

Aus dem Institut für Ernährungswirtschaft und Verbrauchslehre
der Christian-Albrechts-Universität zu Kiel

Shocks, Income Diversification and Welfare in Developing and Transition Countries

Dissertation

zur Erlangung des Doktorgrades

der Agrar- und Ernährungswissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

M.Sc. Toman Omar Mahmoud

aus Dannenberg an der Elbe

Kiel, 2010

Dekanin

Prof. Dr. Karin Schwarz

Erster Berichterstatter

Prof. Dr. Awudu Abdulai

Zweiter Berichterstatter

Prof. Dr. Christian Henning

Tag der mündlichen Prüfung

15. Juli 2010

Gedruckt mit Genehmigung
der Agrar- und Ernährungswissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Diese Arbeit kann als pdf-Dokument unter
<http://eldiss.uni-kiel.de/>
aus dem Internet geladen werden.

Table of Contents

| | |
|---|-----|
| Summary | 1 |
| Introduction | 3 |
| Does AIDS-Related Mortality Reduce Per-Capita Household Income? Evidence from Rural Zambia..... | 9 |
| Orphanhood and Critical Periods in Children’s Human Capital Formation: Long-Run Evidence from North-Western Tanzania | 30 |
| Few Opportunities, Much Desperation: The Dichotomy of Non-Agricultural Activities and Inequality in Western Kenya | 60 |
| Shocks, Structural Change, and the Patterns of Income Diversification in Burkina Faso | 86 |
| Explaining Labor Market Inactivity in Migrant-Sending Families: Housework, Hammock, or Higher Education?..... | 105 |
| The Economics of Human Trafficking and Labor Migration: Micro-Evidence from Eastern Europe | 127 |
| Conclusion..... | 157 |

Summary

Rural life in developing and transition countries is prone to risks. Unanticipated shocks such as droughts, fluctuating prices or diseases pose a heavy burden on rural welfare and frequently push rural households into poverty. Households, however, do not just undergo shocks. In the absence of functioning formal insurance and credit markets, they have developed a remarkable set of strategies to manage risks ex-ante and cope with shocks ex-post. This dissertation takes a welfare perspective to analyze how rural households adapt their livelihoods to the multitude of risks and shocks.

The first part looks at risk-coping strategies and analyzes the socio-economic consequences of AIDS-related mortality in rural sub-Saharan Africa. Chapter 2 shows that households in rural Zambia are able to stabilize their per-capita incomes after the death of a prime-age member. Adjustments of income-generating activities as well as household size and composition are likely to explain this finding. Yet, risk-sharing arrangements appear to spread the burden beyond directly afflicted households. Households that have not lost a member but support afflicted households by fostering orphans incur considerable reductions in per-capita incomes. Overall, however, the impact of AIDS-related mortality on monetary welfare appears to be rather limited. Nevertheless, the death of an adult member could still affect other dimensions of household welfare. Chapter 3 demonstrates for North-Western Tanzania that the age at parental bereavement has important consequences for children's long-term capital accumulation in terms of both health and education. These effects, however, depend on the gender of the deceased parent. Preferences of the surviving parent partly protect same-sex children from the detrimental effects of orphanhood, suggesting that risks are not shared equally within households.

The second part of the dissertation explores risk-management strategies and investigates the income diversification patterns of farm households in sub-Saharan Africa. Chapter 4 analyzes the dichotomy of the non-agricultural sector in Western Kenya and the resulting poverty and inequality implications. The results show that only rich households are able to overcome the entry barriers into high-return activities. Low-return activities, however, are not concentrated among the poor. They are pursued by households across the entire income distribution, possibly reflecting the high risks associated with high-return activities. The chapter also provides evidence that engagement in high-return, but not low-return non-agricultural activities is associated with increased agricultural productivity. Apparently, high-return activities play an important role in triggering positive interactions between different income strategies and may hence contribute to sustainable livelihood success. Chapter 5 examines the determinants of diversification in Burkina Faso over the shock-prone period between 1994 and 2003. Diversification into non-agricultural activities appears to be motivated by insurance motives. During the severe drought in 1997/1998 households earned increased incomes from migration and local non-agricultural activities. The poorest households were hit particularly hard, often being forced to sell livestock. Yet, the

patterns of diversification also reflect structural change offering better opportunities in the non-farm sector.

The final part of the dissertation concentrates on Eastern Europe and looks at the welfare implications of international migration for those who stay behind and the migrants themselves. Chapter 6 examines the reasons for reduced labor supply of migrant-sending households in Moldova. The findings do not support the common view that decreased labor market activity is the result of remittances-driven leisure consumption. Instead, the departure of a migrant appears to raise remaining members' productivity in home production. In addition, young adults in migrant families are substantially more likely to pursue higher education. This effect may be due to remittances relieving credit constraints or migration-induced incentives for additional investments in human capital. Chapter 7 takes into account that migration itself is a very risky activity. It investigates the economic drivers of human trafficking, which refers to a situation in which a migrant has been recruited under false promises and is forced to work for no or little pay by means of coercion. Based on a household survey on human trafficking from Belarus, Bulgaria, Moldova, Romania and Ukraine, the analysis finds that the individual risk of falling victim to human trafficking increases with the size of regional migration flows. The reasons are lower recruitment costs for traffickers and, to a lesser extent, more negative self-selection into migration.

Together, these findings illustrate that rural households in the developing world have adapted their livelihoods in various ways to cope with shocks and reduce their exposure to risks. While these strategies help households to temporarily smooth income or consumption, they are likely to perpetuate poverty and reduce economic growth. Coping strategies such as the depletion of productive assets or the withdrawal of children from school have detrimental and persistent effects on future welfare and vulnerability. Likewise, diversification into low-risk low-return activities prevents capital accumulation or savings. Consequently, measures that protect households from risks, e.g. microinsurance schemes, should play a prominent role in rural development strategies.

Chapter 1

Introduction

Rural households in developing and transition countries face many risks. Climatic shocks such as drought, flooding, or pestilence can substantially reduce crop and livestock income. Economic shocks, in particular fluctuating input and output prices, add to the vulnerability of rural households' income base. Besides, rural livelihoods are subject to political and institutional shocks like conflict or arbitrary treatment by government authorities. In transition countries, the change in production systems characterized by market liberalization and the reallocation of productive assets has also severely impacted on rural livelihoods. Finally, there are several household-specific types of shocks: Illness, death, or crime can impose a heavy burden on a household's income-generating activities. Both the frequency and severity of shocks have not been declining over the past decades. Rather, globalization and the increasing interdependencies of previously segmented markets, the spread of diseases like HIV/AIDS as well as global warming keep risks in rural areas high (Dercon, 2005).

Unlike in developed countries, formal insurance and credit markets are often incomplete or missing altogether in developing countries, in particular in rural areas. Thus, households have to fall back on informal measures to smooth their income or consumption (Murdoch, 1995). These measures basically fall into two categories. First, households can manage risks *ex-ante*. By choosing income portfolios with low positive covariances between different income-generating activities, they are able to mitigate the adverse impacts of shocks before they actually occur. For instance, farmers could engage in non-agricultural activities including migration to reduce their exposure to agricultural risks. Second, households can cope with risks *ex-post*. Various coping strategies, such as asset depletion, changes in the allocation and supply of labor as well as informal borrowing within the extended family or community, help shock-hit households to temporarily stabilize their consumption levels.

Overall, risks can be classified as covariate or idiosyncratic. Covariate risks such as drought or commodity price shocks impinge on the total population of a region or community. By contrast, idiosyncratic risks like sickness only affect particular individuals or households of the population under consideration. In general, idiosyncratic risks are insurable within a community because not every member would be affected at the same time and require transfers (Townsend, 1994). Covariate risks, however, are not compatible with community insurance schemes. The inability to pool risks resulting from the fact that covariate shocks hit all members simultaneously renders these schemes bound to fail (Murdoch, 1999). In such cases, households have to rely on transfers from outside of the community, e.g. remittances sent by migrants or on personal savings including the sale of assets to smooth their consumption. Nevertheless, in case of severe economy-wide shocks the effectiveness of self-insurance through asset depletion might also be

limited. Asset markets may poorly function as the number of asset sellers is likely to exceed the number of asset buyers by far, thus depressing asset prices (Dercon, 2005).

Similarly, ex-ante strategies to manage risks have their limits. Poor farmers who want to diversify into profitable non-agricultural activities face considerable entry constraints (Dercon and Krishnan, 1996). In addition, high-return activities often exhibit high risks. As a result, households may have to resort to low-risk and low-return activities and forego income in order to reduce the overall risk exposure of their income portfolio (Rosenzweig and Wolpin, 1993).

On the whole, rural households in developing and transition countries have developed a remarkable set of strategies to deal with risks and shocks in the absence of functioning formal insurance and credit markets. While both risk management and risk coping reduce households' vulnerability to severe hardship, however, they cannot completely shield from risks (Dercon, 2005).

Against this general backdrop, this dissertation takes a welfare perspective to analyze how rural households cope with shocks and adapt their livelihoods to the multitude of risks. The first part looks into the consequences of AIDS-related mortality on various dimensions of household welfare in rural sub-Saharan Africa. The second part analyzes income diversification patterns of rural households in sub-Saharan Africa. The third and final part concentrates on Eastern Europe and examines the welfare implications of migration for household members left behind and the migrants themselves.

HIV/AIDS poses a relatively new and severe health risk. Since the arrival of the virus, HIV prevalence rates have increased dramatically in many developing countries. Sub-Saharan Africa is particularly affected and has seen AIDS become the prime cause of adult death within just two decades (WHO, 2008). In contrast to other common diseases, AIDS first and foremost hits prime-age adults. It thus kills people at the peak of their productivity and earning capacity and disrupts many families.

Chapter 2 evaluates the impact of AIDS-related mortality on per-capita incomes of surviving household members. The loss of a prime-age adult is likely to reduce total household income. Yet, the effect on *per-capita* income is not clear a priori as the number of consuming members in the household is also reduced. *Ceteris paribus*, the direction of the effect depends on whether the deceased adult was a net producer or net consumer for the household. In addition, as discussed above, households do not just undergo shocks; they try to mitigate adverse welfare impacts. To stabilize per-capita incomes, afflicted households may adjust their income-generating activities or change the composition and size of the household. For instance, they could work longer and harder hours to substitute for the labor previously supplied by the late member. Otherwise, they could call migrants back to the household to increase the number of net producers or reduce the number of net consumers by sending children away to members of the extended family.

A major problem in identifying the causal impact of AIDS-related mortality on household welfare is that health shocks are typically not distributed randomly across the population. With

HIV being primarily transmitted sexually, AIDS is a behavioral disease and may thus cause endogenous selection into prime-age mortality. Another problem is that risk-sharing agreements between households may blur the distinction between the treatment and control group. To the extent that non-afflicted households support afflicted households by sending adults or fostering children, the welfare implications of AIDS-related deaths are likely to be underestimated. These problems are dealt with by using difference-in-difference matching techniques and excluding households from the control group if members departed or joined for reasons related to AIDS. The analysis is based on a large nationally representative longitudinal survey from Zambia that tracks rural households over the period 2001-2004.

Chapter 3 considers some non-monetary consequences of prime-age mortality. Building on the finding that losing a parent has detrimental effects on human capital accumulation, it aims to identify critical periods in child development by examining how the impact of parental loss differs with age at bereavement. There is ample theoretical ground to expect impact heterogeneity across the age at which a child loses a parent. The impacts of psychological trauma, income loss, and changed preferences of the (new) caregivers are all likely to vary with children's age.

The problem of unobserved heterogeneity is addressed by controlling for a wide range of pre-orphanhood characteristics and demonstrating that *future* age at bereavement is not associated with baseline outcomes. An alternative identification strategy is based on family fixed effects and exploits the exogenous variation created by the fact that siblings experience parental death simultaneously, but at different ages. The analysis deals with both dimensions of human capital, health and education, and takes a long-run perspective. It is based on a longitudinal household survey from the predominantly rural Kagera region in North-Western Tanzania that spans thirteen years from 1991 to 2004.

While the first part of this dissertation is concerned with risk-coping, the second part deals with risk-management strategies of rural households. As argued above, income diversification can help households to reduce their vulnerability to income shocks. In rural areas, many farmers diversify into non-agricultural activities. Such activities include both local self- or wage employment and migration. Locally employed members directly contribute non-agricultural income to their households' budget, whereas migrants usually send remittances on a regular base or help out when a shock hits their households of origin. Under normal circumstances, agricultural incomes are not highly correlated with income from local non-agricultural employment. During severe crisis, however, agricultural and non-agricultural incomes may move together as a result of demand and supply linkages in spatially segmented markets (Czukas et al., 1998). Thus, the effectiveness of diversification might be reduced when it is needed most. By contrast, remittances, in particular from urban or international migrants, are unlikely to be subject to the same shocks and may hence provide a more reliable form of insurance in case of emergency (Rapoport and Docquier, 2006).

Chapter 4 analyzes the dichotomy of the local non-agricultural sector and the resulting poverty and inequality implications. Apart from insurance motives, rural households' diversification patterns can be looked at as survival- or opportunity-led. Poor rural households without a sufficient agricultural asset base may be forced to diversify into non-agricultural activities to sustain a living. Such activities are characterized by low returns, often below those in agriculture. By contrast, richer households may use their higher asset endowments to enter profitable niches in the non-agricultural sector. These activities typically exhibit entry barriers which the poor are not able to overcome.

The resulting dichotomy of the non-agricultural sector implies a U-shaped relationship between household wealth and the share of income earned from non-agricultural employment. Survival-led diversification should increase the incomes of the poor and hence decrease inequality. The reverse should be true for opportunity-led diversification. These considerations are tested using a household survey from Kakamega district, a poor rural region in Western Kenya.

Chapter 5 investigates the changing patterns and determinants of income diversification of farm households in Burkina Faso between 1994 and 2003. Burkina Faso is an exceptionally shock-prone country. During the period under consideration, the country experienced two severe droughts (1997/1998 and 2001) and a political crisis in Côte d'Ivoire (2002), the main destination of Burkinabe migrants. Another though rather positive shock for rural households was the CFA Franc devaluation in 1994. The analysis relies on three nationally representative household surveys from 1994, 1998, and 2003. Thus, the surveys were well-timed to capture the effects of adverse shocks on households' income portfolios.

The last part of this dissertation is dedicated to international migration, with a focus on transition countries in Eastern Europe. The transition from a centrally planned to a market economy has fundamentally changed the livelihoods of rural households in the former Soviet Union. Above all, the privatization of agricultural assets and the liberalization of agricultural markets in combination with the abolition or reform of previously existing safety nets have (temporarily) increased insecurity and vulnerability of many farmers. These circumstances as well as the (re-)integration with Western Europe and Russia's economic crisis in 1998 have induced many households to send a migrant abroad (Mansoor and Quillin, 2007).

Chapter 6 examines the reasons for reduced labor market activity of household members left behind by labor migrants. The receipt of remittances has often been argued to lower the work effort and labor supply of migrant-sending households. However, labor market inactivity is not necessarily due to remittances-driven leisure consumption. The departure of a migrant implies that two helping hands might be missing for household duties, child care or work on the household farm. Thus, individuals from migrant families might substitute for the labor of the migrants if they were previously engaged in home production. In addition, younger adults in migrant-sending households could be more likely to engage in higher education. This effect may be due to remittances relieving credit constraints or additional incentives for education as returns

to education are often higher abroad and migrants directly transmit information on jobs in their destinations.

The relative importance of these explanations is assessed with a recent household survey from Moldova. Because of self-selection into migration, any observed differences in inactivity patterns between migrant and non-migrant households might have existed already before migration. This concern is addressed by using a novel instrumental-variable approach that exploits the substantial diversity of first emigrants' destination choices. The instrument is based on the degree to which initial regional migrant networks in the different destination countries differentially exposed Moldovan regions to economic performance and job opportunities abroad, thus affecting the subsequent departure of new migrants.

Chapter 7 focuses on the welfare of the migrant, taking into account that due to the enormous information asymmetries involved migration itself is a very risky activity. It analyses the economic drivers of human trafficking, which refers to a situation in which a migrant has been recruited under false promises and is forced to work for no or little pay by means of coercion. Arguably, human trafficking is the worst possible outcome of migration. Trafficked migrants experience a dramatic loss of material and non-material welfare and are typically unable to send remittances to their households. The main hypothesis is that human trafficking is closely interwoven with regional migration flows. The departure of many migrants from a given area is likely to cause adverse cumulative effects on both the demand side (criminal agents recruiting for employers abroad) and the supply side (vulnerable individuals willing to work abroad) of the human trafficking market, thus increasing the risk of being trafficked. In particular, mass migration from a region may lower recruitment costs for traffickers and increase the degree of negative self-selection into migration. The analysis aims to disentangle the role of demand- and supply-side factors and builds on a household survey on human trafficking from Belarus, Bulgaria, Moldova, Romania, Ukraine and an additional household survey on migration from Moldova.

Literature

- Czukas, K., M. Fafchamps and C. Udry (1998). Drought and Saving in West Africa: Are Livestock a Buffer Stock? *Journal of Development Economics*, 55(2): 237-305.
- Dercon, S. (Ed.) (2005). *Insurance Against Poverty*. Oxford: Oxford University Press.
- Dercon, S., and P. Krishnan (1996). Income Portfolios in Rural Ethiopia and Tanzania: Choices and Constraints. *Journal of Development Studies*, 32(6): 850-875.
- Mansoor, A., and B. Quillin (2007). *Migration and Remittances: Eastern Europe and the former Soviet Union*. Washington D.C.: The World Bank.
- Murdoch, J. (1999). Between the State and the Market: Can Informal Insurance Patch the Safety Net? *World Bank Research Observer*, 14(2): 187-207.
- Murdoch, J. (1995). Income Smoothing and Consumption Smoothing. *Journal of Economic Perspectives*, 9(3): 103-114.
- Rapoport, H., and F. Docquier (2006). The Economics of Migrants' Remittances. In S.C. Kolm and J. Mercier Ythier (Eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*. Amsterdam: North Holland.
- Rosenzweig, M., and K. Wolpin (1993). Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India. *Journal of Political Economy*, 101(2): 223-244.
- Townsend, R.M. (1994). Risk and Insurance in Village India. *Econometrica*, 62(3): 539-591.
- WHO (2008). *The Global Burden of Disease: 2004 Update*. Geneva: WHO.

Chapter 2

Does AIDS-Related Mortality Reduce Per-Capita Household Income? Evidence from Rural Zambia¹

Toman Omar Mahmoud²
Kiel Institute for the World Economy

Rainer Thiele
Kiel Institute for the World Economy

Abstract

Without knowledge about the causal impact of HIV/AIDS on poverty, the benefits of current programs to fight the epidemic cannot be fully assessed. Yet, evidence on the impoverishing effect of HIV/AIDS is still scarce. This paper evaluates the impact of AIDS-related mortality on per-capita incomes of surviving household members, using a nationally representative sample of rural households from Zambia. Using difference-in-difference matching techniques, we find that the death of a prime-age member has no significant effect on per-capita household income. This result continues to hold when we control for spillover effects by excluding households from the control group if members departed or joined for reasons related to AIDS. However, non-afflicted households hosting orphans incur considerable reductions in per-capita incomes. A likely explanation for the limited income effects is that surviving household members pursue a mix of income and demographic coping strategies that stabilize income in the short to medium run.

Keywords: HIV/AIDS, prime-age mortality, per-adult equivalent income, spillovers, difference-in-difference, propensity score matching, Zambia

JEL classification: I31, J19, C14, C23

¹ We are highly indebted to Thomas S. Jayne from Michigan State University's Food Security Research Project for sharing the data with us. We also thank Awudu Abdulai, Arnab K. Basu, Antony Chapoto, Jann Lay and participants of the PEGNet Conference in Berlin, the Annual Conference of the Research Committee on Development Economics in Hanover, the Annual Congress of the Verein für Socialpolitik in Kiel and a workshop on 'HIV/AIDS and development in Zambia' in Lusaka for helpful comments and discussions.

² Corresponding author: toman.mahmoud@ifw-kiel.de

2.1 Introduction

Within two decades, AIDS has become the leading cause of death on the African continent and the sixth highest cause of mortality worldwide (WHO, 2008; UNAIDS, 2007). In 2007 alone about 1.5 million Africans died due to AIDS and almost 2 million got infected with HIV. With 22 million, or two thirds, of the global total of 33 million people infected with HIV, sub-Saharan Africa is the epicenter of the AIDS epidemic (UNAIDS, 2008a). The consequences of AIDS-related deaths are not comparable to those from most other diseases because AIDS hits adults at the peak of their productivity and earning capacity. It disables and kills the people on whom families rely for their livelihood. Consequently, it has been widely conjectured that AIDS may constitute a severe impediment to Africa's socio-economic development. In turn, development agencies, international organizations and national governments are spending billions of dollars each year to respond to the epidemic. The arrival and scaling-up of antiretroviral treatment has dramatically increased health expenditures in highly affected countries. From 2005 to 2008 alone, global spending on HIV/AIDS programs increased from US\$ 7.9 billion to US\$ 13.8 billion (UNAIDS, 2009).

Yet, the impact of AIDS-related mortality on household welfare as measured by per-capita income or expenditure is not clear a priori. While AIDS-related mortality may lower total household income through the loss of productive family members, the size of the household is also reduced by the death of a member. *Ceteris paribus*, the direction of the effect depends on whether the deceased member was a net consumer or a net producer for the household. The death of a net producer would reduce per-capita household income of surviving members, whereas the death of a net consumer would actually raise per-capita income. In addition, AIDS-afflicted households may cushion adverse welfare effects through adjusting the income-generating activities of remaining members (income coping) or the composition and size of the household (demographic coping).

Income-coping strategies include changes in the supply and allocation of household labor. In rural areas, non-farm employment may be given up to fill labor shortages on the farm. Case studies for Rwanda (Donovan et al., 2003) and Kenya (Yamano and Jayne, 2004) show for example that off-farm income of vulnerable and asset-poor households is at particular risk if an adult member dies. By contrast, for the Kagera region in Tanzania, Beegle (2005) finds no significant changes in labor supply of individuals in households having experienced a prime-age death. Another income-coping strategy is the sale of assets. Yamano and Jayne (2004) as well as Chapoto and Jayne (2008) find that prime-age mortality has only a minor impact on livestock, but greatly reduces the value of small animals such as goats and sheep in rural Kenya and Zambia, respectively.

Demographic coping may involve attempts to attract additional net producers, e.g. by letting migrants return to support the household, or to reduce the number of net consumers, e.g. by sending children away to members of the extended family. Demographic coping has been shown

to exist in the African context, with a strong variation across countries as well as the position and sex of the deceased (Mather et al., 2004). At one extreme, afflicted households in Kagera, Tanzania were able to maintain their household size (Ainsworth et al., 1995). In Uganda, by contrast, household size declined by about one person after a prime-age death, indicating that no new members could be attracted (Menon et al., 1998). Based on Demographic and Health Surveys from 21 countries across Africa, Beegle et al. (2009) also show that many single orphans do not stay with the remaining parent, but are rather sent away to live with other relatives.

Despite the theoretical ambiguity and the policy relevance of the topic, surprisingly little empirical knowledge exists about the effects of AIDS deaths on per-capita welfare of African households (Beegle and de Weerd, 2008).³ It is important to look at the *per-capita* implications of prime-age mortality, as the impact on *total household* welfare is not ambiguous. Examining the causal relationship between HIV/AIDS and poverty is essential for a meaningful cost-benefit analysis of current efforts to tackle the epidemic and prioritization of public health interventions. This paper aims to make a contribution to filling this gap by evaluating the changes in per-capita household incomes associated with AIDS-related mortality in rural Zambia, distinguishing between the deaths of household heads and spouses, who are likely to be net producers, as well as other adult household members, who might rather tend to be net consumers. Our methodological focus is on causal interpretation and we use a large nationally representative longitudinal survey that tracks about 5420 Zambian households over the period 2001-2004.

To minimize the selection bias that potentially arises because AIDS-related mortality is likely to be the endogenous outcome of individual behavior, we employ a difference-in-difference propensity score matching (DID-PSM) estimator. While PSM removes the selection bias due to observed differences between afflicted and non-afflicted households, the DID estimator differences out the time-invariant unobserved heterogeneity between the two groups.

To the extent that demographic coping plays a significant role, spillover effects may cause a further bias in the estimated impact of prime-age mortality. For instance, households without a deceased member may experience welfare losses when taking care of children sent away by AIDS-afflicted households. These effects have been neglected in the previous literature. We assess their relevance by testing whether non-afflicted households with AIDS-related demographic changes fare worse in per-capita income terms than their counterparts without such changes. In estimating the welfare effects of prime-age mortality, we also exclude households from the control group if members departed or joined for reasons related to AIDS.

Our paper is related to two previous empirical studies by Beegle et al. (2008) and Chapoto and Jayne (2008). Beegle et al. (2008) share our objective of evaluating the impact of prime-age mortality on the per-capita welfare of surviving household members. Yet they do so for one specific region characterized by high HIV prevalence, Kagera district in Tanzania, which makes it

³ Beegle et al. (2008) for Kagera, Tanzania, as well as Mather and Donovan (2008) for Mozambique are notable exceptions.

difficult to extrapolate their results to the national level. We use the same dataset as Chapoto and Jayne (2008), but their analytical focus is different from our study. Most notably, they look at how AIDS-related deaths affect various (farm and non-farm) income sources rather than household income per capita. Furthermore, Beegle et al. (2008) as well as Chapoto and Jayne (2008) rely on parametric approaches as compared to the non-parametric PSM applied here.

The remainder of the paper is structured as follows. Section II describes some key characteristics of the dataset, while Section III explains the method used to perform the estimations. The results of the empirical analysis are presented in Section IV. The paper closes with some concluding remarks.

2.2 Data and Setting

We use a nationally representative longitudinal dataset of 5420 rural farm households from Zambia surveyed in 2001 and 2004. As a very poor country with high HIV prevalence rates, Zambia makes an interesting case to study the effects of AIDS. Three in four rural Zambians live below the national poverty line (World Bank, 2007) and about 15 percent of adults aged between 15 and 49 are HIV-positive (UNAIDS, 2008b). As a result, life expectancy at birth has fallen to less than 41 years, making it the second lowest in the world (UNDP, 2007). This naturally raises the question to what extent AIDS has contributed to the low levels of socio-economic outcomes in Zambia.

The household survey was conducted by the Central Statistical Office in conjunction with the Ministry of Agriculture and Cooperatives and Michigan State University. In addition to standard socio-economic information of the household and its members, the survey paid particular attention to capturing a household's income generation process and its demographic development including the death of members. It not only recorded all deaths of household members after 2001, but also asked whether households had experienced the loss of a member between 1996 and 2001. To the best of our knowledge, the survey design makes it one of the most representative and comprehensive longitudinal household surveys available from Africa to examine the socio-economic effects of prime-age mortality. In particular, the combination of both the large sample size and relatively high HIV prevalence rates in Zambia provides a sufficiently large subsample of households afflicted by adult mortality. It allows examining the impact of a low-frequency event such as prime-age death with meaningful statistical power.

With a re-interview rate of 78 percent (6922 households were interviewed in 2001), the rate of attrition in the sample turns out to be fairly high.⁴ To the extent that it reflects household dissolution caused by the death of an adult member, attrition may lead to underreporting of prime-age mortality and bias estimations of the socio-economic consequences: households that dissolve are likely to be hit harder by adult mortality than those that remain intact. Using the

⁴ It should be noted, however, that the re-interview rate rises to 89 percent if one excludes attrition caused by enumerators not returning to several enumeration areas (Chapoto and Jayne, 2008).

same dataset, Chapoto and Jayne (2008) apply inverse probability weights to control for potential attrition bias. They find that this correction has little effect on the magnitude of the estimated impact of mortality, which suggests that attrition bias is not a major problem.

Although the surveys collected detailed information on mortality, epidemiological information on the cause of death is not available. For this reason, we follow the standard approach and use disease-related prime-age (15-59 years) mortality as a proxy for AIDS mortality. Comparing provincial HIV prevalence rates from antenatal clinics with adult mortality rates from the survey also used in this study, Chapoto and Jayne (2008) illustrate that the large majority of prime-age deaths are indeed likely to be AIDS-related.

Around ten percent of the households re-interviewed in 2004 experienced at least one prime-age adult death between the two survey years. We reduce this sample in various ways to arrive at treatment and control groups best suitable to identify the effects of adult mortality. First, households are only included if no member had died before the baseline survey was conducted, and if no chronically ill members were present in the household in 2001, because AIDS-related death is typically preceded by prolonged illness.⁵ This correction is made to ensure that 2001 incomes are not contaminated in the sense of already being lowered by AIDS-related adjustments. Second, we only consider the death of individuals who were recorded as household members in the baseline survey, eliminating from the sample the group of more than 150 households that experienced the death of a member who returned to the household to seek terminal care after the baseline survey. The main reason for doing so is that in households suffering the death of a long-term resident the deceased member contributed to initial income in 2001, whereas this was not the case for households incurring the death of a member who had not been a resident at the time of the baseline survey. As a consequence, the impact of prime-age mortality on welfare may be more severe for the former than for the latter (Chapoto and Jayne, 2008). We end up with a treatment group of 221 households, which are fairly equally distributed across income quartiles (Table 2.1). Among these households, 210 reported one dead member, nine reported two dead members, and two reported three dead members. Another 465 households incurred the death of a prime-age adult between 1996 and 2001. Again, afflicted households are fairly evenly spread over the income distribution.

When it comes to evaluating the impact of prime-age mortality on household welfare, either consumption or income can be employed as an indicator. While consumption is usually measured with a higher degree of precision, income changes may give a better indication of the household's future prospects and the sustainability of a household's adopted strategies to cope with the loss of an adult member. For instance, households may be able to temporarily smooth consumption by selling assets, but lower their income-generating capacity through such measures. In our case, consumption data are not available, but the survey offers very detailed information on the

⁵ Overall, 548 households reported a prime-age death in the period 1996-2001. The total number of households with a chronically ill member was 580 in 2001.

income sources of a household. Total yearly income is computed as the sum of the value of agricultural production (including production for home consumption), livestock income, and off-farm income including remittances.⁶ Per-adult equivalent incomes are computed using total household income and the adult equivalent conversion factors provided by the Ministry of Finance and National Planning. In principle, welfare could also be measured in per-capita terms. However, we prefer the use of per-adult equivalent units in order to capture age and sex-specific differences in need and economies of scale in consumption. On average, yearly per-adult equivalent incomes in AIDS-afflicted households slightly fell over the period 2001-2004 (Table 2.1).

TABLE 2.1: NUMBER OF HOUSEHOLDS REPORTING PRIME-AGE MORTALITY

| | Total | Quartile 1 | Quartile 2 | Quartile 3 | Quartile 4 | Income change | |
|-------------------------|-------|------------|------------|------------|------------|--------------------|---------|
| | | | | | | 2001-2004 (in ZMK) | |
| | | | | | | Mean | SD |
| <i>Period 2001-2004</i> | | | | | | | |
| any prime-age death | 221 | 57 | 60 | 44 | 60 | -26379 | 1059819 |
| death of head/spouse | 124 | 28 | 35 | 28 | 33 | -7852 | 1195724 |
| death of other members | 105 | 30 | 26 | 18 | 31 | -58859 | 833998 |
| death of male members | 101 | 18 | 32 | 22 | 29 | 27930 | 866045 |
| death of female members | 130 | 40 | 31 | 24 | 35 | -75022 | 1157807 |
| <i>Period 1996-2001</i> | | | | | | | |
| any prime-age death | 465 | 138 | 101 | 109 | 117 | -36684 | 770837 |
| death of head/spouse | 96 | 27 | 23 | 20 | 26 | -140071 | 775322 |
| death of other members | 375 | 112 | 82 | 89 | 92 | -9133 | 764037 |
| death of male members | 218 | 66 | 51 | 46 | 55 | -90790 | 828569 |
| death of female members | 277 | 77 | 61 | 69 | 70 | 1887 | 692275 |

All monetary values are given in 2004 Zambian Kwacha (ZMK). Mean 2001 per adult equivalent incomes of the four income quartiles are ZMK 91,669; ZMK 217,212; ZMK 422,337; and ZMK 1,347,059.

As argued above, the income effect of a death also depends on whether the deceased was a net consumer or a net producer for the household. However, the dataset does not allow identifying individual income and consumption levels. Instead, we try to account for the previous position of the deceased in the household by distinguishing between the deaths of a household head or his/her spouse and other adult members as well as between the deaths of male and female adults.⁷

2.3 Methodology

In estimating the causal impact of prime-age mortality on household welfare two problems have to be taken into account. First, with sexual transmissions accounting for the vast bulk of HIV

⁶ All monetary values are given in 2004 Zambian Kwacha (ZMK).

⁷ To keep the number of observations for each kind of prime-age death large enough, we abstain from further disaggregation. Note that the ratio of head's/spouse's and other members' death is different for the two periods because for the period 2001-2004 the reference was made to the head at the time of the death, while for the period 1996-2001 the reference was made to the head at the time of the interview in 2001. No deaths of heads should therefore be reported for the period 1996-2001, but some households nevertheless did so, probably confounding current with former (deceased) heads.

infections in Africa (Walker et al., 2003), AIDS is a behavioral disease. This may give rise to endogenous selection into treatment (prime-age mortality being the treatment). Second, to the extent that demographic coping leads to negative spillover effects on the control group of households without a deceased member, the causal effect of AIDS-related deaths will be underestimated.

2.3.1 Estimation Method

As a solution to the problem of endogenous selection, we employ the DID-PSM estimator. While PSM removes the selection bias due to observed differences between afflicted and non-afflicted households, the DID estimator controls for time-invariant unobserved characteristics such as risk attitudes and sexual behavior. A number of different issues have previously been studied using the DID-PSM approach. Recent examples include the impact of subsidies on exporting activity (Görg et al., 2008) and the schooling effects of conditional cash transfers (Behrman et al., 2009).

The general idea of matching is to find a comparison group that shares the same characteristics as the treatment group, but did not receive treatment. However, there is a trade-off between the potentially wide range of characteristics one might want to match on and the chances of finding matches which would be identical across all characteristics. In addition, it is not clear how each characteristic should be weighted. Rosenbaum and Rubin (1983) suggest the use of PSM to solve this dimensionality problem. Performing PSM involves matching treated to non-treated households based on similarity of their predicted probability of receiving the treatment (their ‘propensity scores’):

$$P(X_i) = Pr(D_i = 1 | X_i) \tag{2.1}$$

where $D_i = 1$ refers to households i receiving a treatment, X_i is a vector of pre-exposure characteristics, and $0 < P(X_i) < 1$. PSM matches pairs on the basis of how close the propensity scores $P(X_i)$ are across the two samples of treated and non-treated observations. Through the use of the propensity scores, PSM balances the observable covariates X between the treatment and the control group and thus selects similar non-treated observations for each of the treated observations.

The assumption needed for the treatment effect to be identified is that selection into treatment is random conditional on the observable covariates X (conditional independence assumption). In other words, all factors which simultaneously determine treatment and outcomes of interest should be observed. If data quality is not sufficient to capture important components of X , then the presence of these unobserved characteristics implies that PSM will be biased.

To more convincingly argue that the assumption of conditional independence is satisfied, we combine PSM with the DID approach. Hence, we compare changes in the levels of the outcome variable rather than the levels themselves. This allows us to purge all time-invariant unobservables from the analysis. Yet, unobservable effects may still bias our results if they are

time-varying and apply differently to treated and non-treated observations. This cannot be ruled out completely, but we are confident that our dataset is rich enough to control for important time-varying factors.

The resulting DID-PSM estimator for the average treatment effect on the treated households (ATT) can be expressed as

$$ATT = \sum_i \lambda_i [(y_{i1} - y_{i0}) - \sum_{j \in c(p_i)} \mu_{ij} (y_{j1} - y_{j0})] \quad (2.2)$$

where the λ_i s are sample weights used to construct the mean impact PSM estimator for the afflicted households i , $(y_{i1} - y_{i0})$ is the difference in these households' income per adult equivalent before and after the death of a member, and $(y_{j1} - y_{j0})$ is the difference in income per adult equivalent for households j of the control group over the same period. The μ_{ij} s are the weights applied in calculating the average outcome of the matched non-afflicted households. In our case, these weights are based on Kernel matching, which is a non-parametric approach that uses weighted averages of all households in the control group to construct the counterfactual outcome.⁸ The average treatment effect is only defined in the region of common support $c(p_i)$, which implies that all observations whose propensity score is lower than the minimum and larger than the maximum of the opposite groups are excluded.

Since our longitudinal dataset consists of just two waves, we can only evaluate absolute changes in per-adult equivalent income between 2001 and 2004. Even in a very poor context like rural Zambia it is likely that these absolute changes depend on initial income levels. In absolute terms, richer households may incur higher fluctuations in income than their poorer counterparts. To avoid matching households with very different income levels, we enforce exact matches within income quartiles.

The propensity score for each observation in the treatment and control group is estimated by using the predicted values from a logit model. Only variables which have an effect on both treatment and outcome and are not affected by the treatment should be included. Building on the existing literature on the determinants of HIV infection, we identify a set of variables simultaneously affecting a household's probability to experience the death of a prime-age member and its income-generation capacity. These variables are constructed using pre-treatment data from the baseline survey. At the household level, we include variables capturing the demographic composition of the household, also accounting for the sex and educational levels of prime-age adults and the household's head (compare Hargreaves and Glynn, 2002; Chapoto and Jayne, 2006). We also control for the importance of different income sources, as involvement in local non-farm activities or migration not only has an effect on income levels, but also increases the risk of infection with HIV (Ainsworth and Semali, 1998; Lurie et al., 2003). As a proxy for social capital we add a dummy indicating whether a household has a relation to the village head (compare Pronyk et al., 2008). The effect of household wealth is captured through land

⁸ Note, however, that our results do not seem to be sensitive to the choice of the matching algorithm. Our findings are robust to using caliper matching. See Smith and Todd (2005) for a discussion of different matching methods.

endowments and a broader asset index⁹ (compare Gillespie et al., 2007). We also include a dummy for radio ownership to proxy access to information. At the community level, we control for the distance of the village to the district town and nearest tarred main road (compare Buvé et al., 2002; Tanser et al., 2000). Besides, we add population density and rainfall levels and variability.¹⁰ The latter variable is thought to account for the general exposure of communities to weather risks and related agricultural income shocks. Finally, by introducing provincial dummies, we also control for the observed regional variety in HIV prevalence. Table A2.1 presents a complete list of the explanatory variables included in the logit model to estimate the propensity score.

A formal requirement for a reliable estimation with PSM is that, under the conditional independence assumption, the relevant covariates are balanced between the treatment group and the comparison group. Lack of balance points to a possible misspecification of the propensity score model. Hence, it is important to verify that the balancing condition is satisfied by the data. We perform three different balancing tests suggested in the literature.

The first balancing test (Sianesi, 2004) re-estimates the propensity score on the matched sample, i.e. only on afflicted households and matched non-afflicted households, and compares the pseudo-R² before and after matching. The pseudo-R² indicates how well the regressors explain the probability of being affected by prime-age mortality. After matching there should be no systematic differences in the distribution of covariates between both groups and hence the pseudo-R² should be fairly low. As shown in the upper part of Table A2.3, the pseudo-R² indeed approaches zero after matching.

The second balancing test (Rosenbaum and Rubin, 1985) calculates the standardized bias before and after matching. For each covariate it is defined as the difference in the means of the treated and (matched) control group as a percentage of the square root of the mean of variances in both groups. In the matched sample the standardized bias should be reduced considerably and be close to zero percent. In our case, PSM reduces the mean bias by about 70 percent, bringing the remaining mean bias down to about five percent (lower part of Table A2.3).

The third balancing test (Dehejia and Wahba, 2002) first divides the observations into strata based on the estimated propensity scores. These strata are chosen so that no statistically significant difference in the mean of the estimated propensity scores remains between the treatment and comparison group observations within each stratum. Then, t-tests are used to test within each stratum for mean differences in the regressors between the observations in the

9 The asset index is based on a principal component analysis of the following variables: Dummy for iron roof, dummy for brick walls, dummy for modern door, dummy for cement floor, and the value of agricultural productive assets.

10 Community variables are taken from supplementary datasets: Distance and rainfall data for the period 1990-2004 were kindly provided by Michigan State University's Food Security Group. Population density is derived from the 2000 Census.

treatment and comparison group. In our propensity score model, mean differences turn out to be insignificant across the board, suggesting that the balancing conditions are satisfied.¹¹

As a final point, it is illustrative to briefly compare PSM to other non-experimental methods of impact evaluation. Like PSM, OLS and IV estimations also rely on the conditional independence assumption. In the case of OLS, this is the assumption of exogeneity of treatment. The IV estimator in turn allows for the possibility of endogenous treatment, but rests on the challenging exclusion restriction. In contrast to these approaches, however, PSM is non-parametric. Hence, it does not impose assumptions of functional forms and error distributions. In addition, PSM also uses a different sample by confining matching to the region of common support thus dropping unmatched observations from the control group (Ravallion, 2008).

2.3.2 Spillover Effects

A unique characteristic of the Zambian dataset is that it allows us to link demographic changes in the composition of households to AIDS-related mortality. The survey included specific questions on why a member (12 years or older) joined or left a household between the two waves, and on whether a household hosted orphans (11 years or younger). We then define the following three groups of individuals who may give rise to an AIDS-related demographic burden for households without a deceased member, thus reducing their per-adult equivalent income:

1. *New members* (12 years and older), who were not in the household in the first wave, but were so in the second wave; they joined the household for the following reasons: having lost parents/being fostered, having been widowed, returning to the household because of sickness¹², needing help without being sick, and marrying a household member¹³. Not included as AIDS-related cases are individuals who joined the household to help with activities, to work for the household, because of old age, and because of divorce or separation.
2. *Departed members* (12 years and older), who were a household member in the first wave, but no longer so in the second wave; they departed from the household to live with other relatives. Individuals who departed to find a job, to establish a new home, to enter marriage, and because of divorce or separation are not counted as AIDS-related cases.
3. *Orphans* (11 years and younger), who lost at least one parent and did not live with the remaining parent. Information for orphans is only available for the 2004 wave, i.e., we do not know whether they were already members of the household in the first (2001) wave.

11 Also compare Table A2.4, which shows simple t-tests for the matched sample within different income quartiles.

12 Death due to AIDS is normally preceded by prolonged and severe illness. Given the transmission nature of HIV, it is likely that the partner of a deceased is also infected. Using the same dataset, Chapoto and Jayne (2006) provide evidence that this is the case in Zambia.

13 Beegle and Krutikova (2008) show for Tanzania that adult death affects the timing of marriage in affected households. In particular girls who lose their father marry at significantly younger ages. In addition, Ueyama and Yamauchi (2009) demonstrate for the case of Malawi that women's marriage age is lowered in general if the local marriage market is characterized by excess adult mortality.

Employing this demographic information, we check in various ways whether spillover effects can be observed in the data. First, based on Dzekedzeke and Fylkesnes (2006), who use DHS and antenatal care clinic data to estimate the HIV prevalence in 2001, we compare households without a deceased member in low-prevalence and high-prevalence provinces.¹⁴ Using the same estimation method described above, we test whether the number of members who joined or left a household due to AIDS-related reasons is higher in regions with high HIV prevalence. Obtaining a significant difference across regions would not only point to spillovers but also suggest that the above definitions indeed provide a reliable proxy for AIDS-related demographic changes. Second, we examine whether households that are indirectly affected by AIDS through demographic changes fare worse in terms of income per capita than their unaffected counterparts. Finally, we control for spillover effects in the estimation of the impact of AIDS-related mortality on per-adult equivalent incomes by excluding households from the control group if members departed or joined for AIDS-related reasons.

Of course, per-adult equivalent incomes of households in the control group are not only subject to demographic spillovers. Income levels are also affected through economic spillovers. High adult mortality rates may alter communities' economic structure, e.g. through general equilibrium effects on local labor markets or the availability of land for cultivation. AIDS may also lower the resilience of communities' social networks to other shocks (de Waal and Whiteside, 2003). However, the indirect nature of these effects renders them not only difficult to identify but also likely to be small. Using the same dataset, Jayne et al. (2006) find only minor effects of prime-age mortality on Zambian communities' crop output, mean income, and income per capita. We are thus confident that economic spillovers do not seriously impinge on the validity of our control group.

2.4 Results

2.4.1 Base Estimates

Before turning to the role of spillover effects, we first present PSM-DID estimates based on the unrefined control group (Table 2.2). This facilitates comparison with previous studies where spillovers were not taken into account. Note that the number of observations in Table 2.2 is lower than in Table 2.1 because a few cases are off the region of common support. Standard errors are computed using bootstrapping with 400 replications.¹⁵ At the aggregate level, we find a

14 Provinces with relatively low HIV prevalence are Northern and Northwestern (8.3 and 9.0 percent, respectively), while the three provinces with the highest prevalence are Copperbelt, Lusaka, and Southern (19.9, 22.0 and 17.6 percent, respectively).

15 Standard errors are bootstrapped for two reasons. First, as suggested by Bertrand et al. (2004), block bootstrap, which is also used here, can be used to correct for inconsistent standard errors in difference-in-difference estimations. Second, conventional standard errors tend to be biased because the estimated variance of the treatment effect should also include the variance due to the estimated propensity score and the imputation of the common support (Caliendo and Kopeinig, 2008). The validity of bootstrapped standard errors in matching estimators has

positive but insignificant effect of prime-age mortality on household per-adult equivalent income. Disaggregating impacts according to the gender and position of the deceased person leads to even higher standard errors of the estimated treatment effects, rendering it impossible to identify any pattern that would point to differences between net producers and net consumers.¹⁶

TABLE 2.2: SHORT-RUN INCOME EFFECTS OF RECENT PRIME-AGE MORTALITY (2001-2004)

| Treatment | Observations | ATT (in ZMK) | Standard error | p-value |
|-------------------------|--------------|--------------|----------------|---------|
| Any prime-age death | 217 | 90676 | 73840 | 0.22 |
| Death of head/spouse | 120 | 47278 | 112550 | 0.67 |
| Death of other members | 96 | -32900 | 115592 | 0.78 |
| Death of male members | 98 | 64596 | 102804 | 0.53 |
| Death of female members | 125 | -11234 | 107407 | 0.92 |

Chapoto and Jayne (2008) provide detailed evidence on the coping mechanisms that may lie behind this ‘non-finding’. As concerns income coping, they show that the death of a male prime-age household head is associated with a modest fall in the value of cattle, while the death of other adult household members has no significant effect on cattle assets. After the death of male and female adults, the value of small animals declines by 36 percent and 18 percent, respectively. Selling small animals thus appears to be a quantitatively important coping strategy used by AIDS-afflicted households. In addition, demographic coping plays a role. While households experiencing the death of the head are partially able to replenish their household size by attracting additional members, the death of other members is associated with a more than one-person decline in household size due to a significant reduction in the numbers of boys and girls. According to Chapoto and Jayne (2008, p. 344), the latter suggests that “in many cases, the deceased adult’s family lived with him or her while the sick person received terminal care at a relative’s home (possibly the older parents) and then left after the person passed away.” When re-estimating the relationship for the present sample that excludes sick persons seeking terminal care, we find a less than one-person decline in household size both for aggregate deaths (Table 2.3) and for all subgroups including other members. In combination, the attraction of new household members and the selling of productive assets may have prevented incomes from falling over the period 2001-2004.

Our core result is corroborated by Mather and Donovan (2008) for Mozambique, a neighboring country that provides a comparable setting of predominantly poor rural households. Mather and Donovan (2008) also find insignificant effects of prime-age mortality on rural households’ per-adult equivalent income irrespective of whether household heads or other members died, and

been questioned by Abadie and Imbens (2008), but their critique does not apply to kernel-based matching employed here.

¹⁶ The relevance of unobserved and time-varying heterogeneity could in theory be assessed with Rosenbaum bounds (Rosenbaum, 2002). Rosenbaum bounds calculate the size of hidden bias required to alter significance levels and hence inference about treatment effects. However, this sensitivity analysis is not meaningful in our case, as we do not find significant treatment effects to start with.

identify a loss of assets and changes in household composition as major factors contributing to this outcome.

Two further studies investigating the welfare impact of prime-age mortality come to different conclusions. For the case of Indonesia, Grimm (2010) shows that indirect coping effects more than offset the direct losses caused by the death of an adult household member, leading to overall gains in consumption per adult equivalent. Notably, surviving members of Indonesian households experiencing prime-age adult deaths are found to rely on increases in labor supply as one coping strategy, a scenario that is unlikely to happen in the rural African context as a result of limited employment opportunities. By contrast, Beegle et al. (2008) estimate for the Kagera region in Tanzania that within the first five years after the death of an adult per-adult equivalent consumption of the surviving household members grew by about seven percent less than consumption levels of non-afflicted households. Nonetheless, average consumption growth continued to be positive for AIDS-afflicted households for the period under consideration (2000-2004). Positive economic conditions may thus have enabled AIDS-afflicted households to uphold or even increase their consumption levels over time without resorting to desperation-led selling of assets.

2.4.2 Controlling for Spillover Effects

The demographic coping of AIDS-afflicted households may translate into significant spillover effects and thereby affect the welfare estimates reported above. We find that on average individuals (12 years and older) more frequently join and leave households without a deceased member for AIDS-related reasons when the households reside in regions with high HIV prevalence, but that these changes in the composition of households do not lead to significant welfare changes (Table 2.3).¹⁷ Having to care for orphans, by contrast, is estimated to lower yearly household income per adult equivalent by roughly US\$ 20 on average. Given the low income levels characterizing rural Zambia, this is a quantitatively important income drop of almost 20 percent. Overall, we thus obtain evidence pointing to the existence of considerable spillover effects.

In a subsequent step, we re-run the DID-PSM regressions using only households without AIDS-related demographic changes as a control group. The results are shown in Table 2.4. For aggregate deaths, the estimated impact of AIDS-related mortality on household incomes turns negative when controlling for spillovers, but continues to be insignificant. Disaggregating households by the position or sex of the member who died again produces absolutely insignificant results. We also check for impact heterogeneity across the initial poverty status of a household. Examining the income effects for each income quartile separately, however, yields no significant effects (available upon request).

¹⁷ To estimate the demographic spillover effects, we also employ the DID-PSM estimator with a similar specification.

