Unlike a standard OLS regression, the tobit model explicitly accounts for the non-linearities that arise as a consequence of corner solutions. However, this comes at the cost of additional parametric assumptions about normality and homoscedasticity of the latent error terms. We first test whether normality and homoscedasticity can be rejected. We then investigate the robustness of the main results when estimated by OLS.

We test normality of the latent variable by testing moment conditions on generalized residuals using the "Outer Product of the Gradient"-approach of Skeels and Vella (1999).[i] Critical values are determined using 500 iterations of the parametric bootstrap of Drukker (2002). We cannot reject the null hypothesis of normally distributed errors (Conditional moment test value=53.19; 10% critical value=54.27). We also test the assumption of homoscedasticity using a lagrange multiplier test against the more general form of $\mathrm{var}(\epsilon|x) = \sigma^2 \exp(x\delta)$, where ϵ is the error term of the latent variable, x is the vector of independent variables and δ is a vector of coefficients (Wooldridge, 2010, ch. 17.5.3).[ii] We cannot reject the null hypothesis of homoscedasticity ($p = 0.97$).

Table A.1, column 1 corresponds to the baseline specification in Table 3, column 2. In column 3, the same model is re-estimated using an OLS model. The OLS model does not require normality and homoscedasticity for consistency. The main problem of using OLS in this context is that it assumes that the expectation of the dependent variable is linear in the independent variables, which can only be true over a limited range of values of the independent variables. This can lead to predicted values of dependent variables can be negative. These shortcomings are similar to those encountered when applying a linear probability model a binary outcome (Wooldridge, 2010, ch. 17.1). To compare coefficient sizes, average partial effects for the tobit model, i.e. $\frac{\partial E(y|x)}{\partial x}$'s, are shown in column 2. Sign, magnitude and significance are consistent between the two estimators with the one exception that the coefficient on non-agricultural income of senders (j) is no longer statistically significant when estimated by OLS (column 3). This exception does not change conclusions about the existence of norm-based land access rules, which were primarily based on the existence of a negative effect of recipient non-agricultural income and secondarily on the positive effect of sender non-agricultural income.

[i] The test is conducted on a version of the model without clustering in order for the standard formulas for generalized moments of the tobit model to apply (Lee and Maddala, 1985). All parameter estimates and significance levels are very similar to the specification which includes clusters.

[ii] This test is also conducted on a non-clustered version of the model.

Another reason why the tobit model may fail is due to the assumption of a single parameter affecting both the decision to transfer and the magnitude of the transfer, conditional on transferring. In the absence of suitable instruments to separately identify the two effects, we take a simpler approach by estimating separate regressions for having a land transfer and the size of the transfer, conditional on participation. This requires the added assumption that the two-part models is that the two residuals are independent (i.e. the unobservables which affect the decision to participate are independent of the unobservables that affect the decision of how much to transfer) and that errors are normally distributed with a constant variance and zero mean. We estimate the truncated normal hurdle model of Cragg (1971) and the lognormal hurdle model of Wooldridge (2010, ch. 17.6.2). The advantage of these models is that the effects of the explanatory variables are allowed to vary between the decision to transfer and the magnitude of the transfer.

Estimation results are reported in Table A.1, column 4 and 5. Apart from the coefficient estimate on receivers' (j) agricultural income in the transfer size regression, signs and significance of the coefficient estimates are similar (this is also the case for the control variables). This result lends credibility to the assumption that the same mechanism is likely to drive both decisions. The positive estimate on senders' agricultural income suggests that households with a higher agricultural income are more likely to participate in the land market, but conditional on transferring less land. The variables of primary interest are the non-agricultural incomes of the sender and the receiver. The result suggests that low-income households are more likely to participate in the market, and to receive more land, conditional on receiving land. In contrast, high-income senders are more likely to participate, whereas no significant difference is found between participating senders in terms of the amount of land they send. In summary, the baseline results are consistent with the same level of significance and sign using these alternative estimation techniques.

References for Appendix A

Cragg, J. G. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica*, 39(5):829–844.

Drukker, D. M. (2002). Bootstrapping a conditional moments test for normality after tobit estimation. *The Stata Journal*, 2(2):125–139.

Lee, L.-F. and Maddala, G. S. (1985). The Common Structure of Tests for Selectivity Bias, Serial Correlation, Heteroscedasticity and Non-Normality in the Tobit Model. *International Economic Review*, 26(1):1–20.

Skeels, C. L. and Vella, F. (1999). A Monte Carlo investigation of the sampling behavior of conditional moment tests in Tobit and Probit models. *Journal of Econometrics*, 92(2):275–294.

Table A.1: Baseline results – estimation technique robustness

	(1)	(2)	(3)	(4)		(5)	
	Tobit	Tobit	OLS	Normal hurdle		Log-normal hurdle	
		Partial effects		Probit	OLS	Probit	OLS (log)
j's non-agri. income	-0.439**	-0.004**	-0.007***	-0.076**	-0.837*	-0.076**	-0.121**
	(0.196)	(0.002)	(0.002)	(0.038)	(0.432)	(0.038)	(0.059)
i's non-agri. income	0.478*	0.005	0.006*	0.098*	-0.028	0.098*	0.055
	(0.264)	(0.003)	(0.004)	(0.053)	(0.210)	(0.053)	(0.056)
j's agri. income	0.683**	0.007*	0.005	0.142**	0.196	0.142**	0.027
	(0.278)	(0.003)	(0.003)	(0.062)	(0.385)	(0.062)	(0.079)
i's agri. income	0.267	0.002	-0.002	0.063	-0.638	0.063	-0.193*
	(0.338)	(0.003)	(0.004)	(0.069)	(0.455)	(0.069)	(0.114)
Observations	87,788	87,788	87,788	87,788		87,138	859

Note: Dependent variable: Amount of land j is allocated from i in hectares. All regressions include control variables and village fixed effects. Standard errors are clustered on the village level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively. Average partial effects are calculated unconditionally on receiving land or not.
Source: Authors' calculations.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*, volume 1 of *MIT Press Books*. The MIT Press.

Appendix B: Additional tables and figures

Table B.1: Descriptive statistics: village-level

	Mean	Std. Dev.	Min.	Max.
Inhabitants	521.23	184.46	130	1,077
Area (hectares)	402.16	353.32	20.77	1751.31
Density (inhabitants per square km)	234.62	182.58	25.75	980.14
Ethnic diversity index (= 0 if homogenous)	0.275	0.235	0	0.837
Illiteracy share	0.451	0.207	0	0.913
No access to electricity	0.970	0.042	0.837	1
No private toilettes	0.372	0.296	0	1
Not improved water access	0.912	0.131	0.431	1
Gini (based on self-declared income per cap.)	0.314	0.112	0.144	0.601
Observations		52		

Note: Density is calculated based on the sum of the cultivated and the inhabited village area.
Source: Authors' calculations.

Table B.2: Village characteristics and pro-poorness of land transfers with donor or recipient fixed effects

Panel A: Dependent variable: donor income (i)						
Village characteristic:	None	Density	Diver-sity	None	Density	Divers-ity
	(1)	(2)	(3)	(4)	(5)	(6)
Total income$_i$	0.005	0.009**	-0.002			
	(0.003)	(0.004)	(0.003)			
High*total income$_i$		-0.010*	**0.012****			
		(0.005)	(0.005)			
Non-agri. income$_i$				0.006*	0.009**	-0.000
				(0.004)	(0.004)	(0.004)
Agri. income$_i$				-0.002	0.004	-0.006
				(0.004)	(0.003)	(0.008)
High*non-agri. income$_i$					-0.008	**0.012****
					(0.006)	(0.005)
High*agri. income$_i$					-0.014*	0.008
					(0.008)	(0.009)
Observations	87,788	87,788	87,788	87,788	87,788	87,788
Additional controls:	Donor characteristics and recipient fixed effects					

Panel B: Dependent variable: recipient income (j)						
Village characteristic:	None	Density	Divers-ity	None	Density	Divers-ity
	(7)	(8)	(9)	(10)	(11)	(12)
Total income$_j$	-0.006***	-0.009***	- 0.006*			
	(0.002)	(0.003)	(0.003)			
High*total income$_j$		0.007*	**0.000**			
		(0.004)	(0.004)			
Non-agri. income$_j$				-0.007***	-0.010***	- 0.007*
				(0.002)	(0.003)	(0.004)
Agri. income$_j$				0.005	0.003	-0.004
				(0.003)	(0.004)	(0.006)
High*non-agri. income$_j$					0.006	**-0.001**
					(0.004)	(0.005)
High*agri. income$_j$					**0.004**	**0.012***
					(0.005)	(0.007)
Observations	87,788	87,788	87,788	87,788	87,788	87,788
Additional controls:	Recipient characteristics and donor fixed effects					

Note: OLS. Dependent variable: Amount of land j is allocated from i. "High" is the indicator variable equal to one if the community characteristic in question (either density or diversity) is higher than the median. Control variables are the same as those of Table 3 Standard errors are clustered on the village level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively. Numbers in bold estimates refer to cases where interaction + main effect is significant at the 5 percent level.
Source: Authors' calculations.

CHAPTER 5

COFFEE PRICE VOLATILITY AND INTRA-HOUSEHOLD LABOR SUPPLY: EVIDENCE FROM VIETNAM

Ulrik Beck* Saurabh Singhal† Finn Tarp‡

Abstract

Volatility in commodity markets poses an acute risk to farmers in developing countries who rely on cash crop agriculture. We combine a time series of international coffee prices with a long-running panel on coffee-growing households in Vietnam to investigate coping mechanisms employed by farmers in a transitioning economy. We find that households cope with lower coffee prices by increasing wage labor of adults with children and adolescents substituting for adults on the farm and in home production. Account of this finding should be taken in formulating and implementing social protection and inclusive growth policies.

We thank seminar participants at the University of Copenhagen and UNU-WIDER for helpful comments, and are grateful for productive and stimulating collaboration with the survey teams from CIEM and ILSSA, Hanoi, Vietnam. The paper builds on ideas from the Master's thesis of Ulrik Beck (Beck, 2014), although the dataset has been expanded, the identification strategy has been revised and the primary research question has been adjusted. All interpretations and any remaining errors are the sole responsibility of the authors.

*Department of Economics, University of Copenhagen

†UNU-WIDER, Helsinki

‡UNU-WIDER, Helsinki and Department of Economics, University of Copenhagen

5.1 Introduction

Volatility in commodity markets poses an inherent risk to small farming households eking out a living in developing countries. Volatility in the agricultural commodity prices has been higher in the 2000s relative to preceding decades, raising concern among policy makers and various international organizations (FAO et al., 2011). In order to formulate policies to insulate farmers from the vagaries of the commodity markets, it is essential to understand household responses and welfare effects that result from commodity price volatility in such risky environments. One would expect that as countries develop, commodity price changes become a smaller threat to the livelihoods of households who produce these commodities. In particular, the need for smallholder farmers to weather shocks by taking up wage work or using child labor should fall as the economy develops, income levels increase, and social protection policies are put in place.