TABLE 2.3: SPILLOVER EFFECTS OF RECENT PRIME-AGE MORTALITY

a) Does the adult equivalent size of households with a deceased member change differently?

| Treatment | Observations | ATT | Standard error | p-value |
|---------------------|--------------|---------|----------------|---------|
| Any prime-age death | 220 | -0.43** | 0.18 | 0.02 |

b) Do more persons (12 years and older) join households without a deceased member for AIDS-related reasons in high-prevalence regions?

| Treatment | Observations | ATT | Standard error | p-value |
|------------------------|--------------|--------|----------------|---------|
| High-prevalence region | 500 | 0.08** | 0.04 | 0.05 |

c) Do more persons (12 years and older) leave households without a deceased member for AIDS-related reasons in high-prevalence regions?

| Treatment | Observations | ATT | Standard error | p-value |
|------------------------|--------------|-------|----------------|---------|
| High-prevalence region | 499 | 0.10* | 0.05 | 0.06 |

d) Do households without a deceased member but with AIDS-related demographic changes fare worse in terms of income?

| Treatment | Observations | ATT (in ZMK) | Standard error | p-value |
|---------------------------------|--------------|--------------|----------------|---------|
| New members (12 and older) | 435 | -43397 | 40850.84 | 0.29 |
| Orphans (11 and younger) | 248 | -95182* | 55014.26 | 0.08 |
| Departed members (12 and older) | 398 | 18210 | 50318.79 | 0.72 |

**/* denote significance at the 5/10 percent level respectively.

TABLE 2.4: SHORT-RUN INCOME EFFECTS OF RECENT PRIME-AGE MORTALITY (2001-2004) CONTROLLING FOR SPILLOVERS¹⁸

| Treatment | Observations | ATT (in ZMK) | Standard error | p-value |
|-------------------------|--------------|--------------|----------------|---------|
| Any prime-age death | 207 | -67771 | 74522 | 0.36 |
| Death of head/spouse | 119 | 8376 | 115020 | 0.94 |
| Death of other members | 87 | -80803 | 131160 | 0.54 |
| Death of male members | 93 | -30852 | 100593 | 0.76 |
| Death of female members | 120 | -52689 | 113779 | 0.64 |

Since we do not find evidence for significant short-run income effects, even after controlling for spillovers, we can also examine the medium-run income effects for the period 2001-2004 resulting from prime-age deaths that occurred before 2001. As noted above, the household survey also recorded the death of a member between 1996 and 2001. Our identifying assumption is that 2001 incomes are not contaminated by short-run AIDS-induced welfare losses. The results shown in Table 2.5 point to the absence of such effects, which is in line with what Beegle et al. (2008) found for the Kagera region in Tanzania. Beegle et al. (2008) suspect that over the medium run factors not related to AIDS increasingly blur any existing difference between afflicted and non-afflicted households.

¹⁸ The respective number of observations in Table 2.4 is slightly different from the one in Table 2.2 because matching is confined to the region of common support.

TABLE 2.5: MEDIUM-RUN INCOME EFFECTS OF LESS RECENT PRIME-AGE MORTALITY (1996-2001) CONTROLLING FOR SPILLOVERS

| Treatment | Observations | ATT (in ZMK) | Standard error | p-value |
|-------------------------|--------------|--------------|----------------|---------|
| Any prime-age death | 449 | -29288 | 45954 | 0.52 |
| Death of head/spouse | 85 | -122950 | 121992 | 0.31 |
| Death of other members | 359 | 5068 | 46239 | 0.91 |
| Death of male members | 203 | -49586 | 61844 | 0.42 |
| Death of female members | 265 | 48548 | 52908 | 0.36 |

2.5 Conclusion

Finding no negative impact of prime-age mortality on the per-capita incomes of surviving household members may appear counterintuitive, in particular if it was the household head or the spouse who died. For the *Zambian* case, one might argue that by employing income as a welfare indicator the true impact is not easily identifiable due to possible measurement error, even though income generation was carefully captured in the household survey. Yet, through income and demographic coping, which has been shown to exist on a significant scale, AIDS-afflicted households may well be able to preserve their per-capita income levels over the short to medium term as our analysis and a similar study for rural Mozambique suggest. This is not to deny the potentially negative impact of prime-age mortality on human capital formation and its transmission across generations (Beegle et al., 2010; Bell et al., 2006). Besides, our analysis focuses on AIDS-related mortality, not morbidity. Yamano and Jayne (2005) and Linnemayr (2010) present evidence that AIDS may already lower household welfare well before the infected household member dies. On a more speculative note, our finding could also reflect the specific context of a very poor rural economy with high subsistence shares, where in a Lewis-type fashion surplus labor on farms implies that no welfare losses occur when the household loses a member. In a similar vein, the distinction between net producers and net consumers may get blurred in such a setting because all household members tend to be involved in the same set of low-return activities.

As concerns policymaking, the finding that surviving household members are able to avoid a drop in per-capita incomes after a prime-age death does not render public support unnecessary if coping involves actions with severe long-run consequences, such as depleting the only assets a household owns or taking children out of school. Our analysis of spillover effects suggests that the targeting of AIDS-related interventions should extend to households not experiencing a death but hosting orphans as they are shown to incur considerable reductions in per-capita incomes.

Literature

- Abadie, A., and G.W. Imbens (2008). On the Failure of the Bootstrap for Matching Estimators. *Econometrica*, 76(6): 1537-1557.
- Ainsworth, M., and I. Semali (1998). Who is Most Likely to Die of AIDS? Socioeconomic correlated of adult death in Kagera region, Tanzania. In M. Ainsworth, L.D. Fransen and M. Over (Eds.), *Confronting AIDS: Evidence from the Developing World*. Brussels: European Commission.
- Ainsworth, M., S. Gosh and I. Semali (1995). The Impact of Adult Deaths on Household Composition in Kagera Region, Tanzania. Unpublished Manuscript, World Bank, Washington, D.C.
- Beegle, K. (2005). Labor Effects of Adult Mortality in Tanzanian Households. *Economic Development and Cultural Change*, 53(3): 655-683.
- Beegle, K., and S. Krutikova (2008). Adult Mortality and Children's Transition into Marriage. *Demographic Research*, 19: Article 42.
- Beegle, K., and J. de Weerd (2008). Methodological Issues in the Study of the Socioeconomic Consequences of HIV/AIDS. *AIDS*, 22(1): 89-94.
- Beegle, K., J. de Weerd and S. Dercon (2010). Orphanhood and Human Capital Destruction: Is there Persistence into Adulthood? *Demography*, 47(1): 163-180.
- Beegle, K., J. de Weerd and S. Dercon (2008). Adult Mortality and Consumption Growth in the Age of HIV/AIDS. *Economic Development and Cultural Change*, 56(2): 299-326.
- Beegle, K., D. Filmer, A. Stokes and L. Tiererova (2009). Orphanhood and the Living Arrangements of Children in sub-Saharan Africa. World Bank Policy Research Working Paper 4889.
- Behrman, J.R., S.W. Parker, and P.E. Todd (2009). Schooling Impacts of Conditional Cash Transfers on Young Children: Evidence from Mexico. *Economic Development and Cultural Change*, 57(3): 439-477.
- Bell, C., S. Devarajan and H. Gersbach. (2006). The Long-Run Economic Costs of AIDS: with an Application to South Africa. *World Bank Economic Review*, 20(1):55-89.
- Buvé, A., K. Bishikwabo-Nsarhaza and G. Mutangadura (2002). The Spread and Effect of HIV-1 Infection in sub-Saharan Africa. *The Lancet*, 359(9322): 2011-2017.
- Caliendo, M., and S. Kopeinig (2008): Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22(1): 31-72.
- Chapoto, A., and T.S. Jayne (2008). Impact of AIDS-Related Mortality on Farm Household Welfare in Zambia. *Economic Development and Cultural Change*, 56(2): 327-374.
- Chapoto, A., and T.S. Jayne (2006). Socioeconomic Characteristics of Individuals Afflicted by AIDS-Related Prime-Age Mortality in Zambia. In S. Gillespie (Ed.), *AIDS, Poverty, and Hunger: Challenges and Responses*. Washington D.C.: IFPRI.
- Dehejia, R.H., and S. Wahba (2002). Propensity Score Matching Methods for Non-Experimental Causal Studies. *Review of Economics and Statistics*, 84(1): 151-161
- Donovan , C., L. Bailey, E. Mpsi and M. Weber (2003). Prime-Age Adult Morbidity and Mortality in Rural Rwanda: Effects on Household Income, Agricultural Production and

- Food Security Strategies. Research Report, Ministry of Agriculture, Livestock and Forestry, Rwanda.
- Dzekedzeke, K., and K. Fylkesnes (2006). Reducing Uncertainties in Global HIV Prevalence Estimates: The Case of Zambia. *BMC Public Health*, 6: 83.
- Gillespie, S., K. Kadiyala and R. Greener (2007). Is Poverty or Wealth Driving HIV Transmission? *AIDS*, 21(S7): S5-S16.
- Görg, H., M. Henry and E. Strobl (2008). Grant Support and Exporting Activity. *Review of Economics and Statistics*, 90(1): 168-174.
- Grimm, M. (2010). Mortality und Survivors' Consumption Growth. *Oxford Bulletin of Economics and Statistics*, 72(2): 146-171.
- Hargreaves, J.R., and J. Glynn (2002). Educational Attainment and HIV-1 Infection in Developing Countries: A Systematic Review. *Tropical Medicine and International Health*, 7(6): 489-498.
- Jayne, T.S., A. Chapoto, E. Byron, M. Ndiyoi, P. Hamazakaza, S. Kadiyala and S. Gillespie (2006). Community-Level Impacts of AIDS-Related Mortality: Panel Survey Evidence from Zambia. *Review of Agricultural Economics*, 28(3): 440-457.
- Linnemayr, S. (2010). Consumption Smoothing and HIV/AIDS: The Case of Two Communities in South Africa. *Economic Development and Cultural Change*, 58(2): 475-506.
- Lurie, M.N., B. Williams, K. Zuma, D. Mkaya-Mwamburi, G.P. Garnett, A. Sturm, M. Sweat, J. Gittelsohn and S. Abdool Karim (2003). The Impact of Migration in HIV-1 Transmission in South Africa: A Study of Migrant and Nonmigrant Men and Their Partners. *Sexually Transmitted Diseases*, 30(2): 149-156.
- Mather, D., and C. Donovan (2008). The Impacts of Prime-Age Adult Mortality on Rural Household Income, Assets, and Poverty in Mozambique. Research Report, Ministry of Agriculture and Rural Development, Republic of Mozambique.
- Mather, D., C. Donovan, T.S. Jayne, M. Weber, A. Chapoto, E. Mazhangara, L. Bailey, K. Yoo, T. Yamano and E. Mghenyi (2004). A Cross-Country Analysis of Household Responses to Adult Mortality in Rural sub-Saharan Africa: Implications for HIV/AIDS Mitigation and Rural Development Policies. International Development Working Paper 82. Michigan State University, East Lansing, Michigan.
- Menon, R., M.J. Wawer, J.K. Konde-Lule, N.K. Sewanlambo and C. Li (1998). The Economic Impact of Adult Mortality on Households in Rakai District, Uganda. In M. Ainsworth, L. Fransen and M. Over (Eds.), *Confronting AIDS: Evidence for the Developing World*. Brussels: European Commission.
- Pronyk, P.M., T. Harpham, L.A. Morison, J.R. Hargreaves, J.C. Kim, G. Phetla, C.H. Watts and D. Porter (2008). Is Social Capital Associated with HIV Risk in Rural South Africa? *Social Science and Medicine*, 66(9): 1999-2010.
- Ravallion, M. (2008). Evaluating Anti-Poverty Programs. In R.E. Evenson and T.P. Schultz (Eds.), *Handbook of Development Economics Vol. 4*. Amsterdam: North-Holland.
- Rosenbaum, P.R. (2002). *Observational Studies*. New York, NY: Springer.
- Rosenbaum, P.R., and D.B. Rubin (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *American Statistician*, 39(1): 33-38.

- Rosenbaum, P.R., and D.B. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1): 41-55.
- Sianesi, B. (2004). An Evaluation of the Active Labor Market Programmes in Sweden. *Review of Economics and Statistics*, 86(1): 133-155.
- Smith, J., and P. Todd (2005). Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators? *Journal of Econometrics*, 125(1/2): 305-353.
- Tanser, F., D. LeSueur, G. Solarsh and D. Wilkinson (2000). HIV Heterogeneity and Proximity of Homestead to Road in Rural South Africa: An Exploration Using a Geographical Information System. *Tropical Medicine and International Health*, 5(1): 40-46.
- Ueyama, M., and F. Yamauchi (2009). Marriage Behavior Response to Prime-Age Adult Mortality: Evidence from Malawi. *Demography*, 46(1): 43-63.
- UNAIDS (2009). *What Countries Need: Investments Needed for 2010 Targets*. Geneva: UNAIDS.
- UNAIDS (2008a). *2008 Report on the Global AIDS Epidemic*. Geneva: UNAIDS.
- UNAIDS (2008b). *2008 Zambia Epidemiological Factsheet on HIV and AIDS*. Geneva: UNAIDS.
- UNAIDS (2007). *AIDS Epidemic Update 2007*. Geneva: UNAIDS.
- UNDP (2007). *Human Development Report 2007/2008*. New York, NY: Palgrave Macmillan.
- de Waal, A., and A. Whiteside (2003). New Variant Famine: AIDS and Food Crisis in Southern Africa. *The Lancet*, 362(9391): 1234-1237.
- Walker, P.R., M. Worobey, A. Rambaut, E.C. Holmes and O.G. Pybus (2003). Sexual transmission of HIV in Africa. *Nature*, 422(7): 679.
- World Bank (2007). *Zambia Poverty and Vulnerability Assessment*. Washington D.C.: The World Bank.
- WHO (2008). *The Global Burden of Disease: 2004 Update*. Geneva: WHO.
- Yamano, T., and T.S. Jayne (2005). Working-Age Adult Mortality and Primary School Attendance in Rural Kenya. *Economic Development and Cultural Change*, 53(3): 619-653.
- Yamano, T., and T.S. Jayne (2004). Measuring the Impacts of Working-Age Adult Mortality on Small-Scale Farm Households in Kenya. *World Development*, 32(1): 91-119.

Appendix

TABLE A2.1: SUMMARY STATISTICS

| | Treated (n=221) | | Control (n=3,289) | | t-statistics |
|-----------------------------------|-----------------|--------|-------------------|--------|--------------|
| | Mean | SD | Mean | SD | |
| income change, 2001-2004 (in ZMK) | -26,379 | 71,291 | -23,500 | 12,004 | 0.06 |
| no 0-14 years | 3.11 | 2.36 | 2.96 | 2.02 | 1.07 |
| no 15-59 male skilled | 0.50 | 0.85 | 0.39 | 0.68 | 2.38 |
| no 15-59 male unskilled | 1.18 | 1.31 | 0.94 | 0.93 | 3.61 |
| no 15-59 female skilled | 0.38 | 0.79 | 0.21 | 0.58 | 4.06 |
| no 15-59 female unskilled | 1.51 | 1.23 | 1.17 | 0.84 | 5.60 |
| no 60 years and older | 0.40 | 0.68 | 0.24 | 0.52 | 4.41 |
| no migrants | 0.39 | 0.83 | 0.37 | 0.80 | 0.39 |
| male head | 0.81 | 0.40 | 0.82 | 0.39 | 0.45 |
| skilled head | 0.22 | 0.42 | 0.23 | 0.42 | 0.31 |
| relationship with village head | 0.37 | 0.48 | 0.36 | 0.48 | 0.18 |
| crop income (%) | 69.12 | 33.87 | 71.77 | 31.88 | 1.19 |
| self-employment income (%) | 13.15 | 25.25 | 12.89 | 24.79 | 0.15 |
| remittances income (%) | 3.74 | 11.08 | 3.15 | 9.72 | 0.86 |
| wage income (%) | 10.58 | 24.75 | 9.14 | 23.56 | 0.88 |
| land per adult equivalent | 0.72 | 0.89 | 0.83 | 1.46 | 1.15 |
| radio | 0.38 | 0.49 | 0.37 | 0.48 | 0.07 |
| asset index | -0.13 | 1.45 | -0.20 | 1.33 | 0.78 |
| distance to main road | 26.83 | 39.80 | 24.89 | 35.30 | 0.78 |
| distance to district town | 32.74 | 22.23 | 34.37 | 22.41 | 1.05 |
| population density | 22.46 | 37.92 | 24.92 | 44.58 | 0.80 |
| mean rainfall | 963.02 | 178.49 | 998.84 | 181.99 | 2.84 |
| rainfall variability | 155.06 | 41.88 | 158.27 | 38.44 | 1.19 |
| Central | 0.15 | 0.36 | 0.10 | 0.30 | 2.47 |
| Copperbelt | 0.04 | 0.19 | 0.07 | 0.25 | 1.86 |
| Eastern | 0.21 | 0.41 | 0.22 | 0.42 | 0.31 |
| Luapula | 0.14 | 0.34 | 0.13 | 0.34 | 0.21 |
| Lusaka | 0.05 | 0.22 | 0.02 | 0.15 | 2.51 |
| Northern | 0.12 | 0.33 | 0.16 | 0.37 | 1.56 |
| Northwestern | 0.03 | 0.18 | 0.07 | 0.26 | 2.34 |
| Southern | 0.13 | 0.33 | 0.12 | 0.33 | 0.29 |
| Western | 0.14 | 0.34 | 0.10 | 0.30 | 1.50 |

Treatment is any prime-age death between 2001 and 2004.

TABLE A2.2: LOGIT ESTIMATION OF THE PROPENSITY SCORE (TREATMENT IS ANY PRIME-AGE DEATH BETWEEN 2001 AND 2004)

| | coef | se |
|--------------------------------|-----------|-------|
| no 0-14 years | -0.070** | 0.028 |
| no 15-59 male skilled | 0.189** | 0.095 |
| no 15-59 male unskilled | 0.201*** | 0.053 |
| no 15-59 female skilled | 0.389** | 0.164 |
| no 15-59 female unskilled | 0.351*** | 0.038 |
| no 60 years and older | 0.497*** | 0.179 |
| no migrants | -0.139 | 0.109 |
| male head | -0.085 | 0.176 |
| skilled head | 0.075 | 0.194 |
| relationship with village head | -0.024 | 0.110 |
| crop income | -0.302 | 0.544 |
| self-employment income | -0.106 | 0.668 |
| remittances income | 0.178 | 0.556 |
| wage income | 0.143 | 0.671 |
| land per adult equivalent | -0.023 | 0.041 |
| radio | -0.066 | 0.098 |
| asset index | -0.084 | 0.068 |
| distance to main road | 0.002 | 0.002 |
| distance to district town | -0.005 | 0.005 |
| population density | 0.001 | 0.001 |
| mean rainfall | 0.000 | 0.001 |
| rainfall variability | -0.004 | 0.003 |
| Central | 0.620** | 0.293 |
| Copperbelt | -0.541*** | 0.141 |
| Eastern | 0.017 | 0.270 |
| Luapula | 0.217 | 0.145 |
| Lusaka | 0.857** | 0.366 |
| Northwestern | -0.502*** | 0.096 |
| Southern | -0.024 | 0.379 |
| Western | 0.455 | 0.320 |

***/**/* denote significance at the 1/5/10 percent level respectively.

TABLE A2.3: PSEUDO R2 AND MEAN BIAS REDUCTION

| | Quartile 1 | Quartile 2 | Quartile 3 | Quartile 4 |
|------------------------------|------------|------------|------------|------------|
| Pseudo R2 before matching | 0.136 | 0.145 | 0.133 | 0.118 |
| Pseudo R2 after matching | 0.016 | 0.019 | 0.02 | 0.022 |
| Mean bias reduction (%) | 78.93 | 68.57 | 67.76 | 63.61 |
| Mean bias after matching (%) | 3.19 | 4.07 | 5.24 | 4.49 |

Treatment is any prime-age death between 2001 and 2004.

TABLE A2.4: T-TESTS FOR MATCHED SAMPLE AFTER PSM

| 2001 Characteristics | Quantile 1 | | | | Quantile 2 | | | | Quantile 3 | | | | Quantile 4 | | | |
|--------------------------------|-----------------------|---------|--------------|-------|-----------------------|----------|--------------|-------|-----------------------|---------|--------------|------|-----------------------|---------|--------------|------|
| | Mean (matched sample) | | t-statistics | | Mean (matched sample) | | t-statistics | | Mean (matched sample) | | t-statistics | | Mean (matched sample) | | t-statistics | |
| | Treated | Control | p-value | | Treated | Control | p-value | | Treated | Control | p-value | | Treated | Control | p-value | |
| no 0-14 years | 2.87 | 2.96 | 0.82 | -0.24 | 2.97 | 3.02 | 0.88 | -0.16 | 3.40 | 3.18 | 0.37 | 0.72 | 3.02 | 3.06 | -0.10 | 0.92 |
| no 15-59 male skilled | 0.20 | 0.25 | 0.60 | -0.52 | 0.36 | 0.34 | 0.92 | 0.10 | 0.53 | 0.52 | 0.09 | 0.93 | 0.82 | 0.83 | -0.06 | 0.95 |
| no 15-59 male unskilled | 1.33 | 1.31 | 0.93 | 0.09 | 0.98 | 1.01 | 0.89 | -0.14 | 1.33 | 1.05 | 0.93 | 0.35 | 1.02 | 1.02 | -0.46 | 0.65 |
| no 15-59 female skilled | 0.18 | 0.16 | 0.79 | 0.27 | 0.44 | 0.40 | 0.80 | 0.25 | 0.37 | 0.35 | 0.10 | 0.92 | 0.43 | 0.58 | -0.79 | 0.43 |
| no 15-59 female unskilled | 1.56 | 1.52 | 0.24 | 0.24 | 1.41 | 1.49 | 0.63 | -0.49 | 1.58 | 1.33 | 0.80 | 0.43 | 1.33 | 1.40 | -0.33 | 0.74 |
| no 60 years and older | 0.42 | 0.39 | 0.80 | 0.26 | 0.36 | 0.34 | 0.92 | 0.10 | 0.28 | 0.29 | -0.08 | 0.94 | 0.45 | 0.40 | 0.37 | 0.72 |
| no migrants | 0.22 | 0.26 | 0.73 | -0.35 | 0.36 | 0.36 | 0.98 | -0.03 | 0.53 | 0.54 | -0.02 | 0.99 | 0.48 | 0.49 | -0.01 | 0.99 |
| male head | 0.73 | 0.72 | 0.96 | 0.05 | 0.76 | 0.79 | 0.73 | -0.35 | 0.88 | 0.87 | 0.15 | 0.88 | 0.85 | 0.87 | -0.36 | 0.72 |
| skilled head | 0.11 | 0.12 | 0.89 | -0.14 | 0.20 | 0.18 | 0.73 | 0.34 | 0.23 | 0.25 | -0.17 | 0.87 | 0.33 | 0.38 | -0.48 | 0.63 |
| relationship with village head | 0.44 | 0.43 | 0.92 | 0.10 | 0.36 | 0.39 | 0.69 | -0.41 | 0.37 | 0.39 | -0.17 | 0.87 | 0.30 | 0.30 | -0.05 | 0.96 |
| crop income (%) | 81.05 | 81.94 | 0.87 | -0.17 | 76.34 | 76.51 | 0.97 | -0.03 | 71.63 | 70.23 | 0.20 | 0.84 | 48.10 | 48.01 | 0.01 | 0.99 |
| self-employment income (%) | 3.38 | 3.81 | 0.88 | -0.15 | 12.73 | 12.01 | 0.86 | 0.18 | 13.31 | 13.37 | -0.01 | 0.99 | 22.52 | 24.38 | -0.31 | 0.76 |
| remittances income (%) | 7.34 | 7.13 | 0.95 | 0.06 | 2.74 | 3.02 | 0.81 | -0.25 | 5.01 | 6.02 | -0.36 | 0.72 | 7.77 | 9.94 | -0.28 | 0.78 |
| wage income (%) | 5.72 | 4.93 | 0.81 | 0.24 | 6.70 | 6.18 | 0.87 | 0.16 | 7.35 | 7.70 | -0.08 | 0.94 | 21.86 | 21.29 | 0.09 | 0.93 |
| land per adult equivalent | 0.36 | 0.38 | 0.80 | -0.26 | 0.80 | 0.92 | 0.68 | -0.42 | 0.76 | 0.79 | -0.16 | 0.87 | 0.95 | 0.99 | -0.16 | 0.88 |
| radio | 0.13 | 0.14 | 0.81 | -0.24 | 0.29 | 0.30 | 0.88 | -0.15 | 0.47 | 0.42 | 0.43 | 0.67 | 0.63 | 0.64 | -0.06 | 0.95 |
| asset index | 0.63 | -0.61 | 0.89 | -0.14 | -0.43 | -0.41 | 0.91 | -0.11 | 0.23 | -0.28 | 0.19 | 0.85 | 0.69 | 0.85 | -0.47 | 0.64 |
| distance to main road | 30.70 | 29.72 | 0.90 | 0.12 | 23.64 | 24.15 | 0.94 | -0.08 | 28.82 | 28.21 | 0.07 | 0.95 | 24.18 | 22.95 | 0.19 | 0.85 |
| distance to district town | 32.68 | 32.87 | 0.96 | -0.05 | 31.66 | 32.81 | 0.79 | -0.27 | 35.46 | 34.40 | 0.23 | 0.82 | 32.02 | 31.64 | 0.10 | 0.92 |
| population density | 18.75 | 19.67 | 0.84 | -0.20 | 18.93 | 19.25 | 0.95 | -0.06 | 20.21 | 21.18 | -0.19 | 0.85 | 31.66 | 30.37 | 0.12 | 0.91 |
| mean rainfall | 962.62 | 969.33 | 0.84 | -0.20 | 1,008.70 | 1,000.80 | 0.82 | 0.23 | 968.56 | 978.38 | -0.24 | 0.81 | 922.90 | 921.08 | 0.06 | 0.95 |
| rainfall variability | 151.53 | 152.11 | 0.94 | -0.07 | 150.67 | 149.79 | 0.90 | 0.12 | 150.24 | 151.91 | -0.22 | 0.83 | 166.23 | 165.02 | 0.08 | 0.94 |
| Central | 0.05 | 0.06 | 0.87 | -0.17 | 0.10 | 0.12 | 0.81 | -0.24 | 0.21 | 0.19 | 0.26 | 0.80 | 0.23 | 0.19 | 0.52 | 0.61 |
| Copperbelt | 0.04 | 0.04 | 0.86 | -0.18 | 0.02 | 0.02 | 0.78 | -0.29 | 0.05 | 0.05 | -0.03 | 0.98 | 0.05 | 0.05 | -0.03 | 0.98 |
| Eastern | 0.20 | 0.21 | 0.92 | -0.10 | 0.19 | 0.19 | 0.97 | -0.04 | 0.21 | 0.21 | -0.05 | 0.96 | 0.25 | 0.23 | 0.27 | 0.79 |
| Luapula | 0.16 | 0.17 | 0.92 | -0.10 | 0.20 | 0.21 | 0.94 | -0.07 | 0.14 | 0.14 | 0.04 | 0.97 | 0.05 | 0.06 | -0.17 | 0.86 |
| Lusaka | 0.02 | 0.01 | 0.84 | 0.21 | - | - | - | - | 0.12 | 0.10 | 0.27 | 0.79 | 0.08 | 0.08 | 0.11 | 0.91 |
| Northwestern | 0.02 | 0.03 | 0.66 | -0.44 | 0.03 | 0.04 | 0.86 | -0.17 | 0.02 | 0.03 | -0.28 | 0.78 | 0.05 | 0.04 | 0.24 | 0.81 |
| Southern | 0.16 | 0.16 | 0.96 | 0.05 | 0.07 | 0.09 | 0.68 | -0.42 | 0.07 | 0.07 | -0.06 | 0.95 | 0.18 | 0.21 | -0.31 | 0.76 |
| Western | 0.20 | 0.18 | 0.82 | 0.23 | 0.19 | 0.15 | 0.58 | 0.55 | 0.07 | 0.07 | -0.09 | 0.93 | 0.07 | 0.08 | -0.36 | 0.72 |

*Treatment is any prime-age death between 2001 and 2004.

Chapter 3

Orphanhood and Critical Periods in Children's Human Capital Formation: Long-Run Evidence from North-Western Tanzania¹⁹

Jens Hagen
University of Oxford

Toman Omar Mahmoud²⁰
Kiel Institute for the World Economy

Natalia Trofimenko
Kiel Institute for the World Economy

Abstract

Losing a parent is a trauma that has consequences for human capital formation. Does it matter at what age this trauma occurs? Using longitudinal data from the Kagera region in Tanzania that span thirteen years from 1991-2004, we find considerable impact heterogeneity across age at bereavement, but less so for the death of opposite-sex parents. In terms of long-term health status as measured by body height, children who lose their same-sex parent before teenage years are hit hardest. Regarding years of formal education attained in young adulthood, boys whose fathers die before adolescence suffer the most. Maternal bereavement does not fit into this pattern as it affects educational attainment of younger and older children in a similar way. The generally strong interaction between age at parental death and sex of the late parent suggests that the preferences of the surviving parent partly protect same-sex children from orphanhood's detrimental effects on human capital accumulation.

Keywords: orphans, health, education, timing of parental death, child development, Tanzania

JEL classification: I10, I21, J19, C23

¹⁹ We thank Kathleen Beegle, Jann Lay, T. Paul Schultz, Aparnaa Somanathan, Andreas Steinmayr, Rainer Thiele, Joachim de Weerd and participants of the Annual Bank's Conference on Development Economics in Cape Town, the Annual Meeting of the European Society for Population Economics in Seville and the Annual Conference of the Research Committee on Development Economics in Hanover as well as participants of a research seminar in Dar Es Salaam for helpful comments and discussions.

²⁰ Corresponding author: toman.mahmoud@ifw-kiel.de

3.1 Introduction

About 150 million boys and girls, eight percent of the child population in developing countries, have lost at least one of their parents (UNICEF, 2009). While HIV/AIDS has dramatically increased prime-age mortality in various parts of the world, above all in sub-Saharan Africa, the high rates of orphanhood can be largely attributed to other diseases, accidents, wars and natural disasters.²¹ Undoubtedly, parental bereavement is a trauma that has long-term consequences for child development. In this paper, we examine at what ages the loss of a parent is most critical for a child's human capital formation.

There is growing evidence that parental death, in particular the death of the mother, is detrimental to children's human capital formation, especially so in terms of educational attainment and health status in the short to medium run. So far, the literature has evaluated the effects of parental loss by comparing orphans and non-orphans, implicitly treating orphanhood as a binary event. We propose that the severity of impacts depends on the timing of parental death and investigate whether the long-term effects of parental death on orphans' human capital accumulation depend on the age at which a child is orphaned. For this purpose, we use longitudinal data from the Kagera region in North-Western Tanzania that follows children in several survey rounds from 1991-2004. Specifically, we estimate reduced-form equations to analyze the impact of age at bereavement on height and educational attainment in early adulthood.

Apart from evaluating the heterogeneous impact of orphanhood, this analysis contributes to the burgeoning literature on critical periods in child development. It has been widely shown that conditions before age five are crucial for human capital development and have persistent impacts on later life (Almond and Currie, forthcoming). Yet, relatively little is known about the relevance of conditions after early childhood, especially when it comes to developing countries. By distinguishing between the timing of paternal and maternal bereavement, this paper also adds to the understanding of the importance of different parental preferences for children's human capital formation.

We tackle the problem of unobserved heterogeneity by using a wide range of pre-orphanhood characteristics, including community fixed effects and showing that *future* age at bereavement is not significantly related to baseline outcomes. We also employ family fixed effects and exploit the plausibly exogenous variation created by the fact that siblings experience parental death simultaneously, but at different stages in their human capital formation. As we are interested in identifying the differential impact of parental death at different ages, we do not compare orphans and non-orphans. Accordingly, our sample is confined to children who lost at least one parent.

²¹ In fact, 41 out of 55 million orphans in sub-Saharan Africa lost their parents to causes other than HIV/AIDS (UNICEF, 2009).

We find considerable evidence that the effect of orphanhood on long-term human capital outcomes depends on the timing of parental death. Impact heterogeneity, however, appears to be conditional on the gender of the deceased parent. Children's growth is particularly stifled when they lose a parent of the same sex before teenage years. Similarly, the death of a father prior to age twelve brings about the largest reductions in years of attained schooling for boys. Only when considering educational attainment in the case of maternal bereavement do we not find heterogeneous impact across age at bereavement for both boys and girls. The generally strong interaction between age at parental death and sex of the late parent suggests that to a degree the preferences of the surviving parent can protect same-sex children from the detrimental effects of orphanhood.

In the following section, we present several mechanisms through which the death of a father or a mother could differentially affect child development. Section 3 provides a brief overview of the literature on the mean impact of parental loss. Methodological issues are discussed in section 4, which also describes the dataset and our estimation samples. Section 5 presents our results and conducts some robustness checks. Section 6 concludes.

3.2 Channels of Transmission

As health and education are complementary investments in human capital accumulation, we discuss them jointly at this point. Our focus is on children in developing countries, particularly in sub-Saharan Africa where the AIDS epidemic has led to exceptionally high rates of parental mortality. We discuss in how far each channel might cause heterogeneous impacts across different ages at bereavement.

3.2.1 Psychological Trauma

Parents are, without doubt, the most important caregivers and attachment figures providing emotional support and consistent nurture. In his works on early child development, Bowlby (1982) stresses the role of the relationship between caregivers, in particular mothers, and their children in shaping subsequent development and personality. Losing a parent during childhood is a traumatic event that shapes child development beyond purely economic consequences. Above all, parental death can have a lasting effect on children's emotional status which, in turn, can keep them from studying or attending school altogether (Gertler et al., 2004). Dozier et al. (2008) review, inter alia, evidence on the relation between the loss of an attachment figure and psychological disorders in adulthood and find a strong correlation between depression and anxiety and early maternal loss. Depressive episodes, in turn, can have a profound effect on both educational performance and physical health (Needham, 2009; Katon, 2003). A report from the National Scientific Council on the Developing Child (2010) concludes that early experience of traumatic events which produce lasting fear and anxiety can have lifelong consequences by disrupting the developing architecture of the brain, particularly in areas involved in emotions and

learning. In the absence of the buffering protection of adult support, such chronic stress in early childhood may even be toxic to the developing brain. Makame et al. (2002) find that orphans in Tanzania's capital Dar Es Salaam have more internalizing problems which put their long-term mental health at risk. For the case of Zimbabwe, Nyamukapa et al. (2008) show that orphanhood is associated with high levels of psychosocial distress. Children who have been orphaned by AIDS are likely to suffer from the direct trauma of parental illness and loss, as well as from stigma and discrimination. Such aggravating effects may result in even worse school performance and attendance/graduation rates (Sharma, 2006).

The nature of impact heterogeneity across age at bereavement is not obvious. In general, attachment during the first two years of life is crucial for later peer relations, effective emotional regulation and behavioral self-control (Deklyen and Greenberg, 2008). Infants and toddlers may, however, not yet fully realize the full significance of parental loss and might therefore be better able to relate to a new set of caregivers. In contrast, school-age children may be capable of grasping the significance of their parent's death but not be emotionally equipped to cope with such a loss. Evidence suggests that the death of a parent before the age of eleven has more severe psychological consequences than parental loss later in childhood or adolescence (Brown and Harris, 1978; Agid et al., 1999). Older adolescents, having lived through puberty, might be more capable of coping with the associated psychological burden. This is in line with Liu et al. (2009) who study the impact of parental absence on children's mental health in Chinese migrant households. They find that children approaching adulthood develop fewer symptoms of anxiety and depression than their younger counterparts. Overall, it is likely that the trauma of parental loss is particularly harmful for children in their early stages of human capital formation.

3.2.2 Lost Income

The loss of an income-generating household member is probably the single most important economic impact of parental death. This is particularly true for deaths due to AIDS which typically kills people in their most productive years of life. With low levels of wealth and the absence of well-functioning credit and insurance markets, households may be seriously compromised to respond to income shocks and smooth consumption after the death of an adult member. Beegle et al. (2008) show that households in Tanzania see their per-capita consumption drop by seven percent within the first five years after the death of a prime-age adult. Severe financial constraints may force households to save on food expenditures and reduce nutritional intake which, in turn, is an important determinant of children's health and cognitive development (Alderman et al., 2009; Currie, 2009). Moreover, reduced household resources can impact orphans' schooling directly by making it impossible to pay for tuition fees and school materials. The severity of these effects is likely to depend on the gender of the deceased parent. The loss of a father may be more critical as the death of a male household head typically brings about the largest reduction in material household welfare (Case and Paxson, 2004).

Insufficient or inferior nutrition in early childhood is particularly detrimental for subsequent health outcomes. This phenomenon has received ample attention in the literature on critical periods in child development (Alderman et al., 2006; Glewwe and King, 2001). Younger children suffer more from malnutrition than older children. The impact on educational attainment is less clear. To reduce schooling expenditures, young children's first enrolment could be deferred or even cancelled. However, older siblings might also be pulled out of school.

3.2.3 Increased Opportunity Costs of Children's Time

Due to insufficient or non-existent public safety nets, credit and insurance markets in many parts of the developing world, children who lose a parent often play an important role in households' strategies to cope with the death of an adult (Subbarao and Coury, 2004; Sharma, 2006). For instance, child labor could be a substitute for the deceased parent's labor supply which would increase the opportunity costs of children's time. Hence, school attendance could not only become less affordable due to the new budget constraint of the household, but also in terms of orphans' forgone earnings and domestic labor (Ainsworth et al., 2005).

It is probable that the degree of substitutability depends on the age of the child at the time of the parental loss (Evans and Miguel, 2007). The sex of the child and the late parent may also play an important role, even more so in societies with strong gender roles. Older sons are clearly a better substitute for their deceased fathers than younger ones, particularly in the context of physically demanding farm work or informal employment. The same logic applies to girls and their mothers: Older girls would have a comparative advantage in household production over their younger sisters. This observation has also been made in the literature on the economics of child labor (Edmonds, 2006). Hence, older children are more likely than their younger siblings to spend time in activities that mitigate the negative impacts of the death of an adult and reduce time in formal education.

3.2.4 Parental Preferences, Child Fostering and Lost Parental Inputs

Following parental death, orphans are typically subject to changed preferences of their (new) caregivers. Children who lose only one parent most often stay with the surviving parent (Beegle et al., 2009). Given that mothers tend to invest more in their children than fathers (Case and Paxson, 2001), the gender of the late parent is likely to affect orphans' human capital outcomes. *Ceteris paribus*, paternal death increases the importance of maternal preferences and may change a household's resource allocation in favor of children. Conversely, the death of a mother may reinforce paternal preferences and hence reduce child investments. Another dimension of differential treatment is the sex of the child. Parental preferences in developing countries are normally believed to be in favor of sons (Strauss and Thomas, 1995). This effect, however, is due to the fact that fathers most often exert more bargaining power on the intra-household allocation of resources. Considering parental preferences separately, mothers favor their daughters over

their sons, while the reverse is true for fathers (Thomas, 1990 and 1994; Lamb, 2010).²² Thus, the loss of a mother would have even more adverse effects on girls than the loss of a father.

Given the importance of conditions in early childhood, young children's health status appears to be most sensitive to changes in parental preferences. Educational outcomes are likely to be affected, too, though it is less apparent whether younger or older children would be hit harder.

Double orphans are usually cared for within the extended family network, mostly by their grandparents or other close relatives (Beegle et al., 2009; UNAIDS et al., 2004). It has been argued that the ever rising numbers of orphans due to the AIDS crisis may overstrain the capacity of traditional coping systems in parts of sub-Saharan Africa and shift the burden of child fostering to households that are not willing or able to invest in orphans' development (Beegle et al., 2009).

As discussed above, the presence of parents is of crucial importance for a child's development. It is well-known from developmental psychology that children learn a lot from their parents both through encouragement and through direct teaching and emulation (WHO, 2004). Leibowitz (1974) argues that quality and quantity of time invested by parents are important for healthy child development and can partly explain the variation in children's cognitive abilities prior to the start of formal education. Foster parents can take over these roles, but they are unlikely to be perfect substitutes for biological parents. They may have reduced incentives to invest time and money in orphans' human capital formation if they expect smaller pay-offs because of the informal, non-binding structure of intergenerational contracts as compared to more binding expectations from the biological children (Sharma, 2006; Ainsworth et al., 2005). Subbarao and Coury (2004) discuss the quality of various fostering arrangements and conclude that the arrangement is the more desirable the closer foster parents' kinship or relationship to the orphan is. In general, however, they find that the lack of nurture and guidance, due to parental loss and inadequate compensatory care, may stifle orphans' socialization. Similarly, the World Health Organization (2004) concludes that the effects of inadequate and disrupted care-giving can persist into adolescence in the form of behavioral disorders, anxiety, and depression.

The fact that socialization is a non-uniform process over the child's age reveals the scope for the effects of this channel to depend on the age at bereavement. On the one hand, older children are increasingly shaped by their parents' attitudes, values and willingness to invest time and resources in their upbringing and education. Thus, they may be less reliant on parental non-material inputs than younger children. On the other hand, being adopted at an earlier age means a higher chance of developing a close relationship with the caregiver (Odenstad et al., 2008; UNAIDS et al., 2004). Teenagers might thus be affected harder than others, if foster parents

²² Similarly, Duflo (2003) documents for the case of South Africa that old-age pensions received by women improved the health status of girls living in the same household, but had no significant effects for boys. Pension paid to men, however, had no such impacts.

anticipate more pay-off to education of younger adopted children as a longer care-giving period results in stronger ties with the orphan.

3.2.5 Overall Impact on Human Capital Formation

The theoretical transmission channels discussed above strongly suggest that the impact of parental death on children's human capital formation depends on the age at which a child loses a parent. In addition, this differential impact is most likely to vary with the gender of the late parent and the child. The relative strength of the presented transmission channels, however, is not known a priori and likely to vary according to the context. As a result, the pattern of the overall effect remains theoretically ambiguous and is potentially non-linear in age at bereavement. Moreover, the importance of the various pathways is likely to differ for the two outcomes of interest, health and educational attainment. Losing non-material parental inputs, for instance, could be more relevant to orphans' adult health outcomes; increased opportunity costs should be primarily decisive for schooling outcomes; and changing parental preferences as well as lost household income are likely to affect both. Ultimately, it is an empirical question whether the aforementioned mechanisms are indeed at work and how they interact with each other.

Ideally, one would like to explicitly model health and educational outcomes by two separate production functions that contain, among other things, measures of parental time and care, monetary resources, measures of children's idiosyncratic endowments and abilities as inputs. However, given that the decisions regarding demand for children's health and education are jointly determined with other important factors influencing these outcomes, estimating such production functions is not feasible. Besides, the channels identified above are partly related and overlap to some extent in their predictions regarding the impacts. In the absence of experimental data, however, it is intrinsically difficult to distinguish between these transmission channels. We hence follow the empirical literature on the effects of orphanhood and employ reduced-form estimates. Accordingly, our analysis effectively examines the overall impact of parental loss on human capital accumulation across age at bereavement, irrespective of the channel through which it may operate.

It is important to bear in mind that all effects measured are net effects, i.e. they already take into account the mitigating impact of coping strategies at the household and community level. The availability of access to formal or informal credit and insurance or support from government, NGOs and the like is context-specific and should also affect the human capital formation of orphans.

3.3 Evidence on Mean Impacts

There is growing evidence that parental death stifles children's development. The empirical literature has primarily focused on orphans' educational attainment. Using data from Indonesia's National Socioeconomic Survey from 1994-1996, Gertler et al. (2004) find that recently orphaned

children are substantially more likely to drop out of school than non-orphans. This effect is most pronounced for children at the transition between primary and lower secondary school and between lower secondary and higher secondary school. Regarding heterogeneity across age, they present some evidence that older children's enrolment is decreased more by the loss of a parent.

In South Africa, mother's death reduces children's likelihood to be enrolled by about two percentage points and depresses school-related expenditures by more than ten percentage points. Based on a longitudinal data from 2001 to 2004, Case and Ardington (2006) also show that given their age maternal orphans have completed about a quarter of a year less schooling. Paternal death, however, does not appear to be a significant determinant of orphan's educational outcomes. Older children (aged 11-16 at the last survey round in 2003-2004) experience larger reductions in completed years of schooling than younger children (aged 6-10), although the difference is not statistically significant. This finding, however, refers to children's age at the time of the interview, not to their age at bereavement.

Using Kenyan longitudinal data from 1998 to 2002, Evans and Miguel (2007) find that parental loss decreases school participation rates of children by seven percentage points. Their estimates suggest no sign of recovery of orphans' school attendance for at least three years after the death, which they interpret as indication of possible long-run effects. In line with other studies, the impact is primarily driven by maternal deaths. Younger children (below age 12 in the year of their parent's death) are somewhat more likely to drop out of school than older children. The authors suggest that this result may be due to the high drop-out rates in primary schools in rural Kenya. If students in higher grades are selected on academic ability, potentially higher returns to schooling might keep them at school.

Despite this evidence for reduced school enrolment or attendance in the short and medium run, it is not obvious whether such effects persist into adulthood. While impacts could be cumulative over the duration of orphanhood, they might as well vanish after a period of mourning and household adjustments (Beegle and de Weerdt, 2008). Yet, long-run evidence is still scarce as very few datasets follow orphans and non-orphans over a sufficient period of time.

An important exception is Beegle et al. (2010) who investigate the long-term impacts of parental loss on educational and body height attainment using the same dataset from North-Western Tanzania as we do. Their sample consists of 718 non-orphaned children aged 6-15 years and surveyed in 1991-1994 who are at least 19 years old by the final survey round in 2004. About 20 percent of the children lost at least one parent prior to age 15 after the baseline. Beegle et al. (2010) find that maternal loss permanently reduces educational attainment by about one year and final body height attained by almost two centimeters. In a related study, Beegle et al. (2006) split the same sample into two groups based on children's age at bereavement to separate younger and older orphans. In general, the impact on schooling and health does not seem to differ for orphans above and below twelve years at parental death. However, maternal death is associated with a larger impact on final height attainment for younger orphans.

Taken together, there is considerable evidence that parental loss, in particular the loss of the mother, has detrimental effects on children's human capital formation. This paper contributes to the presented set of studies on the mean impact of orphanhood by analyzing at what ages parental death is most critical to children's human capital formation in the long run.

3.4 Empirical Strategy

3.4.1 Data and Setting

With a per-capita income of about US\$1100 adjusted for purchasing power parity and ranking 152 out of 179 countries on the Human Development Index, Tanzania is one of the world's poorest countries. While the national prevalence rate of HIV among individuals aged 15-49 has shown a slight downward trend from well over seven percent in the mid-1990s to just over six percent in 2007, the absolute number of AIDS orphans continued to rise from an estimated 610,000 in 2001 to around 970,000 in 2007 (UNAIDS, 2008). The total number of orphans irrespective of the cause of parental death is significantly higher at around 2.6 million in 2007.

Our analysis draws upon data from the Kagera Health and Development Survey (KHDS). The Kagera region is located in North-Western Tanzania and was one of the first areas in Africa to be hit by the HIV/AIDS epidemic. It thus provides a well-suited setting for the analysis of the long-run impact of parental death. The original dataset (KHDS-I) was collected between 1991 and 1994 and includes information on 915 households who were interviewed up to four times at 6-7 months intervals. In 2004, a large effort was undertaken to track and re-contact as many surveyed individuals as possible for a repetition of the survey referred to as KHDS-II. Individuals were followed even if their original households had dissolved, if they had left the household, or if they had migrated to neighboring regions or countries. Thanks to the rigorous tracking effort, the re-contact rate is as high as 93 percent and 82 percent on the household and individual level, respectively. The longitudinal dataset thus covers up to 13 years and contains up to five observations per individual.²³

As explained below, we use slightly different samples to analyze the long-run effects of orphanhood on height and educational attainment. This different choice is made to maximize sample size and variation across age at bereavement on the one hand and minimize the possibility of catch-up in terms of human capital accumulation on the other hand. Both samples are based on the group of children who were non-orphaned at the first interview during KHDS-I (1991-1994), but lost their mother or father previous to the re-interview in KHDS-II (2004) and before the age of 18. The cut-off point at age 18 follows the most common definition of orphanhood (UNAIDS et al., 2004; Skinner et al., 2006).

²³ The World Bank (2004) provides detailed information on KHDS-I and KHDS-II.

3.4.2 Height Attainment

Although health clearly is a multidimensional phenomenon, body height is frequently taken as a marker of long-term health status. It has been shown to capture chronic malnutrition and under-nutrition as well as frequent or chronic illness during childhood and adolescence (Case and Paxson, forthcoming; O'Donnel et al., 2008). Thus, if losing a parent really has an adverse impact on children's overall health in addition to what coping strategies can mitigate, this impact is likely to show up in orphans' final height.

Apart from capturing health and nutrition status during childhood, it is widely acknowledged that body height is also correlated with important determinants and indicators of individual welfare later in life. Almond and Currie (forthcoming) review the related literature and conclude that health and socio-economic status during childhood have a significantly positive impact on educational and labor market outcomes in adulthood. This underlines the central role of health in human capital and its predictive power for other outcomes relevant to individual welfare.

To maximize sample size, we use a relative measure of height referred to as height-for-age z-score. An individual's height-for-age z-score is the difference between her height and the average height in her specific reference group, that is, individuals of the same sex and age, expressed in standard deviations of this reference group. Accordingly, height-for-age z-scores can take any real values, typically in the range of -5 to 5. Z-scores have the advantage of being standardized by gender and age in such a way that the z-score of an individual who follows a normal growth path does not follow a time trend. Thus, including adolescents who are still in the process of physical growth does not generate a problem of truncation because their z-scores can still take arbitrary values and are not confined to a certain range due to their lower age.

Considering all orphans who were aged below 18 at baseline (first interview during KHDS-I), our health sample contains 369 individuals. Since some of them did not reach adulthood at the time of KHDS-II in 2004, however, catch-up growth among the younger children in our sample cannot be excluded. Yet, as noted by Case and Paxson (forthcoming), there is considerable correlation between height at any given age and final height attained. Moreover, a large fraction of our sample should have lived through the adolescent growth spurt which coincides with the onset of puberty and reaches its peak at age 12 for girls and age 14 for boys (Ulijaszek et al., 1998). We are thus confident that there is only limited potential for catch-up growth in our sample.

The average z-score is about -1.6, which implies that the individuals in our sample are considerably smaller than their counterparts in the UK reference group. This can be understood as evidence of a generally poorer health and nutritional status during childhood and adolescence. It is important to note that the choice of the reference group should not have an impact on our results. The reference group merely serves as an anchor affecting the absolute values, but not the variation of z-scores. Across populations, growth patterns are predominantly determined by environmental factors and individuals from different ethnic backgrounds would on average be

equally tall if they grew up under similar circumstances (Bogin, 1999; Ulijaszek et al., 1998; Steckel, 1995). Using a different reference group would only result in a parallel shift of the distribution of observed z-scores as the shape of growth paths across age is predominantly determined by biological factors and thus unlikely to vary significantly across countries.

3.4.3 Educational Attainment

The sample used to analyze the long-run effect of age at bereavement on educational attainment only includes those orphans who were aged below 18 at baseline and at least 19 years old at the time of KHDS-II in 2004. After this age, a significant catch-up in terms of formal education is unlikely to occur. Most adolescents have left school by this age and re-enrolment after having entered the labor force is rare. This restricts our education sample to 184 orphans and leaves us with no children who lost a parent prior to 6 years of age. This limited age range is a shortcoming of our education sample since we want to look at heterogeneity across different ages at bereavement and pre-school years could be of particular interest.

Primary schooling in Tanzania is completed after grade seven, lower secondary schooling after grade 11 and upper secondary schooling after grade 13 (Ainsworth et al., 2005). While statutory enrolment age is seven, about 84 percent of all children actually started school significantly later in the mid-1990s, when our baseline data were gathered (World Bank, 1999). Average enrolment age was about ten and nine years for boys and girls, respectively. Thus, the youngest individuals in our sample (aged 19 in 2004) could in principle still have been visiting school at the time of KHDS-II. However, given the fact that only seven percent of secondary-school-age children were enrolled in secondary school in the mid-1990s (World Bank, 1999), this possibility is unlikely to be relevant. In fact, only seven individuals of our sample were still at school in 2004.

We follow Beegle et al. (2010) and take the highest grade completed in 2004 as our dependent variable. We refer to this variable as *years of schooling*, counting each grade as one year. Ideally, one would investigate the effects of parental loss on standardized test scores or even cognitive ability. Educational attainment and a more holistic measure of education can vary largely between adolescents who have completed the same grade. However, given the data at hand, we have to resort to years of schooling as the arguably best proxy for educational attainment available.

3.4.4 Econometric Setup

We take advantage of the longitudinal dimension of the dataset by regressing final outcomes (KHDS-II, 2004) on age at parental loss and a set of pre-orphanhood characteristics, which are taken from the first interview during KHDS-I (1991-1994). If available, we include information on orphans' month of birth and the month in which parental death occurred in order to use as much variation in age at bereavement as the dataset allows for.

Our econometric setup is summarized by the following equation:

$$Y_{i2004} = c + a_1 A_i + a_2 A_i^2 + \nu D_i + x_{ijt} \chi + f_{ikt} \psi + a_{i2004} \omega + \beta_1 [A_i * P_i] + \gamma_1 [A_i * M_i] + \delta_1 [A_i * P_i * M_i] + \kappa P_i + \lambda M_i + \mu [P_i * M_i] + \beta_2 [A_i^2 * P_i] + \gamma_2 [A_i^2 * M_i] + \delta_2 [A_i^2 * P_i * M_i] + \varepsilon_{i2004}$$

where Y_{i2004} is the outcome of interest, height-for-age z-score or highest grade completed, of individual i in 2004 and A_i is the age at bereavement. Squared terms of A_i allow for potentially non-linear effects across A_i , thus helping to identify critical periods of parental death. The possibly more adverse impact of double orphanhood on a child's human capital formation is captured by the dummy variable D_i .

Pre-orphanhood characteristics x_{ijt} at the household and individual level are measured at baseline. Depending on the identification strategy used, we add baseline fixed effects f_{ikt} at the community or family level. The last term represents control for cohort effects via age dummies a_{i2004} in 2004.

To determine whether the impact of parental loss differs by the gender of the deceased (or surviving) parent, one could generate a variable for both age at maternal and paternal death. However, the fact that most children in our sample are single orphans (79 percent and 83 percent in the health and education sample, respectively) would create many missing values for these variables. We circumvent this problem by only considering age at *first* bereavement and interacting A_i and its squared term with a dummy variable P_i indicating whether it was the father who died first. As the effects of orphanhood are likely to vary for boys and girls, we also add interactions with a dummy variable for boys (M_i). Given the limited sample size, this procedure is preferred to splitting the sample into female and male subsamples.

ε_{i2004} is an individual-specific independent and identically distributed error term. The above model is estimated using Ordinary Least Squares.²⁴ Standard errors are clustered at the household level to take into account that outcomes of children living in the same household are likely to be correlated.

3.4.5 Identification

We follow two complementary strategies to identify the effect of age at bereavement on human capital formation. First, we propose that, controlling for pre-orphanhood characteristics, the timing of parental loss is exogenously distributed across children's ages. Second, we add family fixed effects to the above equation and exploit the variation created by the fact that siblings lose their parent at the same time, but at different periods in their human capital formation.

Orphanhood as such is unlikely to be a random event. In Africa, HIV is a disease that is predominantly sexually transmitted and hence subject to selection into risky sexual behavior or

²⁴ For years of education censoring to the left is a potential source of concern as this variable can only take positive values if a child was ever enrolled in school. However, only 15 orphans in the education sample have never attended school. Given this low share and the difficulties of interpreting interaction terms in non-linear models (Ai and Norton, 2003), we follow Beegle et al. (2010) and use linear regression models.

dependencies (Beegle and de Weerdt, 2008). If factors attributable to parental self-selection are correlated with children's health and educational outcomes, however, any observed differences between orphans and non-orphans might have already existed before orphanhood. Nonetheless, this problem does not apply to our analysis. Being interested in the ages at which orphanhood is most critical, we do not compare orphans and non-orphans; our sample only includes children who lost at least one parent.

Still, this does not exclude the possibility that the *timing* of parental death is endogenously determined. There might be confounding factors that affect both age at bereavement and children's human capital accumulation. In this case, observed differences in the dependent variables between children who were orphaned at younger age and those who lost their parents later in their development process might be driven by pre-orphanhood characteristics, too. For instance, children from initially richer households could not only have completed more years of schooling at the time of parental death. *Ceteris paribus*, their age at parental loss could also have been higher on average because more resources could have been devoted to parents' terminal care. A similar argument can be made when considering coping strategies. Arguably, initially richer households are better equipped to cope with the death of an adult. They can more easily afford to send their children to school and forgo their potential earnings or labor supply at home. Their higher ability to safeguard (food) consumption may also positively influence orphans' health.²⁵ Even after controlling for such observable pre-orphanhood characteristics, the alleged impact of timing of parental death may still be driven by unobservable factors. While it is conceivable that such factors are related to selection into orphanhood as such, it is much less obvious why they should be systematically associated with age at bereavement.

Nonetheless, we address the concern of unobserved heterogeneity by regressing baseline outcomes of interest on *future* age at bereavement and the full set of baseline controls (compare Beegle et al., 2010; Case and Ardington, 2006). Only if timing of parental loss is correlated with omitted variables we do not control for would we expect future age at bereavement to be a significant predictor of baseline outcomes. The underlying assumption is that such unobserved variables would adversely affect outcomes already before orphanhood. Yet, Table A3.3 in the appendix shows that the timing of parental death is not significantly related to height or educational attainment. We are hence confident that age at bereavement is sufficiently exogenous to warrant identification, at least after controlling for important pre-orphanhood characteristics.

An alternative identification strategy is to compare siblings by using family fixed effects. This approach effectively neutralizes any observed and unobserved confounding factors at the family level. Before parental death, siblings share the same environment including genetic factors, extended family, family norms, neighborhoods as well as parents' general attitudes towards education and health. However, they experience the loss of a parent at different stages in their

²⁵ Lundberg et al. (2000) discuss households' coping strategies and sources of external funding following an adult death in Kagera. However, nothing suggests that external support in Kagera during the time of the survey benefited children differently across age at bereavement.

development. We exploit this plausibly exogenous variation for identification. In addition, this approach takes care of sample selection bias. To the extent that selection into orphanhood is related to family-level factors, the fixed-effects results for our orphan sample are valid for the entire child population in Kagera.²⁶

The use of family fixed effects, however, ignores differential treatment of children *within* households. First of all, human capital outcomes might have been affected by the date of birth, as the economic condition of a given household is likely to change over time. Depending on their age difference, children of the same parents might have grown up under circumstances more or less favorable for their human capital formation. Birth order may also play a role. First-born children are often found to fare better in terms of both health (e.g. Horton, 1988) and educational attainment (e.g. Black et al., 2005). Such discrimination against later-born siblings might be explained by increased competition for parental resources. In the case of parental death, however, first-born children may also bear a disproportionately large burden, as older children are more likely to take over some of the tasks that were previously performed by the deceased parent. To address these complications, we include a dummy marking first-born children. We also add a full set of age (in 2004) dummies to control for cohort effects. While the age dummies pick up general trends in living conditions, the indicator of being first-born captures heterogeneity between siblings related to order and timing of birth.

The elimination of any confounding factors at the family level is a clear advantage of the siblings analysis. However, it comes at the cost of reduced sample size. Under this approach, the sample only includes orphans with at least one sibling. This lowers the number of observations to 267 in the health and 104 in the education sample. By contrast, the other proposed identification strategy uses the full sample. Relying on variation between families, it is also more robust to concerns regarding differential treatment of children within families.

3.4.6 Control Variables

Our regression models include various baseline characteristics at the individual and household level. These are meant to control for confounding pre-orphanhood differences and capture heterogeneous adoption of coping strategies with adverse impact on human capital accumulation. The family fixed effects model draws on a more parsimonious specification and only considers individual characteristics.

A household's longer term welfare status is approximated by a wealth index which is the first principal component of a set of important household assets and durable consumption goods. These comprise per-adult equivalent land holdings, ownership of a motor vehicle, the number of rooms in the dwelling as well as indicators of housing quality.²⁷

²⁶ A similar argument is made by van den Berg et al. (2009).

²⁷ Housing quality is captured by the availability of electrical lighting, the type of roof, wall and floor materials, access to a private toilet, a private or public water source and the type of water source.

Given the importance of parental inputs and parents' attitude for children's schooling and health outcomes, we control for parents' own educational background (Kenkel, 1991; Thomas et al., 1991). Specifically, we include dummy variables indicating whether mothers and fathers have attended school and whether they were present in the household at baseline. The latter should also provide a proxy for the level of nurture given by each parent. Unfortunately, numerous missing values impede including parents' highest grade completed in our regressions which would give a more detailed picture of the family's educational background.

To capture competition for parental resources as well as a household's demographic possibilities to cope with the loss of an adult, we add measures of the size and composition of the household. Potential differences in preferences for children's well-being before orphanhood are proxied by the gender of the household head.

We also try to control for genetic factors which should play an important role in determining children's body height. A natural step would be to include both mother's and father's anthropometrics as control variables. Because of many missing values for father's height, however, we follow Beegle et al. (2010) and only include mother's height.²⁸ Doing so should at least partially capture genetic endowments and might be more reliable than using father's measures who may not always be the biological father.

Orphans' baseline health and educational outcomes are included to further isolate the impact of parental death from effects of unobserved background characteristics that might influence children's human capital formation. In the health regressions, we include orphans' baseline height-for-age z-scores to capture individual health histories and heterogeneous endowments before losing a parent. In the education regressions, we use a relative measure of years of schooling defined as the difference between a child's years of schooling at baseline and the years of schooling a child of her age would have attained had she been enrolled at age seven, the statutory enrolment age in Tanzania. Compared to an absolute measure, this schooling-for-age variable does not follow a time trend and should hence be less correlated with age at bereavement.

Community fixed effects are added to eliminate the confounding impact of community characteristics that have not changed over the 13-year period covered by the survey. Certainly, communities develop over time and infrastructure including the accessibility and quality of schools as well as health facilities is likely to change. However, in order to bias the parameter estimates of interest, these characteristics would have to have changed differentially across orphans' average age at parental death in different communities. This case seems very unlikely.

²⁸ Another potentially important control variable is parents' age at baseline. Many missing values, however, prevent us from doing so. Still, running the regression on the reduced sample for which mother's age is available does not change our results. This may be due to the fact that parents' age is not perfectly correlated with children's age. Some parents have their first child earlier than others, parents may be of different ages, and they typically have several children in the first place.

Besides, the set of age dummies take care of general trends or shocks that occur at the regional or national level.

3.5 Results

The following two sub-sections present the main findings from our econometric analysis of orphans' height and educational attainment. Due to the multiple and non-linear interaction effects the interpretation of the regression results is most easily done graphically. Full regression tables are provided in the appendix. Besides examining the general impact of age at bereavement on human capital outcomes, we distinguish between four types of orphans: male and female paternal orphans as well as male and female maternal orphans. For each type we evaluate the impact of age at bereavement on human capital outcomes at the mean or most typical values of the other independent variables, depending on whether they are continuous or binary. Together with the confidence intervals, these predicted values show at which ages parental loss is most critical for the outcome under consideration. In addition, the differential impact of orphanhood can be assessed by the slope of the fitted line, which is given by the marginal effect of age at bereavement:

$$\partial Y_{i2004} / \partial A_i = a_1 + 2a_2 A_i + \beta_1 P_i + \gamma_1 M_i + \delta_1 [P_i * M_i] + 2\beta_2 [A_i * P_i] + 2\gamma_2 [A_i * M_i] + 2\delta_2 [A_i * P_i * M_i]$$

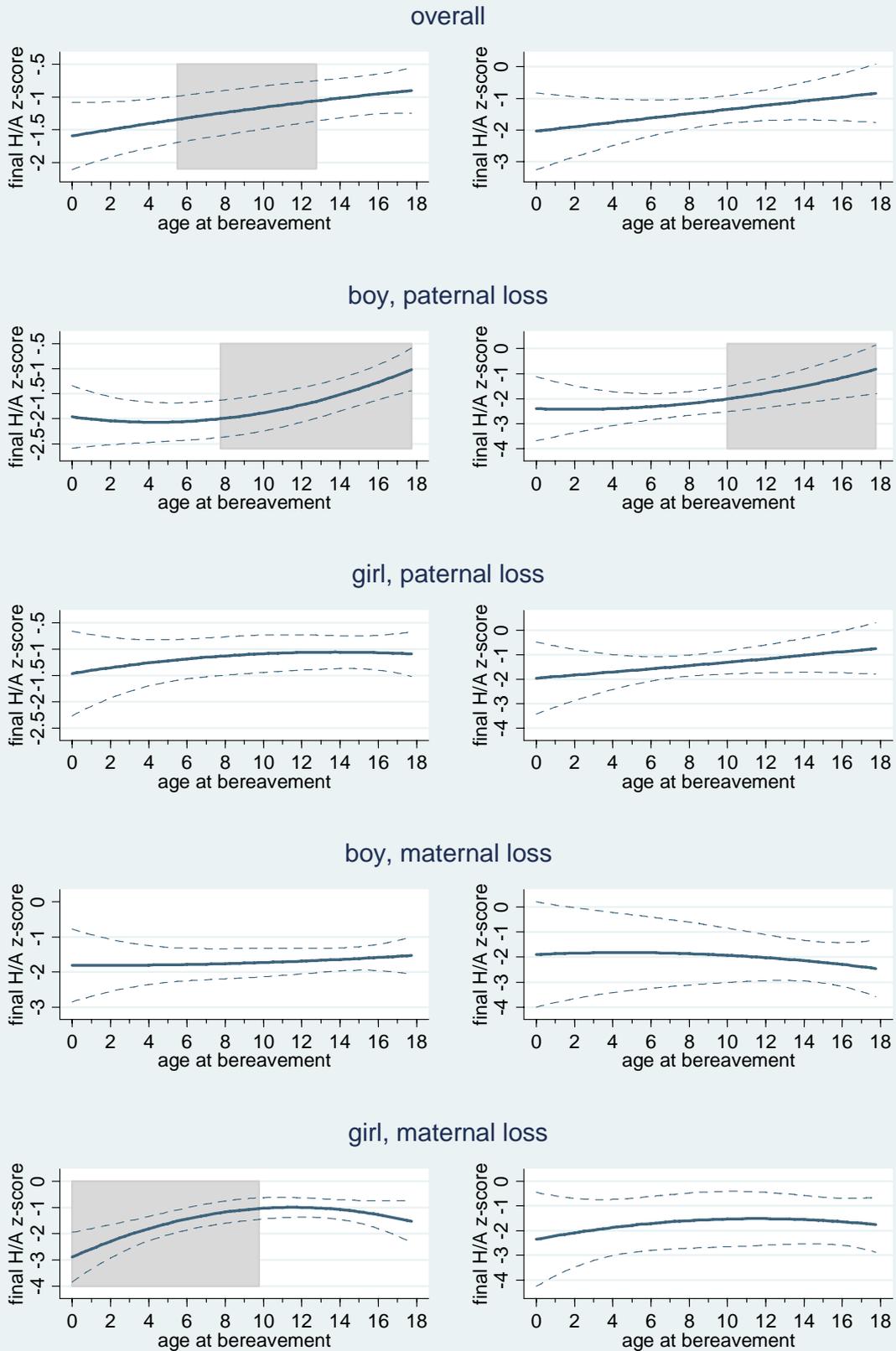
Age periods in which the marginal effect is significantly different from zero indicate when and how the relationship between age at paternal loss and health and education later in life changes. An upward-sloping curve would imply that parental death is more detrimental to younger children, while the reverse would be true in the case of a downward-sloping curve.

Since we do not perform a comparison with the counterfactual scenario of both parents being alive, our estimation results should not be confused with the impacts of parental loss as such. Our focus is on investigating at which ages parental loss is most critical to children's human capital formation. However, children who lost their first parent around the age of 18, our cut-off point used for defining orphanhood, may serve as a reference category. Becoming an orphan at this age should no longer have a major effect on a child's human capital formation. By the age of 18, body growth is generally completed (Bogin, 1999) and children in Tanzania at the end of the 1990s had typically reached their final educational attainment (World Bank, 1999).

3.5.1 Height Attainment

Figure 3.1 summarizes our findings on heterogeneous impacts of parental death on children's long-term health outcomes. For the different kinds of orphans, the left column presents the results from the regressions with community fixed effects, while the right column presents the results from the siblings analysis. Overall, both identification strategies produce very similar results.

Figure 3.1: Body Height Attainment in Early Adulthood



Notes - column 1: predicted values (community fixed effects), column 2: predicted values (family fixed effects); dashed lines: 90% confidence interval; shaded areas: marginal effect of age at bereavement significantly different from zero.

Estimating the relationship between age at bereavement and height-for-age z-scores without interactions with the gender of the late parent and the child, the death of a parent appears to be most critical to body growth when children are young. However, this general impression conceals considerable impact heterogeneity across the different types of orphans.

For boys, paternal loss has the most adverse impacts on their long-term health status early in their lives. Boys whose father died before their teenage years achieve height-for-age z-scores that are on average about one unit lower than those of boys who lost their father close to reaching adulthood. The predicted values suggest that the detrimental impact of age at bereavement on height attainment is relatively flat until age eight. In fact, the marginal effect of age at bereavement is not significantly different from zero before this age. After age eight, however, the unfavorable health impact of paternal orphanhood significantly decreases with age at bereavement. Up to the age of 18, the marginal effect of age at paternal loss is significantly positive for boys. Basically the same observations can be made for brothers. Boys who lose their father in early ages grow significantly shorter than their older brothers. This finding is in line with the medical literature suggesting that pre-adolescent growth is generally more sensitive to environmental factors than adolescent growth (Bogin, 1999). It also corresponds to the theoretical predictions from various transmission channels. Outcomes resulting from psychological trauma and lost income such as chronic stress and malnutrition stifle young children's development the most. The result can also be explained by the change of parental preferences. The loss of a father and the increasing importance of maternal preferences might imply that boys lose their privileged status and receive relatively less investments. Given the importance of early child development, the discontinuation of positive discrimination may be particularly harmful for boys' health formation when they are young.

As regards girls, the long-run health impact of paternal death does not appear to depend on the age at bereavement. Although the curve is slightly upward-sloping the marginal effect of age at paternal loss is not significantly different from zero. Psychological and income-related factors are unlikely to account for the divergent effects of father's death on girls and boys, as they should apply regardless of the sex of the child. However, as indicated above, the death of a male household head may change a household's resource allocation in favor of girls and thus compensate for income losses and distress. Young girls may particularly benefit from a growing role of maternal preferences.

Our findings for the death of a mother mirror those for the death of a father. While age at maternal bereavement does not seem to affect boys' long-term health status, girls suffer most in terms of height attainment when they lose their mother prior to age ten. In contrast to paternally orphaned boys, however, the marginal effect of age at bereavement for maternally orphaned girls is positive before teenage years and then becomes insignificant. A girl whose mother passes away around her first birthday attains height-for-age z-scores in young adulthood which are almost two units lower than those of a girl whose mother survives until she becomes a young teenager. These effects are not significant at the family level though. Younger and older sisters do not appear to

be differentially affected by their mother's death. Again, the change in parental preferences may explain why maternal loss has heterogeneous impact across age at bereavement for girls, but not for boys. The lack of maternal safeguarding of sufficient resource allocation towards girls is likely to be most adverse for health formation in early childhood.

The estimates of the other explanatory variables are mostly in line with expectations. Orphans' baseline height-for-age z-score and mothers' height are statistically highly significant predictors of higher 2004 z-scores. Thus, children who are relatively short at a given age are also likely to reach a relatively low height in adulthood. Parental education and a household's baseline wealth status, however, are not significantly associated with children's long-term health.

3.5.2 Educational Attainment

Figure 3.2 shows our findings on the relationship between age at bereavement and years of schooling attained in early adulthood. Again, the two identification strategies yield highly comparable results. It is important to keep in mind that the education sample does not include children who lost a parent prior to age six.

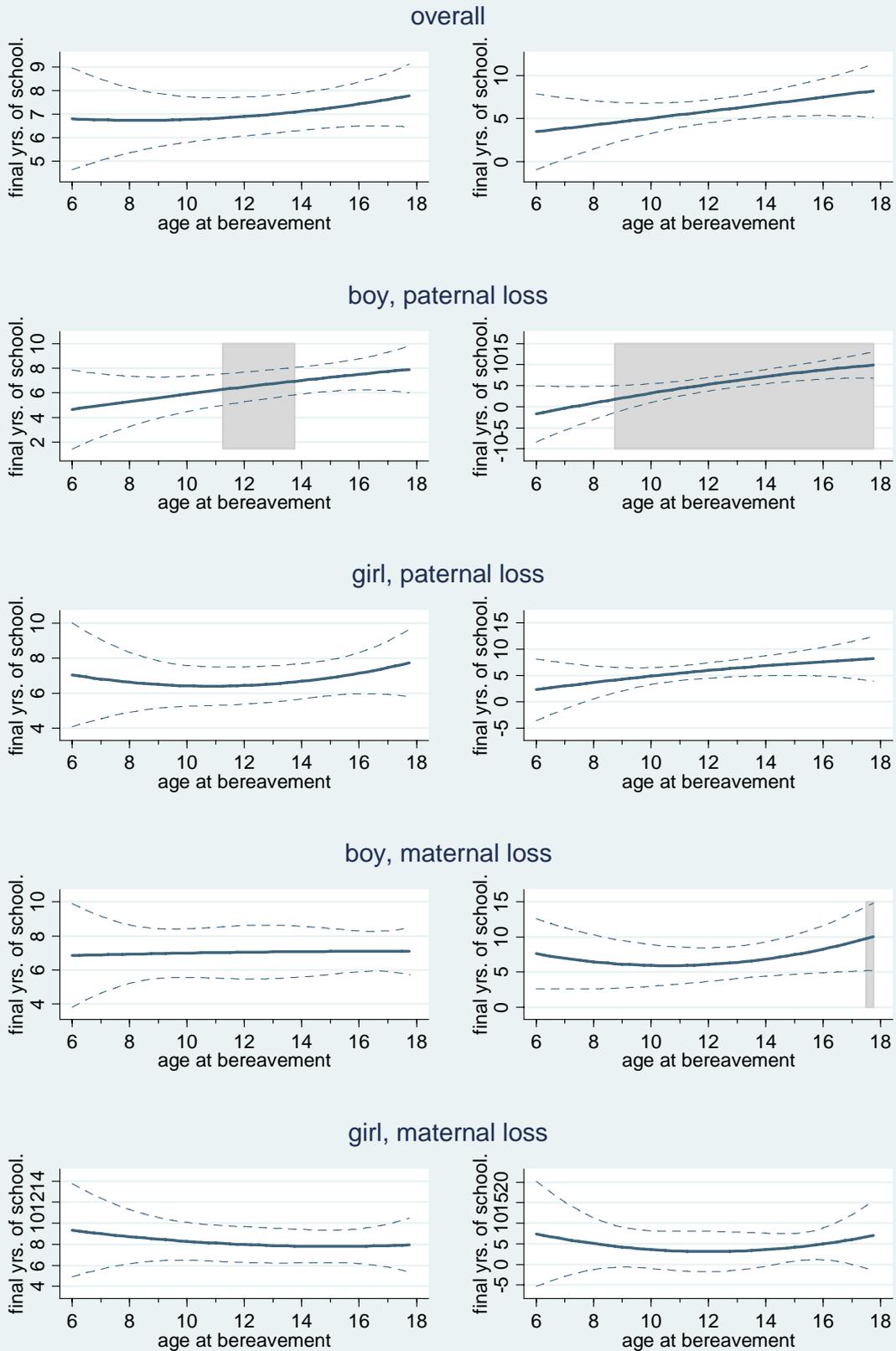
Overall, children who lose their parents earlier in life seem to attain lower levels of schooling in early adulthood. However, the marginal effect of age at bereavement is not significantly different from zero and the general specification without interactions terms masks differential impacts of orphanhood.

Concerning paternal death, the results resemble those for long-term health status. Educational outcomes of boys are most severely affected when their father dies before adolescence. This effect is particularly pronounced for brothers. Boys who lose their father at age six attain up to five years less of schooling than their older brothers. Age at paternal bereavement, however, does not significantly matter for girls' educational level in young adulthood. As in the case of health, maternal preferences seem to protect younger girls from the adverse income effects that are typically associated with the death of the male household head. With respect to opportunity costs, we do not find evidence that older boys or girls spend less time in formal education than younger children following the death of their father. This result might be due to the fact that men often engage in wage employment and can hence not easily be substituted.

For maternal death, the picture looks different. Neither for boys nor for girls do we find heterogeneous impact across age at maternal bereavement. In particular, there is no evidence for increased opportunity costs of children's time as especially older daughters should be a relatively good substitute for female adult labor in domestic tasks.

As regards other covariates, children with relatively high educational levels at baseline have attained more years of schooling in early adulthood. So have children from initially richer households. Maternal education also improves orphans' long-run educational attainment. Fathers' schooling status, however, does not play a significant role.

Figure 3.2: Educational Attainment in Early Adulthood



Notes - column 1: predicted values (community fixed effects), column 2: predicted values (family fixed effects); dashed lines: 90% confidence interval; shaded areas: marginal effect of age at bereavement significantly different from zero.

3.5.3 Robustness Checks

Despite the high re-contact rates of 93 percent and 82 percent on the household and individual level, respectively, non-random attrition poses a potential source of bias. We address this problem by predicting the probability of an individual to be re-interviewed in 2004 based on the same baseline characteristics used above. We then use the inverse of these probabilities as weights in our regressions. Our results, however, are fully robust to this weighting procedure (available upon request). This finding does not come as a surprise given the low overall attrition rates and similar findings by Beegle et al. (2010) who also find that their results on the human capital impact of orphanhood as such remain unchallenged after accounting for potential attrition bias.

Another issue of concern is the role of parental morbidity. If the death of a parent was preceded by a long period of severe illness, orphans' health and educational status at baseline could already be contaminated by their parents' chronic disease. For instance, AIDS-related deaths, which are common in the study area, typically follow up to two years of opportunistic diseases like diarrhea, tuberculosis and acute respiratory infections (Morgan et al., 2002).²⁹ Such episodes inflict a heavy financial and emotional burden upon affected households and could adversely affect human capital accumulation of orphans before the actual death. Since we do not reliably observe the cause of parental death, we control for parents' health status by including a dummy variable indicating whether they reported to have been chronically ill during the six months prior to the baseline survey. Again, our results remain virtually unchanged (available upon request) and are thus unlikely to be confounded by earlier parental morbidity. This finding is also in line with the previous observation that future age at bereavement is not significantly related to baseline height or educational attainment.

3.6 Conclusion

There is ample theoretical ground to suspect that the age at which a child loses her parent(s) affects the magnitude of the resulting impacts on her overall human capital formation. Using longitudinal data from North-Western Tanzania, we provide empirical evidence in support of this view. Overall, however, the impact of orphanhood is less heterogeneous for the death of opposite-sex parents. Preferences of the surviving parent or the foster family seem to partly make up for income losses and psychological stress which typically accompany the loss of a parent. This finding also indicates that risks are not shared equally within households.

Our results may inform policy interventions to take an age-specific approach. Yet, optimal targeting of particularly vulnerable orphans remains challenging in the context of broad-based support schemes as the adverse effects of orphanhood not only vary with age at bereavement, but also with the sex of the deceased parent and the type of human capital. More research on

²⁹ 41 percent and 31 percent of the orphans in the education and health sample, respectively, have lost a parent up to two years after their baseline interview.

critical periods for child development in developing countries is needed in order to design targeted interventions that provide favorable conditions for human capital formation after age five. Such interventions may be crucial for reaping the benefits of previous parental investments and improving a child's chances for later success in life.

Literature

- Agid, O., B. Shapira and J. Zislin (1999). Environment and Vulnerability to Major Psychiatric Illness: A Case Control Study of Early Parental Loss in Major Depression, Bipolar Disorder and Schizophrenia. *Mol Psychiatry*, 4(2): 163-172.
- Ai, C., and E.C. Norton (2003). Interaction Terms in Logit and Probit Models. *Economics Letters*, 80(1): 123-129.
- Ainsworth, M., K. Beegle and G. Koda (2005). The Impact of Adult Mortality and Parental Deaths on Primary Schooling in North-Western Tanzania. *Journal of Development Studies*, 41(3): 412-439.
- Alderman, H., H. Hoogeveen and M. Rossi (2009). Preschool Nutrition and Subsequent Schooling Attainment: Longitudinal Evidence from Tanzania. *Economic Development and Cultural Change*, 57(1): 239-260.
- Alderman, A., J. Hoddinott and B. Kinsey (2006). Long Term Consequences of Early Childhood Malnutrition. *Oxford Economic Papers*, 58(3): 450-474.
- Almond, D., and J. Currie (forthcoming). Human Capital Development Before Age Five. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Volume 4*. Amsterdam: Elsevier.
- Beegle, K., J. de Weerdt, and S. Dercon (2010). Orphanhood and Human Capital Destruction: Is There Persistence Into Adulthood? *Demography*, 47(1): 163-180.
- Beegle, K., and J. de Weerdt (2008). Methodological Issues in the Study of the Socioeconomic Consequences of HIV/AIDS. *AIDS*, 22(1): 89-94.
- Beegle, K., J. de Weerdt, and S. Dercon (2008). Adult Mortality and Consumption Growth in the Age of HIV/AIDS. *Economic Development and Cultural Change*, 56(2): 299-326.
- Beegle, K., J. de Weerdt, and S. Dercon (2006). Orphanhood and the Long-Run Impact on Children. *American Journal of Agricultural Economics*, 88(5): 1266-1272.
- Black, S., P.J. Devereux and K.J. Salvanes (2005). The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics*, 120(2): 669-700.
- Beegle, K., D. Filmer, A. Stokes and L. Tiererova (2009). Orphanhood and the Living Arrangements of Children in sub-Saharan Africa. World Bank Policy Research Working Paper 4889.
- van den Berg, G., P. Lundborg, P. Nystedt and D.-O. Rooth (2009). Critical Periods During Childhood and Adolescence: A Study of Adult Height Among Immigrant Siblings, IZA Discussion Paper 4140.
- Bogin, B. (1999). *Patterns of Human Growth*. Cambridge: Cambridge University Press.
- Bowlby, J. (1982). *Attachment and Loss: Volume 3: Loss, Sadness and Depression*. New York, NY: Basic Books.
- Brown, G.W., and T.O. Harris (1978). *The Social Origin of Depression: A Study of Psychiatric Disorder in Women*. New York, NY: Free Press.
- Case, A., and C. Ardington (2006). The Impact of Parental Death on School Outcomes: Longitudinal Evidence from South Africa. *Demography*, 43(3): 401-420.

- Case, A., and C. Paxson (forthcoming). Causes and Consequences of Early Life Health. *Demography*.
- Case, A., and C. Paxson (2004). Orphans in Africa: Parental Death, Poverty, and School Enrollment. *Demography*, 41(3): 483-508.
- Case, A., and C. Paxson (2001). Mothers and Others: Who Invests in Children's Health? *Journal of Health Economics*, 20(3): 301-328.
- Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature*, 47(1): 87-122.
- Deklyen, M., and M. Greenberg (2008). Attachment and Psychopathology in Childhood. In J. Cassidy and P. Shaver (Eds.), *Handbook of Attachment*. New York, NY: Guilford Press.
- Dozier, M., K. Stovall-McClough and K. Albus (2008). Attachment and Psychopathology in Adulthood. In J. Cassidy and P. Shaver (Eds.), *Handbook of Attachment*. New York, NY: Guilford Press.
- Duflo, E. (2003). Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *World Bank Economic Review*, 17(1): 1-25.
- Edmonds, E. (2006). Understanding Sibling Differences in Child Labor, *Journal of Population Economics*, 19(4): 795-821.
- Evans, D., and E. Miguel (2007). Orphans and Schooling in Africa: A Longitudinal Analysis. *Demography*, 44(1): 36-57.
- Gertler, P., D. Levine and M. Ames (2004). Schooling and Parental Death. *Review of Economics and Statistics*, 86(1): 211-225.
- Glewwe, P., and E. King (2001). The Impact of Early Childhood Nutritional Status on Cognitive Development: Does the Timing of Malnutrition Matter? *World Bank Economic Review*, 15(1): 81-113.
- Horton, S. (1988). Birth Order and Nutritional Status: Evidence from the Philippines. *Economic Development and Cultural Change*, 36(2): 341-354.
- Katon, W.J. (2003). Clinical and Health Services Relationships between Major Depression, Depressive Symptoms, and General Medical Illness. *Biological Psychiatry*, 54(3): 216-226.
- Kenkel, D. (1991). Health Behavior, Health Knowledge, and Schooling. *Journal of Political Economy*, 99(2): 287-305.
- Lamb, M.E. (Ed.) (2010). *The Role of the Father in Child Development*. Hoboken, NJ: Wiley.
- Leibowitz, A. (1974). Home Investments in Children. *Journal of Political Economy*, 82(2): 111-131.
- Liu, Z., X. Li and X. Ge (2009). Left Too Early: The Effects of Age at Separation. *American Journal of Public Health*, 99(11): 2049-2054.
- Lundberg, M., M. Over and P. Mujinja (2000). Sources of Financial Assistance for Households Suffering an Adult Death in Kagera, Tanzania. *South African Journal of Economics*, 68(5): 948-984.
- Makame, V., C. Ani and S. Grantham-McGregor (2002). Psychosocial Well-being of Orphans in Dar-Es-Salaam, Tanzania. *Acta Paediatrica*, 91(4): 459-465.
- Morgan, D., C. Mahe, B. Mayanja, J.M. Okongo, R. Lubega and J.A.G. Whitworth (2002). HIV-1 Infection in Rural Africa: Is There a Difference in Median Time to AIDS and Survival Compared with That in Industrialized Countries? *AIDS*, 16(4): 597-603.

- National Scientific Council on the Developing Child (2010). Persistent Fear and Anxiety Can Affect Young Children's Learning and Development. Working Paper 9.
- Needham, B. (2009). Adolescent Depressive Symptomatology and Young Adult Educational Attainment: An Examination of Gender Differences. *Journal of Adolescent Health*, 45(2): 179-186.
- Nyamukapa, C.A., S. Gregson, B. Lopman, S. Saito, H.J. Watts, R. Monasch and M.C.H. Jukes (2008). HIV-Associated Orphanhood and Children's Psychosocial Distress: Theoretical Framework Tested with Data from Zimbabwe. *American Journal of Public Health*, 98(1): 133-141.
- Odenstadt, A., A. Hjerna, F. Lindblad, F. Rasmussena, B. Vinnerljung and M. Dalen (2008). Does Age at Adoption and Geographic Origin Matter? A National Cohort Study of Cognitive Test Performance in Adult Inter-Country Adoptees. *Psychological Medicine*, 38(12): 1803-1814.
- O'Donnel, O., E. van Doorslaer, A. Wagstaff, and M. Lindelow (2008). *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation*. Washington, D.C.: World Bank.
- Sharma, M. (2006). Orphanhood and Schooling Outcomes in Malawi. *American Journal of Agricultural Economics*, 88(5): 1273-1278.
- Skinner, D., N. Tsheko, S. Mtero-Munyati, M. Segwabe, P. Chibatamoto, S. Mfecane, B. Chandiwana, N. Nkomo, S. Tlou, G. and Chitiyo (2006). Towards a Definition of Orphaned and Vulnerable Children. *AIDS and Behavior*, 10(6): 619-626.
- Steckel, R.H. (1995). Stature and the Standard of Living. *Journal of Economic Literature*, 33(4): 1903-1940.
- Strauss, J., and D. Thomas (1995). Human Resources: Empirical Modeling of Household and Family Decisions. In H. Chenery and T.N. Srinivasan (Eds.), *Handbook of Development Economics, Volume 3A*. New York, NY: Elsevier.
- Subbarao, K., and D. Coury (2004). *Reaching Out to Africa's Orphans: A Framework for Public Action*. Washington, D.C.: World Bank.
- Thomas, D. (1994). Like Father, Like Son; Like Mother, Like Daughter: Parental Resources and Child Health. *Journal of Human Resources*, 29(4): 950-988.
- Thomas, D. (1990). Intrahousehold Resource Allocation: An Inferential Approach. *Journal of Human Resources*, 25(4): 635-664.
- Thomas, D., J. Strauss and M. Henriques (1991). How Does Mother's Education Affect Child Height? *Journal of Human Resources*, 26(2): 183-211.
- Ulijaszek, S., F. Johnston and A. Preece (1998). *The Cambridge Encyclopedia of Human Growth and Development*. Cambridge: Cambridge University Press.
- UNAIDS (2008). *Report on the Global HIV/AIDS Epidemic 2008*. Geneva: UNAIDS.
- UNAIDS, UNICEF and USAID (2004). *Children on the Brink 2004: A Joint Report on New Orphan Estimates and a Framework for Action*. New York, NY: USAID.
- UNICEF (2009). *The State of the World's Children*. New York, NY: UNICEF.
- WHO (2004). *The Importance of Caregiver-Child Interactions for the Survival and Healthy Development of Young Children: A Review*. Geneva: WHO.

World Bank (2004). User's Guide to the Kagera Health and Development Survey Datasets, mimeo.

World Bank (1999). *Tanzania: Social Sector Review*. Washington, D.C.: World Bank.

Appendix

TABLE A3.1: NUMBER OF ORPHANS BY SAMPLE AND CATEGORY

| | Education sample | Health sample |
|----------|------------------|---------------|
| Maternal | 80 | 169 |
| Paternal | 136 | 267 |
| Double | 32 | 77 |
| Total | 184 | 369 |

Notes - The gender distribution is very even. Boys represent 49.5 percent and 48.8 percent of the education and health sample, respectively. Entries in one column do not sum to the total because double orphans are also included as both maternal and paternal orphans.

TABLE A3.2: SUMMARY STATISTICS

| | Health sample (n=369) | | | | Education sample (n=184) | | | |
|--------------------------------|-----------------------|-----------|-------|--------|--------------------------|-----------|-------|-------|
| | Mean | Std. dev. | Min | Max | Mean | Std. dev. | Min | Max |
| height-for-age z-score in 2004 | -1.56 | 1.04 | -4.31 | 1.49 | - | - | - | - |
| years of schooling in 2004 | - | - | - | - | 5.99 | 2.68 | 0.00 | 13.00 |
| age at bereavement | 10.90 | 4.44 | 0.75 | 17.92 | 13.59 | 2.93 | 6.67 | 17.92 |
| father died first | 0.70 | 0.46 | 0.00 | 1.00 | 0.69 | 0.46 | 0.00 | 1.00 |
| double orphan | 0.21 | 0.41 | 0.00 | 1.00 | 0.18 | 0.38 | 0.00 | 1.00 |
| male | 0.49 | 0.50 | 0.00 | 1.00 | 0.49 | 0.50 | 0.00 | 1.00 |
| height-for-age z-score | -1.76 | 1.34 | -4.98 | 3.28 | - | - | - | - |
| schooling for age | - | - | - | - | -2.08 | 1.88 | -9.00 | 1.00 |
| first-born | 0.28 | 0.45 | 0.00 | 1.00 | 0.33 | 0.47 | 0.00 | 1.00 |
| mother present at baseline | 0.76 | 0.43 | 0.00 | 1.00 | 0.70 | 0.46 | 0.00 | 1.00 |
| father present at baseline | 0.69 | 0.46 | 0.00 | 1.00 | 0.71 | 0.46 | 0.00 | 1.00 |
| # hh members 0-18 years | 5.17 | 2.29 | 0.00 | 19.00 | 5.24 | 2.32 | 1.00 | 19.00 |
| # hh members 19-60 years | 2.31 | 1.19 | 0.00 | 8.00 | 2.30 | 1.27 | 0.00 | 8.00 |
| # hh members 61+ years | 0.51 | 0.58 | 0.00 | 2.00 | 0.57 | 0.60 | 0.00 | 2.00 |
| male hh head | 0.82 | 0.39 | 0.00 | 1.00 | 0.84 | 0.37 | 0.00 | 1.00 |
| asset index | -0.07 | 1.53 | -1.67 | 7.46 | -0.03 | 1.42 | -1.62 | 5.86 |
| mother attended school | 0.73 | 0.44 | 0.00 | 1.00 | 0.64 | 0.48 | 0.00 | 1.00 |
| father attended school | 0.92 | 0.28 | 0.00 | 1.00 | 0.91 | 0.28 | 0.00 | 1.00 |
| age in 2004 | 18.11 | 4.69 | 9.00 | 30.00 | 22.63 | 2.80 | 19.00 | 30.00 |
| height of mother | 125.67 | 63.78 | 0.00 | 173.20 | - | - | - | - |
| height of mother missing | 0.20 | 0.40 | 0.00 | 1.00 | - | - | - | - |

TABLE A3.3: PREDICTING BASELINE OUTCOMES WITH FUTURE AGE AT BEREAVEMENT AND A FULL SET OF BASELINE COVARIATES

| Baseline outcome | Estimation method | Coefficient | Std. err. | P-value | Observations |
|------------------------|-------------------|-------------|-----------|---------|--------------|
| <i>Health</i> | | | | | |
| height-for-age z-score | OLS | -0.009 | 0.023 | 0.695 | 369 |
| <i>Education</i> | | | | | |
| years of schooling | OLS | -0.018 | 0.032 | 0.571 | 184 |
| years of schooling | Tobit | -0.025 | 0.036 | 0.499 | 184 |
| enrolled | Probit | 0.013 | 0.043 | 0.755 | 184 |

TABLE A3.4: EFFECTS OF AGE AT BEREAVEMENT ON BODY HEIGHT ATTAINMENT IN EARLY ADULTHOOD

| | Community fixed effects | | | | Family fixed effects | | | |
|---|-------------------------|-------|-------------------|-------|----------------------|-------|-------------------|-------|
| | no interactions | | with interactions | | no interactions | | with interactions | |
| | coef | se | coef | se | coef | se | coef | se |
| age at bereavem. | 0.037** | 0.015 | 0.019 | 0.039 | 0.067 | 0.061 | 0.008 | 0.057 |
| age at bereavem.*male | | | 0.001 | 0.047 | | | -0.054 | 0.080 |
| age at bereavem.*father died | | | -0.007 | 0.039 | | | 0.062 | 0.052 |
| age at bereavem.*father died*male | | | 0.064 | 0.050 | | | 0.098 | 0.082 |
| (age at bereavem.) ² | -0.001 | 0.002 | -0.014** | 0.007 | -0.000 | 0.004 | -0.006 | 0.012 |
| (age at bereavem.) ² *male | | | 0.015+ | 0.009 | | | 0.003 | 0.013 |
| (age at bereavem.) ² *father died | | | 0.012+ | 0.008 | | | 0.007 | 0.013 |
| (age at bereavem.) ² *father died*male | | | -0.007 | 0.010 | | | 0.003 | 0.015 |
| male | -0.560** | 0.096 | -0.707** | 0.235 | -0.500** | 0.129 | -0.448 | 0.408 |
| father died | | | -0.074 | 0.215 | | | 0.276 | 0.783 |
| father died*male | | | -0.031 | 0.255 | | | -0.219 | 0.441 |
| double orphan | 0.151 | 0.120 | 0.159 | 0.119 | -0.019 | 0.285 | -0.020 | 0.278 |
| baseline H/A z-score | 0.345** | 0.042 | 0.348** | 0.043 | 0.378** | 0.077 | 0.355** | 0.073 |
| first-born | 0.113 | 0.114 | 0.135 | 0.113 | -0.211 | 0.172 | -0.178 | 0.194 |
| mother present at baseline | 0.175 | 0.262 | 0.149 | 0.271 | | 0.680 | | |
| father present at baseline | 0.148 | 0.133 | 0.134 | 0.136 | | | | |
| # hh members 0-18 years | 0.006 | 0.025 | 0.004 | 0.023 | | | | |
| # hh members 19-60 years | -0.034 | 0.045 | -0.025 | 0.051 | | | | |
| # hh members 61+ years | 0.199** | 0.098 | 0.190* | 0.097 | | | | |
| male hh head | 0.063 | 0.133 | 0.057 | 0.138 | | | | |
| asset index | -0.061+ | 0.041 | -0.065+ | 0.040 | | | | |
| mother attended school | 0.079 | 0.111 | 0.075 | 0.110 | | | | |
| father attended school | 0.090 | 0.182 | 0.140 | 0.192 | | | | |
| mother's height | 0.047** | 0.008 | 0.043** | 0.008 | | | | |
| mother's height missing | 7.660** | 1.295 | 7.030** | 1.238 | | | | |
| constant | -9.738** | 1.385 | -8.652** | 1.339 | -3.188** | | -3.276** | 0.665 |
| age in 2004 dummies | | yes | | yes | | yes | | yes |
| Number of observations | | 369 | | 369 | | 267 | | 267 |
| Adjusted R2 | | 0.403 | | 0.412 | | 0.342 | | 0.368 |

**/*/+ denote significance at the 5/10/15 percent level, respectively. Standard errors are clustered at the household level.

TABLE A3.5: EFFECTS OF AGE AT BEREAVEMENT ON EDUCATIONAL ATTAINMENT IN EARLY ADULTHOOD

| | Years of schooling in 2004 | | | | | | | |
|---|----------------------------|-------|-------------------|-------|----------------------|-------|-------------------|-------|
| | Community fixed effects | | | | Family fixed effects | | | |
| | no interactions | | with interactions | | no interactions | | with interactions | |
| | coef | se | coef | se | coef | se | coef | se |
| age at bereavem. | 0.123 | 0.091 | -0.054 | 0.189 | 0.408 | 0.293 | 0.370 | 0.645 |
| age at bereavem.*male | | | 0.070 | 0.206 | | | 0.113 | 0.726 |
| age at bereavem.*father died | | | 0.208 | 0.245 | | | 0.067 | 0.695 |
| age at bereavem.*father died*male | | | 0.036 | 0.269 | | | 0.331 | 0.764 |
| (age at bereavem.) ² | 0.012 | 0.023 | 0.018 | 0.058 | 0.002 | 0.040 | 0.119 | 0.216 |
| (age at bereavem.) ² *male | | | -0.020 | 0.077 | | | -0.038 | 0.224 |
| (age at bereavem.) ² *father died | | | 0.009 | 0.072 | | | -0.138 | 0.209 |
| (age at bereavem.) ² *father died*male | | | -0.013 | 0.089 | | | 0.026 | 0.247 |
| male | -0.284 | 0.439 | -0.774 | 0.971 | 0.376 | 0.738 | 3.231 | 2.373 |
| father died | | | -1.237 | 1.066 | | | 3.256 | 2.717 |
| father died*male | | | 1.071 | 1.379 | | | -3.105 | 2.359 |
| double orphan | -0.418 | 0.599 | -0.377 | 0.575 | 0.510 | 1.097 | 0.528 | 0.846 |
| schooling for age | 0.290** | 0.143 | 0.325** | 0.153 | 0.532** | 0.245 | 0.779** | 0.240 |
| first-born | 0.142 | 0.510 | 0.158 | 0.543 | 0.098 | 0.782 | 0.038 | 0.878 |
| mother present at baseline | 0.262 | 0.446 | 0.184 | 0.480 | | | | |
| father present at baseline | -0.585 | 0.450 | -0.244 | 0.510 | | | | |
| # hh members 0-18 years | 0.003 | 0.073 | 0.049 | 0.081 | | | | |
| # hh members 19-60 years | -0.079 | 0.143 | -0.113 | 0.164 | | | | |
| # hh members 61+ years | -0.430 | 0.380 | -0.332 | 0.417 | | | | |
| male hh head | 0.502 | 0.661 | 0.283 | 0.774 | | | | |
| asset index | 0.778** | 0.202 | 0.704** | 0.268 | | | | |
| mother attended school | 0.900** | 0.420 | 0.940** | 0.442 | | | | |
| father attended school | 0.324 | 0.741 | 0.313 | 0.877 | | | | |
| constant | 5.290** | 1.362 | 6.013** | 1.258 | 7.353** | 1.901 | 5.719** | 2.906 |
| age in 2004 dummies | | yes | | yes | | yes | | yes |
| Number of observations | | 184 | | 184 | | 104 | | 104 |
| Adjusted R2 | | 0.096 | | 0.412 | | 0.342 | | 0.368 |

**/*/+ denote significance at the 5/10/15 percent level, respectively. Standard errors are clustered at the household level.