Rural households in developing countries engage in a number of production activities. In addition to farming, households may have off-farm wage employment, operate household enterprises, tend to livestock, and produce household goods. The labor required for these activities is spread over various household members. In the absence of well-developed credit markets, shocks that close one avenue of income will typically lead to the reallocation of labor of household members as the household attempts to maintain consumption. What is the extent of such intra-household spillovers? In particular, what is the burden borne by children and adolescents? Using a long panel of rural Vietnamese households, we exploit exogenous variation in the international price of coffee, a key export commodity of Vietnam, to causally identify the effects of temporary income fluctuations on welfare as well as intra-household labor allocation decisions, with a particular focus on the use of child and adolescent labor.

Adult and child labor supply overlap in a variety of activities, and the reallocation of time across various work alternatives and members in the face of an income loss depend, in part, on the degree of substitutability between the two. For example, Duryea et al. (2007) find that 16-year-old girls are more likely to enter the labor market in the event of their fathers becoming unemployed in Brazil. In related work from Peru, Field (2007) finds that land-titling reduced child labor as titling freed up adult labor, which has comparative advantage in security provision, from guard duties and enabled them to substitute for children in the labor market. While most of the existing literature has studied the labor supply response of adults and children in isolation, we examine the labor supply adjustment behavior of adults and children in the same families, thereby allowing us to provide a better picture of the dynamics of intra-household labor supply decisions. As the Vietnamese economy develops, an important aspect of its transformation is a shift in the focus from primary to secondary and tertiary education. While the literature typically considers the labor market responses of those aged above 14 years together as a single group, we separate out the adolescents, who may attend higher secondary school (aged 15-19 years) from the adults (aged 20-54).

This paper aims to contribute to the literature on labor substitution patterns between adults and children in the event of fluctuations in household income. We combine, as already noted, a long-running panel dataset with a time-series of international coffee prices to identify the channels through which coffee growing households in Vietnam mitigate the effects of coffee price volatility. We first show that international prices are indeed strongly transmitted to small coffee growers (in the form of farm-gate prices) and turn next to our main analysis. It emerges that coffee farmers are not able to perfectly smooth consumption when faced with a fall in the price of coffee, even though the households rely on several strategies to counteract the negative price change. Coffee price shocks are strongly associated with ex-post (re)distribution of work within the household as well as countercyclical wage employment. The effect on wage employment is driven by working age adults, and to a lesser degree, by adolescents. While children do not engage in wage employment, we find that children and adolescents substitute for adults on the farm in low price periods.

The burden borne by children is of particular concern as participation in child labor can have severe adverse long-term consequences for human capital accumulation (Beegle et al., 2009).[1] But in the context of commodity prices, it is not clear if changes will work to increase or decrease the reliance on child labor. Consider a household consisting of adults and children deciding how to allocate the labor of both adults and children among different activities. The activities for adults include working on the farm, non-farm activities such as wage work and leisure, while children can spend their time working on the farm, attending school or on leisure.[2] Assume further that households do not consume their own production. In many contexts, this is not a reasonable assumption (as in the case of rice), but for coffee, which is sold in the market, this is a reasonable approximation. Assume further that labor is a bad in the parental preference function (Basu and Van, 1998). A decrease in the price of the agricultural commodity will decrease the return to child labor tending to lower child labor supply (i.e., the substitution effect). On the other hand, it will also decrease household income, which will increase child labor due to the income effect. For adults as well, the decrease in the price of the agricultural commodity lowers the return to working on the farm (substitution effect). Adults have two options: They can try to switch into a relatively better-paying occupation such as wage work, or they can increase leisure – or rely on a combination of both. However, the loss of income may in itself reduce time spent on leisure and increase work either on- or off-farm. Therefore, the combination of income

[1] See Basu and Van (1998) for a theoretical model and Edmonds (2007) for a recent review of the literature on child labor.

[2] That children do not engage in the wage market is consistent with our data, which is introduced in Section 5.2.

and substitution effects means that the overall direction of impact for both children and adults is theoretically ambiguous a priori.

This study sits within a broader literature that investigates the welfare impacts of commodity price shocks in an increasingly globalized world (Goldberg and Pavcnik, 2007). As supply chains stretch across the world, shocks at one end reverberate to the other end with important distributional consequences (see Lederman and Porto (2016) for a review). For example, Danzer and Grundke (2014) find that a sudden spike in cotton prices in 2010-11 led to greater labor force participation among women in rural Tajikistan. In the context of Vietnam, Edmonds and Pavcnik (2005, 2006) find that an increase in rice prices, resulting from removal of trade restrictions, lead to increased off-farm work for adults and decreased the reliance on child labor among rice producers. Our paper addresses a broader set of issues, and two key distinctions should be kept in mind. First, we consider changes in the price of coffee instead of rice. Smallholder farmers typically consume rice, which adds additional channels through which household behavior is affected by prices. Coffee, on the other hand, is almost entirely produced for sale and the estimated effects are not confounded by substitution and income effects through consumption. Second, while Edmonds and Pavcnik (2005, 2006) use data collected over the years 1993-98, our period of analysis covers the period 2006-14. Vietnam has experienced rapid development in recent decades – in 2013, GDP per capita was four times that of 1993. It is therefore of interest to explore whether price changes had similar effects in the 1990s and in the post-2000 period or whether behavior changed in the process.

Our study is also related to the research on coping mechanisms in the face of income shocks. The literature has identified various other coping mechanisms such as wage employment, migration, sale of assets and livestock, increase in household debt or a reduction in savings (e.g., Paxson, 1992; Kochar, 1999; Dercon, 2002; Bussolo et al., 2007; Klasen et al., 2013). Adhvaryu et al. (2015) identify switching to a greater reliance on household enterprises as an additional channel. Furthermore, while we examine the ex-post response to price shocks, households may also undertake ex-ante adjustments to perceived future income uncertainty (Rose, 2001; Fitzsimons, 2007).

Finally, the evidence on coping mechanisms employed by households in the Central Highlands of Vietnam can be used both more generally and to inform policy when faced with future coffee price booms or busts. In particular, we show that adolescents completing lower-secondary education at around 14 years of age are vulnerable to transient income shocks with implications for the formation of human capital in the face of price fluctuations. Account should be taken hereof in the future provision of safety nets and the pursuit of inclusive growth.

The rest of the paper is organized as follows. Section 5.2 provides background on coffee production in Vietnam and presents the data used in the paper. Section 5.3 outlines the empirical strategy. The main results are presented in Section 5.4 along with a variety of robustness checks while Section 5.5 explores the implications of the main results further. Section 5.6 concludes.

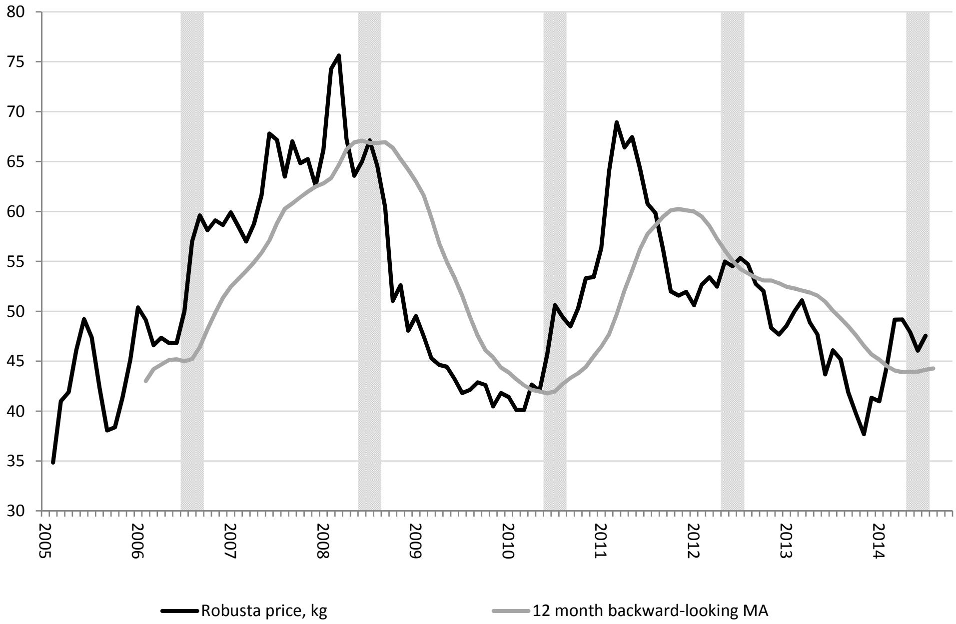
5.2 Background and data

Coffee production in Vietnam has increased very significantly since the beginning of the 1990s. In 1989, Vietnam produced 1.2 percent of the world's coffee output. Ten years later, Vietnam was the second largest supplier in the world, producing 12.4 percent of the world coffee supply. So far, the majority of the coffee grown has been of the "robusta" variety. Compared to "arabica", which is the other main variety, robusta is more resilient to variations in growing conditions, more robust to weather shocks and typically has higher yields, but also fetches a lower price in the international market. More than 85 percent of Vietnamese coffee is grown by smallholder farmers. In 2006, coffee accounted for around 17 percent of the country's commodity exports and provided livelihoods for around four million people (Luong and Tauer, 2006).

International coffee prices are quite volatile and are driven by a mix of climatic conditions in the largest coffee-producing countries, expectations about future prices, changes in demand and interest rates as well as speculation (Deaton, 1999). This means that future coffee prices are difficult to predict, and for the smallholder farmer who cultivates coffee, prediction is virtually impossible. The fluctuations in the world price for robusta over time are illustrated in Figure 1, along with a 12-month backward-looking average. The vertical bars indicate the months where the data collection for the survey employed in this paper took place. For the year preceding the survey rounds of 2006, 2010 and 2014, coffee prices were relatively low, while coffee prices were higher for the year preceding the survey rounds of 2008 and 2012. The changes in prices are substantial: The price in the 12 months preceding the 2008 survey was almost 50 percent higher than the price in 2006; the 2012 price was around 30 percent higher than the 2010 price. The same pattern of price changes over time can be found in the unit prices recorded at household-level detailed in Table 1 – while prices rose from 2006 to 2008 and again from 2010 to 2012, they fell from 2008 to 2010 and again from 2012 to 2014.