FIGURE A3.1: SCATTER PLOT AND HISTOGRAMS OF HEIGHT-FOR-AGE Z-SCORES AND AGE AT BEREAVEMENT

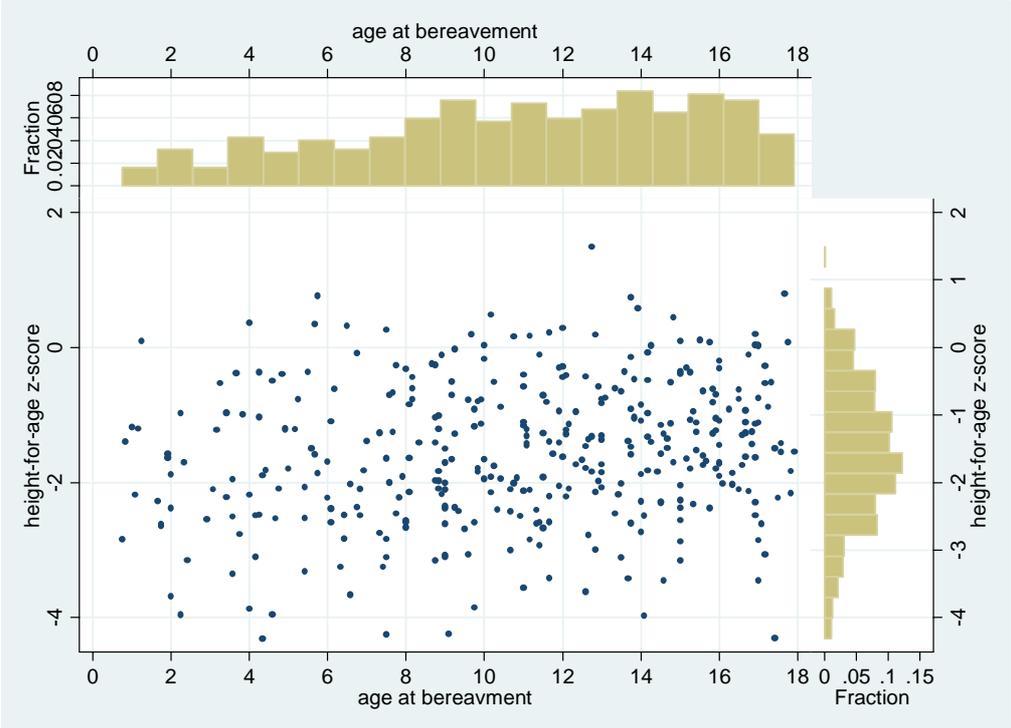
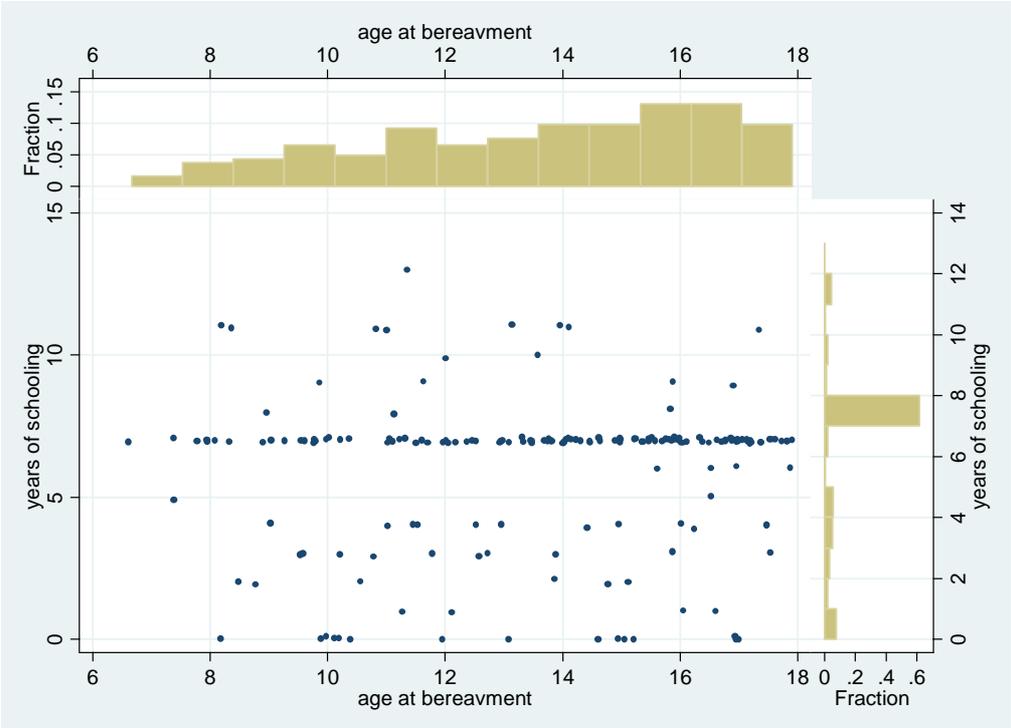


FIGURE A3.2: SCATTER PLOT AND HISTOGRAMS OF YEARS OF SCHOOLING AND AGE AT BEREAVEMENT



Note: To enhance the visibility of multiple observations per grade, this scatter plot introduces minor random variation in the discrete variable *years of schooling*.

Chapter 4

Few Opportunities, Much Desperation: The Dichotomy of Non-Agricultural Activities and Inequality in Western Kenya³⁰

published 2008 in *World Development*, 36(12): 2713-2732

Jann Lay
German Institute of Global and Area Studies
University of Göttingen

Toman Omar Mahmoud³¹
Kiel Institute for the World Economy

George Michuki M'Mukaria
Institute for Development Studies, Nairobi

Abstract

Using data from Western Kenya, we confirm the existence of a dichotomous non-agricultural sector. The poverty and inequality implications of the differently motivated diversification strategies only partly correspond to expected patterns. While high-return activities are indeed confined to richer households, low-return activities constitute an important income source for households across the entire income distribution. Finally, we examine the wider implications of our findings for rural livelihoods. We find that only engagement in high-return non-agricultural activities is significantly associated with increased agricultural productivity. It seems that such high-return activities play a key role in triggering cumulative effects of relative livelihood success.

Keywords: income diversification, non-agricultural activities, inequality, agricultural productivity, sub-Saharan Africa, Kenya

JEL classification: Q12, O17, I31

³⁰ We thank Awudu Abdulai and Rainer Thiele and participants of the AERC-Cornell Conference in Nairobi and the IARIW-SAIM Conference in Kathmandu for helpful comments and discussions.

³¹ Corresponding author: toman.mahmoud@ifw-kiel.de

4.1 Introduction

A number of studies have shown that rural households in sub-Saharan Africa derive their incomes from a variety of sources with non-agricultural activities accounting for a substantial share of total income.³² Despite the importance of non-agricultural activities for rural farm households, we still know little about the impact of such activities on the distribution of income and, hence, on poverty.

Broadly, one may classify diversification strategies of rural households as survival-led or opportunity-led. It has been observed that poor households with low asset endowments embrace multiple livelihoods, in particular engagement in non-agricultural activities, to ensure survival. These households are forced to diversify mainly because they lack sufficient agricultural assets to sustain subsistence (Reardon and Taylor, 1996; Haggblade et al., 2005). Returns to these activities may well be below those in agriculture. At the same time, richer rural households with higher asset endowments will choose to diversify their livelihoods to maximize returns on their assets. Such activities will have at least the same returns as agricultural activities and exhibit entry barriers that the poor are not able to overcome.

The existence of these two types of non-agricultural activities implies a U-shaped relationship between the share of income derived from non-agricultural activities and household wealth as well as household income. The poverty and distributional impact of non-agricultural incomes should hence be ambiguous: Survival-led engagement in non-agricultural activities should be inequality-decreasing through increasing the incomes of the poorer parts of the population and hence reduce poverty. Opportunity-led diversification, however, should increase inequality and have a minor effect on poverty, as it may be confined to non-poor households. Some authors have pointed to this ambiguity (Ferreira and Lanjouw, 2001; Haggblade et al., 2005), but only few, e.g. Elbers and Lanjouw (2001) for rural Ecuador, explicitly address the relationship between different diversification strategies, on the one hand, and poverty and inequality outcomes, on the other. This paper intends to fill this gap by providing evidence from sub-Saharan Africa. We test for the dichotomy of the non-agricultural activities and examine their distributional consequences using household survey data from Kakamega district in Western Kenya that can be considered as representative for the densely populated rural areas of many parts of Eastern sub-Saharan Africa.

We proceed as follows. In the first section we shortly review the theoretical and empirical literature on non-agricultural activities and their poverty and distributional implications. Then, we provide evidence on the incidence and characteristics of the non-agricultural economy in the study region. Subsequently, we present the results of the choice model and, based on this typology of non-agricultural incomes, examine the poverty and inequality implications. We then briefly examine the relationship between the different non-agricultural activities and agricultural

32 Reardon (1997), Reardon et al. (1998), Ellis (2000), Haggblade et al. (2005).

productivity. The last section concludes with policy implications and an outlook for future research.

4.2 The Rural Non-Agricultural Economy: Theory and Empirics

The non-agricultural economy involves employment outside the realm of direct soil cultivation and cattle breeding and includes activities such as services, construction, mining, commerce, manufacturing and processing. Such activities are often pursued through self-employment, but there is also a non-agricultural wage labor market, although this market is typically small in the rural sub-Saharan African context. The contribution of these activities to household income in the developing world in general and sub-Saharan Africa in particular is substantial. Haggblade et al. (2005) observe that non-agricultural income contributes between 30 to 45 percent of rural household incomes in the developing world. Reardon et al. (1998) put this share at 42 percent for sub-Saharan Africa, while Reardon (1999) gives estimates of 32 percent and 40 percent for Asia and Latin America, respectively. Ellis (2000) reports somewhat higher figures from case studies in sub-Saharan Africa in a range of 30 to 50 percent. Jayne et al.'s (2003) estimates for five sub-Saharan African countries lie between 8 percent for Ethiopia and 40 percent for Kenya.

4.2.1 Low- vs. High-Return Activities and Drivers of Participation

Rapid population growth and the related pressure on the natural resource base, in particular land, have been identified as major causes for the rise of non-agricultural activities in sub-Saharan Africa.³³ Other driving forces are supply factors, such as technological advances and the expansion of educational attainment, as well as demand shocks, including higher per capita incomes and increased demand for non-food goods and services (Reardon, 1997).

In order to understand how these variables affect the participation in and patterns of non-agricultural employment and the related incomes, it is useful to differentiate between survival-led and opportunity-led diversification into non-agricultural activities.

When non-agricultural diversification is pursued to ensure survival, for example because of land constraints, it is also referred to as distress-push diversification (Islam, 1997; Reardon et al., 2000; Barrett, Reardon and Webb, 2001). Such diversification will be in low-return non-agricultural activities and may be an indication that the non-agricultural sector is absorbing labor that cannot be employed in agriculture. In contrast, rural households may face new opportunities outside agriculture because of technological advances, the intensification of links with markets outside the local economy, or local engines of growth, such as commercial agriculture or proximity to an urban area.³⁴ If non-agricultural income diversification can be traced back to such factors, it is also regarded as demand-pull diversification.

33 Bryceson and Jamal (1997), Barrett et al. (2000), Bryceson (2002).

34 Reardon (1999), Haggblade et al. (2002).

While the aggregate prevalence of the specific type of non-agricultural diversification in a region (or country) will hence be driven by meso (or macro) determinants, household characteristics will decide on the individual household's diversification strategy. The literature has stressed asset availability and educational endowments as key participation determinants of non-agricultural diversification (Abdulai and CroleRees, 2001; Abdulai and Delgado, 1999; Barrett, Reardon and Webb, 2001; Escobal, 2001; Matsumoto et al., 2006). Whereas entry barriers to low-return diversification should be low, they can be considerable for high-return activities (Abdulai and CroleRees, 2001). In the presence of underdeveloped credit markets, the latter typically require sufficient cash income, in particular from livestock, cash cropping, and/or remittances, both for initial investment and as working capital (Reardon et al., 2000; Barrett, Bezuneh and Aboud, 2001). Skill requirements may impose another important entry constraint (Abdulai and Delgado, 1999; Dercon and Krishnan, 1996; Reardon, 1997). Some high-return non-agricultural activities such as skilled wage employment are restricted to those with formal education.³⁵

Non-agricultural diversification constitutes an important means to deal with risk and smooth income and consumption in rural areas. This is not surprising since agricultural livelihoods are often subject to great uncertainty. In such an environment, diversification aims at lower covariate risk between different household activities to smooth consumption (Bryceson, 1999; Dercon, 1998; Francis and Hoddinott, 1993). For our discussion, it is useful to distinguish between ex-ante risk management and ex-post risk coping strategies. Engagement in high-return non-agricultural activities represents an ex-ante risk management strategy, as it is unlikely that entry barriers can be easily overcome after a negative shock. In contrast, low-return non-agricultural diversification will figure prominently as an ex-post coping strategy, i.e. households will relocate labor towards these activities after they have been hit by a negative agricultural shock, typically a weather shock. Yet, in particular poorer household may also be willing to accept lower returns than in agriculture ex-ante in exchange for lower covariate risk.

While rural households can reduce risks by venturing into non-agricultural activities, risk considerations may also play a role when deciding between different types of non-agricultural activities. If high-return activities are more risky than low-return activities, households able to overcome possible entry barriers may engage in both types of non-agricultural activities according to their risk preferences.

The empirical literature on the rural non-agricultural economy has emphasized the drivers of participation in these activities at the individual, household, and community level. Some of the empirical contributions have distinguished between low- and high-return activities in doing so. Studies in sub-Saharan Africa, Asia, and Latin America have confirmed that the level of formal education is positively correlated with participation in non-agricultural activities, in general, and high-return activities, in particular (Abdulai and Delgado, 1999; Ferreira and Lanjouw, 2001; Lanjouw, 2001; van den Berg and Kumbi, 2006). Land and other productive assets have also

35 Corral and Reardon (2001), Lanjouw (2001), Reardon (1997).

been demonstrated to be important determinants of different types of diversification strategies (Abdulai and CroleRees, 2001; Seppala, 1996; Elbers and Lanjouw, 2001; Marenya et al., 2003). For instance, Elbers and Lanjouw (2001) show that land scarcity is a driving force of participation in a low-return non-agricultural activity while more landholdings seem to provide collateral for investment in high-return non-agricultural businesses. Another household level factor correlated with participation in non-agricultural activities is the size and structure of the household (Corral and Reardon, 2001; Reardon, 1997; van den Berg and Kumbi, 2006). On the one hand, a larger household size enables households to supply more labor to non-agricultural activities, since sufficient family members remain at home to meet labor demands for agricultural subsistence. On the other hand, however, large households are also more likely to suffer from land constraints. Thus, they might find themselves forced to engage in non-agricultural activities to ensure survival. As regards community level determinants, most empirical studies confirm an important role for physical and institutional infrastructure, such as paved roads, efficient communication facilities and provision of rural electrification (Abdulai and Delgado, 1999; Jalan and Ravallion, 2002; Lanjouw and Feder, 2001).

4.2.2 Non-Agricultural Incomes, Poverty and Inequality

There are limited empirical accounts of the relationship between farm households' income composition and inequality, in particular for sub-Saharan Africa.³⁶ Moreover, most existing studies do not distinguish between different types of non-agricultural activities (e.g. Adams, 2002). From the above discussion, it has become apparent that the equity impact of non-agricultural employment depends on the type of activity. Early work on the informal sector (ILO, 1972) claims that, given their intrinsic characteristics, such as easy entry, non-agricultural activities will decrease income inequalities, particularly through self-employment. While low-return activities undertaken by poorer households should hence be inequality-decreasing, high-return activities may well increase inequality, as they tend to reinforce asset inequalities. Accordingly, Haggblade et al. (2005) observe that because of the differing equity impact of various segments of these activities, their overall effect on income distribution remains mixed. Thus, depending on the nature and motivation of non-agricultural activity undertaken, aggregate non-agricultural earnings improve equity in some instances, while they aggravate inequality in others.

This explains the differing results of empirical studies on the equity impact of non-agricultural activities. In fact, most empirical studies tend to find that non-agricultural incomes go primarily to the better-off so that higher non-agricultural incomes (as opposed to more non-agricultural income earners) are associated with higher income inequality. For example, case studies on Burkina Faso by Reardon et al. (1992) and on Ecuador by Elbers and Lanjouw (2001) find that

³⁶ Some exceptions that explicitly address the poverty and inequality implications of non-agricultural activities in sub-Saharan Africa include Jayne et al. (2003), Kijima et al. (2006) on Uganda and van den Berg and Kumbi (2006) on Ethiopia. From a more comprehensive livelihood perspective, the four studies summarized by Ellis and Freeman (2004) examine the link between income portfolios and welfare outcomes. A case study from Kenya can be found in Freeman et al. (2004).

the income share from non-agricultural sources increases with per capita income. For Mexico, de Janvry and Sadoulet (2001) show non-agricultural wages to be inequality-increasing, while non-agricultural self-employment tends to decrease inequality. Some empirical studies show an inequality-decreasing effect of non-agricultural activities, e.g. by Norman et al. (1982) on rural households in Northern Nigeria, Adams (2002) in Egypt, or van den Berg and Kumbi (2006) in Ethiopia.

It is likely that these seemingly contradictory findings could be reconciled through an analysis of the underlying type of non-agricultural activities. Once this is understood, the conditions that drive the prevalence of one type of non-agricultural activity or another should be addressed, as they eventually represent the fundamental causes of the inequality implications. According to Reardon et al. (1998) such conditions include the proximity to urban markets, physical and market infrastructure, resource endowments and the distribution of productive resources within rural areas. In the following, we will (i) assess which kind of activities prevail in the poverty-ridden context that we have studied and (ii) examine whether the poverty and distributional consequences correspond to the patterns one could expect under the specific conditions in the study region.

4.3 The Patterns of Non-Agricultural Activities in Kakamega

The data for our analysis come from a household survey which was conducted in Kakamega district, a densely populated and poor rural area of Western Kenya, in the last quarter of the year 2005. In this year, Kakamega district had a population of about 700,000 with a mean population density of approximately 500 persons per square kilometer. The largest urban agglomeration in the region, the district capital Kakamega Town, has a population of about 85,000, which only partly resides in a strictly urban setting. The region is considered to have good agricultural potential, yet population growth and further subdivision of land has led to land fragmentation and extremely small farm sizes in parts of the district. These developments have also caused high rates of out-migration.³⁷

The survey was designed with a focus on economic activities of rural households in both agriculture and non-agricultural sectors. This focus allows for a very detailed analysis of these activities taking into account regional specificities. Yet, it also constrains the subsequent empirical analysis to those characteristics and leaves out many social and cultural aspects that certainly play a role in determining activity choices and hence welfare outcomes. Only some of those aspects can be captured in a proximate way in the following empirical setup, for example the socio-culturally determined control over household resources.

³⁷ Detailed descriptions of similar rural livelihoods in nearby Kisumu district with a focus on migration can be found in Francis (2000). The study also contains insightful accounts of employment histories in the non-agricultural sector. For more details on Kakamega district see Government of Kenya (2002). A recent analysis of migration using the same data as this paper is provided by Giesbert (2007).

The survey used a two-stage sampling technique and covered 375 households with a total of 1950 household members, which were spread over 20 clusters. One cluster, however, was excluded from our sample as it was mainly inhabited by teachers who had been sent to the study region.³⁸ Six of the remaining clusters were located in urban or peri-urban areas. Nevertheless, households in these clusters derive a substantial amount of their income from agriculture-related activities and therefore form part of our sample. It is important to note that this paper focuses on locally conducted non-agricultural activities. This is why non-agricultural activities, as used subsequently, exclude migration, another important diversification strategy of rural households in the region.

In the following, we give an overview of the structure of household income and non-agricultural employment patterns in the study region, taking into account the different character of low-return and high-return non-agricultural activities. In contrast to other authors, e.g. Ferreira and Lanjouw (2001) who define high-return non-agricultural activities as those whose monthly returns are above the poverty line our definition of such activities takes a different approach. On the one hand, activities based on self-employment are considered to be high-return activities if the household enterprise employs at least one hired worker or two household members. Given the rural character of the Kakamega district, such-defined enterprises can well be assumed to generate higher incomes than traditional agricultural activities. On the other hand, the definition of wage-based high-return activities draws on specific sectors which typically exhibit entry constraints. In our view, these entry barriers should not only encompass special skills or asset requirements, but also such seemingly simple hurdles like a clean and healthy appearance, which some poor households may well not be able to overcome. Accordingly, the following sectors offer high-return wage employment in the study region: repair of motor vehicles, medical services, hair dressing and beauty, churches/NGOs/international organizations, and hotels and restaurants. This definition is bolstered by the fact that high-return wage employment can only be found as primary occupation and not as a secondary one for all individuals in the sample. It is important to note that we exclude employment in the public sector from our analysis. Entry barriers in this segment are likely to be very different from those in other high-return activities, as public employment is often arranged by nepotistic and corrupt structures.

All remaining forms of non-agricultural employment, i.e. household enterprises which are run by one household member only and wage employment in non-agricultural sectors other than those mentioned above, constitute the group of low-return non-agricultural activities. With this definition of low-return and high-return non-agricultural employment, we believe to adequately reflect the idea of survival-led and opportunity-led income diversification.

Table 4.1a shows the participation rates of households in different types of activities. It reveals that households in Kakamega district earn income from a variety of activities.³⁹ As can be

38 Though non-agricultural employment comprises teaching activities, the observed patterns in this cluster are not compatible with the idea of rural income diversification.

39 Our analysis includes both primary and secondary non-agricultural employment.

expected for a rural region, the vast majority of the households (86 percent) work at least partly in agriculture. Still, 46 percent engage in low-return and 20 percent in high-return non-agricultural activities. In addition, migration is a common phenomenon in the area. A fourth of all households reports income in the form of remittances.

Table 4.1b presents a matrix of agricultural and non-agricultural income-generating activities of households. The rows show in which activities households are engaged in addition to the activities indicated in the columns. Accordingly, households on the diagonal line do not diversify their income sources.

Out of all farming households, only 26 percent are fully specialized in agriculture, while about 55 percent diversify into non-agricultural activities, primarily into low-return activities. With a participation rate of 26 percent, migration is also an important diversification strategy of agricultural households. Non-agricultural income constitutes the sole income source for roughly 17 percent of all households. This relatively high proportion is principally due to the fact that our sample still includes the urban and peri-urban clusters to give a more complete overview of the income-generating activities in the study region. Moreover, the table illustrates that more than a third of the households engaged in high-return activities in the non-agricultural sector also pursue some low-return activity. Among the households with non-agricultural activities, about 18 percent also report receiving remittances.

TABLE 4.1A: PARTICIPATION OF HOUSEHOLDS IN INCOME-GENERATING ACTIVITIES

| | Agriculture | Local non-agricultural activities | | | Public employment | Migration (remittances) | No activity |
|-------------------------------------|---------------|-----------------------------------|---------------|--------------|-------------------|-------------------------|-------------|
| | | All | Low-return | High-return | | | |
| Households with respective activity | 314 86.03% | 214 58.63% | 166 45.48% | 73 20.00% | 26 7.12% | 90 24.66% | 4 1.10% |

The total number of households is 365.

TABLE 4.1B: INCOME DIVERSIFICATION STRATEGIES OF HOUSEHOLDS

| | Agriculture | Local non-agricultural activities | | | Public employment | Migration (remittances) |
|-------------------------|---------------|-----------------------------------|---------------|--------------|-------------------|-------------------------|
| | | All | Low-return | High-return | | |
| Agriculture | 82 26.11% | 174 81.31% | 134 80.72% | 56 76.71% | 22 84.62% | 82 91.11% |
| All non-agr. activities | 174 55.41% | 36 16.82% | | | 6 23.08% | 39 43.33% |
| Low-return activities | 134 42.68% | 166 77.57% | 21 12.65% | 25 34.25% | 5 19.23% | 28 31.11% |
| High-return activities | 56 17.83% | 73 34.11% | 25 15.06% | 6 8.22% | 2 7.69% | 15 16.67% |
| Public employment | 22 7.01% | 6 2.8% | 5 3.01% | 2 2.74% | 3 11.54% | 9 10% |
| Migration (remittances) | 82 26.11% | 39 18.22% | 28 16.87% | 15 20.55% | 9 34.62% | 3 3.33% |
| Total | 314 | 214 | 166 | 73 | 26 | 90 |

Due to the fact that a number of households are involved in more than two income-generating activities, the percentage shares do not add up to 100 percent.

Since the focus of this paper is on rural income diversification, we now confine our sample to households which have access to at least half an acre of land and engage in agricultural activities. This step makes the sample more likely to include only those households which diversify out of agriculture and not the ones which have some limited supplementary agricultural activities. Virtually all excluded households are from urban or peri-urban areas.

Based on this sample, we compile a detailed profile of non-agricultural activities in Kakamega district, which clearly reveals that the rural non-agricultural sector is dominated by low-return activities and provides relatively little space for high-return activities. In addition, most non-agricultural activities take the form of household enterprises. In total, we find 136 such enterprises in the sample, 99 of which belong to the low-return segment. This compares to 37 wage-employed individuals, out of which 19 are in the low-return segment. Thus, in the case of wage employment the frequency of low-return and high-return activities seems to be roughly equal.

As regards high-return wage employment, all recorded activities belong to the service sector. Most individuals work with churches, NGOs or international organizations, followed by hotels and restaurants. Interestingly, the same number of men and women are engaged in these activities, suggesting that both sexes have equal access to them. In contrast, low-return wage employment seems to favor men as it often requires physical strength though most activities again belong to the service sector. Only four out of the 19 individuals in this segment of non-agricultural employment are women. The most frequent low-return wage activities include security, food production, and retailing. Women, however, are solely active in retailing, house-help, and informal services.

Micro and small businesses are involved in a fairly wide range of activities, primarily retailing, informal services such as shoe-shining and washing, boda-boda transportation⁴⁰, and construction. Only in the sphere of retailing can we find a concentration of both low-return and high-return household enterprises. This suggests that low-return and high-return businesses operate in relatively segmented markets. Informal services and boda-boda transportation are exclusively provided by low-return enterprises, whereas formal services and food production are clearly dominated by high-return businesses. Some manufacturing activities can be observed in non-agricultural self-employment. These comprise food production, carpentry, and the manufacturing of textile products.

We now turn to the analysis of the determinants of rural non-agricultural employment in the study region. First, we inspect the shares of income from agricultural and non-agricultural activities by basic characteristics of the household head and the household itself. Subsequently, we estimate a multivariate choice model of participation in different types of non-agricultural activities.

40 Boda-boda is a bicycle taxi in East Africa (from English border-border). The bicycle rider can also be called boda-boda.

Table 4.2 presents income shares from different sources tabulated by some key determinants of participation in non-agricultural activities.⁴¹ Overall, income from non-agricultural activities constitutes a major source of income in rural areas of Kakamega district, accounting for 23 percent of total household income on average. Yet, this share is below the averages that have been found in similar studies for sub-Saharan Africa.

TABLE 4.2: SHARES OF INCOME FROM AGRICULTURAL AND NON-AGRICULTURAL ACTIVITIES

| | Households in respective category | Income from | | | | |
|--|---|--|--------------------------|---------------------------|-------------|------------------|
| | | All non- agricultural activities | Low-return activities | High-return activities | Agriculture | Other sources |
| All | 100% | 23% | 16% | 7% | 67% | 10% |
| <i>Age of household head</i> | | | | | | |
| -25 | 6% | 40% | 28% | 12% | 53% | 7% |
| 25-35 | 25% | 29% | 20% | 9% | 64% | 7% |
| 35-45 | 22% | 29% | 20% | 9% | 59% | 11% |
| 45-55 | 18% | 18% | 11% | 6% | 67% | 16% |
| 55- | 29% | 12% | 9% | 4% | 78% | 9% |
| <i>Sex of household head</i> | | | | | | |
| male | 72% | 27% | 18% | 9% | 65% | 9% |
| female | 28% | 12% | 10% | 3% | 74% | 14% |
| <i>Level of formal education of household head</i> | | | | | | |
| no formal education | 26% | 16% | 14% | 3% | 74% | 9% |
| uncompleted primary school | 46% | 25% | 18% | 7% | 71% | 4% |
| completed primary school | 18% | 24% | 16% | 8% | 67% | 9% |
| completed secondary school and higher | 10% | 24% | 12% | 12% | 53% | 23% |
| <i>Land size in acres</i> | | | | | | |
| 0.5-1.5 | 38% | 27% | 21% | 6% | 62% | 11% |
| 1.5-3.5 | 42% | 19% | 13% | 6% | 71% | 10% |
| 3.5- | 20% | 23% | 12% | 10% | 69% | 9% |
| <i>Location</i> | | | | | | |
| rural | 88% | 21% | 15% | 6% | 70% | 9% |
| urban | 12% | 35% | 18% | 17% | 48% | 17% |

We find that in all age groups of the household head agriculture remains the dominant income source, as it always accounts for at least half of all household income. As the age of the household head rises, however, the share of agricultural income increases significantly while the share of income from non-agricultural employment drops sharply. This observation is true for both low-return and high-return non-agricultural activities though the share of low-return income is consistently higher than the share of high-return income. We may relate this pattern to the fact

41 The household's total gross monthly income is computed as the sum of all income from wage or self-employment in the low and high-return non-agricultural sector, farm income including the imputed value of unsold farm produce as well as other income sources such as remittances or pensions. Income from employment in the public sector is also subsumed under other income.

that under traditional land subdivision and inheritance norms older household heads have better claim to land resources (Jayne et al., 2003). This gives them a head start when it comes to agricultural activities, whereas younger household heads will have to embrace non-agricultural strategies to secure their livelihoods.

Looking at the sex of the household head, the table shows that households with a female head earn considerably lower income shares from both types of non-agricultural activities. Given the lower number of adult members in female-headed households and the numerous tasks of their heads in agriculture, housekeeping and child-rearing, these households' ability to engage in non-agricultural employment is likely to be limited.

The share of non-agricultural income does not rise with the level of formal education and accounts for roughly 25 percent of total household income. With 16 percent, however, this share is considerably lower in the case of household heads without any formal education at all. Yet, the actual relationship between non-agricultural income and educational attainment is clouded by the dichotomy of the non-agricultural sector.⁴² The share of low-return non-agricultural income falls with the educational level of the household head, whereas the reverse is true for the share of high-return non-agricultural income.

The tabulation of the share of total non-agricultural income with total land holdings generates a U-shaped relationship. Whereas households with low endowments of land earn about 27 percent of their income from non-agricultural activities, this share falls to 19 percent for medium-endowed households, and then rises again to 23 percent for households that are highly endowed with land. A separate inspection of the shares of income derived from low-return and high-return activities again reveals the two-pronged diversification behavior. The share of income from low-return activities drops sharply with increasing landholdings. The opposite effect is observable for the income share of high-return activities. For many households in sub-Saharan Africa, land is a key asset and serves multiple uses including cultivation, sustaining livestock, storing wealth, and providing collateral for financial credit. The findings hence support the notion that declining farm sizes and related declines in soil fertility force land-poor households to diversify into non-agricultural employment to ensure survival. At the same time, higher land endowments may enable households to diversify into high-return activities as land may serve as collateral for credit or simply provide higher cash flows from agriculture for the necessary start-up capital.

Even in our restricted sample, the share of non-agricultural income considerably varies with the location of the household. Whereas rural households earn about 21 percent of their income in the non-agricultural sector, the corresponding figure climbs to 35 percent for households in peri-urban areas. In particular income from high-return activities plays a more important role for

⁴² The observation that the income share from agricultural activities steadily decreases with educational attainment can be explained by the income earned from employment in the public sector, which is included in income from other sources. As can be clearly seen, the share of the latter rises strongly with educational attainment. Households whose head has a secondary or higher education degree earn as much as 30 percent of their income in the public sector.

households in those places. This lends credence to the notion that rural non-agricultural activities are closely linked to the infrastructural benefits in the urban areas, which also provide access to markets and linkages to the formal sector.

With the results of the univariate analysis above in mind, we now turn to a multivariate choice model to shed more light on the possible determinants of engagement in the rural non-agricultural sector. This allows us to consider individual choice determinants beyond the household head's characteristics and to test whether the univariate results also hold once we control for other factors. We estimate a multinomial logit model where we allow individuals (not households) to choose between the two types of non-agricultural activities and staying on the farm. The results are presented in Table 4.3. The table reports odds ratios of low-return and high-return non-agricultural activities vis-à-vis agricultural ones in the first two columns and the odds of choosing high- vs. low-return activities in the last column. In line with the theory outlined above and existing empirical work, the set of explanatory variables comprises individual characteristics, such as age, gender, position in the household and formal education, and household composition variables. It also includes household assets, such as land and livestock, and a dummy indicating whether or not a household receives remittances. In addition, the distance to the nearest access road and an urban dummy are used as proxies for access to infrastructure and markets. We expect the diversification behavior of sugarcane farmers to be different from other farm households as the period between the cash flows from sugarcane harvests can be longer than three years.⁴³ Given this cycle and lacking access to financial markets, we expect sugarcane farmers to invest their considerable cash income in non-agricultural activities, particularly in the high-return segment, in order to smooth their income. Sugarcane farmers, however, may also be forced into the low-return segment once the last harvest's cash has been consumed.

Since the two types of non-agricultural activities might be more similar than staying on the farm, we test for independence of irrelevant alternatives using the Hausman test. The null hypothesis, i.e. the odds of choosing between alternatives a and b are affected by the existence of alternative c, can be rejected at the one percent level. Hence, the test does not reject our hypothesis of the different characteristics of low-return and high-return non-agricultural activities.

Most effects of the included explanatory variables have the expected sign, but some variables do not turn out to be significant. Yet, the analysis also yields some unexpected results that we will try to explain below. The effect of age corresponds to expectations, as younger individuals are more likely to be engaged in both low- and high-return non-agricultural activities rather than in agriculture. The reported odds ratios of 0.97 and 0.95 imply that the odds of being engaged in non-agricultural activities decrease by three percent for low-return and five percent for high-

43 On average, it takes twenty-four months before a commercial sugarcane crop is harvested. Even after harvest, it may take more months before payments are actually made to the farmers.

return activities for a unit change in the predictor, i.e. for one additional year of age. According to the estimation, age does not seem to affect the choice between low- and high-return activities.

The position of a member in the household is quite important in choosing a non-agricultural activity. Household heads are considerably more likely to take up a high-return activity as compared to both a low-return or agricultural activity. Yet, agricultural activities still seem to be more attractive to them than engaging in the low-return segment of the non-agricultural sector. The latter effect, however, is not significant. This finding may suggest that the control over resources, as typically exerted by the household head, could be crucial for building up the necessary start-up capital to invest into high-return activities. Hence, the decision to diversify into such activities may not solely be taken at the household level, but also reflect individual considerations and possibilities. In view of that, it is not surprising that household composition variables with the exception of the number of young adults do not appear to be important for the decision to diversify.

With regard to gender, we find that females are more than twice as likely to engage in the high-return non-agricultural sector as in agriculture. Yet, this observation is only close to being significant. In contrast, being female does not seem to be associated with involvement in low-return activities. Interestingly, we find different effects for individuals who live in female-headed households. Individuals from these households are much less likely to participate in high-return non-agricultural employment as compared to agriculture. If they participate in non-agricultural activities, they do so in the low-return segment.

In contrast to expectations, education seems to have a significant positive impact on entering low-return non-agricultural employment only. Although the odds ratios indicate a stronger effect of education on the probability of taking up a high-return activity rather than being in agriculture, these results are not significant. Taken together, the effects of education, derived from multivariate analysis, are quite different to those from the univariate analysis above. Apparently, education is correlated with other observed characteristics, for which we already control in our estimation framework. Thus, the additional impact of education seems to be limited when it comes to provide the skills necessary to grasp the few business opportunities available in the study area. Barrett et al. (2006) make very similar observations about the difficulties of school leavers to find remunerative employment in Vihiga district that neighbors Kakamega. They relate this observation to the considerable decline in Kenyan public sector employment and infrastructure in the region.

Our analysis considers two types of assets, land and livestock holdings. To better capture the potentially non-linear effects of landholdings, we split the households into three groups: households with landholdings between 0.5 and 1.5 acres, between 1.5 and 3.5 acres, and larger than 3.5 acres. We use households in the second group as reference category. We find that the likelihood of engaging in low-return non-agricultural vs. agricultural activities is almost twice as high for individuals in households with few landholdings. As hypothesized, we find the opposite

effect for diversification into high-return activities, although this effect is imprecisely measured and hence not statistically significant. Individuals in households with large landholdings are more than twice as likely to engage in the high-return segment of the non-agricultural sector. This supports the view that land endowments play a key role in explaining both survival-led and opportunity-led diversification strategies. Land size, however, turns out to be insignificant, when considering the odds of engaging in high vs. low-return activities. It is also interesting to note that not having a title deed is also a significant predictor for choosing low-return instead of agricultural employment. More secure land rights hence seem to provide an incentive for people to engage more heavily in agriculture. The effects for livestock holdings are very similar to the ones observed for land holdings. Higher values of livestock seem to increase the odds of diversifying into high-return activities. The reverse is true in the case of low-return activities. All in all, the asset variables clearly support our hypothesis on the dichotomous character of the non-agricultural sector.

Since migration is a wide-spread phenomenon in Kakamega district, remittances may be commonly used to invest into non-agricultural enterprises. Our results, however, suggest that receiving remittances is not associated with either form of non-agricultural activities.

For sugarcane farmers, we find the expected effects. Growing a profitable crop, they are less likely to involve in the non-agricultural sector. Yet, the period that has passed since the last sugarcane harvest seems to have a significant positive impact on participation in either type of non-agricultural activity.

The effects of infrastructure are less clear-cut. Whereas being located in an urban area appears to considerably increase the likelihood of high-return non-agricultural employment, the distance to the nearest access road does not seem to play a role in the decision to involve in non-agricultural activities. Yet, better access to markets also implies that it is easier to sell agricultural produce, thereby making agriculture a more profitable activity.

The investigation into the patterns of non-agricultural activities in Kakamega district has shown that income diversification is a widely observed phenomenon and that non-agricultural income contributes significantly to farm households' income. We observe that households primarily pursue low-return non-agricultural activities. The results of the multinomial logit model support the hypothesized dichotomy of non-agricultural activities and point to the existence of entry barriers to high-return activities.⁴⁴ We now turn to the analysis of the distributional implications of these different diversification strategies and test for the existence of virtuous spirals linked to high-return activities.

⁴⁴ Yet, the results are less clear-cut than one might expect from the univariate analysis that was based on household head characteristics. It should be borne in mind though that our sample is relatively small and that we estimate a simple reduced form model without controlling for potential endogeneity of some variables.

TABLE 4.3: MULTINOMIAL LOGIT MODEL OF RURAL NON-AGRICULTURAL EMPLOYMENT

| | Low-return vs. agriculture | High-return vs. agriculture | High- vs. low- return |
|-----------------------------------|-------------------------------|--------------------------------|--------------------------|
| age | 0.97** (-2.18) | 0.95** (-2.20) | 0.97 (-1.01) |
| female | 0.90 (-0.37) | 2.66 -1.49 | 2.96 -1.64 |
| head | 0.65 (-1.25) | 5.99** -2.37 | 9.16*** -2.86 |
| uncompleted primary | 1.46 -1.00 | 5.83 -1.63 | 3.98 -1.24 |
| completed primary | 2.16** -2.06 | 4.96 -1.44 | 2.29 -0.73 |
| completed secondary and higher | 2.23* -1.67 | 2.39 -0.69 | 1.07 -0.05 |
| number of children (0-4) | 0.90 (-0.88) | 0.86 (-0.69) | 0.96 (-0.18) |
| number of children (5-14) | 1.04 -0.45 | 1.00 (-0.02) | 0.95 (-0.26) |
| number of adults (15-24) | 0.94 (-0.60) | 1.20 -1.48 | 1.27* -1.76 |
| number of adults (25-60) | 0.83 (-1.17) | 0.93 (-0.28) | 1.12 -0.43 |
| number of old people (>60) | 1.08 -0.23 | 0.49 (-1.04) | 0.45 (-1.09) |
| female household head | 1.25 -0.74 | 0.31* (-1.78) | 0.25** (-2.03) |
| landsize (0.5-1.5 acres) dummy | 1.81* -1.89 | 1.14 -0.25 | 0.63 (-0.85) |
| landsize (>3.5 acres) dummy | 1.51 -1.15 | 2.39 -1.55 | 1.58 -0.77 |
| title deed | 0.55* (-1.92) | 1.15 -0.31 | 2.11 -1.46 |
| livestock value (1,000 KSh) | 1.00 (-1.60) | 1.01 -1.61 | 1.01** -2.55 |
| remittances dummy | 1.05 -0.15 | 0.90 (-0.19) | 0.86 (-0.26) |
| sugar cane dummy | 0.45* (-1.89) | 0.40 (-1.48) | 0.88 (-0.19) |
| sugar cane period | 1.04*** -2.64 | 1.05*** -2.65 | 1.01 -0.36 |
| distance to nearest access road | 1.00** (-2.19) | 1.00 (-0.10) | 1.00 -1.53 |
| urban | 0.57 (-1.21) | 2.64* -1.66 | 4.66** -2.2 |
| Observations | 437 | 437 | 437 |
| Wald chi2(42) | 95.06 | | |
| Log pseudo-likelihood | -310.59 | | |
| Pseudo R2 | 0.14 | | |

***/**/* denote significance at the 1/5/10 percent level respectively. Robust z-statistics in parentheses.

4.4 Poverty, Inequality and Virtuous Cycles

Table 4.4 shows the participation rates in non-agricultural activities and the respective income shares by per adult equivalent expenditure quintiles. Note that the income share is calculated only for those households that engage in the non-agricultural sector. Surprisingly, participation in low-return activities is not concentrated among poor households. In fact, participation is even lowest among households in the poorest quintile, peaks in the second and third quintile, and is still around 40 percent in the two highest quintiles. Yet, despite relatively low participation, the income share derived from low-return activities is by far the highest for households in the poorest quintile. With increasing living standards, this share declines considerably. As regards high-return activities, barriers seem to effectively exclude the poorest households from such diversification strategies. Only starting from the second quintile do households pursue high-return activities. Participation in the high-return sector as well as the derived share of income then increase strongly with higher consumption levels.

TABLE 4.4: PARTICIPATION RATES AND INCOME SHARES OF LOCAL NON-AGRICULTURAL ACTIVITIES BY EXPENDITURE QUINTILES

| Expenditure Quintile | Participation | | | Average share of non-agricultural income | | | Median share of non-agricultural income | | |
|----------------------|---------------|------------|-------------|--|------------|-------------|---|------------|-------------|
| | All | Low-return | High-return | All | Low-return | High-return | All | Low-return | High-return |
| bottom | 38% | 36% | 2% | 45% | 43% | 3% | 47% | 43% | 0% |
| 2nd | 63% | 50% | 16% | 38% | 31% | 7% | 31% | 31% | 0% |
| 3rd | 55% | 45% | 18% | 44% | 32% | 12% | 36% | 24% | 0% |
| 4th | 61% | 39% | 25% | 39% | 24% | 15% | 38% | 14% | 0% |
| 5th | 62% | 42% | 29% | 41% | 17% | 24% | 37% | 14% | 0% |

To better understand the inequality implications of the different types of non-agricultural activities, we decompose the Gini coefficient of income inequality by income source, using the approach described in Lerman and Yitzhaki (1985) and in Stark et al. (1986). The decomposition allows us to determine the impact that a marginal change in a particular income source would have on overall inequality. The results are reported in Tables 5a and 5b. The last column of the two tables refers to the point change in the Gini that would be brought about by a one percent increase in the respective income source. Three additional elements are included in the result tables: The share of each income source in total income, the Gini of the income source, and the correlation of income from the respective source with a household's per adult equivalent total income rank. Table 4.5a considers both diversifying and non-diversifying households whereas Table 4.5b only looks at diversifying households.

Despite the significantly lower participation rates for high-return activities, both tables show that the total income share of high-return activities is not much lower than that of low-return activities. The previous observation that participation in the high-return sector increases with living standards is reflected in the strong correlation of high-return income with the rank of per adult equivalent total income. The source Ginis (0.88 and 0.93) demonstrate that income from

high-return activities is the most inequitably distributed source of income. Accordingly, we find that a percentage change in income associated with high-return activities brings about a remarkable rise in inequality (Gini elasticities of 0.106 and 0.062).

In contrast, the source Ginis for income from low-return activities (0.70 and 0.83) are markedly lower. This echoes the fact that income from low-return activities constitutes an important income source for households across the entire income distribution, which can also be seen by the relatively low correlation of low-return income with the rank of total income. Consequently, the effect of a marginal increase in low-return income on overall inequality is small (Gini elasticities of 0.009 and 0.022).

TABLE 4.5A: INEQUALITY DECOMPOSITION BY SOURCES, ALL HOUSEHOLDS

| Income source | Income share | Gini coefficient of income source | Correlation with rank of total income | Elasticity of overall Gini coefficient |
|---------------|--------------|-----------------------------------|---------------------------------------|--|
| Low-return | 19% | 0.83 | 0.68 | 0.02 |
| High-return | 14% | 0.93 | 0.80 | 0.06 |
| Agriculture | 50% | 0.45 | 0.78 | -0.16 |
| Other income | 17% | 0.90 | 0.82 | 0.08 |
| Total income | | 0.51 | | |

Other income includes income from public employment.

TABLE 4.5B: INEQUALITY DECOMPOSITION BY SOURCES, ONLY DIVERSIFYING HOUSEHOLDS

| Income source | Income share | Gini coefficient of income source | Correlation with rank of total income | Elasticity of overall Gini coefficient |
|---------------|--------------|-----------------------------------|---------------------------------------|--|
| Low-return | 28% | 0.70 | 0.66 | 0.01 |
| High-return | 20% | 0.88 | 0.77 | 0.11 |
| Agriculture | 44% | 0.41 | 0.70 | -0.15 |
| Other income | 8% | 0.91 | 0.74 | 0.04 |
| Total income | | 0.45 | | |

Other income includes income from public employment.

Agricultural income, in contrast, stands out as the most equitably distributed source of income. A marginal increase of agricultural income would even result in a sizeable reduction of overall inequality. This mirrors the earlier finding that for the poorest quintile overall participation rates in the non-agricultural sector are particularly low.

We now examine the impact of changes in a particular income source on different poverty measures. For this purpose, we increase a household’s income from the respective sources by 50 percent to create sizeable poverty effects. We then calculate the corresponding per adult equivalent amount of this additional income and add it to the actually observed per adult equivalent expenditure level. Table 4.6 summarizes the results. The first and the second column show the poverty headcount and the average normalized poverty gap before increasing the respective incomes. The third and fourth columns present the resulting changes in the poverty measures after the simulated income increases.

The table illustrates that the Kakamega district is indeed a very poor region. The overall headcount ratio is about 76 percent with an average normalized poverty gap of 45 percent. On average, households involved in non-agricultural activities fare considerably better. However, the

dichotomy of the non-agricultural sector is strongly reflected in poverty outcomes. Whereas 72 percent of people living in households which are engaged in low-return activities are below the poverty line, the headcount for their counterparts in high-return activities is considerably lower, but still amounts to 63 percent. In addition, the average normalized poverty gap is markedly smaller for the latter households. The poverty headcount of low-return diversifiers is thus comparable to the headcount of all households. Yet, the intensity of poverty, measured by the poverty gap, is much lower for households that are active in the low-return non-agricultural sector. This again reflects the fact that the poorest households engage relatively less in this sector.

TABLE 4.6: POVERTY EFFECTS OF A 50 PERCENT INCREASE IN NON-AGRICULTURAL INCOME

| | Before | | After | |
|---|--------|-------|-------|-------|
| | P0 | P1 | P0 | P1 |
| <i>Increase of non-agricultural income by 50%</i> | | | | |
| all households | 75.96 | 45.26 | -3.14 | -3.24 |
| only households involved in non-agricultural activities | 70.35 | 37.04 | -5.52 | -5.70 |
| <i>Increase of low-return non-agricultural income by 50%</i> | | | | |
| all households | 75.96 | 45.26 | -2.42 | -2.18 |
| only households involved in low-return activities | 72.28 | 37.94 | -5.67 | -5.11 |
| <i>Increase of high-return non-agricultural income by 50%</i> | | | | |
| all households | 75.96 | 45.26 | -0.33 | -1.08 |
| only households involved in high-return activities | 63.31 | 31.20 | -1.80 | -5.97 |

Changes are reported as percentage points. The poverty line is defined on per adult equivalent expenditures. It is based on the rural poverty line used for the 1997 poverty assessments, KSh 1239 for rural and KSh 2648 for urban areas (Welfare and Monitoring Survey of that year), and adjusted to reflect inflation. Comparing maize and beans prices in urban and rural areas of the study region, we assume the urban price level in Kakamega district to be 25 percent above the rural price level. The resulting poverty lines are then KSh 1843 for rural and KSh 2304 for urban areas, respectively.

A 50 percent increase in incomes from all non-agricultural activities would result in a 3.14 percentage point decrease in the overall poverty headcount and a 3.24 percentage point decrease in the poverty gap. The bulk of this poverty reduction would be attributable to higher earnings from low-return activities which alone would reduce the headcount by 2.42 and the poverty gap by about 2.18 percentage points. The corresponding figures for incomes from the high-return sector are 0.33 and 1.08 percentage points, respectively. The relatively large poverty reduction potential of low-return activities becomes even more apparent when examining the poverty effects for households involved in the low-return and high-return non-agricultural sector separately. A 50 percent rise in incomes from low-return activities would reduce the poverty headcount by 5.67 and the average normalized poverty gap by 5.11 percentage points for households which engage in the low-return sector. This compares to 1.80 and 5.97 percentage points if the same exercise is undertaken for households with high-return activities. The larger simulated poverty impact for incomes from low-return activities reflects the participation rates and income shares observed above. These activities constitute an important source of income throughout the entire income distribution.

A more detailed view on the distributional consequences can be obtained by looking at growth incidence curves, which plot the growth impact of a 50 percent rise in non-agricultural income on per adult equivalent income by per adult equivalent income vintiles (Figures A4.1 and A4.2 in the appendix). Considering all households, Figure A4.1 shows that by and large increasing incomes from low-return activities would be rather pro-poor. Among the poorer half of the sample, per adult equivalent income increases by roughly eight percent. Only from the tenth vintile upwards, does the growth rate consecutively drop to reach less than four percentage points for the highest expenditure vintile. In contrast, the growth incidence curve for high-return activities is almost strictly pro-rich.⁴⁵ Whereas the poorest households are virtually excluded from this segment of the non-agricultural sector, the richest vintiles would experience a per adult equivalent income growth rate of about five percentage points. This low growth rate reflects the relatively low participation rates and low income shares even among the very rich. All in all, households in the middle of the income distribution would profit most from an overall increase in non-agricultural incomes.

The picture changes, in particular for low-return activities, when only considering diversifying households (Figure A4.2). Then, rising incomes from low-return activities would be strictly pro-poor while an increase in incomes from high-return activities would be strictly pro-rich. However, given the relative dominance of the low-return sector in the study region, the growth impact of high-return activities would be much lower than the corresponding growth impact of low-return activities. Thus, in the case of diversifying households a rise in total non-agricultural income would be in favor of the poor, above all the very poor diversifiers who could boost their incomes by more than 25 percent. Most households in the interior parts of the income distribution would see their per-adult equivalent income grow by roughly the same rate of just under 20 percent.

The analysis and simulation in this section represents a very simplistic and purely statistical exercise that is useful to illustrate the de facto distribution of non-agricultural income. Yet, it ignores many facets of the complex relationships shaping rural livelihoods. In particular, it does not account for possible interactions between agricultural and non-agricultural activities that have been stressed in the livelihood literature. It has been noted that high non-agricultural earnings are often associated with high agricultural productivity (Evans and Ngau, 1991; Ellis and Freeman, 2004). This link may be explained by non-agricultural income providing households with the cash necessary to purchase agricultural inputs such as fertilizer or improved seeds as well as hired wage labor. The resulting improvement in production technologies would then increase agricultural yields. This, in turn, would raise the resources available to further invest into non-agricultural activities, thus triggering a virtuous spiral of income.

⁴⁵ We define a growth pattern as strictly pro-poor (pro-rich) when the growth incidence curve is monotonically decreasing (increasing). In the present case, the curve is 'almost pro-rich' as it falls slightly for the richest vintile.

Departing from the dichotomy of the non-agricultural sector, we expect the agricultural productivity implications to depend on the type of non-agricultural activities pursued. Clearly, high-return activities would give rise to this welfare-enhancing cumulative process. Low-return activities, however, due to their survival-led character, would most likely not generate sufficient cash income to initiate such a virtuous spiral. Their impact on agricultural productivity might therefore be limited. Similar to pure agricultural households, low-return diversifiers may find themselves in a poverty trap, since high-return activities constitute one of the few opportunities for resource-poor household to accumulate assets and eventually exit poverty.

To test for the existence of such virtuous circles, we examine the relationship between different diversification patterns and agricultural productivity by regressing the value of agricultural output per acre on two dummies indicating a household's engagement in high-return and low-return non-agricultural activities respectively. Control variables include the characteristics of the household head, household composition, land holdings, and location as well as some basic inputs into agricultural production such as household farm labor or the use of fertilizer and pesticides. Table 4.7 reports the results.

TABLE 4.7: OLS REGRESSION OF AGRICULTURAL YIELDS

| | Log (value of agricultural production per acre) | |
|--|--|---------|
| high-return non-agr. activity | 0.52** | (2.49) |
| low-return non-agr. activity | 0.09 | (0.68) |
| female household head | 0.38*** | (2.96) |
| uncompleted primary (hh head) | 0.30* | (1.81) |
| completed primary (hh head) | 0.37** | (2.38) |
| completed secondary and higher (hh head) | 0.20 | (0.80) |
| number of children (0-4) | 0.01 | (0.08) |
| number of children (5-14) | -0.03 | (-0.52) |
| number of old people (>60) | 0.02 | (0.20) |
| household size | 0.04 | (0.99) |
| land size per adult equivalent | -0.18*** | (-3.53) |
| land size (0.5-1.5 acres) dummy | 0.39*** | (2.63) |
| land size (>3.5 acres) dummy | -0.29 | (-1.47) |
| number of hh farm workers | 0.15 | (1.49) |
| fertilizer dummy | 0.39* | (1.76) |
| pesticides/herbicides dummy | 0.54** | (2.51) |
| distance to nearest market | -0.04** | (-2.06) |
| urban | -0.14 | (-0.65) |
| constant | 5.59*** | (14.75) |
| Observations | 271 | |
| F(19, 251) | 3.27 | |
| R ² | 0.19 | |

***/**/* denote significance at the 1/5/10 percent level respectively. Robust t statistics in parentheses.

Most variables have the expected sign and many turn out to be significant. Like other studies (e.g. Bardhan, 1973), we also find a negative relationship between agricultural yields and the size of land holdings. As expected, high-return activities are strongly and significantly associated with

higher agricultural productivity. This is even true although we already control for the utilization of fertilizer and pesticides, two important agricultural cash inputs that might be purchased with the income earned from high-return activities.⁴⁶ In contrast, low-return activities do not appear to have productivity effects. This finding supports the view that there are cumulative and reinforcing relationships between assets, opportunities, and income sources (Ellis and Bahigwa, 2003). Yet, as we have shown, it is important not to attribute this relative livelihood success to engagement in non-agricultural activities in general. It is only the high-return segment of the non-agricultural sector that potentially gives rise to the positive dynamics described above. This segment, however, accounts for a small portion of the sector in our study area.

4.5 Conclusion

This paper empirically confirms the existence of a dichotomous non-agricultural sector in a fairly typical poor rural East African context. Low-return activities with some but low entry barriers account for the bulk of local non-agricultural activities, while only few households have the skills and assets to engage in activities that offer higher returns. We provide evidence that involvement in such high-return activities may trigger virtuous spirals as exemplified by positive spillovers to agricultural productivity. Higher income from high-return activities consequently tends to aggravate existing inequalities, as these activities are confined to richer households. In contrast, engagement in low-return activities does not seem to affect agricultural productivity, an indication that this sector does not provide a path out of poverty. Finally, our findings suggest that some extremely poor households are even excluded from the latter activities, which makes them particularly vulnerable to shocks that frequently affect agriculture in this climatic zone.

The dominance of low-return activities is possibly the finding that calls most for additional explanation. If engagement in low-return activities is driven by desperation, as implied by the concept of survival-led diversification and backed by our analysis of determinants, we would expect these activities to be pursued primarily if not exclusively by poorer households. This is not the case and, in fact, participation rates for low-return activities of richer households are comparable to those of poorer ones.

One might argue that these results are owed to Kakamega district being a particularly poor region, where just too many households are not able to overcome the entry barriers to high-return activities. However, we find a number of richer households in both low- and high-return activities. These households may not be able to employ their available resources more productively than in low-return activities. This implies that households do not only face asset constraints or other types of entry barriers. Possibly, demand for non-agricultural products that are produced by high-return activities in small spatially segmented markets may just be too limited. Furthermore, the simultaneous diversification into low- and high-return activities may

⁴⁶ This finding may also be explained by unobservables that affect both the capacity to engage in high-return activities and agricultural yields and that eventually give rise to virtuous circles.

reflect the high risk being associated with high-return activities, which these households compensate by venturing into low-return activities. Another explanation for our observation of rich households in low-return activities might be that some individuals do only have limited access to household resources. Our empirical analysis supports this view, showing that the relevant decision unit is not necessarily the household.

Of course, one has to be careful in drawing too far-ranging conclusions from findings from a specific region. In addition, the static “snapshot” character of our analysis limits its contribution towards understanding the mechanisms that would explain the emergence of either type of non-agricultural activity. Research on rural livelihoods is already trying to examine such dynamics (e.g. Barrett, 2004), but both cyclical movements and long-run changes in non-agricultural participation have been somewhat neglected in empirical research.⁴⁷

Along similar lines, our work hints at a shift of research focus away from the analysis of household behavior and household level constraints towards a closer examination of meso- or macro-level determinants of structural transformation in rural areas, in particular in sub-Saharan Africa. It is well known that poor rural households face important asset and skill constraints. Yet, if we want to understand the structural transformation and the emerging of the non-agricultural economy, we need to look beyond the household level and also investigate the meso- or macro-level drivers and facilitators of change, such as demographic pressure, resource degradation, technical change, urbanization processes, and rural infrastructure.

Because of the limited amount of cultivable land, the demographic developments in East Africa will necessarily force people out of agriculture. Despite the effects of HIV/AIDS, the population of Kakamega district is projected to grow from 660,000 to 780,000 between 2005 and 2015⁴⁸ – a district where average farm size already stands at less than a hectare to meet the subsistence needs of more than five people. Under such circumstances, the prospects for growth and poverty reduction will crucially depend on the performance of the non-agricultural sector.

47 There are some important exceptions. Reardon and Taylor (1996) and Kijima et al. (2006) focus on short-term dynamics. In general, contributions from non-economists have more often assessed medium to long-run changes in rural livelihoods (Bryceson, 2002; Francis, 2000)

48 Population projections for Kakamega were provided by Wondimu Kenea, PhD student at the University of Dortmund, Germany.

Literature

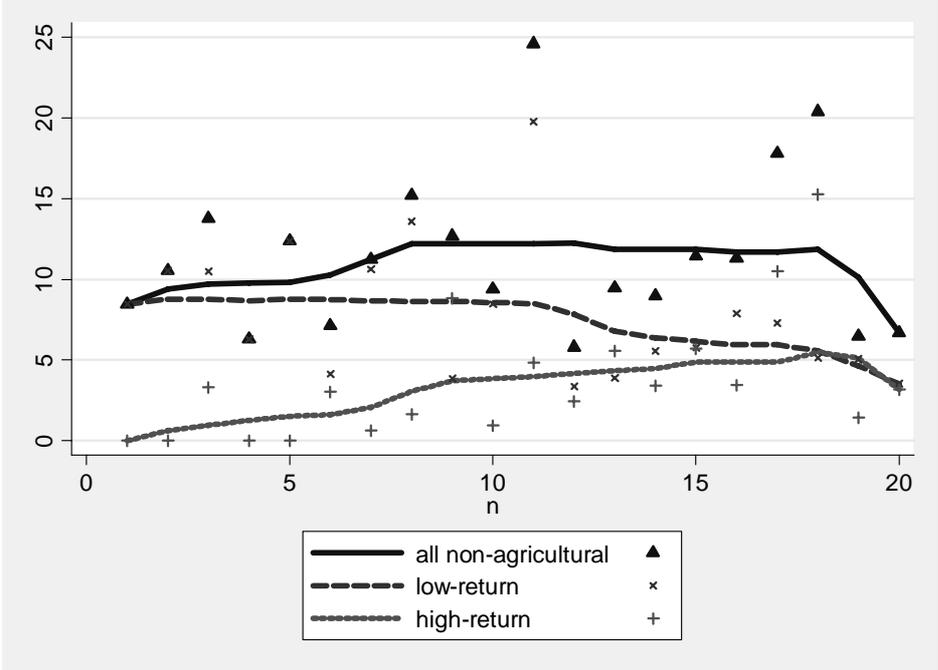
- Abdulai, A., and A. CroleRees (2001). Determinants of Income Diversification among Rural Households in Southern Mali. *Food Policy*, 26(4): 437-452.
- Abdulai, A., and C.L. Delgado (1999). Determinants of Nonfarm Earnings of Farm-Based Husbands and Wives in Northern Ghana. *American Journal of Agricultural Economics*, 81(1): 117-130.
- Adams, R.H. (2002). Non-agricultural Income, Inequality, and Land in Rural Egypt. *Economic Development and Cultural Change*, 50(2): 339-363.
- Bardhan, P.K. (1973). Size, Productivity, and Returns to Scale: An Analysis of Farm-Level Data in Indian Agriculture. *Journal of Political Economy*, 81(6): 1370-1386.
- Barrett, C.B. (2004). Rural Poverty Dynamics: Development Policy Implications. Paper prepared for invited presentation at the 25th International Conference of Agricultural Economists. Durban, South Africa.
- Barrett, C.B., M. Bezuneh and A. Aboud (2001). Income Diversification, Poverty Traps and Policy Shocks in Cote d'Ivoire and Kenya. *Food Policy*, 26(4): 367-384.
- Barrett, C.B., T. Reardon and P. Webb (2001). Non-Agricultural Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics and Policy Implications. *Food Policy*, 26(4): 315-331.
- Barrett, C.B., M. Bezuneh, C.D. Clay and T. Reardon (2000). Heterogeneous Constraints, Incentives and Income Diversification Strategies in Rural Africa. Report to USAID/BASIS CRSP, Cornell University and Michigan State University, Ithaca and East Lansing.
- Barrett, C.B., P.P. Marenja, J. Mpeka, B. Minten, F. Murithi, W.O. Kosura, F. Place, J.C. Randrianarisoa, J. Rasambainarivo and J. Wangila (2006). Welfare Dynamics in Rural Kenya and Madagascar. *Journal of Development Studies*, 42(2): 248-277.
- van den Berg, M., and G.E. Kumbi (2006). Poverty and the Rural Nonfarm Economy in Oromia, Ethiopia. *Agricultural Economics*, 35(3): 469-475.
- Bryceson, D.F. (2002). Multiplex of Livelihoods in Rural Africa: Recasting the Terms and Conditions of Gainful Employment. *Journal of Modern African Economics*, 40(1): 1-28.
- Bryceson, D.F. (1999). African Rural Labor: Income Diversification and Livelihood Approaches: A Long Term Development Perspective. *Review of African Political Economy*, 26(6): 171-189.
- Bryceson, D.F., and V. Jamal (Eds.) (1997). *Farewell to Farms: De-Agrarianisation and Employment in Africa*. Aldershot: Ashgate.
- Corral, L., and T. Reardon (2001). Rural Nonfarm Incomes in Nicaragua: Patterns, Determinants and Implications. *World Development*, 29(3): 427-442.
- Dercon, S. (1998). Wealth, Risk and Activity Choice: Cattle in Western Tanzania. *Journal of Development Economics*, 55(1): 1-42.
- Dercon, S., and P. Krishnan (1996). Income Portfolios in Rural Ethiopia and Tanzania: Choices and Constraints. *Journal of Development Studies*, 32(6): 850-875.
- Elbers, C., and P. Lanjouw (2001). Intersectoral Transfer, Growth and Inequality in Rural Ecuador. *World Development*, 29(3): 481-496.

- Ellis, F. (2000). The Determinants of Rural Livelihood Diversification in Developing Countries. *Journal of Agricultural Economics*, 51(2): 289-302.
- Ellis, F., and G. Bahiigwa (2003). Rural Livelihoods and Poverty Reduction in Uganda. *World Development*, 31(6): 997-1013.
- Ellis, F., and H.A. Freeman (2004). Rural Livelihoods and Poverty Reduction Strategies in Four African countries. *Journal of Development Studies*, 40(4): 1-30.
- Escobal, J. (2001). The Determinants of Nonfarm Income Diversification in Rural Peru. *World Development*, 29(3): 497-508.
- Evans, H.E., and P. Ngau (1991). Rural-Urban Relations, Household Income Diversification and Agricultural Productivity. *Development and Change*, 22(3): 519-545.
- Ferreira, F.H.G., and P. Lanjouw (2001). Rural Non-Agricultural Activities and Poverty in the Brazilian Northeast. *World Development*, 29(3): 509-528.
- Freeman, H.A., F. Ellis and E. Allison (2004). Livelihoods and Rural Poverty Reduction in Kenya. *Development Policy Review*, 22(2): 147-173.
- Francis, E. (2000). *Making a Living: Changing Livelihoods in Rural Africa*. London: Routledge.
- Francis, E., and J. Hoddinott (1993). Migration and Differentiation in Western Kenya: A Tale of Two Sub-Locations. *Journal of Development Studies*, 30(1): 115-145.
- Giesbert, L. (2007). Seeking Opportunities: Migration as an Income Diversification Strategy of Households in Kakamega District in Kenya. Working Paper 58, GIGA, Hamburg.
- Government of Kenya (2002). *Kakamega District Development Plan 2002-2008*. Nairobi: Government Printer.
- Haggblade, S., P. Hazell and T. Reardon (2005). The Rural Non-Agricultural Economy: Pathway Out of Poverty or Pathway In? Proceedings of the research workshop 'The Future of Small-Farms', June 26-29, 2005, UK, International Food Policy Research Institute, Washington D.C.
- Haggblade, S., P. Hazell and T. Reardon (2002). Strategies for Stimulating Poverty-Alleviating Growth in the Rural Nonfarm Economy in Developing Countries. Discussion Paper 92, Environment and Production Technology Division, International Food Policy Research Institute and World Bank, Rural Department, Washington D.C.
- International Labor Office (1972). *Employment, Incomes and Equality: A Strategy for Increasing Productive Employment in Kenya*. Geneva: ILO.
- Islam, N. (1997). The Nonfarm Sector and Rural Development. Discussion Paper 22, Food, Agriculture and Environment, International Food Policy Research Institute, Washington D.C.
- Jalan, J., and M. Ravallion (2002). Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China. *Journal of Applied Econometrics*, 17(4): 329-346.
- de Janvry, A., and E. Sadoulet (2001). Income Strategies among Rural Households in Mexico: The Role of Off-Farm Activities. *World Development*, 29(3): 467-480.
- Jayne, T.S., T. Yamano, M. Weber, D. Tschirley, R. Benfica, A. Chapoto and B. Zulu (2003). Smallholder Income and Land Distribution in Africa: Implications for Poverty Reduction Strategies. *Food Policy*, 28(3): 253-275.

- Kijima, Y., T. Matsumoto and T. Yamano (2006). Nonfarm Employment, Agricultural Shocks, and Poverty Dynamics: Evidence from Rural Uganda. *Agricultural Economics*, 35(3): 459-467.
- Lanjouw, P. (2001). Non-Agricultural Employment and Poverty in Rural El Salvador. *World Development*, 29(3): 529-547.
- Lanjouw, P., and G. Feder (2001). Rural Nonfarm Activities and Rural Development: From Experience towards Strategy. Rural Development Strategy Background Paper 4, World Bank, Washington D.C.
- Lerman, R.I., and S. Yitzhaki, (1985). Income Inequality by Income Source: A New Approach and Application to the United States. *Review of Economics and Statistics*, 67: 151-156.
- Marenya, P.P., W. Kosura-Olouch, F. Place and B. Christopher (2003). Education, Nonfarm Income, and Farm Investment in Land-Scarce Western Kenya. BASIS BRIEF 14, University of Wisconsin-Madison, Madison.
- Matsumoto, T., Y. Kijima and T. Yamano (2006). The Role of Local Nonfarm Activities and Migration in Reducing Poverty: Evidence from Ethiopia, Kenya, and Uganda. *Agricultural Economics*, 35(3): 449-458.
- Norman, D.W., E.B. Simons and H.M. Hays (1982). *Farming Systems in the Nigerian Savannah: Research and Strategies for Development*. Boulder: Westview Press.
- Reardon, T. (1999). *Rural Non-Farm Income in Developing Countries*. Rome: FAO.
- Reardon, T. (1997). Using Evidence of Household Income Diversification to Inform the Study of Rural Non-Agricultural Labor Market in Africa. *World Development*, 25(5): 735-748.
- Reardon, T., and J.E. Taylor (1996). Agro-Climatic Shock, Income Inequality, and Poverty: Evidence from Burkina Faso. *World Development*, 24(5): 901-914.
- Reardon, T., C. Delgado and P. Malton (1992). Determinants and Effects of Income Diversification amongst Farm Households in Burkina Faso. *Journal of Development Studies*, 28(1): 264-296.
- Reardon, T., J.E. Taylor, A. Stamoulis, P. Lanjouw and A. Balisacan (2000). Effects of Non-Agricultural Employment on Rural Income Inequality in Developing Countries: An Investment Perspective. *Journal of Agricultural Economics*, 51(2): 266-288.
- Reardon, T., A. Stamoulis, M.E. Balisacan, J. Berdegue and B. Banks (1998). The State of Food and Agriculture, Part III, Rural Non-Agricultural Income in Developing Countries. Rome: Food and Agriculture Organization.
- Seppala, P. (1996). The Politics of Economic Diversification: Re-Conceptualizing the Rural Informal Sector in South-East Tanzania. *Development and Change*, 27(3): 557-578.
- Stark, O., J.E. Taylor and S. Yitzhaki (1986). Remittances and Inequality. *Economic Journal*, 96(3): 722-740.

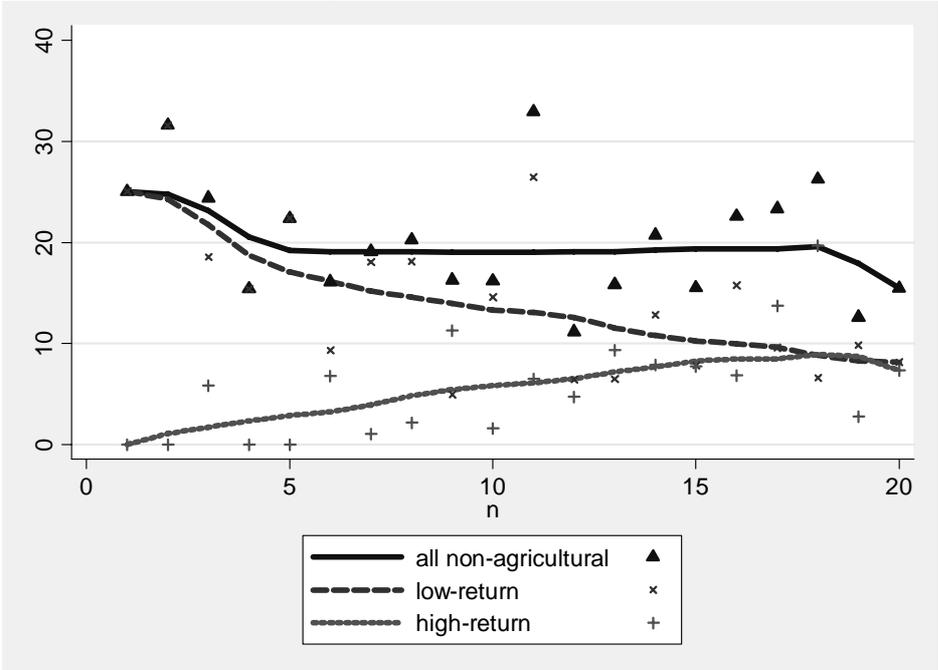
Appendix

FIGURE A4.1: SMOOTHED GROWTH INCIDENCE CURVE, 50 PERCENT INCREASE IN NON-AGRICULTURAL INCOMES, ALL HOUSEHOLDS



Vertical axis: per adult equivalent income growth in percentage points
 Horizontal axis: per adult equivalent expenditure vintiles

FIGURE A4.2: SMOOTHED GROWTH INCIDENCE CURVE, 50 PERCENT INCREASE IN NON-AGRICULTURAL INCOMES, ONLY DIVERSIFYING HOUSEHOLDS



Vertical axis: per adult equivalent income growth in percentage points
 Horizontal axis: per adult equivalent expenditure vintiles

Chapter 5

Shocks, Structural Change, and the Patterns of Income Diversification in Burkina Faso⁴⁹

published 2009 in *African Development Review*, 21(1): 36-58

Jann Lay⁵⁰

*German Institute of Global and Area Studies
University of Göttingen*

Ulf Narloch

University of Cambridge

Toman Omar Mahmoud

Kiel Institute for the World Economy

Abstract

This paper investigates changes in income diversification patterns for the case of Burkina Faso between 1994 and 2003. Contrary to common beliefs, our empirical analysis shows that rural households are not increasingly diversifying their income portfolios. Beyond insuring against and coping with weather shocks, diversification behavior reflects structural change. Higher returns to agricultural activities, in particular in the cotton and livestock sectors, appear to be the root cause for less non-agricultural diversification and some of our findings hint at better opportunities in the non-farm sector. Yet, average returns in the non-farm sector appear to remain relatively low and migration increasingly turns into a desperation-led strategy. Overall, structural change seems to be biased in favor of richer households. Regarding responses to droughts, we confirm earlier findings, especially that the poorest households are hit particularly hard being forced to sell livestock, often their only asset.

Keywords: income diversification, shocks, non-agricultural activities, migration, cotton, Burkina Faso

JEL classification: Q12, O17, I31

49 We thank Michael Grimm for helpful comments and discussions and GTZ for providing very helpful assistance during a research visit in Burkina Faso.

50 Corresponding author: lay@giga-hamburg.de

5.1 Introduction

A number of studies have illustrated the diversity of rural livelihoods in sub-Saharan Africa and shown that income diversification serves as a coping strategy in risky agricultural environments. Still, relatively little is known about changes in patterns of income diversification in the medium to long run. Farming typically constitutes the main source of livelihood for rural households and leaves them vulnerable to crop failures in drought-prone environments, in particular in the Sahel zone. Still, it has been demonstrated that households insure themselves against harvest shortfalls through engagement in activities outside crop farming (e.g. Reardon et al., 1992). A common insurance mechanism is the selling of livestock, but households also diversify into local non-agricultural activities or by sending household members elsewhere (Reardon and Taylor, 1996).

Such an insurance view on off-farm income diversification is clearly too narrow. When diversification is pursued to ensure survival, for example because of land constraints, it is also referred to as distress-push or desperation-led diversification (Islam, 1997; Reardon et al., 2000; Barrett et al., 2001). In contrast, rural households may face opportunities outside agriculture, e.g. due to higher demand for non-food products (Reardon, 1997). In this case, non-agricultural income diversification is also regarded as demand-pull or opportunity-led diversification. For Asia, country case studies show a dramatic increase of the share of non-farm income in high-growth countries like Thailand. This opportunity-led diversification is accompanied by a shift towards more diversified farming systems with a more important role for high-value cash-crops and livestock (Otsuka and Yamano, 2006). While it seems to be common ground that sub-Saharan rural households also increasingly diversify their livelihoods – yet mainly driven by desperation (Bryceson, 2002a) – systematic evidence on the patterns of change over longer time periods is rather scarce. Accordingly, most existing studies on income diversification in the sub-Saharan context rely on cross-sectional or short-term panel data.⁵¹

This paper intends to fill this gap. More specifically, we examine income diversification patterns of farm households in rural Burkina Faso between 1994 and 2003. After structural adjustment in the early 1990s, the country experienced fairly high per capita income growth that went along with substantial poverty reduction in rural areas (Grimm and Günther, 2007a), although the economy was recurrently hit by droughts and other external shocks. In particular, Burkinabe households were affected by the political crisis in Côte d'Ivoire, its major destination for temporary work migration.

Our empirical analysis indicates that rural households are not increasingly diversifying their income portfolios in Burkina Faso. In addition, the motivations for diversification are subject to change and our findings do not support the view that desperation-led diversification is (or becomes) the rule in poor African countries. Rather, a nuanced picture emerges: Better opportunities in agriculture, in particular in the cotton and livestock sectors, allow households to

51 Exceptions include Block and Webb (2001) and Lemi (2006) on Ethiopia.

specialize in these activities although richer households are more likely to do so. Non-farm activities become somewhat more opportunity-led. Yet, average returns in the non-farm sector appear to remain relatively low and migration increasingly turns into a strategy pursued by more desperate households. In addition, farmers' responses to the 1997/98 drought are very similar to the pattern that could be observed in earlier work by Reardon and Taylor (1996) for the early 1980s. The poorest households are forced to sell livestock, often their only asset, which may keep them in a poverty trap. In the remainder of the paper, we first shortly review the literature on the determinants and effects of diversification, with a focus on distinct motivations. We then present some more detailed country background and an empirical analysis of changes in income portfolios for the case of Burkina Faso. The paper closes with some concluding remarks.

5.2 Determinants and Effects of Income Diversification

A large number of rural households in sub-Saharan Africa rely on earnings generated outside home agriculture (Otsuka and Yamano, 2006; Reardon, 1997). Besides soil cultivation and livestock breeding, non-agricultural activities in the local non-farm sector and non-local income generation via migration are important income sources, so that “diversification is the norm” (Barrett and Reardon, 2000). Such behavior is difficult to reconcile with economies of scale in any of the activities and comparative advantage, which would imply a specialization in one activity. Yet, it may be explained by the following factors that can be categorized into insurance, desperation and opportunity motives.

If credit and insurance markets are incomplete and formal safety nets are missing – which is the case for most parts of rural sub-Saharan Africa – households will choose to diversify their income sources in order to smooth consumption (Murdoch, 1995). Under climatically unstable conditions rainfall shocks play a major role in activity decisions. In regions with high rainfall variability households may opt to diversify incomes ex-ante to manage the risk of possible shocks. Alternatively, household may be pushed into diversification ex-post as soon as harvest shortfalls occur (Reardon et al., 1992). For instance, they may be forced to sell their livestock (Dercon, 1998, Reardon and Taylor, 1996). Alternatively, households may send members to engage in markets at other places, where returns are not correlated to the returns at home (Lucas and Stark, 1985).

It has been emphasized in the sub-Saharan context that households may be forced to diversify if land endowments become too small to sustain subsistence following subdivision at inheritance (Bryceson and Jamal, 1997; Barrett et al., 2000a). This may hold particularly for female-headed households, which often have limited access to land due to poor enforcement of land property rights (Chapoto et al., 2006). Such desperation-led diversification may also be observed in areas with generally low agricultural potential (Matsumoto et al., 2006).

In contrast, agriculturally favored regions may exhibit higher levels of opportunity-led non-farm diversification through intersectoral linkages (Haggblade et al., 1989, Liedholm and Kilby, 1989).

In these regions, demand for non-food products may be higher (consumption linkage) and food processing can become an important non-farm activity (production linkage). Certainly, the existence of such local non-farm opportunities is linked to access to markets (de Janvry et al., 1991).

The capacity for diversification is often shaped at the household level. Obviously, entry barriers are higher for opportunity-led diversification, but even desperation-led diversification may be constrained. Household decisions for migration are facilitated by home-village networks at the destination place (Hampshire, 2002). Educational attainment determines earning possibilities off the home-farm (Abdulai and Delgado, 1999; de Janvry and Sadoulet, 2001; Yúnez-Naude and Taylor, 2001). Lucrative non-farm activities often require some capital that asset-poor households are not able to provide (Abdulai and CroleRees, 2001; Woldenhanna and Oskam, 2001; Dercon and Krishnan, 1996). For instance, launching a non-farm enterprise requires business expertise and some capital, even though most rural non-farm activities are labor-intensive (Norcliffe, 1983). Barrett (1997) reveals that many households lack access to lumpy assets (e.g. machinery and other equipment) to enter profitable niches in the non-farm sector. Where credit markets exist, livestock assets can serve as collateral (Fafchamps et al., 1998). When credit is not available, households need liquid assets for such investments, so that cash-cropping can be an important means for overcoming entry constraints. Alternatively, the necessary cash capital can come from remittances (Lucas, 1987; Francis and Hoddinott, 1993).

All these conditions that constrain or foster both opportunity- and desperation-led income diversification – at the household, regional, or national level – are subject to change. Rapid population growth and the related pressure on the natural resource base are often identified as the major cause for the rise of non-agricultural activities and migration in sub-Saharan Africa (Bryceson and Jamal, 1997; Barrett et al., 2000b; Bryceson, 2002b). While shortages of fertile land may hence push households increasingly into diversification, other factors at the household-level, in particular the expansion of education, may enable them to earn higher returns from local non-farm activities or migration. In addition, technological advances, better infrastructure, the intensification of links with markets outside the local economy, local engines of growth, such as commercial agriculture or proximity to an urban area, or simply per capita growth and increased demand for non-food goods and services can lead to a growing non-farm sector (Abdulai and Delgado, 1999; Abdulai and CroleRees, 2001; Reardon, 1999; Haggblade et al., 2002).

The structural reform programs that many African countries have undergone in the late 1980s and early 1990s have certainly shaped the patterns of growth and household portfolios in these countries since then. Possibly, reforms had their most pronounced impact on income portfolios through altering the returns in agricultural activities, in particular in cash crop sectors, which often had been subject to heavy regulation. Furthermore, macroeconomic stability that has been recorded by most reform countries since the mid-1990s may have strengthened farmers' trust into markets thereby contributing to increased market participation.

Changes in the patterns of income portfolio will also have important poverty and inequality implications. Again, the existing evidence typically builds on cross-sectional or short-term panel data. All in all, there seems to be a positive correlation between wealth and diversification in rural Africa (Barrett et al., 2001). If diversifying incomes go mainly to the better-off, growth in these sectors is inequality-increasing (Elbers and Lanjouw, 2001). However, others find that non-farm income is concentrated among the poor, so that an increase in these incomes is eventually pro-poor (van den Berg and Kumbi, 2006; Adams, 2002). These seemingly contradictory findings can possibly be reconciled by taking into account the distinct motivations for diversification. Accordingly, Lay et al. (2008) find the expansion of opportunity-led diversification to have little impact on poverty, while growth of income from desperation-led activities significantly reduces poverty. In such a case, the overall impact of diversification on the income distribution remains mixed.

It is against the background of this variety of empirical findings that we analyze Burkina Faso's income diversification patterns and their determinants.

5.3 Income Diversification and Inequality in Burkina Faso: Patterns and Changes

5.3.1 Country Setting: Shocks and Change

We use three country-wide representative household surveys for Burkina Faso, Enquete Prioritaires (EP) that were conducted in 1994 (EP I), 1998 (EP II), and 2003 (EP III). Between 1994 and 2003, annual per capita growth averaged 2 percent due to a wide range of economic reforms and rural poverty rates fell from 63.4 percent to 53.3 percent (Grimm and Günther, 2007a).

In this West-African country rainfall levels are low and vary over regions. Drought periods occur regularly, like in 1997/98 and 2000/01. Harvest shortfalls often involve an explosion of staple prices, as for example after the 1997/98 drought. This price hike contributed to the high levels of poverty in 1998, when the headcount rose to 68.7 percent, as most of the rural poor are net buyers of food (Grimm and Günther, 2007a). Agriculture has a strong seasonal character with a short rainy season from May to August, so that the agricultural labor force faces a long lean season (Delgado and Ranade, 1987). Therefore, there is a tradition of seasonal migration to agriculturally more favorable regions in Côte d'Ivoire, where household members temporarily work on cocoa and coffee farms (Konseiga, 2007). Furthermore, quite a few Burkinabe households rely on remittances from relatives who have settled down permanently in Côte d'Ivoire. However, in 2002 the ongoing political crisis in Burkina Faso's neighboring country gave rise to a civil war, which affected these labor movements. In both 1998 and 2003, Burkinabe households hence faced substantial adverse shocks that should have influenced their income portfolio choices.

Another though rather positive shock, at least for rural households, was the CFA Franc devaluation in 1994. The devaluation favored Burkina's main export sector, cotton, which became the main driver of agricultural growth in the period under consideration. The share of cotton farmers increased significantly from 8 percent in 1994 to 24 percent in 2003.⁵² Yet, change in agriculture was not limited to the cotton sector: Despite little or no improvement in overall agricultural productivity, market integration of staple-food farmers also improved considerably. Furthermore, the livestock sector grew rapidly. Own calculations based on the survey data reveal that cattle ownership increased from 40 percent of rural households in 1994 to 67 percent in 2003, with the average number of cattle conditional on ownership remaining constant at approximately 8.5.

These developments in the agricultural sector as well as the external shocks shaped the patterns of income diversification that we analyze below. In the Burkinabe context, diversification and its welfare impact have been investigated earlier by Reardon et al. (1992) and Reardon and Taylor (1996). These studies were based on a panel dataset for the cropping seasons 1983/84 and 1984/85 and focused on household responses to drought.⁵³ They show that poor households depend mainly on crop income and lack access to non-farm income that could cushion the effect of harvest failures (Reardon et al., 1988; Reardon et al., 1992; Reardon and Taylor, 1996). As a result, poor households turn to selling their livestock during periods of drought, thereby increasing inequality and possibly trapping these households in poverty (Reardon and Taylor, 1996). As the 1998 survey data should also capture the impact of the 1997/1998 drought, the subsequent analysis also sheds light on the question of whether Reardon and Taylor's findings still hold 13 years later.

We first provide a simple descriptive analysis of the patterns of income diversification for 1994, 1998, and 2003. These observations allow us to identify some broad patterns of change that will then be put to scrutiny and refined in a multivariate framework.

5.3.2 Pattern of Income Diversification in Burkina Faso

Table 5.1 presents a matrix of participation in different income portfolios. We consider participation in crop-farming, livestock selling, non-farm activities, such as wage and self-employment, and migration within the country and to Côte d'Ivoire. The columns show the different types of farm households and the rows further classify them according to their engagement in non-agricultural activities. Accordingly, the first row illustrates that in 1994 most of the sample households were engaged in cropping (62 percent), whereas only a few were engaged in livestock selling as their only farm activity. About 40 percent of the farmers derived

52 Compare Grimm and Günther (2007b). All cotton in Burkina is bought by the parastatal Sofitex, which guarantees a fixed price at the beginning of the season. While the gains from the devaluation therefore mainly benefited Sofitex, cotton producers were protected from the subsequent decline in world market prices between 1998 and 2003.

53 The survey covers only 150 households, 50 in each agro-climatic region, which somewhat impairs the significance of the results.

their income from both agricultural activities. In 1998, the drought year, farmers heavily moved into livestock selling and the share of those who only sold livestock increased significantly, most likely reflecting complete harvest failure. In 2003, agricultural portfolios were similar to 1994.

TABLE 5.1: INCOME DIVERSIFICATION MATRIX: PARTICIPATION RATES IN INCOME-GENERATING ACTIVITIES

| | | Agricultural activities | | | | | | | | |
|-----------------------------|------------------------|-------------------------|-----------|----------------------|----------|-----------|----------------------|----------|-----------|----------------------|
| | | 1994 | | | 1998 | | | 2003 | | |
| | | cropping | livestock | cropping & livestock | cropping | livestock | cropping & livestock | cropping | livestock | cropping & livestock |
| Non-agricultural activities | total | 0.62 | 0.01 | 0.37 | 0.35 | 0.09 | 0.56 | 0.64 | 0.03 | 0.33 |
| | no non-agricultural | 0.21 | 0.01 | 0.11 | 0.15 | 0.04 | 0.18 | 0.32 | 0.01 | 0.15 |
| | non-farm | 0.25 | 0.00 | 0.14 | 0.11 | 0.02 | 0.17 | 0.23 | 0.01 | 0.13 |
| | migration | 0.08 | 0.00 | 0.05 | 0.04 | 0.01 | 0.10 | 0.05 | 0.00 | 0.03 |
| | non-farm and migration | 0.08 | 0.00 | 0.07 | 0.04 | 0.01 | 0.12 | 0.04 | 0.00 | 0.02 |

In each year, the (4x3) diversification matrix shows pure crop households at the top left, e.g. 21 percent in 1994, while households at the bottom right fully diversify, deriving income from cropping, livestock selling, the non-farm sector and migration, e.g. 7 percent in 1994. The degree of diversification is highest in the shock year, 1998, but households seem to become more specialized, when 2003 is compared to 1994. In 1994 non-agricultural diversification was dominated by non-farm activities although quite a large share of households (15 percent) receives income from both remittances and local non-farm activities. For 1998, rising numbers at the bottom right of the matrix indicate increased diversification through livestock selling and migration. After 1998, migration decreases significantly, as does non-farm participation. Even compared to 1994, income portfolios of farm households in Burkina Faso have become less diversified.

Table 5.2 provides information on participation in and income shares (average for all households as well as conditional on participation) from different activities by per capita expenditure quintiles. The crop income shares went down considerably in 1998 due to harvest shortfalls. While rates of participation in crop farming recovered by 2003, the income share remained at the low levels of 1998. It might be that this reflects biases resulting from differences in data collection. Whereas the EP I and EP II collected annual income data, the EP III only recorded income in the last month. As the EP III survey was undertaken in a pre-harvest month, income from cropping is most likely to be underreported. This seasonal bias in the data prevents us from putting too much trust into the income shares for 2003 and we focus on participation rates instead. More specifically, underreporting of crop income is very likely to lead to an upward bias in other shares. We are, however, confident that we can minimize these biases for participation rates (see data description in the appendix).

Participation in livestock selling and income shares from this source rose significantly from 1994 to 1998. This increase was highest in the lowest expenditure quintiles implying that pressure on poor households to sell their livestock, often their only asset, was particularly strong. In 2003, overall participation rates were again at the same level as in 1994, but now richer households

participated more in income generation from livestock selling – possibly indicating a tendency towards commercialization of this sector.

TABLE 5.2: PARTICIPATION RATES AND INCOME SHARES FROM DIFFERENT ACTIVITIES

| Quintile | Cropping | | | Livestock selling | | | Non-farm activities | | | Migration | | |
|--|----------|------|------|-------------------|------|------|---------------------|------|------|-----------|------|------|
| | 1994 | 1998 | 2003 | 1994 | 1998 | 2003 | 1994 | 1998 | 2003 | 1994 | 1998 | 2003 |
| <i>Participation rates</i> | | | | | | | | | | | | |
| 1 | 0.98 | 0.89 | 0.99 | 0.38 | 0.67 | 0.34 | 0.48 | 0.45 | 0.37 | 0.27 | 0.40 | 0.12 |
| 2 | 0.99 | 0.91 | 0.98 | 0.38 | 0.68 | 0.35 | 0.51 | 0.49 | 0.47 | 0.28 | 0.36 | 0.14 |
| 3 | 1.00 | 0.90 | 0.98 | 0.38 | 0.68 | 0.34 | 0.56 | 0.48 | 0.44 | 0.27 | 0.34 | 0.15 |
| 4 | 0.99 | 0.95 | 0.97 | 0.40 | 0.63 | 0.36 | 0.55 | 0.45 | 0.44 | 0.30 | 0.31 | 0.14 |
| 5 | 0.98 | 0.93 | 0.94 | 0.35 | 0.57 | 0.39 | 0.58 | 0.46 | 0.44 | 0.30 | 0.27 | 0.16 |
| Total | 0.99 | 0.91 | 0.97 | 0.38 | 0.65 | 0.36 | 0.54 | 0.47 | 0.43 | 0.28 | 0.33 | 0.14 |
| <i>Share in total income</i> | | | | | | | | | | | | |
| 1 | 0.75 | 0.48 | 0.61 | 0.07 | 0.23 | 0.14 | 0.11 | 0.13 | 0.11 | 0.04 | 0.09 | 0.05 |
| 2 | 0.77 | 0.53 | 0.58 | 0.06 | 0.20 | 0.14 | 0.10 | 0.12 | 0.16 | 0.03 | 0.08 | 0.05 |
| 3 | 0.77 | 0.55 | 0.58 | 0.06 | 0.19 | 0.15 | 0.11 | 0.12 | 0.14 | 0.03 | 0.06 | 0.05 |
| 4 | 0.74 | 0.58 | 0.59 | 0.06 | 0.16 | 0.16 | 0.12 | 0.11 | 0.12 | 0.04 | 0.06 | 0.04 |
| 5 | 0.70 | 0.60 | 0.53 | 0.06 | 0.15 | 0.20 | 0.15 | 0.13 | 0.12 | 0.04 | 0.04 | 0.05 |
| Total | 0.75 | 0.55 | 0.58 | 0.06 | 0.19 | 0.16 | 0.12 | 0.12 | 0.13 | 0.04 | 0.07 | 0.05 |
| <i>Shares conditional on participation</i> | | | | | | | | | | | | |
| 1 | 0.74 | 0.54 | 0.62 | 0.25 | 0.34 | 0.43 | 0.23 | 0.30 | 0.31 | 0.26 | 0.18 | 0.45 |
| 2 | 0.74 | 0.59 | 0.58 | 0.19 | 0.30 | 0.41 | 0.21 | 0.27 | 0.29 | 0.15 | 0.23 | 0.40 |
| 3 | 0.75 | 0.61 | 0.58 | 0.19 | 0.29 | 0.46 | 0.21 | 0.25 | 0.28 | 0.12 | 0.21 | 0.35 |
| 4 | 0.72 | 0.61 | 0.60 | 0.20 | 0.27 | 0.45 | 0.24 | 0.26 | 0.27 | 0.14 | 0.21 | 0.38 |
| 5 | 0.69 | 0.64 | 0.55 | 0.24 | 0.28 | 0.53 | 0.28 | 0.30 | 0.30 | 0.15 | 0.19 | 0.37 |
| Total | 0.72 | 0.60 | 0.58 | 0.21 | 0.29 | 0.47 | 0.24 | 0.28 | 0.29 | 0.15 | 0.22 | 0.38 |

Participation in non-farm activities was negatively affected by the drought and participation rates dropped significantly. This appears to contradict the perception of non-farm activities as an insurance strategy. Yet, it may just reflect that households use this strategy to insure ex-ante by maintaining some level of non-farm activity. It seems plausible that returns to non-farm activities fall during the drought since local demand for non-agricultural products and services is likely to decline when a region is hit by a drought. Non-farm activities still serve as an insurance mechanism, as the associated losses may still be lower than those in cropping activities. Interestingly, households from upper quintiles withdraw disproportionately from this activity. At the same time, income shares conditional on participation increased strongest in the lowest quintiles. It may be that the poorest households were forced to stick to this income source whereas wealthier households were in the position to withdraw from the non-farm sector as a response to falling returns. Participation rates further dropped in 2003 and this decline of the non-farm sector affected all household groups equally. Surprisingly, shares from this income source were not significantly higher than in 1994 or 1998, as should be expected as a result of underreported crop incomes. Thus, it seems that non-farm income shares shrank in reality.

Overall participation rates in migration went up in 1998, disproportionately in the lowest quintiles, while they decreased in the highest quintile. This suggests that the most desperate

households turned to this strategy in order to mitigate the drought impact. By 2003, migratory activities had declined significantly mainly due to the Ivorian crisis, affecting all households to the same extent.

In sum, these observations suggest the following broad patterns of changes in income portfolios: First, in the drought year livestock selling and migration were coping mechanisms for dealing with harvest shortfalls. Mainly the most desperate households seemed to turn to these strategies. Non-farm activities did not serve as coping mechanism through increased participation in that sector, but income shares rose in shock years. Second, between 1994 and 2003 income portfolios became less diversified. Overall participation in livestock selling remained constant, but participation rates over expenditure quintiles reversed, with richer households becoming more likely to sell livestock. The decline of non-farm activities and of migration affected all quintiles equally.

5.3.3 Determinants of Income Portfolio Diversification

We estimate a multivariate probit model to illustrate the importance of specific determinants of diversification behavior. Furthermore, the multivariate framework allows us to more rigorously assess the determinants of change. In the model, households choose to engage in livestock ownership, livestock selling, non-farm activities, migration, and cotton farming. The reason for distinguishing livestock selling and ownership in the choice model is to disentangle the determinants of the former, which often represents an ex-post reaction, from the latter that can be seen as an ex-ante diversification strategy for insurance against shocks.⁵⁴ Both of course capture livestock breeding and selling as a commercial activity. The model is estimated on the pooled cross-sections, using year interaction terms for most variables to allow for changes in coefficients over years.

At the regional or community level, we incorporate market access, the availability of a village-based migration network, and rain-variables at the provincial level.⁵⁵ We include mean rainfall to proxy for agricultural potential, rain variability to proxy for the risk of harvest shortfalls and the deviation from average rainfall levels in the respective year for capturing the effect of agricultural performance. At the household level, household composition variables and household head characteristics, like education and age are used.

Table 5.3 presents the estimation results. The hypothesis that the correlations between the error terms of each equation are all zero can be rejected at a very high significance level, so that that the multivariate probit model seems to be appropriate. The residuals of the models for livestock owning and selling are highly correlated. The positive and significant sign proves that these two strategies are complementary, which is not surprising by definition. All pair-wise correlations

54 More than 40 percent of the sample households own livestock, but do not generate any income by selling it.

55 Burkina Faso is divided into 13 regions, subdivided into 45 provinces. A detailed description of all the included variables is provided in Appendix 2.

between the error terms of the migration and other activity choices considered turn out to be significantly negative. This suggests that migration is an important substitute for the other income-generating strategies. In principle, some of these strategies could also be complementary, e.g. if remittances could be used by the household to set up a non-farm activity. Although these effects may still be present, our results imply that they are not strong enough. Given the high average number of children in Burkinabe households, the departure of an adult member is likely to make it too difficult for the remaining members to pick up a non-crop activity, which are often quite labor-intensive. In contrast, the error terms indicate important complementarities between activities in the cotton and livestock sectors. Livestock holdings, especially in the form of cattle, play an important role in Burkina Faso, indicating social status (Nianogo and Somda, 1999) and serving as a store of wealth (Fafchamps et al., 1998). It seems that households able to engage in the attractive cotton sector use these returns to invest in livestock. Doing so, they were also able to take part in livestock selling. Yet, the correlation coefficient between cotton and livestock ownership is considerably higher than between cotton and livestock selling, so that cotton farmers predominantly hold livestock for storing wealth and not for selling it. Although cotton could provide the necessary start-up capital for non-farm activities, the correlation of the respective error terms does not support this view. For cotton farmers it is not attractive to engage in the non-farm sector as returns relative to cotton are low.

Most controls at the household level show the expected signs. Several household composition and size variables are significantly linked to diversification strategies. While the availability of male, female or child labor may favor engagement in a certain activity, more members, in particular prime age adults, may also put pressure on the household's resource base, forcing it into non-crop activities.⁵⁶ The positively significant coefficients of all age and gender groups on livestock ownership as well as livestock selling may indeed reflect both effects. Non-farm activities appear to be undertaken mainly by households with a relatively high number of female adults, while many male adults in a household increase the likelihood of engaging in migration or the cultivation of cotton. The decision to migrate is negatively affected by the number of children, since parents would not want to leave their children behind. Accordingly, the number of old members is positively linked to it, since they may step in and take care of children. The number of female adults is negatively linked to migration, as they would be left in a weak position in the Burkinabe society (Diallo, 2002) after the departure of male household members.

Older household heads are more likely to send household members elsewhere, while younger household heads tend to diversify more into non-farm activities and especially cotton. At least in 1994, better educated households tend to participate more in non-farm activities and migration as well as cotton. However, the education coefficients vary significantly across the years, a finding that we interpret in more detail below.

⁵⁶ Note that household composition variables may be jointly determined with activity choices, as households' fertility choices may depend on future diversification plans (e.g. migration).