The majority of Vietnamese coffee production takes place in the Central Highlands region. Once coffee trees are planted, it is costly to switch to other types of crops. Coffee trees have a life span of more than 50 years, and it is a labor-intensive task to cut the trees down to make plots suitable for other types of agricultural production. In 2004, the cost of uprooting a single coffee tree was estimated at 2,000 Vietnamese Dong (VND), or around 0.15 US dollars (Giovannucci et al., 2004). This corresponds to VND 5,200 in 2014 prices, which amounts to half of the average per capita daily food expenditure of our sample. Since the farmers in our sample have on average about one hectare of land dedicated to coffee production, and assuming plantations

Figure 1: World market coffee prices, '000 real June 2014 VND/kg



Note: The horizontal bars show the three-month periods of the collection of the VARHS survey rounds.
Source: Authors' illustration based on ICO (2015) and World Bank (2015).

Table 1: Descriptive statistics

	All years		2006	2008	2010	2012	2014
	Mean	SD			Mean		
Household size	4.88	1.74	5.06	5.03	4.91	4.78	4.75
HH head is Kinh[1]	0.61	0.49	0.69	0.59	0.60	0.61	0.61
Food expenditures[2]	1540	1055	1402	1406	1503	1802	1485
Asset index	0.52	1.38	-0.15	0.07	0.41	0.85	0.97
Unit price of coffee/SD	1.26	0.48	1.22	1.43	1.06	1.36	1.19
Area of plots, m2	17334	16470	15993	16626	17391	18328	17454
Coffee share of land	0.48	0.41	0.60	0.40	0.47	0.49	0.51
Coffee produced, kg	2316	3922	1608	1763	2547	2482	2821
HH sold coffee[1,3]	0.99	0.11	0.99	0.99	0.98	0.99	0.99
HH has HH business[1]	0.16	0.37	0.18	0.13	0.19	0.18	0.12
HH has wage work[1]	0.58	0.49	0.59	0.48	0.62	0.61	0.59
Child work..							
in agriculture[1]	0.20	0.40	0.28	0.23	0.31	0.12	0.14
in wage work[1]	0.01	0.08	0.01	0.00	0.01	0.01	0.01
in HH business[1]	0.01	0.07	0.02	0.00	0.01	0.00	0.00
No. of observations	2355		209	516	515	562	553

1: Binary variable
2: Real 2014 '000 VND. 28-day period
Source: Authors' calculations based on VARHS.

consist of 1,100 trees per hectare (D'haeze et al., 2003), it is a very costly task to stop coffee production. This means that we do not expect many farmers to abandon coffee production when coffee prices are low. The inability to adjust the area dedicated to coffee trees in the short term means that households must instead use other strategies, such as reallocation of labor, when faced with low coffee prices. Further, since coffee is a cash crop, households consume very little of their own production. This implies that on the consumption side, household responses are almost entirely through changes in household income and not through changes in relative prices of consumption items. These two factors make this setting very useful for identifying the impact of agricultural price shocks on on- and off-farm responses.

The data used for this paper comes from five waves of the Vietnam Access to Resources Household Survey (VARHS hereafter; see CIEM, 2007, 2009, 2011, 2013, 2015). This panel survey has been collected in the months of May-September every second year since 2006 among rural households from 12 provinces of Vietnam.[3] Additional households were added in 2008. In 2012, the sample was again expanded to compensate for the aging of the panel sample relative to the existing household structure in Vietnam giving us an unbalanced panel of households. The VARHS includes three provinces in the Central Highlands – Dak Lak, Dak Nong, and Lam Dong. It is the households in these three provinces, which constitute our main empirical sample. We further restrict our main sample to households that appear in at least two rounds and that report having harvested coffee at least once over the period 2006-14. Data on households residing in the remaining nine provinces is used for a robustness check in Section 5.4.3.

The sample of households varies from 209 households in 2006 to 562 households in 2012. The total sample consists of 2,355 household-year observations. Table 1 shows descriptive statistics for the total sample, as well as by year. The information in the survey gives a telling picture of contemporary rural life in Vietnam's Central Highlands. 39 percent of households belong to ethnic minority households. While this share is high compared to the rest of the country, it is typical of the Central Highlands. According to the latest population census, 35 percent of the population in the Central Highlands region belonged to a minority ethnic group GSO 2010. The VARHS includes a brief module on the value of consumption of some important food items. The average real food expenditure during the 28-day period before the survey is VND 1.54 million, measured in June 2014 prices.[4] However, there has been substantial improvement over the period. This is perhaps best illustrated using the asset index, which is a composite index of a range of different asset indicators, the relative weights of which are determined using principal component analysis.[5] There has been a substantial

[3]The 2006 round was collected in July, August and September; the 2008 and 2010 rounds were collected in June, July and August; the 2012 and 2014 rounds were collected in May, June and July.

[4]Due to outliers in the food expenditure variable, we have opted to winsorize this variable by replacing all values above the 99th percentile with the value of food expenditures at the 99th percentile.

[5]The asset index is based on the following variables: the number of cows, buffaloes, pigs, color TVs, video/DVD machines, telephones, motorcycles, bicycles, cars, pesticide sprayers; the square meterage of

increase in the index from 2006 to 2014, going from a mean of -0.15 to a mean of 0.97. Of the households in the sample, most depend on coffee for a substantial share of their agricultural production; on average, households dedicate slightly under half of their land to coffee production. Since 2008, this variable has been increasing.[6] These increases correspond well with the increases in the amount of coffee produced. The average household produced 2,316 kg of coffee, and production generally increased between each survey round, except from 2010 to 2012 where there is a small drop. Farm sizes are large compared to other areas of Vietnam, with a mean farm size of 1.7 hectares (17,000 square meters). Many households supplement the agricultural income with either a wage job or through a household enterprise; on average, 16 percent engage in a household enterprise and 58 percent engage in wage work. Most wage work consists of unskilled jobs in agriculture and forestry (65 percent of job observations in the sample), followed by unskilled jobs in mining and construction (8 percent).

Table 1 also gives the farm-gate price the household received per kilo of coffee sold, in June 2014 prices. Again, in order to ease interpretation of subsequent regressions, we normalize this variable by dividing by its standard deviation.[7] Prices were highest in 2008 and 2012. This is consistent with the evolution of the world market coffee prices, as illustrated in Figure 1.

In order to investigate the effect of coffee price shocks we combine the VARHS survey data with monthly robusta coffee prices available from the International Coffee Association, converted into real 2014 VND (ICO, 2015). This data is shown in Figure 1. In order to take into account that it is not the coffee price at the time of the interview that matters, but rather the coffee prices faced by the household in the year preceding the survey (the time period over which the survey asked questions), we use a 12-month backward-looking moving average. Thus, the relevant coffee price for a household surveyed in month m is the average coffee price over the months from m-13 to m-1. Our key explanatory variable is constructed by dividing this lagged coffee price by its standard deviation over the survey period (2006-2014).

In order to investigate intra-household responses, we also work with a panel that consists of individuals from households that make up the main sample described above. This results in an unbalanced panel of 10,022 observations across rounds of individuals who appear at least twice over the course of the survey period in the same age group. This sample is split into children (2,246 observations), adolescents (1,733 observations) and adults (6,043 observations). We define child labor as work done by household members who are 14 years of age or younger. This is in keeping with the

the household's house; and indicator variables equal to one if the household has access to an improved water supply, to a modern form of lighting and to improved toilet facilities. The index is normalized to have a mean of zero and the weights are based on the full survey sample (12 provinces).

[6]The drop in the share of land dedicated to coffee production from 2006 to 2008 is primarily caused by the expansion of the sample in 2008.

[7]Due to outliers in this variable, we have winsorized it by replacing all values above the 99th percentile with the value at the 99th percentile.

International labor Organization (ILO) definition of "child" for employment purposes and also reflects the minimum legal working age in Vietnam of 15 years (ILO, 1973; MoLISA, 2013). A fifth of households employ child labor in agriculture. This has declined substantially over time, particularly over 2010-12. In contrast, only around one percent of households utilize child labor in wage employment and in household businesses.

There is some suggestive evidence that wage labor is not a preferred occupation among our sample of coffee-growing households. This can be seen by comparing agricultural households who have at least one member participating in wage labor, compared to those who do not. The mean per capita food expenditure of coffee-growing households who participate in wage labor is VND 311,000 per capita, compared to VND 390,000 for those who do not. The same pattern is evident for the asset index, with the mean asset index for those who participate in wage labor being 0.37 and 0.73 for those who do not. Further, based on the households in all 12 provinces of the 2012 VARHS wave, Markussen et al. (2014) find that wage workers have lower levels of subjective well-being than those who work on their own farm, even when income is controlled for. This is indicative of households preferring to stay out of wage labor if possible. Therefore, we expect that in the setting we are considering here, households or members of households will only switch to wage labor when they are forced to do so, i.e., when returns from working on their own farm are no longer sufficient to sustain household welfare.

As in all panel surveys, some amount of attrition is inevitable. However, we find a very low attrition rate in our sample in the range of 0-1.6 percent from round to round. This is partly a testament to the high quality of the survey but also a consequence of the relative stability of conditions in rural Vietnam. Rural-to-urban migration is lower than in many other countries in the region.

5.3 Estimation strategy

In order to evaluate the effects of coffee prices on household outcomes we estimate the following household-level equation using a linear model with fixed effects:

$$y_{it} = \alpha + \beta p_t + \delta_c t + \lambda x_{it} + \eta_i + \epsilon_{it} \tag{1}$$

where y_{it} is the outcome of interest for household i at time t. p_t is the international coffee price which varies by month. x_{it} is a vector of additional control variables. We include indicator variables that control for household exposure to shocks in the preceding two years (since the previous survey). Specifically, we control for three kinds of shocks: natural disasters (floods, droughts etc.), pest attacks and health shocks, i.e., the death of a household member or serious illness. We also include household size, household size squared and the asset index. As discussed earlier, p_t is the average international robusta coffee price in the 12 months preceding the survey, normalized by dividing by its standard deviation. The coefficient β measures the marginal effect of a one standard deviation increase in the international coffee price. Since the individual farmer does not influence the international coffee prices, we interpret this as causal. All household-specific time-invariant characteristics are captured by the household fixed effect η_i.[8] In order to avoid any spurious correlation caused by underlying time-variant effects, we include a linear time trend variable. To allow for spatial heterogeneity in time trends, we use province-specific time trends (δ_c). We cluster standard errors at the level of the commune to allow household decisions to be correlated within a commune as well as over time.

In order to study individual-level outcomes, we employ a linear model with fixed effects at the individual level:

$$y_{jt} = \alpha + \beta p_t + \delta_c t + \lambda x_{it} + \eta_j + \epsilon_{jt} \tag{2}$$

where j denotes an individual of household i. All remaining notation is as in equation (1). Decisions of members of the same household are expected to be correlated. However, as we continue to cluster errors at the commune level, we make no assumptions about the correlation structure of within-household error terms.

One may be concerned that differential returns to education explain the labor allocation decisions for children. Glewwe and Jacoby (2004) argue that school enrollment in Vietnam is driven by increases in income rather than changes in the returns to education. Further, in all our specifications, we include province-specific time trends to control for unobserved province-specific time-varying factors.

Similarly, our results could overestimate the direct effect of coffee prices on schooling and labor market decisions related to children if the supply of schooling also changed with the price of coffee. This could happen, for example, if the provision of local schooling depended on tax revenue. While school funding is nominally allocated to the provinces, and to the districts within each province, from the state budget, the reality is more complicated as funding is often complemented by local funding which is dependent on tax revenues (Cobbe, 2011).[9] It is therefore possible that such general equilibrium effects are contained in the results related to behavior and schooling

[8]Fixed effects at the level of the household or individual are important if household or individual characteristics are correlated with coffee prices and also affect the outcome variables independently. Since the world market coffee price is exogenous to the household, it is possible that fixed effects are not important, as argued by Edmonds and Pavcnik (2005). However, even though the coffee price is exogenous to the individual household, the time-varying control variables may not be. If these are correlated with time-invariant characteristics of the household or individual, this will bias *all* coefficient estimates. For this reason, we believe fixed effects are sufficiently important that we include them in all regressions.

[9]Taxes where revenue accrues at least partly to the provincial governments include land taxes, VAT, and personal income taxes (Shukla et al., 2011).

outcomes of children and adolescents. Finally, one may speculate that macroeconomic trends other than international coffee prices could affect both labor supply decisions and coffee prices. While the inclusion of time trends addresses this concern to some extent, we also present a variety of alternative estimation strategies in Section 5.4.3, which also discusses a series of other potential threats to identification.

5.4 Results

5.4.1 Coffee prices and household welfare

We begin by analyzing the effect of international coffee prices on households in Vietnam. These results are presented in Table 2. The first column shows, as expected, that the international price is highly correlated with the farm-gate price received by farmers. An increase in the international price of one standard deviation increases the farm-gate price by 0.13 standard deviations indicating that international prices strongly affect observed farm-gate prices. While one might have expected an even higher correlation, other things such as local climatic and crop disease conditions affect the farm-gate price. Further, intermediate traders, and transaction costs more generally, absorb a share of the price changes (Fafchamps and Hill, 2008).

Next, we examine if coffee price shocks affect household welfare. Column 2 shows that the monthly household food expenditure is positively affected by coffee prices such that a one standard deviation increase in coffee prices increases the monthly food expenditure by VND 47,950 (or 3.1 percent from a mean of VND 1,540,000). This

Table 2: Coffee prices and household welfare

	(1) HH unit price/SD	(2) Food Exp.	(3) Asset index	(4) Wage work
Coffee price/SD	0.128***	47.953***	0.049**	-0.037***
	(0.008)	(16.752)	(0.020)	(0.008)
Constant	0.088	-343.564	-2.243***	0.358***
	(0.122)	(266.773)	(0.254)	(0.132)
Province time trend	Yes	Yes	Yes	Yes
HH size	Yes	Yes	Yes	Yes
Shocks	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes
Mean of dep. var.	1.26	1540.09	0.52	0.58
N	1922	2355	2355	2355

Note: Coffee price and HH unit price are both in real prices. Food expenditure is in real 2014 '000 VND. Wage work is an indicator variable equal to one if the household participated in wage work in the last year. Standard errors in parentheses are clustered at the commune level. The unit price can only be calculated for households that report selling coffee in a given year. This reduces the sample in column 1. Estimates of control variables can be found in Table B.1 in the Appendix. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

indicates that households are not able to perfectly smoothen consumption when faced with price volatility in the coffee market. Similarly, column 3 shows that the index of durable assets is also significantly and positively affected by the international coffee price. Overall, these results indicate that volatility in international prices translates into transitory income changes, affecting household food expenditure and asset ownership.

Lastly, we examine whether international coffee prices affect the probability that coffee-growing households participate in the labor market. These results are presented in column 4. Coffee prices negatively affect the probability that at least one member of the household engages in off-farm wage employment. The effect is economically meaningful; a one standard deviation increase in coffee prices results in a 3.7 percentage point drop in the propensity to engage in wage work. Given a mean of 58 percent, this translates into a 6.3 percent decrease. This strong negative effect indicates that wage work is used as a coping mechanism – in periods when coffee prices are low, household members engage in wage work in order to sustain income. This in turn has implications for the allocation of labor within the household. As the incentives for the distribution of labor within the household change, what is the burden borne by children and adolescents? In order to examine the effects of coffee prices on intra-household labor supply, we now turn to an individual-level analysis.

5.4.2 Coffee prices and individual labor supply

In order to explore individual-level responses to transitory income shocks, we construct three age groups: children (aged 6-14 years); adolescents (aged 15-19); and working age adults (aged 20-54). The children age group is consistent with the age group used to define child labor and the adolescent age group corresponds with the typical age of senior high-school students in Vietnam. We examine intra-household labor complementarities and substitution across three primary activities: off-farm wage employment, farm work, and housework. Equation 2 is estimated separately for each age group and work category and the results are presented in Table 3. Since we see a very low incidence of wage employment among children (1 percent on average), we do not examine this work category for children.

Panel A examines whether coffee price shocks affect the probability of engaging in wage employment. Coffee price shocks strongly affect the probability that adolescents and adults engage in wage employment. We find that a one standard deviation increase in the international price decreases the probability of undertaking wage work by 3.2 percentage points for adolescents and by 6.1 percentage points for adults, corresponding to decreases of 33 percent and 19 percent in the propensity to work, respectively. This provides evidence of countercyclical wage employment among the adolescents and working-age adults such that when coffee prices are high adolescents and working age adults are less likely to engage in wage employment.

Table 3: Coffee prices and intra-household labor responses by age group

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel A: Wage work			
Price/SD		-0.032	-0.058
		(0.007)***	(0.005)***
Mean of dep. var.		0.09	0.30
Panel B: Agricultural work			
Price/SD	-0.045	-0.061	-0.011
	(0.013)***	(0.017)***	(0.004)***
Mean of dep. var.	0.24	0.59	0.74
Panel C: Housework			
Price/SD	-0.000	0.002	0.035
	(0.015)	(0.011)	(0.007)***
Mean of dep. var.	0.53	0.70	0.68
N	2246	1733	6043
Province time trend	YES	YES	YES
HH size, asset index	YES	YES	YES
Shocks	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

Note: Each coefficient represents a separate regression. The panels are named after the outcome variables. Within each age category, the number of observations is the same in all regressions. Standard errors in parentheses are clustered at the commune level. Estimates of control variables can be found in Table B.2 in the Appendix. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

In Panel B, we examine the decision to work on the farm. We find that an increase in coffee prices significantly reduces the probability that children and adolescents work on the farm: A one standard deviation increase in price reduces the probability of doing farm work by 4.5 percentage points for children and by 6.1 percentage points for adolescents. This corresponds to reductions of 19 percent and 10 percent, respectively. Column 3 also indicates a small decline of adults working on the farm (1.1 percentage points, which corresponds to a 1.4 percent decrease). The results indicate that in periods of low prices, adults try to supplement household income by engaging in wage employment while the children substitute for them on the farm. Adolescents are vulnerable to price shocks as well: Their supply of labor to both wage employment and farm work increases.

Finally, we turn to participation in housework in Panel C. We find that engagement of children and adolescents in housework is not sensitive to movements in international coffee prices. On the other hand, adults are more likely to take part in housework when prices are high. This ties in with the results noted for wage employment and farm work above – when coffee prices are high, adults are less likely to spend time outside the house (in wage employment) and are more likely to have time available for housework.

The finding that adults increasingly participate in wage work in the face of low prices differs from Edmonds and Pavcnik (2006), who found that an increase in the rice price increased wage work participation in Vietnam. Whether it is our focus on coffee instead of rice, or the difference in terms of temporal and spatial coverage of the two datasets is not clear. A priori, we expected the effect to be at least in the same direction for the two crops since the large majority of farmers were net-sellers of rice also in the 1990s. Arguably, Edmonds and Pavcnik (2006) may have picked up a specialization effect for rice farmers that does not exist for coffee farmers, as coffee cultivation is already a specialized enterprise.

To summarize, within-household labor decisions are affected by changes in the coffee price. When prices fall, we find increased participation in wage work for adults, and they become less involved in housework. The labor supply of adolescents is also sensitive to changes in the coffee price as we find increased participation rates in both wage work and agriculture. This is a potential problem if it means that adolescents are pulled out of upper school in order to work for a wage and on the farm. Perhaps even more worrying, we find an increase in child labor on the farm when coffee prices are low.

5.4.3 Robustness checks

We now test the robustness of our main results reported in Table 3 to a variety of alternative assumptions.

The primary concern related to identification is that even though the main results include province-specific linear time trends in the analysis, it is possible that other macroeconomic trends, which are not linear in time but rather correlated with the coffee price, are driving the results. This is a possibility given the high correlation between international commodity price series (Deaton and Laroque, 1992; Byrne et al., 2013). We conduct two robustness checks in order to address this concern.

We first use farm-gate prices in place of the international world market coffee price. We have already shown how the farm-gate price is affected by world market price, and this check directly exploits between-household variations in price changes, even though these households were making decisions under the same macroeconomic conditions. The regression we run is the following:

$$y_{jt} = \alpha + \beta p_{it}^{farmgate} + \delta_c t + \lambda x_{it} + \eta_j + \epsilon_{jt} \tag{3}$$

Where $p_{it}^{farmgate}$ is the observed farm-gate price of coffee for household i at time t, divided by the standard deviation of unit prices. A potential issue is that farm-gate prices are endogenous to household choices if quality of coffee varies. The inclusion of individual-fixed effects alleviates this concern to some degree, but if households can influence the farm-gate price, e.g., by adjusting labor supply and other inputs, thereby affecting quality and price, the problem remains. The perennial nature of coffee trees alleviates this concern to some degree: Quality is not singularly determined by inputs over the last year. Note also that farm-gate prices are only available for those household-year observations where the household sold coffee. This means that the sample is somewhat reduced, compared to Table 3.

Results are presented in Table 4. Results are similar to those of the main specification of Table 3: For wage work, we continue to find negative effects for both adolescents and adults. For agricultural work, we continue to find negative effects for children and adolescents. We find no significant effects on housework using this approach.