TABLE 5.3: DIVERSIFICATION DETERMINANTS, MULTIVARIATE PROBIT ESTIMATION FOR POOLED CROSS-SECTIONS OF 1994, 1998, AND 2003

| | (1) | (2) | (3) | (4) | (5) |
|---|------------------------|----------------------|------------|------------|------------|
| | Livestock ownership | Livestock selling | Migration | Non-farm | Cotton |
| children < 10 | 0.0162*** | 0.0133*** | -0.0106*** | 0.0061*** | 0.0069*** |
| male children 10-15 | 0.0167*** | 0.0165*** | -0.0173*** | -0.0017 | 0.0043 |
| female children 10-15 | 0.0135*** | -0.0001 | -0.0075 | 0.0127** | -0.0005 |
| male adults | 0.0307*** | 0.0253*** | 0.0864*** | 0.0032 | 0.0157*** |
| female adults | 0.0244*** | 0.0137*** | -0.0198*** | 0.0248*** | -0.0001 |
| old > 64 | 0.0215*** | -0.0039 | 0.0269*** | -0.0038 | -0.0102** |
| age of head | -0.0004* | -0.0004 | 0.0016*** | -0.0021*** | -0.3635*** |
| education of head | -0.0403** | -0.0214 | 0.0491** | 0.0793*** | 0.0446*** |
| education of head*1998 | 0.0128 | -0.0786* | -0.0493* | 0.0743* | -0.0166 |
| education of head*2003 | -0.0123 | -0.1025** | -0.0600** | 0.0694* | -0.015 |
| poor accessibility of markets | 0.0693*** | 0.0841** | -0.0723* | -0.0747** | 0.0421 |
| poor accessibility of markets*1998 | -0.1063*** | -0.1125** | 0.0393 | -0.0633 | -0.0396 |
| poor accessibility of markets*2003 | -0.0551 | -0.0671 | 0.0375 | -0.0041 | -0.0511 |
| mean rain | -0.0132*** | -0.0497*** | -0.0232*** | -0.0237*** | 0.0198*** |
| rain variability | 0.0017* | 0.0015 | 0.0027** | -0.0004 | -0.0005 |
| rain deviation | 0.0001 | -0.0006* | -0.0004** | 0.0002 | 0.0005** |
| Sahel zone | 0.0650*** | 0.1659*** | -0.1501*** | -0.1535*** | -0.1113*** |
| Sahel zone*1998 | 0.0371 | 0.0254 | -0.1094*** | 0.0308 | 0.0079 |
| Sahel zone*2003 | 0.0234 | 0.3115*** | 0.2496*** | 0.0988* | |
| Sub-Sahel zone | 0.0764*** | 0.2520*** | 0.0017 | -0.0166 | -0.0802*** |
| Sub-Sahel zone*1998 | 0.0167 | 0.0323 | 0.0671* | 0.0082 | -0.0181 |
| Sub-Sahel zone*2003 | -0.0643** | -0.1475*** | -0.0196 | -0.0963** | 0.0136 |
| South Sudanese zone | -0.0401 | -0.0895* | -0.0235 | 0.1107*** | 0.0880** |
| South Sudanese zone*1998 | -0.0163 | 0.0029 | -0.0178 | -0.0663 | -0.0108 |
| South Sudanese*2003 | 0.0580** | 0.2955*** | 0.0683** | 0.0961 | -0.0161 |
| village migration network | | | 0.1568*** | | |
| year dummy 1998 | 0.1268*** | 0.2855*** | -0.023 | -0.0142 | 0.0951*** |
| year dummy 2003 | 0.1347*** | 0.0162 | -0.0928*** | -0.0344 | 0.1331*** |
| <i>Correlation between error terms</i> | | | | | |
| rho 21 | 0.6553*** | | rho 42 | -0.0051 | |
| rho 31 | -0.1533*** | | rho 52 | 0.0480** | |
| rho 41 | -0.0111 | | rho 43 | -0.0360** | |
| rho 51 | 0.1743*** | | rho 53 | -0.1787*** | |
| rho 32 | -0.0681*** | | rho 54 | -0.0264 | |
| Likelihood ratio test of $\rho_{21} = \rho_{31} = \rho_{41} = \rho_{51} = \rho_{32} = \rho_{42} = \rho_{52} = \rho_{43} = \rho_{53} = \rho_{54} = 0$ $\chi^2(10) = 2229.88$ Prob > $\chi^2 = 0.0000$ | | | | | |

***/**/* denote significance at the 1/5/10 percent level respectively. Marginal effects are computed holding all other variables constant at their mean. Effects of dummy variables: discrete change from 0 to 1. Effects of rain variables consider 10 percent changes. The Sahel zone dummy interacted with the 2003 dummy is a perfect predictor for non-participation in cotton and is hence dropped. Standard errors are clustered at the locality level.

At the community level, migrant networks play an important role in facilitating migration. As expected, good access to markets increases the likelihood of participating in non-farm activities and migration. Yet, if markets are far and farm and non-farm products cannot be sold easily, livestock ownership seems to be an alternative strategy for income generation. Accordingly, livestock selling is also more likely to occur where markets are inaccessible. Another explanation might be that livestock selling is the only way of income generation when organized market places are absent, as livestock may just be traded among neighboring households not requiring a market place. In fact, this would explain the switch in the sign of the coefficient in 1998: Trade

between neighbors may collapse in case of a shock that affects all villagers simultaneously, which implies a more important role for markets for livestock trading.

In such a diverse country in agro-climatic terms, diversification patterns of course differ over regions. In the Sahel and Sub-Sahel zone, where soil quality is poor, many farm households do not depend on cropping including cotton but are herders. Moreover, in the Sahel region diversification into non-agricultural activities is rare. A small population scattered over vast areas would render most non-farm activities unprofitable. By contrast, rural households in the South-Sudanese belt, which comprises tree-covered and cotton regions, are more likely to engage in non-farm activities and the cultivation of cotton, but are less likely to sell livestock.

The negative effect of average rainfall in all activity choices except cotton indicates that households in provinces with low levels of rainfall are forced to diversify into non-cropping activities. As expected, livestock owning serves as an ex-ante diversification strategy that is most relevant in provinces with high rainfall variability. Accordingly, households turn to selling livestock when rain falls short of long-run averages. Similar to livestock ownership, migration tends to be more prevalent in provinces with high rainfall variability serving as an insurance mechanism. In addition, migration appears to occur in response to negative shocks, as the number of households receiving remittances increases with the short-run rain deviations.

So far, the multivariate analysis illustrates that diversification in Burkina Faso is mainly pursued as an insurance strategy. In particular, rainfall variables indicate that participation in non-cropping activities takes place in provinces with the poorest agro-climate. Both migration and livestock selling appear to serve as direct responses to negative shocks. In line with the above descriptive findings on participation, non-farm activities are an ex-ante diversification strategy that is pursued in better connected and more densely populated areas. Unfortunately, we lack data on landholdings to investigate more rigorously the hypothesis of diversification being driven by desperation.

From these more general patterns, we now turn to changes over time. The year dummies include year-specific effects and time trends that are not captured by other controls. Both the 1998 and 2003 time dummies are large and significant for livestock ownership and cotton reflecting the abovementioned expansion of these sectors. The 2003 year dummy is negative for migration reflecting the negative effects of the unstable situation in Côte d'Ivoire.

Education becomes more important in non-farm participation, possibly reflecting better opportunities and higher returns, but also higher barriers to better remunerated non-farm activities. Thus, it seems that there is a shift to more sophisticated, skill-based activities for generating income from this sector. For migration, we observe the opposite: the positive effect of education is offset in 1998 and 2003. This points to migration being less determined by opportunities at the destination place, but rather by desperation at home.

5.5 Conclusion

Most existing studies of income diversification in rural sub-Saharan Africa provide a snapshot view on household behavior. Often, these snapshots yield pessimistic views on non-farm activities as being mainly driven by desperation. Furthermore, in countries with unfavorable and unstable climatic conditions, as for example Burkina Faso, the literature has stressed income diversification as a means to smooth income and cope with recurrent droughts. Our empirical analysis shows that diversification behavior is indeed dominated by insurance motivations. Yet, there are also signs of more profound structural changes. We observe a tendency for non-farm activities to become more opportunity-led, while poorer households seem to resort to migration although overall migration flows decline significantly because of the political crisis in Côte d'Ivoire.

This paper hence argues that a more comprehensive view on income diversification (beyond insuring or coping with risks) even in shock-prone environments is needed. Clearly, the analysis of three cross-sections has important limitations. In particular, it is difficult to establish causal relationships and to isolate the effects of shocks. Nonetheless, our rather descriptive approach based on imperfect data yields interesting insights into the relationship between growth, income diversification, and inequality in a shock-prone agricultural environment.

Had we examined only household reactions to the 1997/98 drought, we might have concluded that the livelihoods of Burkinabe farmers had not changed much between the mid-1980s and the mid-1990s. Although the non-farm sector shields households against crop shortfalls, it is also negatively affected by the drought and households resort to selling livestock to smooth consumption or ensure survival. This may turn out to be particularly harmful to very poor households that are deprived of their asset base. While these patterns can still be observed in Burkina Faso, there are also signs of structural change that are reflected in income portfolio choices. Between 1994 and 2003 income portfolios become less diversified, mainly since agricultural activities, in particular cotton and livestock, appear to offer higher returns. In addition, some of our findings hint at better opportunities in the non-farm sector. Both findings reflect a more important role of a household's comparative advantage in income portfolio choices, which implies that diversification is increasingly opportunity-led. This seems to be a favorable development that can be attributed to economic reforms to some extent. Yet, as outlined by Barrett et al. (2001, p. 326), "the key point is that the wealthy have greater freedom to choose among a wider range of options than do the poor." In fact, some of our findings illustrate that this (increasingly) holds for the Burkinabe context. Low returns in non-farm activities go hand in hand with higher entry barriers and migration seems to be more desperation-led. Furthermore, richer households appear to be in a better position to participate in livestock and cotton. Finally, one should not be overly optimistic with regards to the increased agricultural specialization of rural Burkinabe households. Such a strategy entails important risks that may not be limited to climatic factors: Between 2004/05 and 2006/07 the parastatal Sofitex recorded

losses of USD 128 million. This puts a big question mark behind the sustainability of Burkina Faso's cotton sector success story.

Literature

- Abdulai, A., and A. CroleRees (2001). Determinants of Income Diversification among Rural Households in Southern Mali. *Food Policy*, 26(4): 437-452.
- Abdulai, A., and C.L. Delgado (1999). Determinants of Nonfarm Earnings of Farm-Based Husbands and Wives in Northern Ghana. *American Journal of Agricultural Economics*, 81(1): 117-130.
- Adams, R.H. (2002). Non-Agricultural Income, Inequality, and Land in Rural Egypt. *Economic Development and Cultural Change*, 50(2): 339-363.
- Barrett, C. B. (1997). Food Marketing Liberalization and Trader Entry: Evidence from Madagascar. *World Development*, 25(5): 763-177.
- Barrett, C. B., and T. Reardon (2000). Asset, Activity, and Income Diversification among African Agriculturalists: Some Practical Issues. Project report to USAID BASIS CRSP, University of Wisconsin-Madison Land Tenure Centre, March 2000.
- Barrett, C.B., T. Reardon and P. Webb (2001). Non-Agricultural Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics and Policy Implications. *Food Policy*, 26(4): 315-331.
- Barrett, C.B., M. Bezuneh and A. Aboud (2000a). Income Diversification, Poverty Traps and Policy Shocks in Cote d'Ivoire and Kenya. Department of Applied Economics and Management, Cornell University, Ithaca.
- Barrett, C.B., M. Bezuneh, C.D. Clay and T. Reardon (2000b). Heterogeneous Constraints, Incentives and Income Diversification Strategies in Rural Africa. Report to USAID/BASIS CRSP, Cornell University and Michigan State University.
- van den Berg, M., and G.E. Kumbi (2006). Poverty and the Rural Non-Farm Economy in Oromia, Ethiopia. *Agricultural Economics*, 35(Suppl.): 469-475.
- Block, S., and P. Webb (2001). The Dynamics of Livelihood Diversification in Post-Famine Ethiopia. *Food Policy*, 26(4): 333-350.
- Bryceson, D.F. (2002a). The Scramble in Africa: Reorienting Rural Livelihood. *World Development*, 30(5): 725-739.
- Bryceson, D.F. (2002b). Multiplex of Livelihoods in Rural Africa: Recasting the Terms and Conditions of Gainful Employment. *Journal of Modern African Economics*, 40(1): 1-28.
- Bryceson, D.F., and V. Jamal (Eds.) (1997). *Farewell to farms: De-agrarianisation and employment in Africa*. Aldershot: Ashgate.
- Chapoto, A., T.S. Jayne and N. Mason (2006). Security of Widows' Access to Land in the Era of HIV/AIDS: Panel Survey Evidence from Zambia. Food Security Research Project. Working Paper 19, Lusaka.
- Delgado, C. L., and C.G. Ranade (1987). Technological Change and Agricultural Labor Use. In J.W. Mellor, C.L. Delgado and M.J. Blackie (Eds.), *Accelerating Food Production in sub-Saharan Africa*. IFPRI, Baltimore: Johns Hopkins University Press.
- Dercon, S. (1998). Wealth, Risk and Activity Choice: Cattle in Western Tanzania. *Journal of Development Economics*, 55(1): 1-42.
- Dercon, S., and Krishnan, P. (1996). Income Portfolios in Rural Ethiopia and Tanzania: Choices and Constraints. *Journal of Development Studies*, 32(6): 850-875.

- Diallo A. (2002). La Position de la femme dans la problématique foncière au Burkina Faso. Paper presented at the Pan-African 'Land and Resource Rights-Workshop' in Cairo, March 2002.
- Elbers, C., and P. Lanjouw (2001). Intersectoral Transfer, Growth and Inequality in Rural Ecuador. *World Development*, 29(3): 481-496.
- Fafchamps, M., C. Udry and K. Czukas (1998). Drought and Saving in West Africa: Are Livestock a Buffer Stock? *Journal of Development Economics*, 55(2): 273-305.
- Francis, E., and J. Hoddinott (1993). Migration and Differentiation in Western Kenya: A Tale of Two Sub-Locations. *Journal of Development Studies*, 30(1): 115-145.
- Grimm, M., and I. Günther (2007a). Growth and Poverty in Burkina Faso. A Reassessment of the Paradox. *Journal of African Economies*, 16(1): 70-101.
- Grimm, M., and I. Günther (2007b). Pro-Poor Growth in Burkina Faso. The Role of Price Shocks. In M. Grimm, A. McKay and S. Klasen (Eds.), *Determinants of Pro-Poor Growth: Analytical Issues and Findings from Country Cases*. London: Palgrave-Macmillan.
- Grimm, M., and I. Günther (2007c). Measuring Pro-Poor Growth When Relative Prices Shift. *Journal of Development Economics*, 82(1): 245-256.
- Haggblade, S., P. Hazell and J. Brown (1989). Farm Non-Farm Linkages in Rural Sub-Saharan Africa. *World Development*, 17(8): 1173-1202.
- Haggblade, S., P. Hazell and T. Reardon (2002). Strategies for Stimulating Poverty-Alleviating Growth in the Rural Nonfarm Economy in Developing Countries. International Food Policy Research Institute EPTD Discussion Paper 92.
- Hampshire, K. (2002). Fulani on the Move: Seasonal Economic Migration in the Sahel as a Social Process. *Journal of Development Studies*, 38(5): 15-36.
- Islam, N. (1997). The Nonfarm Sector and Rural Development. Food, Agriculture and Environment Discussion Paper 22. Washington D.C: International Food Policy Research Institute.
- de Janvry, A., and E. Sadoulet (2001). Income Strategies among Rural Households in Mexico: The Role of Off-Farm Activities. *World Development*, 29(3): 467-480.
- de Janvry, A., M. Fafchamps and E. Sadoulet (1991). Peasant Household Behavior with Missing Markets: Some Paradoxes Explained. *Economic Journal*, 101(409): 1400-1417.
- Konseiga, A. (2007). Household Migration Decisions as Survival Strategy: The Case of Burkina Faso. *Journal of African Economies*, 16(2): 198-233.
- Lay, J., G. Michuki M'Mukaria and T. Omar Mahmoud (2008). Few Opportunities, Much Desperation: The Dichotomy of Non-agricultural Activities and Inequality in Western Kenya. *World Development*, 36(12): 2713-2732.
- Lemi, A. (2006). The Dynamics of Income Diversification in Ethiopia: Evidence from Panel Data. University of Massachusetts Boston, Economics Department Working Papers 3.
- Liedholm, C., and P. Kilby (1989). The Role of Non-Farm Activities in the Rural Economy. In J.G. Williamson and V.R. Anchamukhi. (Eds.), *The Balance Between Industry and Agriculture in Economic Development*. London: Macmillan.
- Lucas, R. (1987). Emigration to South Africa's Mines. *American Economic Review*, 77(3): 313-330.
- Lucas, R., and O. Stark (1985). Motivations to Remit: Evidence from Botswana. *Journal of Political Economy*, 93(5): 901-918.

- Matsumoto, T., Y. Kijima and T. Yamano (2006). The Role of Non-Farm Activities and Migration in Reducing Poverty: Evidence from Ethiopia, Kenya, and Uganda. *Agricultural Economics*, 35(Suppl.): 449-458.
- Murdoch, J. (1995). Income Smoothing and Consumption Smoothing. *Journal of Economic Perspectives*, 9(3): 103-114.
- Nianogo, J. N., and J. Somada (1999). Diversification et intégration inter-spécifique dans les élevages ruraux au Burkina Faso. *Biotechnology, Agronomy, Society and Environment*, 3(3): 133-139.
- Norcliffe, G. (1983). Operating Characteristics of Rural Non-Farm Enterprises in Central Province, Kenya. *World Development*, 11(11): 981-994.
- Otsuka, K., and T. Yamano (2006). Introduction to the Special Issue on the Role of Nonfarm Income in Poverty Reduction: Evidence from Asia and East Africa. *Agricultural Economics*, 35(Suppl.): 393-397.
- Reardon, T. (1999). *Rural Non-Farm Income in Developing Countries*. Rome: FAO.
- Reardon, T. (1997). Using Evidence of Household Income Diversification to Inform the Study of Rural Non-Agricultural Labor Market in Africa. *World Development*, 25(5): 735-748.
- Reardon, T., and J.E. Taylor (1996). Agro-Climatic Shock, Income Inequality, and Poverty: Evidence from Burkina Faso. *World Development*, 24(5): 901-914.
- Reardon, T., C. Delgado and P. Malton (1992). Determinants and Effects of Income Diversification amongst Farm Households in Burkina Faso. *Journal of Development Studies*, 28(1): 264-296.
- Reardon, T., P. Matlon and C. Delgado (1988). Coping with Household-Level Food Insecurity in Drought-Affected Areas of Burkina Faso. *World Development*, 6(9): 1065-1074.
- Reardon, T., J.E. Taylor, A. Stamoulis, P. Lanjouw and A. Balisacan (2000). Effects of Non-Agricultural Employment on Rural Income Inequality in Developing Countries: An Investment Perspective. *Journal of Agricultural Economics*, 51(2): 266-288.
- Woldenhanna, T., and A. Oskam (2001). Income Diversification and Entry Barriers: Evidence from the Tigary Region of Northern Ethiopia. *Food Policy*, 26(4): 351-365.
- Yúnez-Naude, A., and J.E. Taylor (2001). The Determinants of Nonfarm Activities and Incomes of Rural Households in Mexico, with Emphasis on Education. *World Development*, 29(3): 561-572.

Appendix

DATA DESCRIPTION

The surveys were conducted by the Institut National de la Statistique et de la Demographie de Burkina Faso in collaboration with the United Nations Development Programme and the World Bank. Each survey covers 8,500 households.

From the approximately 5,800 rural households per year we exclude households that fully specialize in non-agricultural activities to focus on diversification patterns out of agriculture. This leaves us with a sample of 5,608 households in 1994, 5,325 households in 1998 and 5,460 households in 2003.

Crop income includes cash income from selling crops and the value of self-consumed crops.

Livestock income contains sales of cattle, goats, sheep and pigs.

Non-farm income is composed of income from wage and self-employment. Only two percent of the households in the sample receive wage income.

Migration only considers remittances from national migration and from migration to Côte d'Ivoire. We expect these migration flows to follow quite similar patterns due to the geographical and cultural proximity of the two countries.

The EP I was undertaken in the post-harvest months of 1994, but the EP II and EP III were conducted in the pre-harvest months of 1998 and 2003 respectively. Even though taking yearly incomes, this difference can lead to seasonal biases, as the latest incomes would be recalled more accurately.

The EP III only presents information about monthly incomes, creating some seasonal bias. We reduce the bias by adjusting crop incomes with the value of self-consumed food items and non-farm incomes with profit of non-farm enterprises in the last 12 months.

TABLE A5.1: VARIABLE DESCRIPTION AND MEAN VALUES

| Variable | Description | 1994 | 1998 | 2003 |
|-------------------------------|--|--------|--------|--------|
| per-capita expenditure | constructed by Grimm and Günther (2007c) based on household consumption data, deflated to the base year 1994 by using urban and rural decile specific price deflators | 56,222 | 43,002 | 64,307 |
| children < 10 | number of household members under 10 years | 4.74 | 4.32 | 3.36 |
| male children 10-15 | number of male household members between 10 and 15 years | 0.88 | 0.87 | 1.37 |
| female children 10-15 | number of female household members between 10 and 15 years | 0.80 | 0.78 | 0.64 |
| male adults | number of male household members between 16 and 64 years; for all migrant households that do not list any potential migrant, this number is adjusted with a factor of 1 to reflect household size before the departure of the migrant; migrants are typically male adults. (Konseiga 2007, Hampshire 2002) | 2.71 | 2.64 | 1.96 |
| female adults | number of female household members between 16 and 64 years | 3.08 | 3.09 | 2.36 |
| old > 64 | number of household members older than 64 years | 0.46 | 0.45 | 0.33 |
| age of head | age in years of household head | 50.66 | 50.39 | 47.76 |
| education of head | dummy taking the value 1 if household head has ever attended any school | 0.08 | 0.05 | 0.08 |
| poor accessibility of markets | dummy taking the value 1 if the next market cannot be reached in less than two hours | 0.07 | 0.13 | 0.07 |
| village network | migration dummy taking the value 1 if the share of migrant households within a ZD is higher than 10%, excluding the respective household itself in the share | 0.86 | 0.81 | 0.63 |

| | | | | |
|---------------------|---|-------|-------|-------|
| mean rain | mean of provincial rainfall within the last four years prior to the harvest year reported in the survey | 60.51 | 66.74 | 65.87 |
| rain variability | standard deviation of provincial rainfall within the last four years prior to the harvest year reported in the survey | 9.23 | 14.58 | 10.39 |
| rain deviation | difference between rainfall in the year of the reported harvest and mean rain relative to mean rain (mean rain again based on the four years prior to the harvest year) | 0.39 | -0.12 | -0.06 |
| Sahelian zone | dummy for households living in the Sahel region | 0.06 | 0.07 | 0.06 |
| sub-Saharan zone | dummy for households living in one of these regions: Nord, Centre-Nord, Est, Plateau Central or Centre | 0.35 | 0.39 | 0.38 |
| North-Saharan zone | dummy for households living in one of these regions: Mohon, Centre-Sud, Centre-Ouest or Centre-Est | 0.38 | 0.37 | 0.38 |
| South-Saharan zone | dummy for households living in these regions: Haut-Bassins, Sud-Ouest or Cascades | 0.21 | 0.17 | 0.18 |
| livestock ownership | dummy for households owning cattle or small stock (sheep, pigs, goats) | 0.79 | 0.92 | 0.90 |
| livestock selling | dummy for households selling cattle, sheep, pigs, goats or donkeys | 0.38 | 0.65 | 0.36 |
| non-farm | dummy for households engaged in wage employment, non-farm enterprises or other non-farm self-employment | 0.54 | 0.47 | 0.43 |
| migration | dummy for households receiving remittances from urban or rural areas from Burkina Faso or from Côte d'Ivoire | 0.28 | 0.33 | 0.14 |
| cotton farmer | dummy for farmers selling cotton | 0.11 | 0.19 | 0.24 |

Chapter 6

Explaining Labor Market Inactivity in Migrant-Sending Families: Housework, Hammock, or Higher Education?⁵⁷

Dennis Görlich

Kiel Institute for the World Economy

Toman Omar Mahmoud⁵⁸

Kiel Institute for the World Economy

Christoph Trebesch

Free University of Berlin

Hertie School of Governance

Abstract

This paper investigates the impact of emigration and remittances on labor market activity and time allocation of migrant-sending families. Departing from the common finding that labor market participation is lower in migrant households, we analyze in a unified framework whether the reasons for labor market inactivity, i.e. leisure consumption, home production and higher education are affected by migration. Based on household data from Moldova, our results challenge the assertion that those who stay behind consume more leisure. Instead, living in a migrant household implies higher probabilities of intra-household labor substitution and a substantially higher likelihood of university enrolment.

Keywords: migration, remittances, labor market participation, time allocation, home production, higher education, Moldova

JEL classification: F22, J22, O15, C35

57 We thank Awudu Abdulai, Michael Clemens, Nicola Coniglio, Vladimir Ganta, Lena Giesbert, David Grigorian, Stephen Jenkins, Namsuk Kim, Krisztina Kis-Katos, Jann Lay, Ghazala Mansuri, Michal Myck, Rainer Thiele, Arne Uhlendorff and participants of the Annual Meeting of the European Economic Association in Budapest, the Annual Meeting of the European Society for Population Economics in Chicago, the Spring Meeting of Young Economists in Hamburg, the Annual Meeting of the European Association of Labor Economists in Oslo, IZA/World Bank Conference in Rabat and research seminars in Berlin, Chisinau and Washington D.C. for helpful comments and discussions.

58 Corresponding author: toman.mahmoud@ifw-kiel.de

6.1 Introduction

Many countries have witnessed large-scale emigration over the past decades. In parallel, remittances by migrants to their families have increased dramatically and become a central component of capital flows to developing countries. Migration and remittances have a strong impact on the development of migrant-sending communities. Amongst others, the departure of migrants and the subsequent receipt of remittances have been found to reduce household poverty levels, improve child health, and spur entrepreneurship (Adams, 2005; McKenzie, 2005; Woodruff and Zenteno, 2007).

Recently, the literature has paid more attention to the effects of migration and remittances on labor markets of source countries. A question of particular interest is how the departure of a household member influences the labor market behavior of those who stay behind. There is growing evidence that individuals in migrant households supply less labor, meaning that they work fewer hours and are less likely to be active on the labor market.⁵⁹ Many studies have attributed this finding to potential disincentive effects of remittances including moral hazard problems.

The economic rationale for this interpretation is that regular transfers from abroad raise recipients' reservation wages and may also lower their work efforts, thus increasing their consumption of leisure. Supposedly, unearned income from remittances makes remaining household members "lazy" (Azam and Gubert, 2006, p. 426) so that they "simply stop working and wait from month to month for the overseas remittance" (Kapur, 2005, p. 152). If such effects were widespread, they would obviously have detrimental consequences for development. In the worst case, emigration and remittances could lead to a culture of dependency in source communities, along with a reduction of productive activities, labor shortages and other adverse economic impacts.

Building on the non-separability of migration and remittances, this paper argues that reduced labor supply in migrant households is not necessarily due to increased leisure consumption. Instead, remaining family members might respond to migration and remittances by allocating more time to home production and higher education. The departure of a migrant implies that two helping hands might be missing for household duties, child care or work on the subsistence farm. Accordingly, individuals in migrant-sending households may provide less labor on the market because they find it more rewarding to get involved in home production activities. In addition, young adults in migrant families might be more likely to engage in higher education, be it due to the flow of remittances relieving credit constraints or migration-induced incentives for additional education. Pursuing college or university studies would then reduce young adults' labor supply.

59 E.g. Funkhouser (1992), Rodriguez and Tiongson (2001), Acosta (2006), Hanson (2005), Amuedo-Dorantes and Pozo (2006), Kim (2007), Grigorian and Melkonyan (2008), Lokshin and Glinskaya (2009), Binzel and Assaad (2009), and Mendola and Carletto (2009).

We contribute to the literature in two ways. First, we use a unified framework to test whether living in a migrant household affects an individual's probability (i) to be active on the labor market, (ii) to be voluntarily inactive (*disincentive effect*), (iii) to engage in home production (*labor substitution effect*), and (iv) to pursue higher education (*education effect*). This approach differs from the existing literature, which has not systematically accounted for the actual reasons for reduced labor supply. Our analysis draws upon an exceptionally rich and large household survey from Moldova, a country which has experienced high rates of outmigration during the past ten years. Second, we use a novel instrumental variable strategy that exploits the substantial diversity of first emigrants' destination choices to address the issue of endogenous selection into migration. Our instrument is based on the degree to which initial regional migrant networks in the different destination countries differentially exposed Moldovan regions to economic performance and job opportunities abroad, thus affecting the subsequent departure of new migrants.

In line with previous studies, we find that persons living in migrant households are significantly less likely to be active on the labor market. However, we cannot attribute this finding to the disincentive effect of remittances. Rather, our results are strongly in favor of the labor substitution and education effect.

The remainder of this article is structured as follows. In section 2, we discuss the theoretical background of our analysis and review the related literature. Section 3 presents the setting and the dataset used for the analysis as well as the econometric strategy. Section 4 summarizes our empirical results and section 5 concludes.

6.2 Theoretical Considerations and Related Literature

This section first reviews the general literature on the labor market impact of emigration and remittances. We then discuss the theoretical considerations behind the disincentive, labor substitution and education effect and review the related empirical literature.

Funkhouser (1992) is one of the first to examine the labor supply effects of migration and remittances. Using simple probit regressions, he finds that the receipt of remittances is a negative determinant of labor force participation in Managua. Rodriguez and Tiongson (2001) investigate the labor market decision of migrant relatives in Manila. They find that having a migrant in the household strongly reduces remaining members' labor supply.

More recent research accounts for the potentially severe problem of endogenous selection into migration. Using an instrumental variable strategy, Acosta (2006) confirms previous findings for the case of El Salvador. Hanson (2005) employs parametric and non-parametric methods and shows that individuals in Mexico with a migrant abroad are less likely to participate in the labor force. Recent years have seen further studies on the effects of migration and remittances on labor supply (Amuedo-Dorantes and Pozo, 2006; Kim, 2007; Grigorian and Melkonyan, 2008; Lokshin and Glinskaya, 2009). However, to the best of our knowledge, the literature has not systematically examined the reasons for reduced labor market activity in migrant households in a unified

framework. Building on the non-separable impacts of migration and remittances (McKenzie and Sasin, 2007), we propose three potential explanations of labor market inactivity.

6.2.1 Disincentive Effect

The theoretical foundation for the disincentive effect follows directly from the neoclassical theory of labor supply. Under the usual assumption that leisure is a normal good, non-labor income raises the reservation wage of a potential worker. If the prevailing market wage is lower than her reservation wage, an individual is voluntarily inactive on the labor market. In our context, remittances from migrants abroad are a source of non-labor income and thus increase the reservation wage of the remaining household members. Assuming identical preferences and the same wage offer to similar persons in migrant and non-migrant households, the disincentive effect then suggests that persons in remittances-receiving households are less likely to participate in the labor market.

There is extensive evidence on the effects of non-labor income on labor supply. For developing countries, Rosenzweig (1980) and Schultz (1990) show that unearned income has a negative effect on labor supply. Bertrand et al. (2003) find that working-age adults reduce their labor supply in response to large pension transfers to elderly household members. Studying the labor market behavior of migrant-sending households, Rodriguez and Tiongson (2001), Acosta (2006) and Grigorian and Melkonyan (2008) suggest that lower labor supply in migrant households can be attributed to increased leisure consumption. Similarly, Fullenkamp et al. (2005) point to disincentive effects on work effort and even moral hazard problems as likely explanations for finding aggregate remittances flows to be negatively correlated with GDP growth.⁶⁰

6.2.2 Labor Substitution Effect

The labor substitution effect can be derived from the neoclassical model of labor supply that allows for the production and consumption of non-market household goods.⁶¹ Under the assumption that home production is subject to diminishing marginal productivity, the absence of the migrant can raise the marginal productivity of home production for those who stay behind. If an individual was engaged in home production before migration, her departure may hence induce remaining household members to reallocate their labor from market work to work at home. *Ceteris paribus*, individuals in migrant households would then allocate more time to home production than their counterparts in non-migrant households.

We are not aware of studies that explicitly estimate the effect of migration and remittances on intra-household labor substitution. Some authors, however, have noted the possibility of

⁶⁰ Azam and Gubert (2005) and Rozelle et al. (1999) take agricultural yields as dependent variable to identify potential disincentive effects of remittances.

⁶¹ Gronau (1977), Juster and Stafford (1991) or Chiappori (1997) show that it is crucial to distinguish between time used for home production and leisure time.

increased intra-household specialization of labor in migrant households (e.g. Acosta, 2006; Hanson, 2005).

6.2.3 Education Effect

The theoretical literature proposes two arguments for how migration and remittances affect educational investments in migrant-sending communities (Mountford, 1997; Stark et al., 1997; Stark and Wang, 2002). First, remittances can alleviate credit constraints, enabling receiving households to send their children to school or university. Second, the brain gain hypothesis proposes that the prospect of future migration can stimulate individuals to invest in human capital, as labor markets in destination countries typically offer higher returns to education than those at home.⁶² Given the direct access to the migrant network, the incentives for educational investments should then be stronger for members of migrant households.

Empirical studies on the impact of migration on education show rather mixed results.⁶³ Hanson and Woodruff (2003), Cox Edwards and Ureta (2003), Mansuri (2006) and Yang (2008) provide evidence for a positive effect of migration and remittances on schooling using data from Mexico, El Salvador, Pakistan and the Philippines, respectively. Batista et al. (2007) and Chand and Clemens (2008) also find a positive impact of emigration on tertiary education. By contrast, Acosta (2006) reports that remittances-receiving households in El Salvador do not invest more in children's human capital. For Mexico, McKenzie and Rapoport (forthcoming) and McKenzie (2005) illustrate that migration might even discourage educational attainment. They argue that most Mexican migrants in the U.S. work in low-skilled jobs, so that young Mexicans have little incentives to invest in higher education. Under these circumstances, returns to education are ultimately higher in Mexico.

Taken together, the disincentive, labor substitution, and education effects define the optimal time allocation strategy. Assuming that prices equal one, an individual would supply labor up to the point where the marginal rate of substitution between consumption and leisure is not only equal to the wage rate, but also to the marginal productivity of home production, as well as to the net marginal return of investing in further education. In this framework, the theoretical arguments from above suggest that, *ceteris paribus*, members of migrant households are likely to allocate less time to market work, but more time to home production, leisure, and higher education. The relative importance of the three effects, however, is difficult to establish *a priori* as it depends on local labor market characteristics, the availability of higher education and other socio-economic factors.

⁶² Under certain conditions, these additional investments in human capital can outweigh the human capital loss brought about by the departure of high-skilled migrants. In this case, emigration would leave the country of origin with higher levels of human capital, leading to a brain gain.

⁶³ Our focus in this short overview is on micro-level studies. Cross-country evidence is provided by Beine et al. (2008).

6.3. Empirical Strategy

6.3.1 Data and Setting

Our analysis draws upon data from Moldova's Labor Force Survey in 2008. The survey contains rich information on individuals' labor market activities as well as their engagement in home production including subsistence farming. It also provides detailed information on migration patterns at the individual level. The Labor Force Survey is nationally representative and with a sample of about 10,000 households the largest household survey in Moldova.

The sample is based on all individuals at working age (18-64 years) who were resident in Moldova at the time of the survey with the exception of early-age pensioners and individuals in compulsory military service. Besides, we do not include households with recently returned migrants as their time allocation might differ from both current migrant and non-migrant households. The sample also excludes households from the only two urban agglomerations in Moldova (Chisinau and Balti). Labor markets in these cities are structurally not comparable to those in the rest of the country. This leaves us with a final sample of 8,817 individuals out of which 1,601 reside in migrant households.

Moldova is a small, landlocked and densely populated country with large parts of the population living in villages or small towns. Having undergone a sharp economic decline after independence in 1991, Moldova remains the poorest nation in Europe. The economy is predominantly based on agriculture and related industries and has not experienced a promising structural change since independence.

In recent years, the country has witnessed large-scale outmigration. The regional economic crisis in 1998 is generally seen as the main trigger for mass emigration from Moldova. Starting from very low levels, the number of migrants has been rising steadily since then. Current estimates suggest that out of Moldova's population of 3.6 million at least half a million people have become migrants (Lücke et al., 2009). Poverty and access to migrant networks are the principal determinants of Moldova's massive emigration flows (Cuc et al., 2005; Görlich and Trebesch, 2008). Parallel to the growing rates of outmigration, remittances have considerably increased in recent years. According to the World Bank, the level of remittances has surpassed a share of 30 percent of GDP in 2008 (Ratha et al., 2009).

Similar to other CIS (Commonwealth of Independent States) migrant-sending countries like Tajikistan or Kyrgyzstan, the typical Moldovan migrant is male (62 percent), relatively young (35 years), comes from rural areas (74 percent) and engages in physical work. The most popular destinations are Russia (57 percent) and Italy (21 percent). Other frequent destinations include Ukraine, Israel, and Turkey. Typically, only one family member works abroad (78 percent) so that large parts of the family stay behind (Tables A6.1 and A6.3 in the appendix; Mansoor and Quillin, 2007).

6.3.2 Definition of Dependent Variables

To test whether reduced labor supply in migrant-sending households is due to the disincentive, labor substitution or education effect, we would ideally observe the time allocation of each individual. Since the Labor Force Survey does not provide information on the amount of time spent in different activities, we have to resort to binary indicators of the four outcomes of interest: labor market activity, voluntary labor market inactivity, home production, and education.

An individual is defined to be active on the labor market if she is in any form of wage employment or self-employment. Farmers are included in this definition as long as they sell at least parts of their agricultural produce on the market. This classification is summarized in the dummy variable “*market work*”.⁶⁴

To capture the disincentive effect, we construct the binary variable “*voluntarily inactive*”. This dummy variable takes the value of one if an inactive person reported voluntary inactivity as the main reason for not wanting to work.⁶⁵ In line with our theoretical considerations, we suppose that these persons’ reservation wages are higher than the wages available to them so that they opt for leisure consumption.

The labor substitution effect is proxied by the dummy variable “*home production*” indicating that an individual was active in family duties, housekeeping or subsistence work on the household farm. According to our theoretical propositions, we assume that individuals engage in home production until they have equalized their marginal productivity to the wage rate.

The education effect is reflected by the binary variable “*education*”, which simply describes whether or not an individual was following any form of higher education. Investigating the education effect for adults is relevant in this context since Moldova features relatively high rates of secondary and tertiary education. Hence, the decision of additional investments in human capital becomes interesting only after the age of 18.⁶⁶

Table 6.1 compares the activities of individuals in migrant and non-migrant households. In line with the empirical literature, living in a household with a migrant abroad is associated with considerably lower rates of labor market activity. Members of migrant families are also much more active in home production and engage more often in higher education. However, there are no indications that voluntary inactivity is significantly more common in migrant than in non-migrant households.

64 Our analysis does not consider unemployed individuals. In the absence of considerable unemployment benefits, official unemployment rates in Moldova are notoriously low. Most individuals would still engage in some kind of work, though they are often underemployed or have to work in the informal sector. In fact, the Labor Force Survey identifies only 150 unemployed persons in our sample.

65 Other answer categories included education or training, family responsibilities, disease and seasonal agricultural reasons. While there might be underreporting of voluntary inactivity, it is not obvious why reporting patterns should systematically differ between members of migrant and non-migrant households.

66 According to the Labor Force Survey, 91 percent of 15 to 17 year olds attended school in 2008.

TABLE 6.1: ACTIVITIES OF INDIVIDUALS IN NON-MIGRANT AND MIGRANT HOUSEHOLDS

| | Members of non-migrant households | Members of migrant households | Test of equality of proportions (probability of rejecting the null hypothesis of equality) |
|----------------------|-----------------------------------|-------------------------------|--|
| market work | 77.33% | 66.46% | 0.00 |
| voluntarily inactive | 3.42% | 4.00% | 0.26 |
| home production | 45.69% | 52.59% | 0.00 |
| education | 9.47% | 15.05% | 0.00 |

6.3.3 Econometric Setup

The starting point for our econometric specification is the following four-equation model:

$$marketwork_i = c_1 + a_1migranthb_i + x_{i1}'\beta_1 + \varepsilon_{1i} \quad (6.1)$$

$$voluntarilyinactive_i = c_2 + a_2migranthb_i + x_{i2}'\beta_2 + \varepsilon_{2i} \quad (6.2)$$

$$homeproduction_i = c_3 + a_3migranthb_i + x_{i3}'\beta_3 + \varepsilon_{3i} \quad (6.3)$$

$$education_i = c_4 + a_4migranthb_i + x_{i4}'\beta_4 + \varepsilon_{4i} \quad (6.4)$$

where $marketwork_i$, $voluntarilyinactive_i$, $homeproduction_i$ and $education_i$ are the observed realizations of the respective latent variables.

Our regressor of interest is the binary variable $migranthb$ denoting whether an individual lives in a migrant household. We deliberately opt for an indicator of having a migrant abroad rather than receiving remittances because we intend to capture the overall impact of migration on remaining household members, i.e. not only the monetary, but also the physical and informational impact of a person abroad.⁶⁷ The vectors x_1 , x_2 , x_3 , and x_4 include other explanatory variables which control for observable differences between migrant and non-migrant households.

As argued above, the decisions to work, to consume leisure, to engage in home production or follow education are not independent of each other. An individual allocates her time to equalize the marginal returns of each activity. As a result, the error terms ε_{1i} , ε_{2i} , ε_{3i} , and ε_{4i} are likely to be correlated and call for a simultaneous estimation approach (Maddala, 1983). Under the assumption that the error terms are drawn from a multivariate normal distribution with a mean of zero and a variance-covariance matrix Σ with values of 1 on the leading diagonal and correlations of the form $\rho_{jk}=\rho_{kj}$ as off-diagonal elements, equations 6.1-6.4 can be jointly estimated with the multivariate probit model. Since the evaluation of the resulting likelihood function involves the computation of multivariate normal distribution functions, it is based on simulated maximum likelihood using the Geweke-Hajivassiliou-Keane smooth recursive simulator (Greene, 2008; Cappellari and Jenkins, 2003 and 2006). The appropriateness of the multivariate probit model can be assessed by testing whether the pair-wise correlations between the error terms are jointly

⁶⁷ Including both $migranthb$ and a variable capturing remittance flows would ignore the non-separability of migration and remittances. According to the Labor Force Survey, 77 percent of Moldovan migrants send money to their families. As stated in McKenzie and Sasin (2007), disentangling the impact of migration and remittances is inherently difficult and remains an unresolved empirical challenge.

significantly different from zero. If this is not the case, equations 6.1-6.4 could be estimated separately with simple univariate probit models.

6.3.4 Identification

Unobserved characteristics between migrant and non-migrant households constitute a major challenge in estimating the causal impact of migration and remittances on the activity choices of those who stay behind. If such characteristics are correlated with both migration and time allocation patterns, any observed differences between members of households with and without a migrant might have already existed before the departure of the migrant. Depending on the direction of self-selection into migration, we would then over- or underestimate the true effect of *migranthb*.

Following Maddala (1983), we estimate a recursive multivariate probit model to address the potentially endogenous nature of the binary variable *migranthb*. The resulting simultaneous-equation model comprises the previous four activity equations and a fifth equation to account for selection into migration:

$$migranthb_i = c_5 + \gamma z_i + x_{i5}'\beta_5 + \varepsilon_{5i} \quad (6.5)$$

Assuming that the error terms of the reduced-form equations 6.1-6.4 and the structural equation 6.5 are multivariate normally distributed, the estimation of the full recursive system follows the same procedures as described above (Greene, 2008).

Identification of recursive multivariate probit models rests on the exclusion restriction. Like in the standard linear IV case, this restriction requires the presence of at least one variable, the instrument z , in the structural equation that is excludable from the reduced-form equations (Maddala, 1983).

Our instrumental variable strategy rests on the differential degree to which Moldova's 35 different districts have been exposed to demand for migrant labor from abroad. The central idea is that economic conditions in the foreign destinations determine job opportunities for new migrants, which already established migrants then channel to their home communities. Migrant networks have been shown to play a pivotal role in shaping migration flows by providing information on jobs abroad and lowering migration costs for subsequent migrants (Munshi, 2003; Winters et al., 2001.; Carrington et al., 1996). This observation also applies to Moldova, where local migrant networks are among the principal determinants of migration (Görlich and Trebesch, 2008). The overall strength of the signal transmitted back home ultimately depends on the size and composition of the established migrant population abroad. *Ceteris paribus*, districts with an initial migrant network at a booming destination would subsequently experience more outmigration than districts with a network of the same size at an economically depressed destination. Over time, economic conditions abroad would exogenously determine growth of a given migration network and ultimately drive current migration flows. Our identifying assumption is that previous district-level exposure to demand for migrant labor from abroad does

not affect time allocation patterns directly, but only indirectly through its impact on the current propensity of migration at the household level.⁶⁸

For each district k , we measure exposure to economic conditions abroad as the weighted sum of destination countries' GDP growth between 1999 and 2007, where the weight for each destination j is given by district k 's prevalence of migration to country j in 1999:

$$exposure_k = \sum_j (growth_{j,1999-2007} * prevalence_{kj,1999}) \quad (6.6)$$

As an alternative to GDP growth, we also consider absolute changes in the unemployment rates for the period 1999-2007. GDP and unemployment data are taken from the World Development Indicators, district migration patterns come from census data.

Table 6.2 summarizes GDP and unemployment statistics for Moldova's six most important destination countries in 1999. Most countries experienced considerable GDP growth and also saw unemployment decline between 1999 and 2007. However, there are remarkable differences between countries. While Russia and Ukraine could boost their GDP by about 75 percent, Italy's economy only expanded by 12 percent. Similarly, Turkey's unemployment rate increased by two percentage points, whereas Russia, Ukraine, but also Italy were able to reduce their unemployment rates by more than five percentage points. These figures illustrate that there has been substantial variation in economic conditions in the different destinations.

Table 6.3 shows district-specific prevalence of emigration to the same countries in 1999. Both overall emigration prevalence and the direction of emigration flows varied tremendously across districts. Although Russia was the most common destination, the share of migrants to Russia was on average lower than 50 percent and in some districts even as low as 20 percent. Moreover, emigrants from different districts often concentrated in different sets of destinations. Based on this initial stock and distribution of migrants, one can expect that different districts have experienced varying degrees of exposure to demand for migrant labor from abroad.

Figures 6.1 and 6.2 demonstrate that our instruments are indeed highly correlated with current migration prevalence. Districts that were previously exposed to higher GDP growth or larger reductions in unemployment abroad report substantially higher shares of households with a migrant abroad.

68 McKenzie and Rapoport (2007) employ a related identification strategy. However, when calculating the push factors from abroad, they only consider the single most important destination city in the US for each Mexican village.

TABLE 6.2: GDP GROWTH AND CHANGES IN UNEMPLOYMENT RATES IN SIX MAJOR DESTINATION COUNTRIES

| | Russia | Italy | Romania | Greece | Ukraine | Turkey |
|--|--------|-------|---------|--------|---------|--------|
| GDP growth 1999-2007 (percent) | 72.52 | 12.31 | 54.11 | 39.54 | 77.40 | 49.25 |
| change in unemployment rate 1999-2007 (percentage points) | -7.40 | -5.60 | -0.40 | -3.60 | -5.20 | 2.20 |

Source: World Development Indicators.

TABLE 6.3: CENSUS-BASED EMIGRATION RATES (PER 10,000 INHABITANTS) TO SIX MAJOR DESTINATION COUNTRIES IN 1999, BY DISTRICT

| District | Russia | Italy | Romania | Greece | Ukraine | Turkey | Overall | Share of migrants to Russia | Total number of migrants |
|-----------------|--------|-------|---------|--------|---------|--------|---------|-----------------------------------|-----------------------------|
| Mun. Chisinau | 7.23 | 6.01 | 3.88 | 3.90 | 1.11 | 0.77 | 32.70 | 22.11% | 2329 |
| Mun. Balti | 32.61 | 4.47 | 4.31 | 5.88 | 1.73 | 0.31 | 56.84 | 57.38% | 725 |
| Anenii Noi | 28.64 | 6.73 | 4.65 | 4.28 | 1.71 | 2.33 | 58.38 | 49.06% | 477 |
| Basarabasca | 20.02 | 14.49 | 2.76 | 13.11 | 5.52 | 2.07 | 66.26 | 30.21% | 192 |
| Briceni | 9.61 | 2.18 | 1.41 | 0.64 | 2.82 | 0.13 | 20.76 | 46.30% | 162 |
| Cahul | 27.59 | 9.98 | 9.90 | 7.55 | 2.18 | 4.19 | 74.56 | 37.01% | 889 |
| Cantemir | 20.50 | 9.00 | 6.17 | 3.50 | 0.33 | 1.00 | 47.50 | 43.16% | 285 |
| Calarasi | 11.06 | 15.72 | 4.66 | 3.20 | 0.93 | 0.40 | 47.15 | 23.45% | 354 |
| Causeni | 24.61 | 5.08 | 4.08 | 1.77 | 3.09 | 0.44 | 42.16 | 58.38% | 382 |
| Cimislia | 20.35 | 23.47 | 4.43 | 3.45 | 1.31 | 0.33 | 76.65 | 26.55% | 467 |
| Criuleni | 10.24 | 6.64 | 3.32 | 3.60 | 0.00 | 0.28 | 28.37 | 36.10% | 205 |
| Donduseni | 23.26 | 1.51 | 2.15 | 0.00 | 3.23 | 0.00 | 33.38 | 69.68% | 155 |
| Drochia | 9.76 | 1.72 | 3.90 | 1.72 | 0.34 | 0.00 | 18.26 | 53.46% | 159 |
| Dubasari | 7.06 | 1.18 | 2.65 | 2.94 | 0.88 | 0.29 | 19.40 | 36.37% | 66 |
| Edinet | 20.52 | 2.33 | 2.95 | 1.47 | 2.21 | 0.49 | 32.56 | 63.02% | 265 |
| Falesti | 23.03 | 2.21 | 3.21 | 1.99 | 1.00 | 0.33 | 32.44 | 70.99% | 293 |
| Floresti | 16.00 | 4.36 | 3.02 | 1.12 | 0.34 | 0.22 | 27.41 | 58.37% | 245 |
| Glodeni | 21.48 | 5.58 | 3.44 | 2.62 | 2.30 | 0.16 | 37.88 | 56.71% | 231 |
| Hîncești | 11.19 | 14.86 | 2.76 | 8.18 | 0.42 | 1.09 | 45.76 | 24.45% | 548 |
| Ialoveni | 13.41 | 6.14 | 3.58 | 1.54 | 0.51 | 1.02 | 33.47 | 40.06% | 327 |
| Leova | 27.42 | 7.05 | 6.86 | 2.35 | 0.78 | 1.57 | 53.47 | 51.28% | 273 |
| Nisporeni | 7.86 | 14.63 | 8.01 | 4.16 | 0.31 | 0.31 | 39.89 | 19.69% | 259 |
| Ocnita | 33.98 | 0.35 | 2.48 | 0.53 | 3.89 | 0.00 | 45.13 | 75.29% | 255 |
| Orhei | 13.50 | 6.62 | 6.88 | 4.64 | 1.29 | 0.34 | 41.63 | 32.44% | 484 |
| Rezina | 16.84 | 2.70 | 6.24 | 1.25 | 0.83 | 0.00 | 32.22 | 52.26% | 155 |
| Rîșcani | 19.01 | 1.87 | 6.19 | 1.87 | 0.14 | 0.43 | 32.11 | 59.19% | 223 |
| Sîngerei | 17.90 | 7.46 | 3.67 | 0.69 | 1.49 | 0.34 | 36.72 | 48.75% | 320 |
| Soroca | 14.00 | 3.05 | 4.63 | 4.11 | 0.95 | 0.21 | 30.43 | 46.02% | 289 |
| Straseni | 11.47 | 9.79 | 6.30 | 5.96 | 0.56 | 0.68 | 44.77 | 25.63% | 398 |
| Soldanesti | 10.42 | 5.21 | 3.08 | 0.95 | 0.24 | 0.00 | 21.79 | 47.83% | 92 |
| Stefan Voda | 8.78 | 2.98 | 3.26 | 0.71 | 2.55 | 0.14 | 21.39 | 41.06% | 151 |
| Taradia | 22.48 | 4.40 | 2.78 | 3.24 | 6.49 | 3.48 | 66.51 | 33.80% | 287 |
| Telenesti | 13.55 | 5.28 | 6.27 | 1.28 | 1.00 | 0.29 | 30.23 | 44.81% | 212 |
| Ungheni | 23.79 | 5.07 | 5.79 | 1.90 | 1.18 | 0.36 | 41.07 | 57.93% | 454 |
| U.T.A. Gagauzia | 16.06 | 0.39 | 0.77 | 0.00 | 1.03 | 9.19 | 29.49 | 54.47% | 459 |

Source: Census of the Republic of Moldova in 2004.

FIGURE 6.1: REGIONAL EXPOSURE TO PREVIOUS GDP GROWTH ABROAD AND CURRENT MIGRATION PREVALENCE

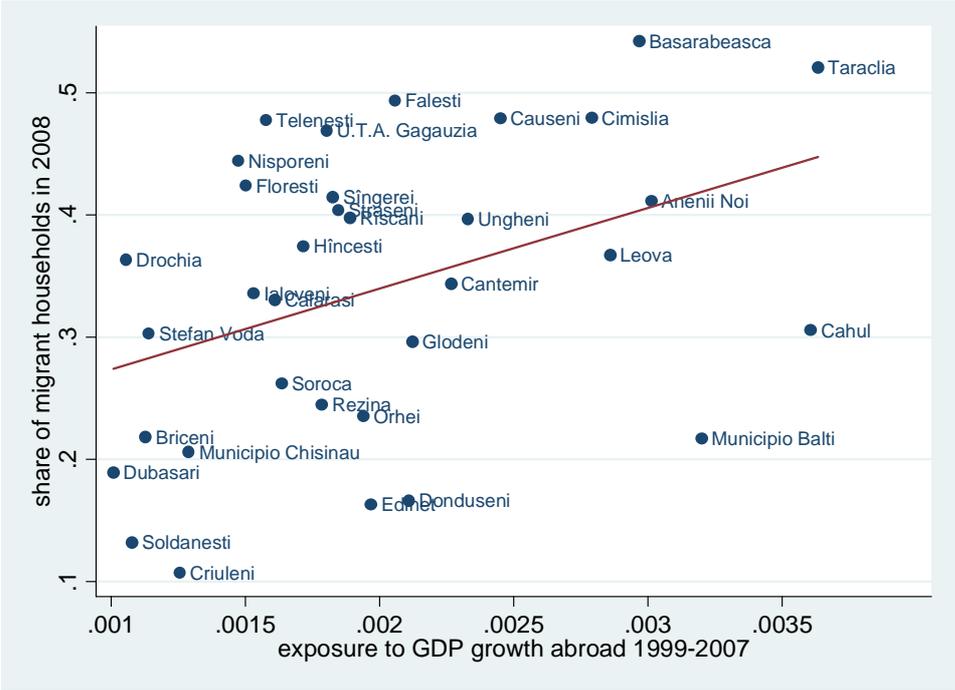
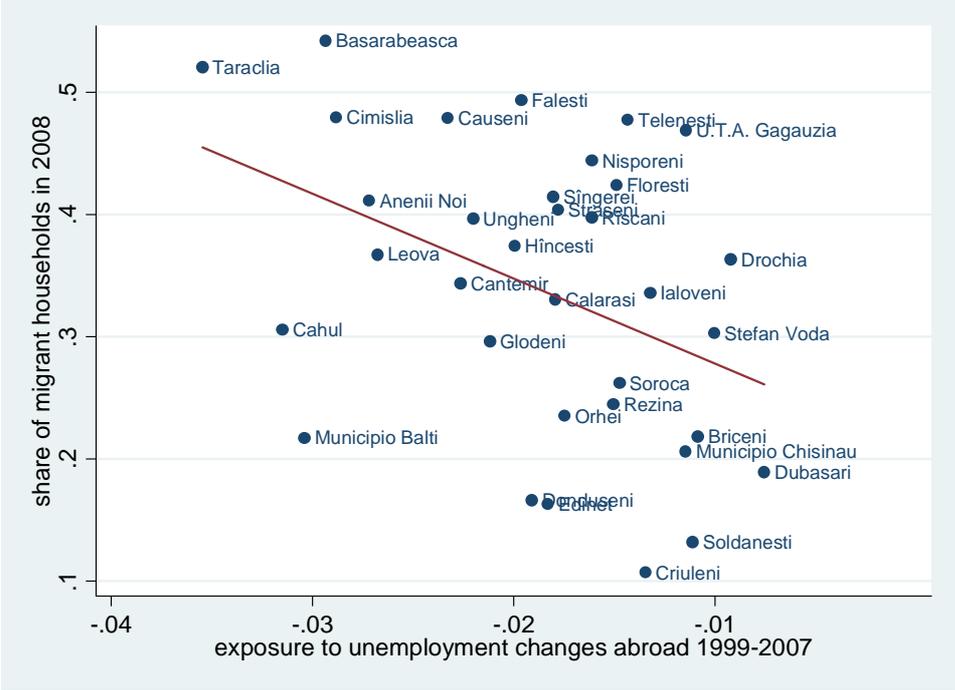


FIGURE 6.2: REGIONAL EXPOSURE TO PREVIOUS UNEMPLOYMENT CHANGES ABROAD AND CURRENT MIGRATION PREVALENCE



A potential concern for the validity of our instrument is that persistent labor market conditions at home might be a determinant of migration in both 1999 and 2008. We argue that this is unlikely to be the case as emigration was still a very minor phenomenon and just starting to take off in 1999. In total, the census registered only 13,067 emigrants. In other words, average migration prevalence was a mere 0.0036 percent in 1999 and even lower in many districts. Also taking into account that these first migrants left for so many different destinations, we make a case that emigration patterns in 1999 were not systematically related to labor market conditions at home, but rather the outcome of other factors like kinship or friendship ties. Nevertheless, to address this concern, our estimations control for a wide range of labor market characteristics at the district and locality level.

To test the robustness of our identification strategy, we also make use of an alternative approach suggested by Wilde (2000). He has theoretically shown that identification of recursive multivariate probit models is possible without the exclusion restriction as long as there is sufficient variation in the data, i.e. each equation contains at least one varying exogenous regressors.

6.3.5 Control Variables

While the identification strategy is designed to deal with unobserved heterogeneity, a number of explanatory variables at the individual, household, community and district level are used to capture observed heterogeneity between migrant and non-migrant households.

Individual characteristics comprise age, education, gender, marital status and position in the household. We control for household composition by including the number of young and older children, working-age adults and elderly members.⁶⁹ A household's income-generating possibilities are proxied by the number of members with higher education and a dummy variable indicating whether the household has a farm. In a given context, this set of individual- and household-level variables should reflect important determinants of an individual's time allocation.

As discussed above, it is essential to control for local characteristics to identify the impact of migration and remittances. These characteristics constitute potentially confounding factors that affect both the propensity of outmigration and the decision to engage in market work, home production or higher education or to be voluntarily inactive. We address this concern by drawing on unusually rich data on the local and regional structure of employment patterns and development disparities provided by the Ministry of Economy and Trade and the National Bureau of Statistics. Specifically, we include the number of non-agricultural enterprises, the distance to the nearest city, the population size and the level of local tax revenues for each locality to capture the intensity of economic activity. We also control for district-level unemployment rates and wage levels. Moreover, we add the share of ethnic Moldovans in each

⁶⁹ As the Labor Force Survey only identifies migrants who are still considered to be member of the household, we treat migrants like household members when computing household composition variables. Our decision is backed by the fact that most migrants leave their families only on a seasonal basis (Lücke et al., 2007).

district to deal with ethnically motivated migration and employment choices.⁷⁰ Finally, we also include the share of working-age individuals who were involved in the activities under consideration. These variables are thought to reflect locality-specific patterns of market and non-market work and voluntary inactivity.

6.4 Results

Table 6.4 displays the marginal effects of migration and remittances on the activity choices of those left behind. Full regression results for the base specification are provided by Table A6.4 in the appendix. The first column presents the results for our base specification, where we instrument migration at the household level with previous exposure of a household's district to GDP shocks abroad. Throughout our specifications, the instrument is highly significant at the one percent level, thus not indicating a weak instrument problem. Most of the estimated coefficients of correlation of error terms between the different activities including migration turn out to be highly significant. In other words, unobserved factors in the different equations are significantly correlated, which suggests that selection into migration is endogenous to our outcomes of interest and that it is more appropriate to estimate the four activities jointly rather than separately.⁷¹

In line with the previous literature, we find that members of migrant households have significantly lower rates of labor market activity. Living in a migrant household decreases the probability of being active on the labor market by more than 10 percentage points.

There is no evidence, however, that reduced labor supply is due to the disincentive effect. Individuals with a family member abroad are not more likely to be voluntarily inactive on the labor market than individuals in non-migrant families. The corresponding marginal effect is very small and far from conventional significance levels.

It rather seems that the labor substitution effect can account for low participation rates in market work. Being member of a migrant household significantly increases the probability of engaging in home production by about 14 percentage points. Thus, members of families with a migrant are presumably more productive in home production than their counterparts in non-migrant households. This finding supports our hypothesis that a household may reallocate labor from market work to work at home in order to take over the home production activities previously performed by the migrant.

70 Although Moldova is relatively homogenous in terms of ethnic composition, there are some regional concentrations of ethnic minorities like the Gagauz in the South-West or the Ukrainians in the North.

71 A formal test of endogeneity corresponds to testing the joint significance of ρ_{51} , ρ_{52} , ρ_{53} and ρ_{54} . The null hypothesis of joint insignificance can be rejected ($\chi^2(4)=475.46$, $\text{Prob}>\chi^2=0.000$). Likewise, pair-wise correlations between the error terms of the four activities are jointly significantly different from zero ($\chi^2(6)=449.14$, $\text{Prob}>\chi^2=0.000$).

TABLE 6.4: MARGINAL EFFECTS OF MIGRATION AND REMITTANCES ON ACTIVITIES OF INDIVIDUALS IN NON-MIGRANT AND MIGRANT HOUSEHOLDS

| | Instrument: exposure to gdp growth abroad (base specification) | Young (18-27) | Without instrument | Instrument: exposure to unemployment changes abroad | Identification through variation in the regressors |
|----------------------|---|---------------|-----------------------|--|--|
| market work | -0.104*** | -0.142* | -0.060*** | -0.103*** | -0.102*** |
| voluntarily inactive | 0.004 | -0.004 | 0.006 | 0.004 | 0.004 |
| home production | 0.139*** | -0.086 | 0.089*** | 0.139*** | 0.139*** |
| education | 0.007** | 0.171* | 0.003* | 0.007** | 0.006* |

***/**/* denote significance at the 1/5/10 percent level respectively. Marginal effects are computed holding all other variables constant at their mean and using the Stata command `mvppred`.

We also find strong evidence for the education effect. Having a family member abroad significantly raises the probability of following higher education by 0.7 percentage points. This effect may appear small, but one has to keep in mind that the sample includes all working-age adults. As enrolment in higher education is likely to be relevant for younger individuals only, we re-run the analysis and restrict the sample to individuals who are aged between 18 and 27 years (column 2).⁷² For these individuals, the education effect turns out to be very strong. Being a young member of a migrant household increases the probability of engaging in higher education by 17 percentage points.⁷³ As discussed above, this positive linkage between migration and education might be explained by the potential of remittances to relieve credit constraints or by migration-related incentives to engage in education. The non-separability of migration and remittances, however, prevents us from discriminating between these two explanations. In contrast to the results for the full sample, the labor substitution effect does not apply to this subsample of young adults. This finding suggests that the time available to young members of migrant households is rather used for higher education and consequently not available for work at home.

For comparison, the third column presents the result of an ordinary multivariate probit estimation that ignores the endogeneity of migration. The marginal effects are generally of lower magnitude, but the ordering and direction as well as significance levels are not affected. The last two columns check the robustness of our findings to the use of our alternative instrument, previous exposure to unemployment changes abroad, and to identification based on the variation in the exogenous regressors as proposed by Wilde (2000). In both cases, our previously reported results remain virtually the same.

One might argue that having a migrant member is just a proxy for having a working member in the household. As our sample is predominantly rural, a migrant and a full-time working member might have a similar impact on the time allocation of other household members. Both would be cash earners and not be able to heavily engage in home production activities. To verify that our

⁷² Our complete sample includes only 22 individuals over the age of 27 who were following higher education.

⁷³ In theory, this finding could also be due to remittances inducing students to take more time to complete their degrees. This scenario, however, seems implausible as there are no differences in the age distribution of students in migrant and non-migrant households.

findings are indeed related to living in a migrant household, we control for the number of full-time salaried or self-employed family members including migrants. The first column of Table 6.5 shows that our results are robust to the inclusion of this additional variable. Apparently, the permanent absence of a migrant working abroad causes stronger effects than a working member who remains at home.

TABLE 6.5: ROBUSTNESS CHECKS, MARGINAL EFFECTS OF MIGRATION AND REMITTANCES ON ACTIVITIES OF INDIVIDUALS IN NON-MIGRANT AND MIGRANT HOUSEHOLDS

| | Control for number of full- time employers | Broader definition of voluntary inactivity |
|----------------------|--|--|
| market work | -0.103*** | -0.090*** |
| voluntarily inactive | 0.008 | 0.004 |
| home production | 0.154*** | 0.135*** |
| education | 0.008*** | 0.007** |

***/**/* denote significance at the 1/5/10 percent level respectively. Marginal effects are computed holding all other variables constant at their mean.

Another issue of concern might be the way in which we measure the disincentive effect. So far, we have proxied the disincentive effect by voluntary inactivity. Yet, it could be that our choice is too narrow. To check this possibility we broaden the definition to also include inactive persons who were not looking for a job because they “did not know how and where to look for a job”, “did not feel professionally ready”, “thought that they would not find a job due to their age”, or “had unsuccessfully looked for a job before.” Column 2 shows that our results are not affected by this change in definition. We still do not find that reduced labor supply is attributable to the disincentive effect.

6.5 Conclusion

Living in a migrant household does not seem to be a systematic determinant of leisure-driven inactivity on the labor market. Rather, we find strong indications that migration increases the probability of being active in home production or higher education. However, our results should be interpreted with caution. We do not show that unearned income from remittances creates no disincentive effects at all. It could well be that individuals in migrant households remain active on the labor market but work less hours to consume more leisure. Moreover, we cannot observe remaining members’ productivity levels in home production and educational activities.

Still, our findings suggest that the concerns of remittances creating a culture of dependence in areas with high rates of outmigration might be overstated. In Moldova, many migrant households seem to use their additional resources to invest in the education of their young adults. Once these individuals complete their degrees, they may have higher income-generating capacities and become less dependent on future remittances flows. Nonetheless, the long-term benefits of higher rates of education for the country are uncertain as those who studied may eventually follow their family members abroad and leave the country.

Literature

- Acosta, P. (2006). Labor Supply, School Attendance, and Remittances from International Migration: The Case of El Salvador. World Bank Policy Research Working Paper 3903.
- Adams, R.H. (2005). Remittances, Poverty, and Investment in Guatemala. In Ç. Özden and M. Schiff (Eds.), *International Migration, Remittances, and the Brain Drain*. Washington D.C.: World Bank/Palgrave Macmillan.
- Amuedo-Dorantes, C., and A. Pozo (2006). Migration, Remittances and Male and Female Employment Patterns. *American Economic Review*, 96(2): 222-226.
- Azam, J.-P., and F. Gubert (2006). Migrants' Remittances and the Household in Africa: A Review of Evidence. *Journal of African Economies*, 15(2): 426-462.
- Azam J.-P., and F. Gubert (2005). Those in Kayes. The Impact of Remittances on Their Recipients in Africa. *Revue Economique*, 56(6): 1331-1358.
- Batista, C., A. Lacuesta and P. Vicente (2007). Brain Drain or Brain Gain? Micro Evidence from an African Success Story. IZA Discussion Paper 3035.
- Beine, M., F. Docquier and H. Rapoport (2008). Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers. *Economic Journal*, 118(528): 631-652.
- Bertrand, M., S. Mullainathan and D. Miller (2003). Public Policy and Extended Families: Evidence from Pensions in South Africa. *World Bank Economic Review*, 17(1): 27-50.
- Binzel, C., and R. Assaad (2009). The Impact of International Migration and Remittances on the Labor-Supply Behavior of Those Left Behind: Evidence from Egypt. DIW Discussion Paper 954.
- Cappellari, L., and S.P. Jenkins (2006). Calculation of Multivariate Normal Probabilities by Simulation, with Applications to Maximum Simulated Likelihood Estimation. *Stata Journal*, 6(2): 156-189.
- Cappellari, L., and S.P. Jenkins (2003). Multivariate Probit Regression Using Simulated Maximum Likelihood. *Stata Journal*, 3(3): 278-294.
- Carrington, W.J., E. Detragiache and T. Vishwanath (1996). Migration with Endogenous Moving Costs. *American Economic Review*, 86(4): 909-930.
- Chand, S., and M. Clemens (2008). Skilled Emigration and Skill Creation: A Quasi-Experiment. Center for Global Development Working Paper 152.
- Chiappori, P.A. (1997). Introducing Household Production in Collective Models of Labor Supply. *Journal of Political Economy*, 105(1): 191-209.
- Cox Edwards, A., and M. Ureta (2003). International Migration, Remittances and Schooling: Evidence from El Salvador. *Journal of Development Economics*, 72(2): 429-461.
- Cuc, M., E. Lundbäck and E. Ruggiero (2005). *Migration and Remittances in Moldova*. Washington, D.C.: International Monetary Fund.
- Fullenkamp, C., R. Chami and S. Jahjah (2005). Are Immigrant Remittance Flows a Source of Capital for Development? *IMF Staff Papers*, 52(1): 55-81.
- Funkhouser, E. (1992). Migration from Nicaragua: Some Recent Evidence. *World Development*, 20(8): 1209-1218.
- Görlich, D., and C. Trebesch (2008). Mass Migration and Seasonality. Evidence on Moldova's Labor Exodus. *Review of World Economics*, 144 (1): 107-133.

- Greene, W.H. (2008). *Econometric Analysis*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Grigorian, D., and T. Melkonyan (2008). Microeconomic Implications of Remittances in an Overlapping Generations Model with Altruism and Self-Interest, IMF Working Paper 08/19.
- Gronau, R. (1977). Leisure, Home Production, and Work - The Theory of the Allocation of Time Revisited. *Journal of Political Economy*, 85(6): 1099-1123.
- Hanson, G.H. (2005). Emigration, Remittances, and Labor Force Participation in Mexico. Mimeo, University of California at San Diego.
- Hanson, G.H., and C. Woodruff (2003). Emigration and Educational Attainment in Mexico. UCSD Working Paper.
- Juster, F.T., and F.P. Stafford (1991). The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement. *Journal of Economic Literature*, 29(2): 471-522.
- Kapur, D. (2005). Remittances: The New Development Mantra? In S.M. Maimbo and D. Ratha (Eds.), *Remittances: Development Impact and Future Prospects*. Washington, D.C.: The World Bank.
- Kim, N. (2007). The Impact of Remittances on Labor Supply: The Case of Jamaica. World Bank Policy Research Working Paper 4120.
- Lokshin, M., and E. Glinskaya (2009). The Effect of Male Migration on Employment Patterns of Women in Nepal. *World Bank Economic Review*, 23(3): 481-507.
- Lücke, M., T. Omar Mahmoud and P. Pinger (2007). Patterns and Trends of Migration and Remittances in Moldova. International Organization for Migration, Geneva.
- Lücke, M., T. Omar Mahmoud and A. Steinmayr (2009). Labor Migration and Remittances in Moldova: Is the Boom Over? International Organization for Migration, Geneva.
- Maddala, G.S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.
- Mansoor, A., and B. Quillin (2007). *Migration and Remittances: Eastern Europe and the former Soviet Union*. Washington D.C.: The World Bank.
- Mansuri, G. (2006). Migration, School Attainment, and Child Labor: Evidence from Rural Pakistan. World Bank Policy Research Working Paper 3945.
- McKenzie, D. (2005). Beyond Remittances: The Effects of Migration on Mexican Households. In Ç. Özden and M. Schiff (Eds.), *International Migration, Remittances, and the Brain Drain*. Washington D.C.: World Bank/Palgrave Macmillan.
- McKenzie, D., and M.J. Sasin (2007). Migration, Remittances, Poverty, and Human Capital: Conceptual and Empirical Challenges. World Bank Policy Research Working Paper 4272.
- McKenzie, D., and H. Rapoport (2007). Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico. *Journal of Development Economics*, 84(1): 1-24.
- McKenzie, D., and H. Rapoport (forthcoming). Can Migration Reduce Educational Attainment? Evidence from Mexico. *Journal of Population Economics*.
- Mendola, M., and G. Carletto (2009). International Migration and Gender Differentials in the Home Labor Market: Evidence from Albania. World Bank Policy Research Working Paper 4900.
- Mountford, A. (1997). Can a Brain Drain be Good for Growth in the Source Economy? *Journal of Development Economics*, 53: 287-303.

- Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *Quarterly Journal of Economics*, 118(2): 549-599.
- Ratha, D., S. Mohapatra and A. Silwal (2009). Migration and Remittance Trends 2009. World Bank Migration and Development Brief 11.
- Rodriguez, E.R., and E.R. Tiongson (2001). Temporary Migration Overseas and Household Labor Supply: Evidence from Urban Philippines. *International Migration Review*, 35(3): 709-725.
- Rosenzweig, M. (1980). Neoclassical Theory and the Optimizing Peasant: An Econometric Analysis of Market Family Labor Supply in a Developing Country. *Quarterly Journal of Economics*, 94(1): 31-55.
- Rozelle, S., J.E. Taylor and A. DeBrauw (1999). Migration, Remittances, and Agricultural Productivity in China. *American Economic Review*, 89(2): 287-291.
- Schultz, P. (1990). Testing the Neoclassical Model of Family Labor Supply and Fertility. *Journal of Human Resources*, 25(4): 599-634.
- Stark, O., and Y. Wang (2002). Inducing Human Capital Formation: Migration as a Substitute for Subsidies. *Journal of Public Economics*, 86(1): 29-46.
- Stark, O., C. Helmenstein and A. Prskawetz (1997). A Brain Gain with a Brain Drain. *Economics Letters*, 55(2): 227-234.
- Wilde, J. (2000). Identification of Multiple Equation Probit Models with Endogenous Dummy Regressors. *Economics Letters*, 69(3): 309-312.
- Winters, P., A. de Janvry and E. Sadoulet (2001). Family and Community Networks in Mexico-U.S. Migration. *Journal of Human Resources*, 36(1): 159-184.
- Woodruff, C., and R. Zenteno (2007). Migration Networks and Microenterprises in Mexico. *Journal of Development Economics*, 82(2): 509-528.
- Yang, D. (2008). International Migration, Remittances, and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks. *Economic Journal*, 118: 591-630.

Appendix

TABLE A6.1: CHARACTERISTICS OF MIGRANTS (N=3,639)

| | Mean | Std. Dev. | Min | Max |
|---------------------|--------|-----------|-----|-----|
| age | 34.623 | 11.017 | 15 | 64 |
| male | 0.622 | 0.485 | 0 | 1 |
| secondary education | 0.678 | 0.467 | 0 | 1 |
| higher education | 0.219 | 0.414 | 0 | 1 |
| hh head or spouse | 0.428 | 0.495 | 0 | 1 |

TABLE A6.2: CHARACTERISTICS OF NON-MIGRANT HOUSEHOLD MEMBERS (N=7,216)

| | Mean | Std. Dev. | Min | Max |
|---|----------|-----------|----------|----------|
| market work | 0.773 | 0.419 | 0 | 1 |
| voluntarily inactive | 0.034 | 0.182 | 0 | 1 |
| home production | 0.457 | 0.498 | 0 | 1 |
| education | 0.095 | 0.293 | 0 | 1 |
| age | 40.599 | 13.074 | 18 | 64 |
| male | 0.486 | 0.500 | 0 | 1 |
| secondary education | 0.648 | 0.478 | 0 | 1 |
| higher education | 0.244 | 0.430 | 0 | 1 |
| hh head or spouse | 0.723 | 0.448 | 0 | 1 |
| hh head or spouse has higher education | 0.347 | 0.476 | 0 | 1 |
| # children 0-6 | 0.248 | 0.548 | 0 | 3 |
| # children 7-17 | 0.564 | 0.830 | 0 | 7 |
| # adult members 18-64 | 2.519 | 0.962 | 1 | 6 |
| # elderly members 65- | 0.116 | 0.358 | 0 | 2 |
| # members with higher education | 0.609 | 0.883 | 0 | 4 |
| farm hh | 0.609 | 0.488 | 0 | 1 |
| # non-agricultural enterprises in locality | 87.302 | 157.696 | 3 | 800 |
| distance to nearest city (in km) | 16.983 | 14.550 | 0 | 66.111 |
| locality per-capita revenues | 0.922 | 0.290 | 0.179 | 2.174 |
| locality population size | 7108.702 | 8175.375 | 930 | 36764 |
| rural | 0.779 | 0.415 | 0 | 1 |
| locality share of employed people | 0.758 | 0.129 | 0.200 | 0.982 |
| locality share of individuals involved in home production | 0.463 | 0.280 | 0 | 0.946 |
| locality share of voluntarily inactive people | 0.035 | 0.050 | 0 | 0.259 |
| district unemployment rate | 0.009 | 0.005 | 0.004 | 0.043 |
| district wage rate | 2103.492 | 232.821 | 1689.200 | 2631.100 |
| share of ethnic Moldovans in district | 83.509 | 19.683 | 4.831 | 97.783 |
| exposure to gdp growth abroad | 0.002 | 0.001 | 0.001 | 0.004 |
| exposure to unemployment changes abroad | -0.019 | 0.006 | -0.035 | -0.008 |

TABLE A6.3: CHARACTERISTICS OF MIGRANT HOUSEHOLD MEMBERS (N=1,601)

| | Mean | Std. Dev. | Min | Max |
|---|----------|-----------|----------|----------|
| market work | 0.665 | 0.472 | 0 | 1 |
| voluntarily inactive | 0.040 | 0.196 | 0 | 1 |
| home production | 0.526 | 0.499 | 0 | 1 |
| education | 0.151 | 0.358 | 0 | 1 |
| age | 38.871 | 13.625 | 18 | 64 |
| male | 0.385 | 0.487 | 0 | 1 |
| secondary education | 0.665 | 0.472 | 0 | 1 |
| higher education | 0.194 | 0.396 | 0 | 1 |
| hh head or spouse | 0.697 | 0.460 | 0 | 1 |
| hh head or spouse has higher education | 0.297 | 0.457 | 0 | 1 |
| # children 0-6 | 0.249 | 0.519 | 0 | 3 |
| # children 7-17 | 0.668 | 0.826 | 0 | 5 |
| # adult members 18-64 | 3.245 | 1.029 | 2 | 7 |
| # elderly members 65- | 0.086 | 0.297 | 0 | 2 |
| # members with higher education | 0.586 | 0.854 | 0 | 4 |
| farm hh | 0.637 | 0.481 | 0 | 1 |
| # non-agricultural enterprises in locality | 58.375 | 119.885 | 3 | 800 |
| distance to nearest city (in km) | 17.534 | 12.707 | 0 | 66.111 |
| locality per-capita revenues | 0.877 | 0.274 | 0.179 | 2.174 |
| locality population size | 5486.258 | 5990.443 | 930 | 36764 |
| rural | 0.850 | 0.357 | 0 | 1 |
| locality share of employed people | 0.731 | 0.139 | 0.200 | 0.982 |
| locality share of individuals involved in home production | 0.498 | 0.262 | 0 | 0.946 |
| locality share of voluntarily inactive people | 0.036 | 0.049 | 0 | 0.259 |
| district unemployment rate | 0.009 | 0.004 | 0.004 | 0.043 |
| district wage rate | 2083.893 | 249.107 | 1689.200 | 2631.100 |
| share of ethnic Moldovans in district | 82.066 | 23.301 | 4.831 | 97.783 |

TABLE A6.4: MULTIVARIATE PROBIT MODEL OF INDIVIDUAL ACTIVITIES

| | (1) Market work | | (2) Voluntarily inactive | | (3) Home production | | (4) Education | | (5) Migrant hh | |
|---|-----------------------|------------|--------------------------|------------|------------------------|------------|-------------------------|------------|--------------------------|------------|
| | coef/se | marg. eff. | coef/se | marg. eff. | coef/se | marg. eff. | coef/se | marg. eff. | coef/se | marg. eff. |
| migrant hh | -0.359*** (0.093) | -0.104 | 0.087 (0.147) | 0.004 | 0.351*** (0.085) | 0.139 | 0.272** (0.138) | 0.007 | | |
| age | 0.205*** (0.011) | 0.049 | 0.007 (0.019) | 3.3E-04 | 0.071*** (0.011) | 0.028 | -0.362*** (0.019) | -0.005 | | |
| age squared | -0.002*** (0.000) | 0.138 | -2.4E-04 (2.3E-04) | 0.004 | -0.001*** (1.3E-04) | 0.113 | 0.004*** (2.5E-04) | -0.007 | | |
| male | 0.332*** (0.036) | 0.086 | 0.042 (0.056) | 0.002 | -0.329*** (0.033) | -0.129 | -0.414*** (0.058) | -0.008 | | |
| secondary education | -0.045 (0.035) | -0.013 | 0.003 (0.060) | 1.1E-04 | -0.049 (0.035) | -0.019 | | | | |
| higher education | 0.619*** (0.061) | 0.134 | -0.062 (0.101) | -0.003 | -0.097* (0.059) | -0.038 | | | | |
| hh head or spouse | 0.101 (0.065) | 0.027 | 0.062 (0.111) | 0.003 | 0.403*** (0.066) | 0.156 | -0.408*** (0.129) | -0.011 | | |
| # children 0-6 | -0.102*** (0.031) | -0.028 | -0.066 (0.062) | -0.003 | 0.376*** (0.034) | 0.149 | -0.735*** (0.074) | -0.007 | -0.040 (0.032) | -0.009 |
| # children 7-17 | 0.059** (0.024) | 0.015 | -0.015 (0.041) | -0.001 | -0.008 (0.021) | -0.003 | 0.069* (0.039) | 0.002 | 0.148*** (0.019) | 0.038 |
| # adult members 18-64 | 0.096*** (0.023) | 0.024 | -0.013 (0.036) | -0.001 | -0.145*** (0.021) | -0.057 | 0.040 (0.043) | 0.001 | 0.435*** (0.016) | 0.126 |
| # elderly members 65- | 0.156*** (0.058) | 0.038 | -0.073 (0.110) | -0.003 | 0.022 (0.056) | 0.009 | -0.044 (0.081) | -0.001 | -0.006 (0.051) | -0.002 |
| # members with higher education | -0.197*** (0.027) | -0.056 | -0.008 (0.049) | -3.4E-04 | 0.004 (0.029) | 0.002 | 0.097** (0.049) | 0.002 | -0.075*** (0.019) | -0.017 |
| farm hh | -0.340*** (0.042) | -0.086 | 0.627*** (0.076) | 0.025 | 0.517*** (0.038) | 0.200 | 0.063 (0.066) | 0.001 | -0.209*** (0.039) | -0.051 |
| # non-agricultural enterprises in locality | 7.4E-06 (2.6E-04) | 1.9E-06 | 5.7E-05 (3.9E-04) | 2.5E-06 | -2.0E-04 (3.0E-04) | -8.1E-05 | -1.1E-04 (4.0E-04) | -2.2E-06 | 0.001** (3.0E-04) | 1.4E-04 |
| distance to nearest city (in km) | 0.002 (0.001) | 4.1E-04 | -0.005*** (0.002) | -2.4E-04 | -0.002 (0.001) | -0.001 | -0.005** (0.002) | -9.8E-05 | -0.001 (0.001) | -1.5E-04 |
| locality per-capita revenues | 0.041 (0.066) | 0.003 | -0.010 (0.106) | -1.2E-04 | 0.054 (0.058) | 0.006 | -0.234** (0.105) | -0.001 | -0.312*** (0.069) | -0.021 |
| locality population size | -8.6E-07 (5.2E-06) | -0.002 | 7.7E-06 (7.8E-06) | 0.003 | 4.9E-07 (5.7E-06) | 0.002 | 3.4E-06 (8.0E-06) | 0.001 | -2.9E-05*** (5.8E-06) | -0.049 |
| rural | 0.240*** (0.076) | 0.067 | -0.206** (0.102) | -0.010 | -0.259*** (0.075) | -0.103 | -0.210* (0.120) | -0.005 | -0.009 (0.078) | -0.002 |
| district unemployment rate | -1.249 (4.287) | -0.003 | -5.333 (6.488) | -0.002 | -3.369 (3.679) | -0.013 | 4.850 (6.586) | 0.001 | 7.098** (3.445) | 0.018 |
| district wage rate | -3.2E-05 (7.4E-05) | -0.002 | 1.2E-04 (1.3E-04) | 0.001 | 1.6E-04** (6.9E-05) | 0.015 | 3.9E-04*** (1.3E-04) | 0.002 | -3.0E-04*** (7.5E-05) | -0.016 |
| share of ethnic Moldovans in district | -0.022 (0.001) | -5.8E-05 | -0.064 (0.001) | -2.8E-05 | -0.154* (0.001) | -0.001 | -0.189 (0.001) | -3.8E-05 | -0.072 (0.001) | -1.7E-04 |
| locality share of employed people | 3.601*** (0.142) | 0.009 | | | | | | | | |
| locality share of voluntarily inactive people | | | 8.841*** (0.393) | 0.004 | | | | | | |
| locality share of individuals involved in home production | | | | | 3.366*** (0.078) | 0.013 | | | | |
| hh head or spouse has higher education | | | | | | | 0.374*** (0.101) | 0.009 | | |
| exposure to gdp growth abroad | | | | | | | | | 79.883*** (27.655) | 0.013 |
| constant | -6.661*** (0.289) | | -2.608*** (0.497) | | -3.431*** (0.262) | | 5.572*** (0.474) | | -1.142*** (0.201) | |

rho21=0.084*; rho31=-0.075*; rho41=0.010; rho51=-0.103; rho32=-0.253***;
rho42=-0.713***; rho52=-0.895***; rho43=0.200***; rho53=-0.130***; rho54=0.219***
log pseudolikelihood = -12647.249
likelihood ratio test of rho21=rho31=rho41=rho51=rho32=rho42=rho52=rho43=rho53=rho54=0
chi2(10)=1189.04; prob>chi2=0.0000
observations 8,817

***/**/* denote significance at the 1/5/10 percent level respectively (robust standard errors). Marginal effects are computed holding all other variables constant at their mean. They reflect the change in the predicted probability for a one-unit change in the continuous regressor under consideration. In case of dummy variables, the marginal effect reflects the discrete change from 0 to 1. For the variables “locality per-capita revenues”, “locality population size” and “district wage rate” marginal effects are calculated for a change by one standard deviation.

Chapter 7

The Economics of Human Trafficking and Labor Migration: Micro-Evidence from Eastern Europe⁷⁴

published 2010 in *Journal of Comparative Economics*, 38(2): 173-188

Toman Omar Mahmoud⁷⁵
Kiel Institute for the World Economy

Christoph Trebesch
Free University of Berlin
Hertie School of Governance

Abstract

Human trafficking is a humanitarian problem of global scale, but quantitative research on the issue barely exists. This paper is the first attempt to analyze the economics of human trafficking and labor migration based on micro data, using unique household surveys from Belarus, Bulgaria, Moldova, Romania, and Ukraine. We find that individual trafficking risks are much higher in regions with large emigration flows. The reasons are lower recruitment costs for traffickers in emigration areas and, to a lesser extent, more negative self-selection into migration. Our results also indicate that illegal migration increases trafficking risks and that better information, e.g. through awareness campaigns, might be an effective strategy to reduce the crime. These findings may help policymakers to better target anti-trafficking efforts.

Keywords: human trafficking, migrant exploitation, illegal migration, migrant networks, Eastern Europe

JEL classification: F22, J61, K42, O17

⁷⁴ We are highly indebted to George Gigauri from IOM for his support. We thank the Editor and an anonymous referee for providing comments that helped to considerably improve this paper. We also thank Awudu Abdulai, Satish Chand, Michael Clemens, Alexander Danzer, Dennis Görlich, Johannes Koettl, Ghazala Mansuri, Andreas Steinmayr, Rainer Thiele and participants of the IZA/World Bank Conference in Rabat, the Annual Meeting of the European Society for Population Economics in Seville and the Annual Meeting of the European Economic Association in Barcelona for helpful comments and discussions.

⁷⁵ Corresponding author: toman.mahmoud@ifw-kiel.de

“Like many things that should have been stamped out a long time ago, slavery, it seems, is alive and well.”

The Economist, 9 March 2005

“The trade in people is surely the most morally repugnant of all the illicit trades that flourish today. But it’s deeply entrenched and interwoven with the world’s ever more complex migration flows.”

Moises Naim in “Illicit” (2005, p. 89)

7.1 Introduction

The exploitation and trafficking of men, women and children is a humanitarian problem of global scale. A recent study estimates that at least 12 million people worldwide are trapped in conditions of forced labor. Around a fifth of these are being exploited as a result of human trafficking (ILO, 2005). These forms of modern day slavery have become one of the most profitable businesses in the world – and one of the most horrifying. Human trafficking is said to be the fastest growing source of income for organized crime and its third most important, exceeded only by drugs and arms trade (Obuah, 2006).

After decades of neglect, public attention on the problem has increased considerably in recent years. Policymakers in international organizations and national governments all over the world as well as NGOs have augmented their efforts to combat trafficking and to raise public awareness on the issue.⁷⁶ There has also been a notable increase in donor funding, e.g. by the European Union and the US government.

Despite the policy relevance, academic research on the topic remains scarce, especially when it comes to economics. The large economic literature on organized crime and illicit activities (e.g. Fiorentini and Peltzman, 1995) has largely ignored the sizeable market of human trafficking and migrant exploitation. The same is true for the rapidly growing field on the economics of migration, which has been notably silent on the issue. As a result, and despite frequent calls for empirical research on the topic,⁷⁷ there is still very little knowledge on human trafficking as an economic phenomenon. One major reason for this is the grave difficulty in gathering reliable and representative data.

⁷⁶ Most prominently, the UN issued the Protocol to Prevent, Suppress and Punish Trafficking in Persons (UN, 2000a) and the Protocol Against the Smuggling of Migrants (UN, 2000b), which entered into force in 2004 and 2003 respectively. In March 2007, the UN also formally launched a Global Initiative to Fight Human Trafficking. The work of counter-trafficking NGOs such as Free the Slaves or La Strada and pioneering books such as Bales (1999) have helped to increase awareness of the problem.

⁷⁷ Gozdziaik and Collett (2005, p.122) conclude that “future research needs to move beyond stating that there is a problem, to more systematic and rigorous data collection and analysis.” Among others, similar calls have been made by Tyldum and Brunovskis (2005), Piper (2005) and Chand (2008).

This paper is the first to analyze the drivers of human trafficking with micro data. Our theoretical argument is based on simple supply and demand considerations and rooted in the literature on the economics of crime and migration. In a nutshell, we argue that trafficking and exploitation are the sad but obvious consequence of migration pressure in a world of closed borders. The growing, but unsatisfied demand for legal migration options has created a breeding ground for criminal organizations and exploitative employers, who have learned to make a profit from people's desire to work abroad. Migration has become a multibillion-dollar industry that can be particularly lucrative for those willing to use physical violence and restraint.

Our focus on the close link between trafficking and labor migration differs from that of many recent media and policy reports.⁷⁸ Most reports point to legislation and law enforcement, poor border controls, bribery and corruption, or insufficient education as drivers of human trafficking. These factors certainly exacerbate the trafficking problem, but may not explain it at its core. We argue instead that it is first and foremost the wish for a better life abroad that puts millions of people at the risk of ending up in exploitative work conditions. Their willingness to depart and to take risks in the migration process can be easily exploited by criminal agents.

In line with the above, we expect the risk of human trafficking to be highest in areas with high rates of emigration. Simply put, "traffickers fish in the stream of migration" (Coomaraswamy, 2001, p. 3), so that more people leaving means more people at risk. In addition, two channels may explain why individual trafficking risks increase with regional emigration rates.

The first channel relates to the demand side of the human trafficking market (criminal agents recruiting for employers abroad). We predict that recruitment costs for traffickers will be lower in areas of large-scale emigration. One reason is that it is easy to find potential victims, simply because so many people leave from there. The other reason is the agglomeration of shadow migration industries. Large migrant flows attract a variety of middlemen offering services such as false documents, smuggling across borders or work procurement abroad. The presence of these agents offers ample opportunities for traffickers as they can collaborate or free-ride on their reputation (Tamura, forthcoming).

The second channel is linked to the supply side of the human trafficking market (vulnerable individuals willing to work abroad). Migrants departing from high migration areas are likely to be negatively self-selected and share characteristics that make them more vulnerable to trafficking. High-emigration areas are often disadvantaged in the first place, so that migrants from there may be more willing to take risks, and possibly even consent to hazardous working conditions abroad. It has also been shown that large-scale emigration can lower the costs of migration, thus fostering the migration of poorer, less educated workers (McKenzie and Rapoport, forthcoming; Beine et al., forthcoming). In sum, we predict trafficking risks to be higher in high-migration areas (i) because traffickers benefit from lower recruitment costs and free-riding opportunities there and/or (ii) because those departing from high-migration areas can be more easily deceived.

78 Compare ILO (2005), UNODC (2006), US State Department (2008).

To test our arguments, we use a novel and unique survey on human trafficking which covers 5513 randomly selected households from 82 regions in Belarus, Bulgaria, Moldova, Romania, and Ukraine. These Eastern European countries are among the most important source countries of human trafficking worldwide (ILO, 2005). Human trafficking is defined as a situation in which an individual travelling abroad was locked and forced to work for no or little pay via means of coercion. As such, trafficking differs importantly from migrant smuggling, a commercial service that normally occurs with the consent of migrants, and from illegal migration, which does not typically involve any forms of exploitation. About 7 percent of the migrant families (108 out of 1563 households) identified by the survey reported to have a trafficking victim among their members. This includes victims of both sexual and non-sexual forms of human trafficking, but not trafficked children.

To supplement our findings, we present an additional case study of Moldova, which has by far the highest prevalence of human trafficking among the five countries under study. The case study is based on a second dataset covering 1679 migrants from 211 localities in Moldova. Although this second survey does not explicitly focus on trafficking, it allows identifying worst forms of migrant exploitation and contains ample additional information on migrants' characteristics including their destination countries.

The key result is in line with our expectations. Individual trafficking risks increase with the size of migration flows from the region (five-country survey) or locality (Moldova case study). Migration prevalence and other proxies of migration pressure turn out to be key predictors of trafficking on the household level. Thus, we find that trafficking tends to occur where migration flows are largest, not in remote regions with no or little migration. This result holds when controlling for selection effects and a broad range of household and migrant attributes as well as characteristics of the locality, region and country. As to the potential channels explaining this finding, we find stronger evidence for demand-side than for supply-side factors. Beyond this, we can report two further main results. First, illegal migration patterns increase trafficking risks. Second, we present evidence that awareness campaigns could play a successful role in reducing trafficking risks. The incidence of trafficking is considerably lower in regions where the awareness of the phenomenon of human trafficking is high.

The rest of the paper is structured as follows: Section 2 defines the main concepts and summarizes the current state of knowledge on human trafficking. It also provides stylized facts on the patterns of migration and trafficking in Eastern Europe. Section 3 sketches our main argument. In section 4, we describe the data and our empirical strategy. Section 5 discusses the results and presents a number of robustness checks. Section 6 presents the case study on Moldova. Section 7 concludes.

7.2 What Do We Know About Human Trafficking?

7.2.1 Definitions

It is crucial to clearly define the concepts that lie at the core of this paper. We follow the UN protocols against human trafficking, which, after years of debate, succeeded in clearly defining and distinguishing these concepts.⁷⁹ In accordance with the UN definition, the main purpose of *smuggling* is to facilitate the illegal entry of a person into another country. Typically, migrants are smuggled with their consent, paying the smuggler for his service. The smuggler gains a material benefit, but does not necessarily exploit the migrant. In contrast, *trafficking* always involves the violation of human rights and severe forms of exploitation. Victims of trafficking are recruited, transported and forced to work by means of coercion. The main purpose of trafficking is *exploitation*, either sexual exploitation or other forms of economic exploitation. While these definitions are helpful from a legal perspective, trafficking, exploitation and smuggling are often interrelated.⁸⁰ Some migrants using smuggling services may depart voluntarily, but end up being exploited. In such cases, smugglers become traffickers.

7.2.2 Determinants of Human Trafficking

A fundamental breeding ground for trafficking and exploitation is the economic situation of people in poorer regions of the world, pushing vulnerable people to emigrate and seek better opportunities abroad (Bruckert and Parent, 2002; Chand, 2008). The large potential gains from migration⁸¹, joint with network and herd effects have generated an unprecedented demand for legal migration to richer countries (World Bank, 2005). However, with most middle and high-income country labor markets shut off, there is only very limited supply of legal employment opportunities abroad. At the same time, there is a constant demand for cheap manual workers and prostitutes in both developed and developing countries.

This situation has fostered the emergence of shadow migration industries offering services such as border crossings and the procurement of illegal work abroad (Orrenius, 1999; Stalker, 2000; Gathmann, 2008). These activities, including its worst forms such as human trafficking, offer large monetary rewards and have to be seen in parallel to other fields of criminal intermediation, e.g. drugs and arms trade. Aronowitz (2001), Salt and Stein (1997) and Schloenhardt (1999) provide a detailed analysis of the business of trafficking and exploitation. In essence, the organizational structure of trafficking networks is similar to that of providers of legal services. Trafficking organizations find and attract people willing to work abroad via advertisements in

79 The exact definitions are given in UN (2000a) and UN (2000b).

80 In this paper, we use the terms 'trafficking' and 'exploitation' interchangeably.

81 For an individual, emigration can be the single most effective strategy to improve economic opportunities and increase income, often by a multiple. The large "place premium" in wages between sending and receiving countries is carefully documented in a recent paper by Clemens et al. (2008).

newspapers, the internet or through employment agencies. They also contact people on an individual basis, often via dispersed recruiters who may have even been trafficking victims themselves.

All available evidence indicates that human trafficking is an exceptionally lucrative business for criminal groups. In a recent study, the ILO (2005) estimates that sexual and labor exploitation yields US\$ 32 billion of profits a year to the actors involved. This corresponds to an estimated US\$ 13,000 of yearly profits for each forced laborer. Another estimated figure is the profit of criminal gangs from sex trafficking alone, which is conservatively estimated at US\$ 5 to 7 billion a year, with Interpol giving a higher estimate of US\$ 19 billion annually (ILO, 2005; US State Department, 2008).

The business can be highly attractive for criminal actors. Start-up costs are small and as Väyrynen (2005) harshly puts it, “people are a good commodity as they do not easily perish, but they can be transported over long distances and can be re-used and re-sold”. Moreover, the risks of detection, prosecution or arrest are much lower compared to other fields of illegal activity such as drug or arms trade. Exploited victims often feel discouraged to denounce their traffickers, as they face the risk of deportation and other legal consequences when contacting authorities in receiving countries (Langberg, 2005; Surtees, 2005). And even if traffickers are arrested, penalties are relatively low compared to other illicit activities (LexisNexis, 2008).

In economic research, two recent theoretical papers have shed some light on the link between migration and its dark side in the form of trafficking or forced labor.⁸² Tamura (forthcoming) focuses on the interaction of human smuggling and trafficking. In his model, migrants hire a smuggler to cross borders and find work abroad. Once migrants depart, it depends on the smuggler’s decision and the profitability of exploitation whether they end up being trafficked or not. In a different setup, Friebel and Guriev (2006) model the market of illegal migration with debt/labor contracts. As most migrants cannot pay for migration costs in advance, criminal intermediaries and smugglers offer loans to potential migrants, which they have to pay back in the destination country. The contracts between migrants and intermediaries can only be enforced in the illegal sector, not in the legal one. Although the enforcement of such contracts can take place through coercion and punishment, the possible risk of exploitation and trafficking by the intermediaries is ruled out.

While theory is scarce, empirical evidence is even scarcer. There is very little systematic knowledge about which households and regions are most vulnerable to exploitation and human trafficking. We are aware of only one econometric study on human trafficking.⁸³ Akee et al.

82 In a further paper, Dessy and Pallage (2006) focus on the tradeoff between trafficking and child labor. In their model, the risk of trafficking acts as a deterrent to sending children to work. They come to the conclusion that successful policy measures against trafficking may lead to an increase in child labor, because parents will be less afraid of letting their children work.