The second robustness check to address the above concern exploits the information in the survey from nine Vietnamese provinces outside the Central Highlands where coffee is not grown. More specifically, we include non-coffee-growing farmers living in nine provinces outside the Central Highlands in the sample and augment equation (2) by including an interaction of price and a dummy indicator for the Central Highlands ($p_t * CH$):[10]

$$y_{jt} = \alpha + \beta p_t + \gamma p_t * CH + \delta_c t + \lambda x_{it} + \eta_j + \epsilon_{jt} \tag{4}$$

The coefficient on the interaction term now captures the effect of temporal variation in the price of coffee for an individual residing in the Central Highlands, relative to individuals outside the coffee-growing region. This allows us to estimate the additional effect of coffee prices in coffee-growing provinces. The underlying assumption is that other time-varying trends, which correlate with the coffee price and affect household behavior in the same way in both Central Highland provinces and non-Central Highland provinces.

This method improves on the disentangling of the causal effect of coffee price changes from other price changes. We note that this is not necessarily the most important estimate from the perspective of a policy maker in a world that is characterized by co-movements in commodity prices. If the correlation structure between commodity prices is constant, the baseline estimates based on (2) are informative about how households will react when coffee prices change. On the other hand, it should be cause for concern if there is no effect at all in the interaction term of (4).

The results of (4) are presented in Table 5. For wage work, we continue to find a negative effect for adults in the Central Highlands. This supports our maintained hypothesis that coffee prices affect coffee growers directly – and over and above any other macroeconomic trends correlated with the coffee price. Turning to agricultural work in Panel B, we find that relative to children and adolescents in non-coffee-growing

[10] The nine provinces outside the Central Highlands are Lao Cai, Phi Tho, Lai Chau, Dien Bien, Ha Tay, Nghe An, Quang Nam, Khanh Hoa and Long An.

Table 4: Coffee prices and intra-household labor responses by age group using farm-gate prices

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel A: Wage work			
Farm-gate price/SD		-0.042	-0.039
		(0.022)*	(0.019)**
Mean of dep. var.		0.09	0.30
Panel B: Agricultural work			
Farm-gate price/SD	-0.062	-0.146	-0.014
	(0.034)*	(0.032)***	(0.011)
Mean of dep. var.	0.24	0.59	0.74
Panel C: Housework			
Farm-gate price/SD	-0.074	-0.032	0.017
	(0.051)	(0.034)	(0.014)
Mean of dep. var.	0.53	0.70	0.68
N	1842	1454	4991
Province time trend	YES	YES	YES
HH size, asset index	YES	YES	YES
Shocks	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

Note: Each coefficient represents a separate regression. The panels are named after the outcome variables. Within each age category, the number of observations is the same in all regressions. Standard errors in parentheses are clustered at the commune level. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

regions, those in the Central Highlands are less likely to work on the farm when coffee prices increase. This supports the results of the main specification that children and adolescents are increasingly used in agricultural work of coffee-farming households when prices are low. Finally, the estimates on housework in Panel C support our main finding that working-age adults are more likely to participate in housework when prices are high.

Another channel through which bias could be introduced in our main estimates is if transitory shocks to household income affect fertility and/or mortality, thereby altering household size, i.e. the dependency ratio, motivating a reallocation of labor within the household. We test this directly by checking whether the probability of a birth or death in the coffee-growing household responds to changes in the price of coffee. We also test whether household size is affected by the coffee price. None of the three variables are significantly affected by coffee prices. These results are presented in Table A.1 in the Appendix. Further, we control for this in the main regressions by including an indicator for whether a household member died or fell seriously ill during the preceding year, as well as household size and household size squared.

Table 5: Coffee prices and intra-household labor responses by age group using all provinces

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel A: Wage work			
Price/SD*CH		-0.010	-0.019
		(0.007)	(0.006)***
Panel B: Agricultural work			
Price/SD*CH	-0.033	-0.061	-0.001
	(0.014)**	(0.018)***	(0.004)
Panel C: Housework			
Price/SD*CH	0.015	0.001	0.049
	(0.016)	(0.013)	(0.009)***
N	11932	9145	38164
Province time trend	YES	YES	YES
HH size, asset index	YES	YES	YES
Shocks	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

Note: Each coefficient represents a separate regression. CH is an indicator variable equal to one if the household resides in the Central Highlands and ever cultivated coffee. The panels are named after the outcome variables. Within each age category, the number of observations is the same in all regressions. Standard errors in parentheses are clustered at the commune level. Estimates of control variables can be found in Table B.3 in the Appendix. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

Next, we check that our results are robust to a series of alternative specifications of (2). First, one may worry that the province-specific trends are not sufficient to control for potential spurious correlation caused by changes in the underlying macroeconomic variables. We alleviate this concern by adding quadratic province-specific time trends to the estimations of equation (2) and do not find results to be sensitive to this change. These results are presented in Table A.2 in the Appendix.

Second, it might be that outcomes are correlated at a higher level than the commune level. More specifically, if labor markets clear at the district level, the labor market participation decisions may be correlated within the district. In order to account for this, we estimate standard errors clustered at the district level. Since we have 28 districts in our sample, which may be considered a small number for clustering, we employ the clustered wild bootstrap-t method of Cameron et al. (2008) to estimate cluster-robust standard errors. As the results reported in Appendix Table A.3 show, we do not find the significance of our results to differ.

Third, it is possible that slight within-year variations in price could spuriously affect the results. We check the robustness of our results to an alternative definition of price shocks. In Appendix Table A.4, we replace our price variable with a binary indicator equal to one for the survey months where the relevant 12-month moving average price for the month is above the median price. This is the case for the survey months in the years 2008 and 2012. Although this price shock indicator is less precise than our preferred price variable, we find that the effects are qualitatively the same except for a significantly negative estimate on the engagement of children in housework.

Fourth, the province-specific time trends may hide within-province variation if local labor or coffee markets exhibit different trends. One such possibility arises from the fact that traders are often assigned licenses at the district level. We check this by re-estimating our models with district-specific trends (28 districts). Once again, these results, reported in Appendix Table A.5, are both qualitatively and quantitatively comparable to those of the main specification.

To conclude, we find that our main conclusions continue to hold using a wide range of alternative identification strategies and definitions of the key dependent variable providing strong robust evidence on the effect of coffee price volatility on intra-household labor allocation decisions.

5.5 Additional analysis

This section explores the implications of our main results further. First, we investigate whether there is any heterogeneity in response between different groups of households and different types of household members. Second, we explore the implications of the reallocation of labor within the household on the schooling of children and adolescents. Finally, we investigate other mechanisms such as credit and household enterprises that could be used by households to smooth consumption in addition to the main channel of wage labor, which was considered in the previous section.

5.5.1 Heterogeneity in response

We begin by exploring the degree to which household labor allocation responses to international coffee prices vary by the characteristics of the household. Table 6 presents the results of heterogeneity of responses along three dimensions: (i) gender; (ii) household's wealth (as measured by the asset index); and (iii) ethnicity (measured by an indicator variable that takes the value 1 if the household head does not belong to the majority Kinh group).

Panel A in Table 6 shows the heterogeneity of the results for wage employment. We find that wage employment of adult females is more sensitive to fluctuations in coffee prices. Females have a lower overall rate of participation in wage employment (13.5 percent) compared to males (21.1 percent), which is, however, more elastic when faced with economic hardship. Consistent with wage work being a coping mechanism, we find that households with a higher asset index are less likely to have to resort to the labor market to tide over low coffee price periods. Further, adolescents and adults in ethnic minority (i.e., non-Kinh) households are more likely to seek wage employment. This

Table 6: Heterogeneous responses

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel A: Wage work			
Price/SD		-0.031	-0.040
		(0.009)***	(0.007)***
Female*Price/SD		-0.002	-0.036
		(0.011)	(0.011)***
Price/SD		-0.043	-0.063
		(0.008)***	(0.005)***
Asset index*Price/SD		0.024	0.013
		(0.006)***	(0.004)***
Price/SD		-0.017	-0.053
		(0.005)***	(0.006)***
Nonkinh*Price/SD		-0.035	-0.012
		(0.012)***	(0.006)**
Panel B: Agricultural work			
Price/SD	-0.051	-0.066	-0.009
	(0.017)***	(0.019)***	(0.005)*
Female*Price/SD	0.012	0.009	-0.004
	(0.015)	(0.016)	(0.006)
Price/SD	-0.045	-0.053	-0.008
	(0.013)***	(0.015)***	(0.004)**
Asset index*Price/SD	-0.001	-0.018	-0.008
	(0.008)	(0.009)**	(0.003)***
Price/SD	-0.037	-0.063	-0.012
	(0.015)**	(0.023)***	(0.005)**
Nonkinh*Price/SD	-0.015	0.004	0.003
	(0.017)	(0.028)	(0.007)
Panel C: Housework			
Price/SD	-0.025	0.028	0.059
	(0.017)	(0.015)*	(0.009)***
Female*Price/SD	-0.012	-0.034	-0.037
	(0.016)	(0.015)**	(0.011)***
Price/SD	-0.030	0.013	0.043
	(0.015)**	(0.011)	(0.005)***
Asset index*Price/SD	-0.017	-0.003	-0.006
	(0.009)*	(0.007)	(0.004)*
Price/SD	-0.040	0.027	0.039
	(0.017)**	(0.013)**	(0.005)***
Nonkinh*Price/SD	0.016	-0.037	0.004
	(0.016)	(0.022)*	(0.007)
N	2246	1733	6043

Note: Each horizontally bordered segment represents a separate regression. The panels are named after the outcome variables. This is regressed on the coffee price as well as interactions with either an indicator for being female, the asset index, and an indicator for being non-Kinh. All regressions include the same controls as in Table 3. Within each age category, the number of observations is the same in all regressions. Standard errors in parentheses are clustered at the commune level. * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Authors' calculations based on VARHS.

lines up well with the existing literature, which has found that non-Kinh households are systematically worse off compared to Kinh households, including in access to credit (Baulch et al., 2004; World Bank, 2012; Singhal and Beck, 2016). Our results suggest that exclusion of non-Kinh households from alternative coping mechanisms has resulted in increased vulnerability to vagaries of the commodity markets.

Looking at heterogeneity in agricultural work in Panel B, we find that households with a higher asset index are more likely to work on the farm when prices are low. This indicates that households with more assets, including larger landholdings, have a greater reliance on family labor during low price periods. This is possibly explained by a decreased reliance on hired labor in periods of low prices. Households with high levels of asset ownership are also more likely to rely on hired labor (see Figure A.1 in the Appendix). Further, we find that households rely less on hired labor when coffee prices are low (results not reported). Taken together, this suggests that in low price periods, coffee growing households with greater wealth substitute hired labor with unpaid family labor. Finally, heterogeneity of responses in the context of household work is examined in Panel C. While adults are more likely to spend time at home in high price periods, this response is weaker for females.

5.5.2 School attainment

The main results imply that a decline in the international coffee prices leads to greater participation in agriculture (by children and adolescents) and in the labor market (by adolescents and adults). This could adversely affect educational attainment of children and adolescents, both directly (dropping out of school) and indirectly (for example, less time for homework may lead to grade repetition). As discussed earlier, theoretically, the net effect of a price increase on child investment (and labor) is ambiguous due to countervailing income and substitution effects. The existing empirical literature finds mixed results, depending on the country, context and type of shock studied. For example, Cogneau and Jedwab (2012) find investments in children to be pro-cyclical to cocoa prices in Cote d'Ivoire. Beegle et al. (2006) find a similar result of transitory income shocks to farming households in Tanzania and Jacoby and Skoufias (1997) uncover negative income effects on school attendance in the Indian ICRISAT villages. Grimm (2011) also identifies a decline in enrollment in response to a decline in income among Burkinabe households. On the other hand, Kruger (2007) finds evidence of counter-cyclicality to temporary fluctuations in coffee prices in Brazil.

We explore the impacts on education outcomes in Table 7. Once again, as this information is not available for 2006, the period of analysis is restricted to 2008-14. In the first column for each age group, the dependent variable takes the value 1 if the child is "currently attending school". We do not find coffee prices to affect school attendance for either age group. Undeniably, attendance does not necessarily translate into learning. We try to capture educational attainment in two ways. In columns a2 and b2, the

Table 7: Education outcomes

	Ages 7-14 years			Ages 15-19 years		
	(a1)	(a2)	(a3)	(b1)	(b2)	(b3)
	Attending school	Grade	Overage	Attending school	Grade	Overage
Price/SD	-0.000	-0.049	0.005	-0.010	-0.076*	0.019
	(0.006)	(0.059)	(0.006)	(0.013)	(0.043)	(0.013)
Constant	0.790***	2.625*	-0.100	1.246***	6.950***	-0.736**
	(0.189)	(1.513)	(0.142)	(0.324)	(0.974)	(0.326)
Province time trend	Yes	Yes	Yes	Yes	Yes	Yes
HH Size	Yes	Yes	Yes	Yes	Yes	Yes
Shocks	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	1725	1725	1725	1367	1367	1367

All regressions include the same controls as in Table 3. Attending school is a an indicator equal to one if the child is currently attending school. Grade is highest grade attained. Overage is an indicator equal to 1 if the person does not have a normal progress in terms of schooling years for his/her age. Education outcomes are not available for the 2006 survey round. Standard errors in parentheses are clustered at the commune level. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

outcome variable is the highest grade attained by the child. Columns a3 and b3 capture age-grade distortion, similar to the method used by Patrinos and Psacharopoulos (1997) and Dammert (2008) where the outcome variable takes the value 1 if the child is overage relative to her grade (as measured by the index).[11] Once again, for both measures we do not find strong evidence that coffee prices have affected educational attainment; only the grade variable is significant at the 10 percent level for the adolescents group. While these measures may not adequately capture human capital accumulation of children in coffee-growing households exposed to severe price volatility, overall the results indicate that some amount of work done by children on the farm (in response to low prices) may be compatible with schooling.

5.5.3 Other coping mechanisms

Faced with volatility in international coffee prices, we have shown that coffee-growing households in Vietnam primarily rely on wage employment and increased labor of children and adolescents in agriculture in order to smooth consumption. We now examine the extent to which these households use other means to cope with the resulting income fluctuations. Results are reported in Table 8.

Operating a household enterprise can be another channel through which households may attempt to diversify income. For example, Adhvaryu et al. (2015) find enterprise

[11]More precisely, we specify the age-grade distortion as $AGD = grade - (age - 7)$. The overage indicator takes the value of 1 if $AGD < 0$. Children usually start school at the age of 6 years, so if they are older than 7 after having completed the first year of primary school then they are over-age for their grade level.

Table 8: Coffee prices, household enterprises, and loans

	(1) HH enterprise	(2) HH took loan	(3) HH took investment loan	(4) HH took other loan
Coffee price/SD	-0.008	-0.023**	0.014	-0.037***
	(0.007)	(0.010)	(0.012)	(0.011)
Constant	0.154	0.540***	0.087	0.444***
	(0.098)	(0.174)	(0.149)	(0.164)
Province time trend	Yes	Yes	Yes	Yes
HH size	Yes	Yes	Yes	Yes
Shocks	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes
Mean of dep. var.	1.26	1540.09	0.52	0.58
N	2355	2146	2146	2146

Note: Credit information is not available for 2006. This reduces the sample in columns (2)-(4). Standard errors in parentheses are clustered at the commune level. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

ownership increases significantly in periods on low coffee prices in Tanzania. In column 1 of Table 8 we do not find any effect of coffee prices on the propensity to engage in household enterprises.

Our finding that household consumption is affected by coffee prices means that smoothing of consumption is less than perfect. With well-functioning credit markets, this would not be the case. It is therefore relevant to ask whether this is caused by credit constraints. In column 2 of Table 8, we find that the coffee price is negatively correlated with an indicator variable of whether the household took a loan in the two years preceding the interview.[12] This indicates that households attempt to smooth consumption by borrowing money in low price years. Splitting the loan indicator on investment loans (i.e., loans with the stated purpose of building a house, buying land or other assets etc.) and other types of loans (mainly consumption loans) supports this conclusion – investment loans are unaffected by coffee price swings and the whole effect is driven by other types of loans.

Finally, households could also respond to dips in international coffee prices by sending members to work in urban areas. A recent paper found this strategy to be important for rural Vietnamese households in order to cope with the loss of income from a particular destructive typhoon, which hit Vietnam in 2009 (Gröger and Zylberberg, 2016). Whether households employ this strategy in response to less disastrous but still impactful price fluctuations is therefore of interest, especially considering that overall, the level of migration in Vietnam remains low. Unfortunately, we have insufficient information in the VARHS data to investigate this channel. The VARHS survey contains

[12]The sample size is smaller as the 2006 VARHS did not collect information on credit that is strictly comparable with subsequent survey rounds.

information on migrants only for the 2012 and 2014 survey rounds, and we find that the proportion of households where a household member had migrated within the last year was 7.3 percent in 2012 and 13 percent in 2014. This is consistent with migration as a coping response, since coffee prices were higher in 2012 than in 2014. On the other hand, table A.1 showed that the size of the household is not sensitive to the coffee price, which provides indirect evidence that migration is not a major issue for this study.

5.6 Conclusion

In this study, we have established that households in the Central Highlands of Vietnam were unable to smooth consumption during the period 2006-14 where they faced highly volatile international coffee prices. We detected both drawdown of assets as well as increased uptake of credit for consumption, and our results show that households attempted to cope with the loss of income via intra-household labor reallocation. More specifically, we have demonstrated that decreases in the coffee price lead to significant increases in wage employment for both adults and adolescents. We also found substantial and significant increases in employment in agriculture by children and adolescents; and these results are robust to a number of different checks.

Concerning the labor supply of children, our results are similar to Edmonds and Pavcnik (2005), even though the spatial and temporal context is different. Since coffee is not consumed by the household to any substantial degree, this is evidence that the effect from a loss of income dominates any potential substitution effects stemming from changes in relative returns to work and other activities. While we did not find evidence that the uptake of agricultural work in low price periods by children affected school attainment, the potentially debilitating effects of such a coping strategy should not be understated. The quality of the learning may well have suffered and there could be negative consequences in dimensions not assessed here such as health and psychological well-being (Friedman and Thomas, 2009; Miller and Urdinola, 2010; Baird et al., 2013).

Historically, international coffee prices have been quite volatile. From a public policy perspective there is a need for social safety nets that protect household consumption in lean years, and help children and adolescents transition safely to adulthood. Access to insurance schemes or cash transfer programs are some of the policy instruments that could insulate households from volatile commodity markets. For example, while the cash transfer scheme "Red de Protección Social" in Nicaragua was not explicitly designed to respond to shocks, Maluccio (2005) finds that the program protected small coffee farmers from a sharp decline in coffee prices in the early 2000s. Not only could such a scheme reduce incentives for labor substitution patterns between children and adults observed in this study. The impact of coffee price fluctuations on food expenditures shows that households are unable to perfectly smooth consumption through the credit markets. By alleviating household credit constraints, such a scheme would also enable farmers to tend to their own coffee trees instead of seeking off-farm wage work, thereby boosting long-term productivity. It follows that policy makers should also consider improving access to credit markets and reducing imperfections in the land and labor markets more generally.

Vietnam has experienced high growth rates in the last decades and is no longer categorized as a low-income-country. In this light, it is revealing that smallholder farmers continue to rely on the labor market and on child labor to cope with price changes of their outputs. The results of this study suggest that it is likely to be socially beneficial to put in place countervailing policy measures to address future coffee-market volatility. However, our labor substitution results suggest that a public works program approach might be of limited benefit, as it would lead to children spending more time on the family farm. Greater stability and predictability for coffee farmers is desirable for welfare and human capital accumulation and in the final analysis most likely for coffee output as well.

References

Adhvaryu, A., Kala, N., and Nyshadham, A. (2015). Booms, busts, and household enterprise: Evidence from coffee farmers in tanzania. Mimeo.

Baird, S., Hoop, J. d., and Özler, B. (2013). Income Shocks and Adolescent Mental Health. *Journal of Human Resources*, 48(2):370–403.

Basu, K. and Van, P. H. (1998). The Economics of Child Labor. *The American Economic Review*, 88(3):412–427.

Baulch, B., Chuyen, T. T. K., Haughton, D., and Jonathan, H. (2004). Ethnic Minority Development in Vietnam: A Socioeconomic Perspective. In Glewwe, P., Agrawal, N., and Dollar, D., editors, *Economic Growth, Poverty, and Household Welfare in Vietnam*. The World Bank, Washington DC.

Beck, U. (2014). The Price of Growing Coffee. Master's thesis, Department of Economics, University of Copenhagen.

Beegle, K., Dehejia, R., and Gatti, R. (2009). Why Should We Care about Child Labor? The Education, Labor Market, and Health Consequences of Child Labor. *The Journal of Human Resources*, 44(4):871–889.

Beegle, K., Dehejia, R. H., and Gatti, R. (2006). Child labor and agricultural shocks. *Journal of Development Economics*, 81(1):80–96.

Bussolo, M., Godart, O., Lay, J., and Thiele, R. (2007). The impact of coffee price changes on rural households in Uganda. *Agricultural Economics*, 37(2-3):293–303.

Byrne, J. P., Fazio, G., and Fiess, N. (2013). Primary commodity prices: Co-movements, common factors and fundamentals. *Journal of Development Economics*, 101:16–26.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, 90(3):414–427.

CIEM (2007). Characteristics of the Vietnamese Rural Economy: Evidence from a 2006 Rural Household Survey in 12 Provinces of Viet Nam. Research report, Central Institute for Economic Management, Hanoi.

CIEM (2009). Characteristics of the Vietnamese Rural Economy: Evidence from a 2008 Rural Household Survey in 12 Provinces of Viet Nam. Research report, Central Institute for Economic Management, Hanoi.

CIEM (2011). Characteristics of the Vietnamese Rural Economy: Evidence from a 2010 Rural Household Survey in 12 Provinces of Viet Nam. Research report, Central Institute for Economic Management, Hanoi.

CIEM (2013). Characteristics of the Vietnamese Rural Economy: Evidence from a 2012 Rural Household Survey in 12 Provinces of Viet Nam. Research report, Central Institute for Economic Management, Hanoi.

CIEM (2015). Characteristics of the Vietnamese Rural Economy: Evidence from a 2014 Rural Household Survey in 12 Provinces of Viet Nam. Research report, Central Institute for Economic Management, Hanoi.

Cobbe, J. (2011). Education, Education Finance, and the Economy. In London, J. D., editor, *Education in Vietnam*. ISEAS Publishing, Singapore.

Cogneau, D. and Jedwab, R. (2012). Commodity Price Shocks and Child Outcomes: The 1990 Cocoa Crisis in Côte d'Ivoire. *Economic Development and Cultural Change*, 60(3):507–534.

Dammert, A. C. (2008). Child labor and schooling response to changes in coca production in rural Peru. *Journal of Development Economics*, 86(1):164–180.

Danzer, A. M. and Grundke, R. (2014). Export price shocks and rural labor markets: The case of cotton in Tajikistan. Mimeo.

Deaton, A. (1999). Commodity Prices and Growth in Africa. *The Journal of Economic Perspectives*, 13(3):23–40.

Deaton, A. and Laroque, G. (1992). On the Behaviour of Commodity Prices. *The Review of Economic Studies*, 59(1):1–23.

Dercon, S. (2002). Income Risk, Coping Strategies, and Safety Nets. *The World Bank Research Observer*, 17(2):141–166.

D'haeze, D., Deckers, J., Raes, D., Phong, T. A., and Minh Chanh, N. D. (2003). Over-irrigation of Coffea canephora in the Central Highlands of Vietnam revisited: Simulation of soil moisture dynamics in Rhodic Ferralsols. *Agricultural Water Management*, 63(3):185–202.

Duryea, S., Lam, D., and Levison, D. (2007). Effects of economic shocks on children's employment and schooling in Brazil. *Journal of Development Economics*, 84(1):188–214.

Edmonds, E. V. (2007). Child Labor. In Schultz, T. P. and Strauss, J. A., editors, *Handbook of Development Economics*, volume 4, chapter 57, pages 3607–3709. Elsevier.

Edmonds, E. V. and Pavcnik, N. (2005). The effect of trade liberalization on child labor. *Journal of International Economics*, 65(2):401–419.

Edmonds, E. V. and Pavcnik, N. (2006). International trade and child labor: Cross-country evidence. *Journal of International Economics*, 68(1):115–140.

Fafchamps, M. and Hill, R. (2008). Price Transmission and Trader Entry in Domestic Commodity Markets. *Economic Development and Cultural Change*, 56(4):729–766.

FAO, IFAD, IMF, OECD, UNCTAD, WFP, The World Bank, WTO, IFPRI, and UN HLTF (2011). Price Volatility in Food and Agricultural Markets: Policy Responses. Policy report, Coordinated by OECD and UNCTAD.

Field, E. (2007). Entitled to Work: Urban Property Rights and Labor Supply in Peru. *The Quarterly Journal of Economics*, 122(4):1561–1602.

Fitzsimons, E. (2007). The Effects of Risk on Education in Indonesia. *Economic Development and Cultural Change*, 56(1):1–25.

Friedman, J. and Thomas, D. (2009). Psychological Health Before, During, and After an Economic Crisis: Results from Indonesia, 1993-2000. *The World Bank Economic Review*, 23(1):57–76.

Giovannucci, D., Lewin, B., Swinkels, R., and Varangis, P. (2004). The Socialist Republic of Vietnam: Coffee Sector Report. Technical Report 29358-VN, The World Bank.

Glewwe, P. and Jacoby, H. G. (2004). Economic growth and the demand for education: is there a wealth effect? *Journal of Development Economics*, 74(1):33–51.

Goldberg, P. K. and Pavcnik, N. (2007). Distributional Effects of Globalization in Developing Countries. *Journal of Economic Literature*, 45(1):39–82.

Gröger, A. and Zylberberg, Y. (2016). Internal Labor Migration as a Shock Coping Strategy: Evidence from a Typhoon. *American Economic Journal: Applied*, Forthcoming.

Grimm, M. (2011). Does household income matter for children's schooling? Evidence for rural Sub-Saharan Africa. *Economics of Education Review*, 30(4):740–754.

GSO (2010). The 2009 Vietnam Population and Housing Census: Major Findings. Technical report, General Statistics Office, Vietnam, Hanoi.

ICO (2015). Coffee Price Statistics. http://www.ico.org/. Accessed October 1 2015.

ILO (1973). Minimum Age Convention. Convention 138, International Labour Organization.

Jacoby, H. G. and Skoufias, E. (1997). Risk, Financial Markets, and Human Capital in a Developing Country. *The Review of Economic Studies*, 64(3):311–335.

Klasen, S., Priebe, J., and Rudolf, R. (2013). Cash crop choice and income dynamics in rural areas: evidence for post-crisis Indonesia. *Agricultural Economics*, 44(3):349–364.

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

Kruger, D. I. (2007). Coffee production effects on child labor and schooling in rural Brazil. *Journal of Development Economics*, 82(2):448–463.

Lederman, D. and Porto, G. (2016). The Price Is Not Always Right: On the Impacts of Commodity Prices on Households (and Countries). *The World Bank Research Observer*, 31(1):168–197.

Luong, Q. V. and Tauer, L. W. (2006). A real options analysis of coffee planting in Vietnam. *Agricultural Economics*, 35(1):49–57.

Maluccio, J. A. (2005). Coping with the "Coffee Crisis" in Central America: The Role of the Nicaraguan Red de Proteccion Social. FCND Discussion Paper 188, IFPRI.

Markussen, T., Fibæk, M., Tarp, F., and Tuan, N. D. A. (2014). The happy farmer: Self-employment and subjective well-being in rural Vietnam. WIDER Working Paper 2014/108, UNU-WIDER.

Miller, G. and Urdinola, B. P. (2010). Cyclicality, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Colombia. *The Journal of Political Economy*, 118(1):113–155.

MoLISA (2013). Promulgating List of Light Works Allowed Using Persons under 15 Years Old to Work. Circular 11/2013/TT-BLDTBXH, Ministry of Labour, Invalids and Social Affairs, Vietnam.

Patrinos, H. A. and Psacharopoulos, G. (1997). Family Size, Schooling and Child Labor in Peru: An Empirical Analysis. *Journal of Population Economics*, 10(4):387–405.

Paxson, C. H. (1992). Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand. *The American Economic Review*, 82(1):15–33.

Rose, E. (2001). Ex ante and ex post labor supply response to risk in a low-income area. *Journal of Development Economics*, 64(2):371–388.

Shukla, G. P., Pham, D. M., Engelschalk, M., and Minh Le, T. (2011). *Tax Reform in Vietnam: Toward a More Efficient and Equitable System*. World Bank, Washington, DC.

Singhal, S. and Beck, U. (2016). Ethnic Disadvantage in Vietnam: Evidence using Panel Data. In Tarp, F., editor, *Growth, Structural Transformation and Rural Change in Viet Nam: A Rising Dragon on the Move*. Oxford University Press, Oxford. Forthcoming.

World Bank (2012). Well Begun, Not Yet Done: Vietnams Remarkable Progress on Poverty Reduction and the Emerging Challenges. Poverty assessment, The World Bank.

World Bank (2015). World Development Indicators 2015. Accessed December 15, 2015.

Appendix A: Additional tables and figures

Table A.1: Do coffee prices affect household size?

	(1) Birth	(2) Death/serious illness	(3) HH size
Coffee price/SD	0.006	-0.009	0.006
	(0.004)	(0.010)	(0.011)
Constant	0.036	0.188**	5.072***
	(0.035)	(0.082)	(0.112)
Province time trend	Yes	Yes	Yes
HH size	Yes	Yes	Yes
Shocks	Yes	Yes[1]	Yes
HH Fixed Effects	Yes	Yes	Yes
N	1922	2355	2355

Note: Coffee price is both in real prices. Standard errors in parantheses are clustered at the commune level. Births are not observable in 2006. This leads to a reduction in the sample of column 1. * significant at 10%; ** significant at 5%; *** significant at 1%.
1: Only includes pest and natural disaster shocks.
Source: Authors' calculations based on VARHS.

Table A.2: Main results with linear and quadratic time trends

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel A: Wage work			
Price/SD		-0.033	-0.064
		(0.007)***	(0.005)***
Panel B: Agricultural work			
Price/SD	-0.053	-0.060	-0.012
	(0.013)***	(0.017)***	(0.004)***
Panel C: Housework			
Price/SD	-0.000	0.007	0.040
	(0.015)	(0.011)	(0.006)***
N	2246	1733	6043
Province time trend	YES	YES	YES
Sq. province time trend	YES	YES	YES
HH size, asset index	YES	YES	YES
Shocks	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

Note: Each coefficient represents a separate regression. The panels are named after the outcome variables. Within each age category, the number of observations is the same in all regressions. Standard errors in parentheses are clustered at the commune level. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

Table A.3: Main results with clustering at the district level

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel A: Wage work			
Price/SD		-0.032	-0.058
		(0.000)***	(0.000)***
		[0.002]***	[0.002]***
Panel B: Agricultural work			
Price/SD	-0.045	-0.061	-0.011
	(0.011)**	(0.009)***	(0.006)***
	[0.036]**	[0.024]**	[0.014]**
Panel C: Housework			
Price/SD	-0.000	0.002	0.035
	(0.986)	(0.875)	(0.000)***
	[0.950]	[0.856]	[0.000]***
N	2246	1733	6043
Province time trend	YES	YES	YES
Sq. province time trend	YES	YES	YES
HH size, asset index	YES	YES	YES
Shocks	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

Note: Standard errors are clustered at district-level. Normal parentheses report clustered p-values. Square brackets report p-values based on 999 iterations of the clustered wild bootstrap-t procedure of Cameron et al. (2008). * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

Table A.4: Main results using indicator variable for high price years

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel a: Wage work			
High price year (=1)		-0.055	-0.084
		(0.015)***	(0.014)***
Panel B: Agricultural work			
High price year (=1)	-0.126	-0.153	-0.030
	(0.025)***	(0.042)***	(0.009)***
Panel C: Housework			
High price year (=1)	-0.078	-0.029	0.050
	(0.032)**	(0.028)	(0.012)***
N	2246	1733	6043
Province time trend	YES	YES	YES
HH size, asset index	YES	YES	YES
Shocks	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

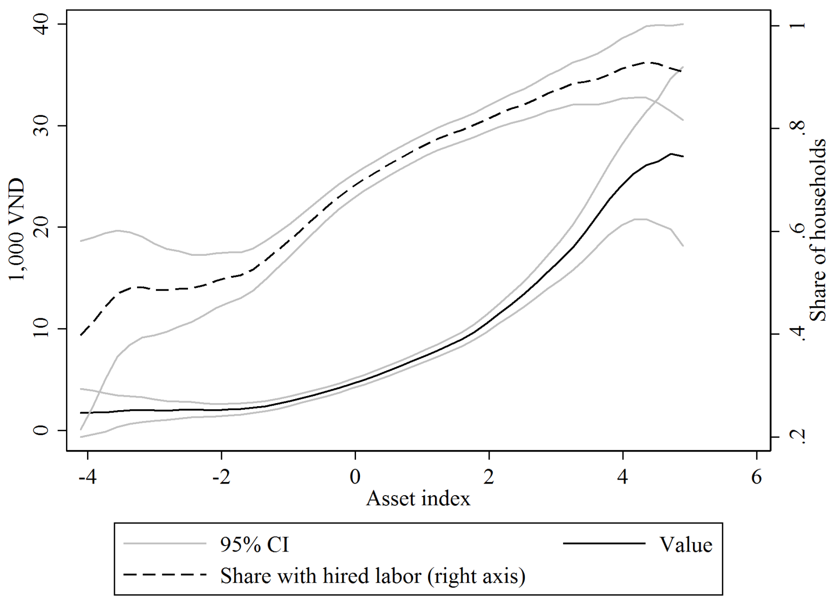
Note: Each coefficient represents a separate regression. The panels are named after the outcome variables. Within each age category, the number of observations is the same in all regressions. Standard errors in parentheses are clustered at the commune level. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

Table A.5: Main results using district time trends

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel A: Wage work			
Price/SD		-0.032	-0.053
		(0.007)***	(0.006)***
Panel B: Agricultural work			
Price/SD	-0.047	-0.062	-0.013
	(0.013)***	(0.017)***	(0.004)***
Panel C: Housework			
Price/SD	-0.004	0.001	0.034
	(0.016)	(0.012)	(0.008)***
N	2246	1733	6043
District time trend	YES	YES	YES
HH size, asset index	YES	YES	YES
Shocks	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

Note: Each coefficient represents a separate regression. The panels are named after the outcome variables. Within each age category, the number of observations is the same in all regressions. Standard errors in parentheses are clustered at the commune level. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

Figure A.1: Relationship between hired labor and asset index



Note: Kernel-weighted local polynomial regressions. The solid line is the local average value of hired agricultural labor over the last year in real 2014 Vietnamese Dong. To avoid undue influence of outliers, values above the 99th percentile were replaced with the value of the 99th percentile. The dashed line is the local share of households who have used agricultural wage labor in the last year. Eight outlier observations with asset index values above 5 were discarded before running these regressions.
Source: Authors' calculations based on VARHS

Appendix B: Control variables of main specifications

Table B.1: Coffee prices and household welfare with control variables

	(1) HH unit price/SD	(2) Food Exp.	(3) Asset index	(4) Wage work
Coffee price/SD	0.128***	47.953***	0.049**	-0.037***
	(0.008)	(16.752)	(0.020)	(0.008)
Constant	0.088	-343.564	-2.243***	0.358***
	(0.122)	(266.773)	(0.254)	(0.132)
Natural disaster (=1)	-0.046	10.532	-0.008	-0.047
	(0.033)	(45.426)	(0.042)	(0.031)
Serious ilness or death	-0.007	121.740*	0.035	-0.010
of HH member (=1)	(0.063)	(63.656)	(0.079)	(0.035)
Pest attack (=1)	0.017	68.117	0.108**	0.056***
	(0.025)	(56.013)	(0.042)	(0.019)
HH size	0.050	275.860***	0.334***	0.126***
	(0.035)	(75.728)	(0.059)	(0.033)
(HH size)^2	-0.003	-5.946	-0.013***	-0.005**
	(0.002)	(6.458)	(0.004)	(0.002)
Province time trend	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes

Note: This table shows the results of Table 2, but including parameter estimates of control variables. Coffee price and HH unit price are both in real prices. Food expenditure is in real 2014 '000 VND. Wage work is an indicator variable equal to one if the household participated in wage work in the last year. Standard errors are clustered at commune-level. The unit price can only be calculated for households that report selling coffee in a given year. This reduces the sample in column 1. * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Authors' calculations based on VARHS.

Table B.2: Coffee prices and intra-household labor responses by age group with control variables

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel a: Wage work			
Price/SD		-0.032	-0.058
		(0.007)***	(0.005)***
Natural disaster (=1)		0.018	-0.019
		(0.031)	(0.022)
Serious ilness or death		0.087	-0.027
of HH member (=1)		(0.042)**	(0.022)
Pest attack (=1)		0.014	0.035
		(0.026)	(0.016)**
Asset index		0.013	-0.002
		(0.010)	(0.007)
HH size		0.077	0.062
		(0.038)**	(0.022)***
(HH size)^2		-0.004	-0.001
		(0.003)*	(0.002)
Panel b: Agricultural work			
Price/SD	-0.045	-0.061	-0.011
	(0.013)***	(0.017)***	(0.004)***
Natural disaster (=1)	0.012	0.066	-0.001
	(0.035)	(0.038)*	(0.009)
Serious ilness or death	-0.032	0.042	-0.011
of HH member (=1)	(0.069)	(0.060)	(0.016)
Pest attack (=1)	0.072	0.099	0.016
	(0.029)**	(0.035)***	(0.012)
Asset index	-0.001	-0.009	0.006
	(0.015)	(0.019)	(0.005)
HH size	0.047	0.283	0.172
	(0.049)	(0.063)***	(0.023)***
(HH size)^2	-0.003	-0.013	-0.008
	(0.004)	(0.004)***	(0.002)***
Panel c: Housework			
Price/SD	-0.000	0.002	0.035
	(0.015)	(0.011)	(0.007)***
Natural disaster (=1)	0.045	0.031	-0.011
	(0.042)	(0.041)	(0.019)
Serious ilness or death	-0.092	0.053	-0.055
of HH member (=1)	(0.047)*	(0.054)	(0.022)**
Pest attack (=1)	0.092	0.077	-0.008
	(0.039)**	(0.031)**	(0.016)
Asset index	0.004	0.000	0.009
	(0.016)	(0.014)	(0.007)
HH size	0.163	0.223	0.154
	(0.046)***	(0.081)***	(0.017)***
(HH size)^2	-0.013	-0.010	-0.006
	(0.003)***	(0.006)*	(0.001)***
Province time trend	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

Note: This table shows the results of Table 3, but including parameter estimates of control variables. Each coefficient represents a separate regression. The panels are named after the outcome variables. Within each age category, the number of observations is the same in all regressions. * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Authors' calculations based on VARHS.

Table B.3: Coffee prices and intra-household labor responses by age group using all provinces

	Ages 6 to 14 years	Ages 15 to 19 years	Ages 20 to 54 years
Panel a: Wage work			
Price/SD*CH		-0.010	-0.019
		(0.007)	(0.006)***
Price/SD		-0.024	-0.040
		(0.003)***	(0.003)***
Natural disaster (=1)		-0.010	0.003
		(0.011)	(0.009)
Serious ilness or death		0.034	0.007
of HH member (=1)		(0.020)	(0.010)
Pest attack (=1)		0.004	0.013
		(0.010)	(0.007)*
Asset index		0.003	0.011
		(0.005)	(0.003)***
HH size		0.044	0.086
		(0.011)***	(0.007)***
(HH size)^2		-0.002	-0.004
		(0.001)***	(0.000)***
Panel b: Agricultural work			
Price/SD*CH	-0.033	-0.061	-0.001
	(0.014)**	(0.018)***	(0.004)
Price/SD	-0.013	-0.009	-0.009
	(0.006)**	(0.005)*	(0.002)***
Natural disaster (=1)	0.041	0.014	0.008
	(0.015)***	(0.019)	(0.006)
Serious ilness or death	-0.003	0.014	-0.003
of HH member (=1)	(0.021)	(0.025)	(0.007)
Pest attack (=1)	0.005	0.017	0.010
	(0.014)	(0.015)	(0.005)*
Asset index	0.006	-0.002	-0.005
	(0.007)	(0.008)	(0.003)*
HH size	0.018	0.174	0.105
	(0.010)*	(0.023)***	(0.008)***
(HH size)^2	-0.000	-0.008	-0.003
	(0.001)	(0.002)***	(0.001)***
Panel c: Housework			
Price/SD*CH	0.015	0.001	0.049
	(0.016)	(0.013)	(0.009)***
Price/SD	-0.014	-0.004	-0.010
	(0.005)***	(0.006)	(0.003)***
Natural disaster (=1)	0.072	0.034	0.018
	(0.020)***	(0.019)*	(0.009)*
Serious ilness or death	-0.042	0.010	-0.011
of HH member (=1)	(0.021)*	(0.027)	(0.009)
Pest attack (=1)	0.073	0.024	0.015
	(0.014)***	(0.016)	(0.006)**
Asset index	0.003	-0.004	0.009
	(0.007)	(0.008)	(0.003)***
HH size	0.079	0.203	0.122
	(0.015)***	(0.020)***	(0.007)***
(HH size)^2	-0.004	-0.009	-0.004
	(0.001)***	(0.001)***	(0.001)***
Province time trend	YES	YES	YES
Individual Fixed Effects	YES	YES	YES

Note: This table shows the results of Table 5, but including parameter estimates of control variables and the uninteracted price variable. Each coefficient represents a separate regression. CH is an indicator variable equal to one if the household resides in the Central Highlands and ever cultivated coffee. The panels are named after the outcome variables. Within each age category, the number of observations is the same in all regressions. * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Authors' calculations based on VARHS.