83 Basu and Chau (2003) provide cross-country evidence on child labor in debt bondage, but do not place their analysis in a migration or trafficking context.

(2007) analyze the determinants of trafficking of children and women in a cross-section of countries. A main result is that the stage of economic development and inequality of incomes between countries are major determinants of trafficking. Contrary to expectations, they find that granting legal status to trafficked victims in host countries and banning prostitution in source countries leads to a higher, not lower, likelihood of trafficking. Despite this pioneering study, much remains to be done to uncover the drivers and mechanisms of human trafficking, particularly on the micro level.

7.2.3 Human Trafficking and Labor Migration in Eastern Europe

There is a small but growing body of literature on the general migration patterns in Eastern Europe.⁸⁴ The countries under study have witnessed large and often seasonal migration flows since the breakdown of the Soviet Union. The main drivers of migration in the area are poor living conditions and a lack of jobs as well as migration networks. There is little evidence for positive self-selection, as most migrants work in low-skilled jobs in sectors such as construction, agriculture or in domestic services, often under dismal conditions. As to destinations, Mansoor and Quillin (2006) estimate that from 1990 to 2006 about 80 percent of emigrants from the former Soviet Union moved within the CIS region, predominantly towards Russia, which is now the second largest recipient of international migrants worldwide. Other main destinations, particular for migrants from Bulgaria and Romania, are countries in the European Union.

Countries of the former Eastern bloc have also become major origins of forced labor and human trafficking in the last two decades. The exploitation of human beings has a long and sad history in the region. In medieval times, Venetian and Genoese merchants established a massive trade with slaves from Eastern Europe and the epoch of “second serfdom” implied highly repressive forms of exploitation (Blum, 1957; Domar 1970). Today, the main destinations of trafficking victims are Western Europe and Russia, but they also include North America, Asia and increasingly the Middle East, in particular Turkey, Israel, Lebanon and the United Arab Emirates (Surtees, 2005; UNODC, 2006). Estimates of the extent of trafficking and forced labor in Eastern Europe vary widely. The report by the ILO (2005) estimates that a minimum of 200,000 people from Eastern Europe and Central Asia have fallen victim to trafficking, while Mansoor and Quillin (2006) cite estimates from the United Nations Population Fund that more than 175,000 persons are trafficked *annually* in Europe and Central Asia.

Media coverage and reports such as UNODC (2006) suggest that Eastern European trafficking victims are mostly young women who are exploited in the commercial sex industry. It is certain that many victims from the region are females forced into prostitution. Yet, the IOM (2006) report and recent more qualitative surveys from Russia and South-eastern Europe (Tyuryukanova, 2005; Surtees, 2005) underline the dominant but neglected issue of male

84 The following stylized facts are based on Bauer and Zimmermann (1999), Kraler and Iglicka (2002), Mansoor and Quillin (2006) and Görlich and Trebesch (2008).

exploitation.⁸⁵ According to these sources, a large share of those trafficked from Central and Eastern Europe are men being exploited in agriculture, construction, or warehouse work.

A further key insight from recent studies in the region is that most of the victims of trafficking depart on a voluntary basis (ILO, 2005; Tyuryukanova 2005). They often enter the destination countries at legal border crossings and with legal documents (Surtees, 2005; ILO, 2005). Typically, victims become subject to coercion and violence only after arrival at the destination, where their life in the shadows and a new environment makes them prone to be lured into seemingly attractive jobs by traffickers. The role of middlemen in the Eastern European trafficking business has also been documented in a small number of studies. According to the ILO (2005), trustworthy social networks can lower the risks of being trafficked, while the use of middlemen increases it. However, it is important to underline that social networks do not assure safety from trafficking. The IOM study by Surtees (2005) illustrates that many victims of sexual exploitation in countries like Albania, Bulgaria, or Macedonia were recruited by friends. Similarly, Laczko and Gramegna (2003) report that in South-eastern Europe 60 percent of victims are recruited by acquaintances, mostly with the promise to find them a job abroad.

In sum, these stylized facts show that the human trafficking flows from the region run in parallel to the general migration flows in Eastern Europe and trafficking affects men and women alike. It also appears that many trafficked victims leave voluntarily, but ill-prepared.

7.3 Human Trafficking as a Side Effect of International Labor Flows

This section discusses our theoretical argument and the related channels at work. In line with the above, we argue that human trafficking is an unavoidable side effect of migration in a world with large income disparities but closed borders. We therefore predict that large migrant outflows will increase the incidence of trafficking in a region. Simply put, more people departing means more people at risk. Assuming a fixed (low) probability that a departing migrant ends up being trafficked, it is straightforward to expect a larger number of trafficking victims to come from regions with more emigration.

Our argument, however, goes beyond this purely mechanical effect. We predict that *individual* trafficking risks increase with the scale of regional emigration. Two channels can explain this. The first channel is linked to the demand side of the human trafficking market, as mass emigration can lower recruitment costs of traffickers. The second channel relates to the supply side of the trafficking market, as those departing from high-emigration regions may share characteristics that make them more prone to be trafficked.

On the demand side, large-scale emigration may create agglomeration and free-riding opportunities for traffickers. Just like legal businesses, criminal activities tend to agglomerate in

⁸⁵ Piper (2005) highlights that there is a particular need for data and empirical research on non-sexual forms of trafficking.

larger markets. This is true for drug gang concentration in cities (Glaeser and Sacerdote, 1999) and applies to many other crime industries that concentrate geographically (Freeman et al., 1996; Zenou, 2003). In the context of trafficking, markets will be larger and more attractive in areas of mass emigration which offer high economies of scale. Most importantly, these regions feature a continuous supply of potential victims, which can considerably ease recruitment. Additionally, traffickers can benefit from the existence of a local shadow migrant industry, in which private agents offer all kinds of work procurement or travel services to emigrants. Related industries have thrived in Northern Africa or at the US-Mexican border, but also in many emigration regions of Eastern Europe. In such settings, traffickers can take advantage of agglomeration effects by collaborating with, or free-riding on, other agents such as people smugglers or travel agencies (Tamura, forthcoming).⁸⁶ The result will be a higher concentration of traffickers and, in turn, higher trafficking risks in such regions.

On the supply side, it is mostly negative self-selection that could explain higher risks in emigration areas. Localities witnessing large-scale out-migration may be particularly disadvantaged in the first place, often suffering from poor economic development and weak institutions. In deprived areas, migration may be the only strategy available to increase economic opportunities and start a better life. This will have implications for the preference set of those departing, as they may be more prone to take risks in the migration process, such as crossing borders illegally or accepting jobs in the informal sector abroad. In extreme cases, migrants from particularly disadvantaged regions may even consent to exploitative working conditions. They may be keen to access so-called “3D” jobs (difficult, dirty, dangerous) because they would do nearly everything just to escape their dismal living conditions at home (Bhabha, 2005). Besides such risk-taking behavior, skills and education may also play a role. The issue of negative or positive self-selection into migration is highly controversial.⁸⁷ However, recent studies agree in that larger migration flows can foster negative self-selection (Beine et al., forthcoming; Chiquiar and Hanson, 2005; McKenzie and Rapoport, forthcoming). Larger networks and an established migration infrastructure lower the costs of migration, thereby increasing the likelihood that more people from lower social classes start to emigrate. These less-qualified late-comers in the migration wave may be more vulnerable to deception schemes and less able or willing to thoroughly prepare their journey. On the supply side we therefore predict trafficking risks to be higher in high-migration areas because those departing from there may be more desperate or simply less qualified to handle the many perils of the migration process.

86 Recent research indicates that the markets of migrant smuggling, work procurement and human trafficking/exploitation are indeed closely intertwined (Leman and Janssens, 2007; Väyrynen, 2005; Aronowitz, 2001; Ruggiero, 1997).

87 Recent years have seen evidence for positive, intermediate and negative self-selection into migration, depending on the data and country investigated. Chiquiar and Hanson (2005), Mora and Taylor (2005), or Adams (2005) find years of schooling and other human capital variables to have a significant *positive* impact on the decision to migrate. Orrenius and Zavodny (2005) find evidence for *intermediate* selection in Mexican migration to the U.S., while Borjas (1987) predicts *negative* self-selection from countries with high earnings inequality such as Mexico.

Stretching even further, one could imagine the first and second channel to mutually reinforce each other. Negative self-selection may be beneficial for traffickers, and pose an additional incentive to agglomerate in regions of mass emigration and large networks. The reason is that negative self-selection may have relevant feedback effects on network composition or network quality. The more low-skilled workers depart, the worse the migrant network might become with regard to information filtering and transmission and in safeguarding from exploitation. As a result, larger, and thereby less-skilled, regional migrant networks might facilitate recruitment.

In conclusion, both supply and demand-side channels predict individual trafficking risks to be higher in areas of large-scale emigration. Note, however, that this reasoning stands in contrast to some of the migration literature. Most authors in the field suggest that risks should be lower in areas of mass emigration, mainly because of the protective role of migrant networks. Community-level networks are often assumed to decrease information asymmetries and facilitate the arrival at the destination abroad, thereby reducing risks (Carrington et al., 1996; Winters et al., 2001; Munshi, 2003; Bauer et al., 2007). We argue, however, that the role of networks may be ambiguous, especially when external agents learn to misuse them. This view is in line with some recent sociological research showing how exploitative employers learn to take advantage of migrant networks. One example is the study on “Networks of Exploitation” (Cranford, 2005) which shows that employers in the Los Angeles janitorial industry rewarded Mexican supervisors to recruit vulnerable workers from their home communities into dismal jobs. Migrant networks were systematically misused to replace unionized workforce with immigrants, who would earn much less and could be more easily exploited. Further studies with related arguments and evidence are Martin et al. (2007) who study the exploitation of immigrants’ arriving in the Chicago South Side, and Zhang and Chin (2002) who track the decisive role of networks in Chinese human smuggling. Thus, it seems that networks on the regional or local level do not necessarily shield from the (low) risk of being trafficked.

7.4 Data and Empirical Strategy

7.4.1 Data

The main body of analysis uses a unique household survey which was commissioned by the International Organization for Migration and covers five Eastern European countries. In addition, we also employ a second household survey from Moldova to check the robustness of our results and provide supplementary evidence. This survey is described in section 6.

Fieldwork for the five-country survey was conducted in Belarus, Bulgaria, Moldova, Romania and Ukraine between August and September 2006. To arrive at nationally representative estimates, the survey relied on a multi-stage stratified sampling scheme.⁸⁸ Accordingly, the survey covers

⁸⁸ See IOM (2006) for more details.

both urban and rural areas and sampled households of all regions in the five countries.⁸⁹ In each country, at least 1000 individuals were interviewed, with slightly larger samples in Bulgaria and Ukraine, resulting in a total sample of 5513 households. The interviews were conducted with the person whose birthday was last among the rest of dwellers.

The largest module in the survey was dedicated to human trafficking across borders. Most importantly, interviewed persons were asked whether persons in their close surroundings (relatives, friends and neighbors), in their close family,⁹⁰ or they themselves had experienced situations which would be classified as trafficking according to the definition above. Since households were randomly selected for the interview, selection bias suffered by other datasets that are based on the number of assisted victims or the number of border apprehensions (Tyldum and Brunovskis, 2005) is avoided here. To the best of our knowledge, the relatively large sample and the survey design make it the most comprehensive and representative micro-level survey on human trafficking available worldwide.

7.4.2 Dependent Variable and Econometric Model

Our dependent variable is a dummy denoting whether the interviewed person reported a victim of human trafficking among close family members.⁹¹ Given the relatively small number of individuals with trafficking in their close family, we combine sexual exploitation and other forms of exploitation. More specifically, the dummy takes the value of one if a close family member had travelled abroad and experienced one of the following situations: (i) “was offered a domestic or nursing job, but was locked and forced to work for no pay”, (ii) “was offered a job at an enterprise, on a construction site, or in agriculture, but was locked and forced to work for no or little pay” (iii) “was offered employment, but the passport was taken away upon arrival to the destination country and was forced to work in the sex business”. In line with the UN definition of human trafficking given above, these three situations have in common that the individual was forced to work by means of coercion.

Overall, 108 individuals indicate a victim of human trafficking in their close family. Out of these, there are 22 reported cases of members being forced to work in the sex business, while 86 cases were associated with non-sexual forms of exploitation. In our sample, most trafficked cases come from Moldova (56 out of 1073). The rest is found in Bulgaria (18 out of 1007), Ukraine (15 out of 1345), Belarus (11 out of 1071), and Romania (8 out of 1017). A very large number of interviewed people had not heard about any case of trafficking among the people they knew (4905 out of 5513).

⁸⁹ Overall, the analysis covers 82 regions.

⁹⁰ The term ‘close family’ is defined as only including parents, children, husband/wife, brothers and sisters of the interviewed person. The survey generally refers to households as identified by close family ties with respect to the interviewee and not by residency.

⁹¹ As the interviewee is by definition part of the close family, we make sure that our dependent variable also includes the case when the interviewed person was a victim of human trafficking him- or herself.

Our data are likely to suffer from some degree of reporting bias. Victims of human trafficking and their relatives might be reluctant to share their experiences, being afraid of stigmatization or problems with authorities at home or abroad. This is especially true for the case of sexual exploitation, which is a highly sensitive issue (Laczko, 2005). To minimize the degree of underreporting, the survey instrument used various ways of reducing the stigma associated with human trafficking. First, instead of asking directly whether a family member had been “trafficked”, the questionnaire rather referred to typical situations of coerced labor. Second, the respondent did not have to reveal the identity of a trafficked family member. Third, if there was a case of sexual exploitation in the family, the interviewed person could still “hide” this experience in a more general answer category of exploitation. Based on these considerations, we believe that the potential under-reporting bias should not seriously invalidate our findings.⁹²

We estimate the determinants of human trafficking on the household level with binary response models. A potential problem in this regard is that human trafficking, although affecting millions of people, is a relatively rare event on the household level. King and Zeng (2001) show that the maximum likelihood estimator of standard logit or probit models is biased when the number of non-events (zeros) is multiple times higher than the number of events (ones). They illustrate that standard binary models underestimate the probability of rare events and do increasingly so as the event gets rarer. As a baseline method, we therefore rely on the rare event logit model suggested by King and Zeng (2001). In essence, their model is based on the basic logit model, but uses a bias-corrected estimator for rare events that generates a lower mean square error for coefficients. Note that this approach is particularly popular in political science (e.g. Fearon and Laitin, 2003; Harff, 2003; Collier and Hoeffler, 2004) but is also increasingly used in research on economic issues (e.g. Hausmann et al., 2005; Leblang and Satyanath, 2008). To test the robustness of our results, standard logit models are also applied.⁹³

7.4.3 Identifying the Channels at Work

Our empirical approach follows the theoretical arguments above. We first test the general mechanical prediction that trafficking is more likely to occur in regions of large-scale emigration. We then focus on the subsample of migrant households to establish whether the *individual* likelihood of being trafficked increases with regional migration prevalence. In a final step, we investigate the channels behind our finding on the link between trafficking and migration flows. More specifically, we use proxies of regional recruitment conditions (demand-side factors) and proxies indicating the skill level or desperation of those departing (supply-side factors) and

92 An unavoidable bias in any database on human trafficking or migrant exploitation is that it only provides information on victims with known whereabouts or those returned. Another problem might be that the survey did not put a focus on marginal social groups, such as Sinti and Roma, among whom trafficking rates could be higher.

93 In light of the large economic and institutional differences between the countries under consideration, we cluster standard errors on the country level throughout the analysis. However, our main results are not affected when clustering on regions.

interact them with regional emigration rates. This will provide some indication on whether mass emigration reinforces the demand and supply-side channels described above.

Our key variable of interest is the regional migration prevalence rate. It is constructed by taking the share of surveyed households in each region that reported to have had a member of the close family abroad in the last 3 years. As an alternative, we employ a proxy of low migration pressure. Concretely, we construct a regional variable capturing the share of interviewed individuals who stated that they could never imagine working abroad (“share of stayers”).⁹⁴ Following our argument, our main proposition is that an increase in migration prevalence in the region should increase the risk of trafficking, while a higher share of stayers (lower migration pressure) should decrease trafficking risks.

On the supply side, one proxy for recruitment conditions in the region is the scale of illegal migration. The indicator used captures the regional proportion of migrants who work illegally abroad. Illegal migration is likely to foster the activities of people smugglers or agents procuring false travel documents, with which traffickers can collaborate. And because shadow migration industries will tend to be larger the more people depart, the recruitment of illegal migrants may be particularly easy in emigration areas. We therefore expect the adverse effect of illegal migration to depend on migration prevalence rates, so that that the interaction term of regional migration and illegality should be positively signed. The same intuition applies when using a dummy variable for illegal migration experience on the household level, which we also interact with migration prevalence. As a second demand-side proxy, we include a variable measuring the level of criminal activity (reported crimes per 10,000 inhabitants in a region) which we take from the Statistical Yearbooks of the respective countries and which we also interact with migration rates. We assume that trafficking will be easier the more criminal a region, especially if this region has large emigration flows. This would predict positive coefficients of both the crime rate variable and its interaction with regional migration prevalence.

Turning to the supply side of the human trafficking market, our focus lays on indicators of desperation and negative self-selection into migration. Because the survey intentionally omitted questions on trafficked family members, it is not possible to rely on education-based measures of selection. Likewise, we cannot construct a measure of network composition. However, the survey allows proxying the degree of desperation-led migration with the share of persons in the region who state that they would do “any kind of work” abroad.⁹⁵ Economic deprivation can cause negative self-selection on its own, but this effect will be exacerbated in areas with high emigration and larger networks (McKenzie and Rapoport, forthcoming). To test this prior we interact the desperation proxy with migration prevalence and expect a positive coefficient. Our intuition is

94 Since this variable could potentially be subject to reverse causality, the share is calculated excluding households with a trafficking case in their close or distant family. Note that the share of people not wanting to leave is negatively correlated with regional migration prevalence rates, at a coefficient of -0.30.

95 The question allowed for ten other answer categories. One can thus reasonably interpret the choice of “any kind of work” as a strong signal of desperation. To avoid endogeneity, this share is calculated excluding persons who live in migrant households or know victims of human trafficking.

that migrant networks will allow an even higher share of low-skilled and vulnerable workers to depart from deprived areas, resulting in more trafficking. In a similar vein, we interact a dummy capturing a household's "poor financial status" with migration prevalence rates.⁹⁶ We expect poorer workers, who might not have migrated in the absence of large migrant networks, to be more prone to fall victim to traffickers.

7.4.4 Other Explanatory Variables

Two variables are used to assess the role of information asymmetries and access to information. First, we code a dummy for households that use the TV, as opposed to other media and social contacts, to get informed about social and political issues. TV use can be seen as a source of relatively objective and high-quality information and a potential channel for anti-trafficking awareness campaigns. We therefore expect a negative coefficient of the TV dummy. Second, we include a regional risk-awareness measure by taking the share of respondents who stated that they had heard of the phenomenon of human trafficking before.⁹⁷ Intuitively, risk awareness is likely to reduce the likelihood of trafficking, which again predicts a negative coefficient of this variable.

Given the heterogeneity across regions and households, it is necessary to control for additional factors that might influence trafficking risks.⁹⁸ We proxy the level of socio-economic development with regional infant mortality rates (per 1000 live births) and the number of physicians (per 10,000 inhabitants). The remoteness of a region is captured by the share of people living in rural localities. On the household level, we use the number of children aged 16 or younger⁹⁹, a dummy for households living in rural areas and another dummy indicating whether households live in the district around their respective capital city. We also include age, gender, marital status, and educational level of the respondent, but only to control for systematic reporting biases along these dimensions. Descriptive statistics of all variables are presented in Table A7.1 in the appendix.

96 More precisely, households indicating that they "have to save money for the basic necessities (food, clothing, footwear)" are coded as poor. Unfortunately, this information is unavailable for households interviewed in Bulgaria and Romania.

97 We try to most closely resemble a migrant's knowledge status before departure and reduce potential endogeneity by excluding respondents who knew a victim of human trafficking or lived in a migrant household when calculating this share.

98 These variables were coded from the Statistical Yearbooks of each of the five countries. Given that our dependent variables are retrospective and in order to minimize potential endogeneity problems, all variables from the Statistical Yearbooks were coded for the year 2004.

99 Due to a modification of the question, the number of children is defined as 14 or younger in Romania.

7.5 Results

7.5.1 Main Results

Table 7.1 displays our main estimation results for the whole sample.¹⁰⁰ First, and most importantly, we find migration prevalence to be the key predictor of human trafficking (column 1). Compared to other explanatory variables, it has a large marginal effect.¹⁰¹ A one percentage point increase in the share of migrant households in a region raises the likelihood of having a trafficked family member by 0.056 percentage points. This may appear small, but one has to keep in mind that trafficking is a rare event. A one percentage point increase in migration prevalence effectively translates into a five *percent* increase in the predicted probability of human trafficking on the household level. Or, when increasing the regional share of migrant households by one standard deviation (from 0.20 to 0.37), the risk of trafficking more than doubles. This finding is confirmed when we use our proxy for low migration pressure, the regional share of people who could never imagine working abroad (*share of stayers*), see column 2. The incidence of trafficking appears to be significantly lower in regions where fewer people intend to leave.

A further main result is on the role of risk perceptions and trafficking. Households in regions with higher awareness of the phenomenon of human trafficking are less likely to have a trafficked family member. The same is true for households which use the TV as main source of information on social and political issues.¹⁰² These findings, although not fully robust, underline the potential benefits of public awareness campaigns to counter human trafficking.

To some surprise, we find that many regional control variables seem to play no role. Living in a region with a large share of rural population, a high rate of infant mortality, or low density of physicians does not increase the likelihood of trafficking. If at all, there is weak evidence that households in regions with higher infant mortality are less likely to report a trafficked family member.

The right part of Table 7.1 shows our main results for the subsample of 1560 migrant households. The analysis for this subsample is of particular relevance, as it allows testing our hypothesis that *individual* trafficking risks rise with regional migration rates. As can be seen, we find that all of our key results hold in this migrant subsample. In particular, columns 3 and 4, show that the regional scale of emigration remains a key predictor of trafficking. Migration prevalence has a positive and significant coefficient, indicating that trafficking risks increase more

100 All of our main findings hold when standard logit procedures are employed. For reasons of brevity, however, these results are available upon request.

101 With the exception of dummy variables, all marginal effects are evaluated at the mean of the independent variables.

102 Arguably, TV use may also be a proxy for household wealth. Note, however, that this result remains unaffected when controlling for a household's financial status (column 5 of Table 7.2).

than proportionally with emigration rates. We also find that our proxy for low migration pressure (“share of stayers”) keeps its significantly negative coefficient.

TABLE 7.1: BASELINE REGRESSIONS

| | (1) | | (2) | | (3) | | (4) | |
|--------------------------------------|--------------------------|-----------------|--------------------------------|-----------------|------------------------------|-----------------|--------------------------------|-----------------|
| | All households (n=5,513) | | | | Migrant households (n=1,560) | | | |
| | Baseline | | With regional share of stayers | | Baseline | | With regional share of stayers | |
| | coef/se | marginal effect | coef/se | marginal effect | coef/se | marginal effect | coef/se | marginal effect |
| migration prevalence | 4.726*** (1.013) | 0.056 | | | 5.108*** (0.784) | 0.158 | | |
| share of stayers | | | -1.391** (0.686) | -0.006 | | | -1.711* (0.947) | 0.004 |
| awareness of trafficking | -1.535* (0.889) | -0.018 | -1.857* (1.081) | -0.017 | -2.674*** (0.921) | -0.083 | -2.932*** (1.089) | -0.055 |
| TV use | -0.381** (0.157) | -0.005 | -0.372** (0.155) | -0.005 | -0.562** (0.237) | -0.020 | -0.522* (0.285) | -0.020 |
| number of children (0-16) | 0.063 (0.071) | 0.001 | 0.071 (0.070) | 0.001 | 0.161 (0.114) | 0.004 | 0.181* (0.108) | 0.005 |
| rural | 0.165 (0.206) | 0.002 | 0.204 (0.203) | 0.003 | -0.032 (0.264) | -0.001 | -0.014 (0.280) | 0.000 |
| capital district | 1.265** (0.616) | 0.024 | 1.394** (0.665) | 0.029 | 0.504 (0.651) | 0.017 | 0.699 (0.713) | 0.027 |
| share of rural population | 0.599 (1.189) | 0.007 | 1.869 (1.186) | 0.024 | 0.813 (1.056) | 0.026 | 2.557** (1.067) | 0.086 |
| infant mortality | -0.071*** (0.021) | -0.001 | -0.040 (0.030) | -0.001 | -0.113*** (0.028) | -0.004 | -0.083** (0.035) | -0.003 |
| density of physicians | -0.010 (0.011) | 0.000 | -0.005 (0.011) | 0.000 | 0.004 (0.021) | 0.000 | 0.014 (0.019) | 0.000 |
| crime rate | 0.000 (0.001) | 0.000 | -0.001 (0.001) | 0.000 | 0.002 (0.002) | 0.000 | 0.000 (0.002) | 0.000 |
| Belarus | -0.208 (0.286) | -0.003 | -1.194*** (0.273) | -0.023 | 0.937** (0.435) | 0.041 | -0.223 (0.437) | -0.097 |
| Bulgaria | 0.767 (0.467) | 0.011 | -0.250 (0.375) | -0.011 | 1.465*** (0.467) | 0.079 | 0.402 (0.431) | -0.007 |
| Romania | -0.768** (0.321) | -0.008 | -1.313*** (0.232) | -0.003 | -0.050 (0.377) | -0.003 | -0.518* (0.304) | 0.014 |
| Ukraine | 0.436 (0.424) | 0.006 | -0.474 (0.488) | -0.012 | 0.985*** (0.332) | 0.045 | 0.112 (0.455) | -0.017 |
| age (respondent) | -0.009* (0.005) | 0.000 | -0.009* (0.006) | 0.000 | -0.007 (0.006) | 0.000 | -0.008 (0.005) | 0.000 |
| male (respondent) | 0.185 (0.116) | 0.002 | 0.192* (0.116) | 0.002 | 0.153 (0.243) | 0.005 | 0.182 (0.245) | 0.006 |
| no or primary education (respondent) | -0.352* (0.212) | -0.004 | -0.334 (0.207) | -0.004 | -0.126 (0.290) | -0.004 | -0.067 (0.283) | -0.002 |
| secondary education (respondent) | -0.079 (0.186) | -0.001 | -0.030 (0.182) | 0.000 | 0.006 (0.310) | 0.000 | 0.066 (0.325) | 0.002 |
| married (respondent) | 0.289** (0.115) | 0.003 | 0.302** (0.120) | 0.004 | 0.138 (0.207) | 0.005 | 0.186 (0.202) | 0.007 |
| Number of observations | 5,513 | | 5,513 | | 1,560 | | 1,560 | |

Results of rare events logit estimation. The dependent variable is a dummy for human trafficking incidence on the household level. ***/**/* denote significance at the 1/5/10 percent level respectively. Country-clustered standard errors in parentheses.

7.5.2 Results on the Channels at Work

We have demonstrated that regional migration flows are associated with a higher individual likelihood of being trafficked. This section presents evidence on the channels at work. As our empirical strategy in this part of the analysis relies on interaction terms, we follow McKenzie and Rapoport (forthcoming) and use linear regression models (OLS) for estimation. The main purpose is to facilitate interpretation of the interaction term coefficients, which is problematic in the non-linear models typically used for binary outcomes, such as logit or probit models (Ai and Norton, 2003).

The left part of Table 7.2 presents results on demand-side effects. The interaction term between the share of illegal migration in the region and regional migration prevalence is positive and significant (column 1). The same is true for the interaction with illegal migration experience in the household (column 2). Both results indicate that illegality has a particular risk-increasing effect in high-migration areas. In line with our arguments above, traffickers seem to take advantage of larger shadow migration industries and illegal migration patterns. However, we do not find any evidence that crime facilitates trafficking in regions with high rates of out-migration. Column 3 shows that the interaction of reported crimes and migration rates in the region is insignificant and negatively signed. This non-finding might be due to measurement error and difficulties in comparing crime statistics across the five countries under investigation. Moreover, we lack alternative measures, such as the share of people convicted or the number of homicides in a region, which might provide a better indicator for relevant criminal activities.

As to potential supply-side factors, the results are displayed in columns 4 and 5. We find that our proxy for negative self-selection in the region (the share of people who would do any kind of job abroad), positively depends on migration prevalence. This is consistent with our hypothesis that migrant networks may foster the departure of more vulnerable types of migrants, particularly in deprived regions. Note, however, that the coefficient of the interaction term is significant only at the 10 percent level and not fully robust to specification changes. On the household level, we find no evidence for negative self-selection effects. The dummy on poor financial status and the interaction is insignificant. It should be kept in mind, however, that the survey does not report any pre-migration assessments of living standards. But one should also recall that other proxies of regional development, such as remoteness, child mortality or the availability of physicians also turned out to be of no importance. Overall, we thus find remarkably little evidence for supply-side factors.

TABLE 7.2: RESULTS ON THE CHANNELS AT WORK

| | (1) | (2) | (3) | (4) | (5) |
|--|------------------------------|-------------------------------|----------------------------|--------------------------------------|----------------------|
| | Demand-side effects | | | Supply-side effects | |
| | Illegal migration (regional) | Illegal migration (household) | Criminal activity (region) | Desperation-led migration (regional) | Poverty (household) |
| | OLS coef/se | OLS coef/se | OLS coef/se | OLS coef/se | OLS coef/se |
| share of illegal migration in region | -0.119 (0.077) | | | | |
| illegal migr. share*migration prevalence | 0.348* (0.195) | | | | |
| illegal migration (household) | | -0.030 (0.023) | | | |
| illegal hh migration*migration prevalence | | 0.176*** (0.058) | | | |
| crime rate in region*migration prevalence | | | -0.001 (0.001) | | |
| desperation-led migr. (% willing to do any job abroad) | | | | -0.208 (0.196) | |
| desperation-led migration*migration prevalence | | | | 1.223* (0.717) | |
| poor financial status of household | | | | | 0.008 (0.051) |
| household poverty*migration prevalence | | | | | -0.008 (0.088) |
| migration prevalence | 0.099 (0.066) | 0.144*** (0.049) | 0.260** (0.105) | 0.137*** (0.040) | 0.259*** (0.032) |
| awareness of trafficking | -0.118** (0.055) | -0.113** (0.058) | -0.142* (0.073) | -0.119** (0.055) | -0.233*** (0.047) |
| TV use | -0.024 (0.015) | -0.021 (0.015) | -0.024* (0.014) | -0.024* (0.015) | -0.039*** (0.011) |
| number of children (0-16) | 0.009 (0.006) | 0.008 (0.006) | 0.009 (0.008) | 0.009 (0.008) | 0.015* (0.008) |
| rural | -0.004 (0.018) | -0.006 (0.017) | -0.005 (0.012) | -0.003 (0.012) | -0.001 (0.020) |
| capital district | 0.018 (0.024) | 0.021 (0.024) | 0.024 (0.027) | 0.025 (0.024) | 0.016 (0.027) |
| share of rural population | 0.035 (0.052) | 0.037 (0.050) | 0.036 (0.053) | 0.033 (0.044) | 0.029 (0.024) |
| infant mortality | -0.003 (0.002) | -0.003 (0.002) | -0.003 (0.002) | -0.003 (0.002) | -0.003** (0.002) |
| density of physicians | 0.000 (0.001) | 0.000 (0.001) | 0.000 (0.001) | 0.000 (0.001) | 0.001* (0.001) |
| crime rate | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| country dummies | yes | yes | yes | yes | yes |
| control for respondent's characteristics | yes | yes | yes | yes | yes |
| Number of observations | 1,560 | 1,560 | 1,560 | 1,560 | 987 |

Results

of ordinary least squares estimation. The dependent variable is a dummy for human trafficking incidence on the household level. ***/**/* denote significance at the 1/5/10 percent level respectively. Country-clustered standard errors in parentheses.

7.5.3 Robustness Checks

To test the robustness of the results presented above, this section conducts a series of additional checks. In a first step, we explicitly account for the fact that human trafficking should only be observable for migrant households. To econometrically allow for this we estimate a probit sample

selection model with a dummy for migrant households as dependent variable in the selection equation and a dummy for households with a trafficked family member in the outcome equation. One way to identify this model is to find a variable that is a significant predictor in the selection equation, but excludable from the outcome equation. Our dataset offers only a limited choice of such instruments. Arguably, the most credible variable is a dummy indicating whether the household has an employed member at home. Our identifying assumption is that having a domestically employed household member relieves a household's migration pressure, but does not directly affect the risk of being trafficked. Having an employed household member turns out to be a highly significant and negative predictor of having a family member abroad.¹⁰³ The results of the probit selection model (columns 1 and 2 of Table 7.3) closely resemble those of the rare events logit estimation. In addition, our estimates do not appear to be biased due to sample selection. The Likelihood Ratio test of independent equations cannot reject the null hypothesis of no correlation between the error terms (p-value 0.56). We are aware that identification may not be fully warranted with this approach. We implicitly assume negative self-selection into migration, as households with unemployed members are more likely to send a migrant. As discussed above, such negative self-selection effects might also affect trafficking risks, especially in high-emigration regions, which raises concerns on whether our instrument is strictly excludable. However, we show that instrument choice does not affect the results in the Moldova case study (section 6). Moreover, an alternative form of identification through functional form, thus relying on the non-linearity of the model. Re-estimating the regressions without an instrument yields similar results, with migration prevalence remaining a significant and substantial determinant of trafficking risks (available upon request). We thus conclude that it is sufficient to separate the sample selection model and focus on the outcome equation as done in Tables 7.1 and 7.2.

Another issue of concern is the role of migrant destinations. It is possible that high migration flows correlate with migration to destinations with poorer law enforcement or weaker anti-trafficking laws, e.g. Russia or Turkey. While the survey does not contain information on migrant destinations, we take advantage of well-known migration patterns across countries to conduct a simple robustness check. Specifically, we divide the sample into sending countries with predominantly "Western" destinations (Romania and Bulgaria) and countries with more "Eastern"-oriented migration flows (Belarus, Moldova, Ukraine). We find that in both groups of countries regional migration prevalence remains a highly significant predictor of human trafficking risks (columns 3 and 4).

In a next step, we restrict our sample to rural households. The rationale for this is that networks are typically thought to have stronger effects in rural areas. However, as shown in column 5, the results remain largely the same. We also find little changes when explicitly controlling for proxies

103 Most evidence on migration in the region indicates that households first try to place a member in the domestic labor market and send a migrant only if this is unsuccessful. Generally, migration in the region is overwhelmingly motivated by the lack of income or job opportunities (88 percent of respondents in the survey state so). In Moldova, for example, only about a third of those migrating were active in the labor market before their departure (Omar Mahmoud and Trebesch, 2009).

of regional institutions. More specifically, we use the share of respondents answering that trafficking in the region mainly occurs (i) because of weak laws and poor law enforcement or (ii) because of corruption. As can be seen in column 6, we find these two proxies to be insignificant.

TABLE 7.3: ROBUSTNESS CHECKS

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|--|--------------------------------|----------------------|----------------------|---------------------------|-----------------------|----------------------------|----------------------|
| | Heckman probit selection model | | Bulgaria and Romania | Belarus, Moldova, Ukraine | Rural households only | With regional institutions | With family network |
| | Trafficking | Migration | Trafficking | Trafficking | Trafficking | Trafficking | Trafficking |
| | coef/se | coef/se | coef/se | coef/se | coef/se | coef/se | coef/se |
| migration prevalence | 2.744*** (0.758) | 1.847*** (0.213) | 3.870*** (0.569) | 4.842** (2.352) | 5.788*** (0.466) | 5.261*** (1.495) | 5.278*** (0.786) |
| weak laws and law enforcement | | | | | | -0.239 (1.992) | |
| corruption | | | | | | -0.126 (1.191) | |
| family network | | | | | | | 0.039 (0.041) |
| awareness of trafficking | -1.180*** (0.430) | -0.081 (0.152) | -0.409 (2.279) | -3.302*** (1.152) | -2.566*** (0.768) | -2.533* (1.302) | -2.647*** (0.860) |
| TV use | -0.231 (0.170) | 0.055 (0.054) | -0.087 (0.442) | -0.864*** (0.175) | -1.065*** (0.216) | -0.568** (0.237) | -0.632*** (0.185) |
| number of children (0-16) | 0.089 (0.064) | 0.061** (0.026) | 0.139*** (0.023) | 0.209* (0.111) | 0.168 (0.175) | 0.157 (0.117) | 0.136 (0.136) |
| rural | 0.014 (0.133) | 0.099** (0.043) | -0.052 (0.345) | -0.086 (0.368) | | -0.027 (0.274) | -0.040 (0.267) |
| capital district | 0.193 (0.253) | -0.021 (0.077) | 1.040** (0.524) | 0.374 (0.779) | 0.931 (1.041) | 0.490 (0.645) | 0.577 (0.742) |
| share of rural population | 0.375 (0.634) | 0.081 (0.199) | 4.079* (2.315) | 0.138 (0.570) | -1.754 (1.448) | 0.763 (1.014) | 0.759 (1.078) |
| infant mortality | -0.048 (0.029) | 0.010 (0.007) | -0.169* (0.096) | -0.069 (0.105) | -0.166*** (0.032) | -0.114*** (0.031) | -0.113*** (0.027) |
| density of physicians | 0.000 (0.010) | 0.000 (0.003) | -0.010 (0.045) | 0.011 (0.025) | -0.037 (0.028) | 0.003 (0.021) | 0.000 (0.021) |
| crime rate | 0.001 (0.001) | -0.001** (0.000) | 0.013*** (0.001) | 0.000 (0.000) | 0.003 (0.004) | 0.002 (0.002) | 0.002 (0.002) |
| country dummies | yes | yes | yes | yes | yes | yes | yes |
| control for respondent's characteristics | yes | yes | yes | yes | yes | yes | yes |
| employed | | -0.115*** (0.043) | | | | | |
| rho | 0.471 (0.864) | | | | | | |
| Number of observations | | 5,513 | 552 | 1,008 | 864 | 1,560 | 1,550 |

Results of rare events logit and Heckman probit selection estimation. The dependent variable in columns 1 and 3-7 is a dummy for human trafficking incidence on the household level. ***/**/* denote significance at the 1/5/10 percent level respectively. Country-clustered standard errors in parentheses.

Finally, we analyze the differing effects of family networks (strong ties) versus community networks (weak ties). Many of the arguments on networks made so far would apply to weak-tie networks, but not to family networks, which are likely to be more protective when it comes to human trafficking. The family network variable available to us simply captures the number of other members in the households that have migrated in the last three years. Family networks

appear to play a protective role. Their effect, however, is far from being statistically significant (column 7). The same is true if we include regional fixed effects (available upon request). However, a more in-depth analysis with better data on family migration patterns is needed to strengthen this conclusion.

7.6 Moldova Case Study

To supplement the analysis and address some limitations of the five-country survey, this section presents a detailed case study of Moldova, the Eastern European country which is most affected by human trafficking. A main difference to the five-country study presented above is the range of available data for Moldova, allowing us to conduct the analysis at less aggregated levels. First, we have access to a household survey with very detailed information on migrant characteristics and their migration patterns including the country of destination. Second, the Ministry of Economy and Trade provides a rich database on socio-economic characteristics of every community in Moldova. We can therefore investigate the drivers of human trafficking on the individual, not household level and include locality-specific migration prevalence rates and control variables.

Individual data come from a specialized household survey on migration and remittances, capturing 3916 households. The survey was commissioned by IOM and executed by the CBSAXA opinion research firm in summer 2008 (see Lücke et al. (2009) for further details). The impressive scale of emigration from Moldova is reflected by the fact that 1127 households in the survey (29 percent) had at least one member abroad. Although the survey instrument does not explicitly focus on trafficking and exploitation, we can take advantage of one question on problems confronted abroad. This question was answered by almost all respondents and allows identifying the worst forms of migrant exploitation analogous to the definition used so far. Specifically, we code migrants as trafficked if (i) they were locked by their employer abroad, (ii) their passport was taken away and/or (iii) they were subject to sexual exploitation. 31 migrants in our sample meet these criteria. This share is notably smaller than in the five-country survey, a fact that can be attributed to survey design, in particular to the way trafficking-related questions were asked, but also to the degree of retrospectiveness, as the question only referred to the survey year and the preceding one (2007 and 2008).

To assure comparability, the Moldovan case study follows the five-country analysis as closely as possible. Some minor differences in the construction of variables are worth mentioning. Above all, instead of controlling for the characteristics of the respondent, we include migrants' personal traits directly. Community migration prevalence is proxied by the share of people who are temporarily absent from a locality. The Moldovan Labor Force Survey shows that this is a meaningful proxy as labor migrants represent more than 80 percent of those temporarily absent. A major advantage of this variable is its representativity at the local level, as it is derived from administrative statistics and not survey-based. In addition to these changes, availability of doctors

is now expressed as the distance to the next doctor and crime rates are replaced by perceived levels of personal security.

Taken together, the baseline results for the Moldovan migrant sample are nearly identical to the ones for the five-country study (column 1 of Table 7.4). This strongly supports our argument and findings from above. Migration prevalence remains by far the most important determinant of trafficking risks. This result holds when using fixed effects for the 35 regions in Moldova, restricting the sample to rural areas (available upon request), or including dummies for migrant destinations (column 2). Quite remarkably, migrants' personal characteristics, such as age, sex and education, do not appear to matter. The same is true for destinations. Neither is migration to Russia or Italy, which together host about 80 percent of Moldovan migrants, associated with increased trafficking risks.

Again, we also investigate the differing role of family and community networks. The Moldovan survey data allow for a better measurement of family networks, as it contains information on the number of family members that migrated prior to the migrant him- or herself. As can be seen in column 3, we now find the family network variable to be significant and negative, indicating that previous migration experience of other household members can help lowering the risk of trafficking and exploitation. This supports the view that our arguments may only apply to weak-tie, but not strong-tie networks.

Robustness of the probit selection model is tested by using an alternative instrument. Specifically, we use occupational information on whether surveyed individuals are employed as construction workers in Moldova or were so prior to departure. Given the building boom in Moldova's main destination countries until 2008 and the resulting demand for cheap construction labor, construction workers have a much higher probability to emigrate (Omar Mahmoud and Trebesch, 2009). At the same time, there is no reason to expect construction workers to face a higher (or lower) risk of trafficking, an intuition that is confirmed by the data. Using this identification approach, we find the key result to hold in the sample selection model (available upon request). As before, our estimates do not appear to be biased due to sample selection.

TABLE 7.4: RESULTS FOR MOLDOVA CASE STUDY

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---------------------|---------------------|---------------------|------------------------------|-------------------------------|---------------------|----------------------|
| | Main results | | | Demand-side effects | | Supply-side effects | |
| | Baseline | With destinations | With family network | Illegal migration (regional) | Illegal migration (household) | Migrant education | Network composition |
| | coef/se | coef/se | coef/se | OLS coef/se | OLS coef/se | OLS coef/se | OLS coef/se |
| migration prevalence | 7.519*** (2.507) | 7.357*** (2.612) | 7.633*** (2.575) | -0.259** (0.123) | 0.017 (0.045) | 0.199 (0.168) | 0.662*** (0.137) |
| destination Russia | | 0.008 (0.753) | | | | | |
| destination Italy | | -0.590 (1.047) | | | | | |
| destination other CIS countries | | -0.063 (0.977) | | | | | |
| other EU countries | | 0.736 (0.712) | | | | | |
| family network | | | -0.931* (0.542) | | | | |
| share of illegal migration (% in region) | | | | -0.040 (0.060) | | | |
| illegal migr. share*migration prevalence (region) | | | | 1.556*** (0.589) | | | |
| illegal migration (migrant) | | | | | 0.003 (0.018) | | |
| illegal migration (migrant)*migration prevalence higher education | | | | | 0.460** (0.209) | | |
| (migrant)*migration prevalence network composition (% highly-educated migrants) | | | | | | 0.115 (0.166) | |
| network composition*migration prevalence (region) | | | | | | | 0.121** (0.051) |
| | | | | | | | -1.818*** (0.528) |
| age (migrant) | 0.011 (0.025) | 0.015 (0.025) | 0.002 (0.026) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| male (migrant) | -0.041 (0.390) | -0.170 (0.394) | -0.148 (0.410) | -0.001 (0.006) | 0.001 (0.006) | 0.000 (0.006) | -0.001 (0.006) |
| secondary education (migrant) | -0.016 (0.587) | 0.035 (0.597) | 0.008 (0.593) | 0.001 (0.008) | 0.001 (0.010) | | 0.001 (0.010) |
| higher education (migrant) | 0.263 (0.716) | 0.272 (0.683) | 0.255 (0.738) | 0.006 (0.011) | 0.007 (0.010) | -0.003 (0.009) | 0.005 (0.011) |
| married (migrant) | -0.370 (0.495) | -0.459 (0.480) | -0.387 (0.474) | -0.008 (0.008) | -0.006 (0.008) | -0.005 (0.008) | -0.005 (0.007) |
| number of children (0-16) | 0.027 (0.247) | 0.041 (0.241) | 0.028 (0.247) | 0.001 (0.004) | 0.002 (0.004) | -0.000 (0.004) | 0.000 (0.003) |
| rural | 0.143 (0.553) | 0.200 (0.550) | 0.175 (0.563) | 0.006 (0.007) | 0.003 (0.005) | 0.005 (0.008) | 0.001 (0.009) |
| capital district | -0.228 (1.016) | -0.353 (1.078) | -0.204 (1.051) | -0.007 (0.011) | -0.004 (0.013) | -0.001 (0.009) | -0.002 (0.019) |
| share of rural population | 0.000 (0.014) | -0.000 (0.015) | 0.000 (0.014) | -0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) | 0.000 (0.000) |
| infant mortality | 0.000 (0.001) | 0.000 (0.001) | 0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) |
| distance to doctor | 0.049 (0.054) | 0.046 (0.053) | 0.049 (0.054) | 0.000 (0.001) | 0.000 (0.001) | 0.000 (0.001) | 0.000 (0.001) |
| crime in locality | -0.321 (0.435) | -0.293 (0.440) | -0.335 (0.432) | -0.002 (0.007) | -0.005 (0.006) | -0.007 (0.008) | -0.002 (0.007) |
| Number of observations | 1,679 | 1,679 | 1,679 | 1,679 | 1,679 | 1,679 | 1,679 |

Results of rare events logit and OLS estimation in a sample of 1679 migrants. The dependent variable indicates if a migrant has been trafficked or not. ***/**/* denote significance at the 1/5/10 percent level respectively. Locality-clustered standard errors in parentheses.

The remaining columns of Table 7.4 presents evidence on why trafficking risks increase with migration prevalence. To test for demand-side forces, we rely on indicators of illegal migration and crime. This time, however, we can use the proportion of migrants who crossed borders illegally as proxy for the extent of illegal migration. Illegal border crossing is arguably a more exogenous indicator than illegal work abroad and might also capture the size of the regional shadow migration industry more accurately. In line with our previous evidence, our results point to the existence of agglomeration effects, as the impact of illegality increases with emigration rates (columns 4 and 5). Again, the existence of criminal activities does not appear to increase the likelihood of trafficking in high-migration areas (available upon request). As regards factors on the supply side, the availability of migrants' educational attainment allows for alternative identification of negative self-selection and the possibility to analyze the effect of network composition. We do not find that larger networks induce particularly vulnerable migrants to leave (column 6). However, network composition, measured as the proportion of migrants with higher education, seems to play a role. Migrants from regions with high fractions of low-skilled migrants face a particularly high risk when networks are large (column 7). This lends some credence to the hypothesis that network quality may decrease with network size and the degree of negative self-selection. However, more research with better data is needed to validate this point. Overall, our findings for Moldova are very similar to the ones of the five-country study and more in favor of demand-side channels.

7.7 Conclusion

Our results for the Eastern European context have several implications. First, policymakers concerned about human trafficking should acknowledge the important role of migration. We find strong indication that traffickers and exploitative employers are more prone to recruit victims from regions with large-scale emigration flows. For this reason, we propose to target awareness campaigns and other anti-trafficking policies to areas where migration flows are high or on the rise, not to remote areas with no or little migration.

Second, our indicative findings on the role of risk awareness and information use may provide some orientation for policy initiatives facing cost-benefit trade-offs. In fact, the results give some reason for optimism that anti-trafficking campaigns and measures to reduce information asymmetries in the migration process can indeed be a promising way to reduce the crime. Besides awareness campaigns, the ILO (2005) suggests to set up labor market information systems on jobs at home and abroad and the general use of model employment contracts.

Third, our results document the close link between illegal migration patterns and trafficking. Migrants who cross borders illegally and who depart from regions with more illegal emigrants face significantly higher risks. One can argue that restrictive immigration policies exacerbate the problem, pushing would-be immigrants into illegality and fostering the emergence of shadow migration industries. However, a moderate increase of legal migration opportunities is unlikely to

solve the problem given the magnitude of global migration pressure (Amin and Mattoo, 2006). In addition, traffickers' incentives would remain unchanged through such measures.

Finally, we find some indication that large community migration networks increase risks and may thus have ambiguous effects. This stands in some contrast to the existing migration literature, which generally advocates a risk-reducing role of networks. However, more research is needed to better understand the role of social networks for migration and trafficking.

As a general conclusion, it seems likely that the market of human trafficking and the number of victims will continue to grow as long as migration pressure remains high.

Literature

- Adams, R.H. (2005). Remittances, Poverty, and Investment in Guatemala. In Ç. Özden and M. Schiff (Eds.). *International Migration, Remittances, and the Brain Drain*. Washington D.C.: World Bank/Palgrave Macmillan.
- Ai C., and E.C. Norton (2003). Interaction Term in Logit and Probit Models. *Economics Letters*, 80(1): 123-129.
- Akee, R., A.K. Basu, A.S. Bedi, and N. Chau (2007). Determinants of Trafficking in Women and Children: Cross-National Evidence, Theory and Policy Implications. Mimeo.
- Amin, M., and A. Mattoo (2006). Can Guest Worker Schemes Reduce Illegal Migration? World Bank Policy Research Working Paper 3828.
- Aronowitz, A.A. (2001). Smuggling and Trafficking in Human Beings: The Phenomenon, the Markets That Drive It and the Organizations That Promote It. *European Journal on Criminal Policy and Research*, 9(2): 163-195.
- Bales, K. (1999). *Disposable People: New Slavery in the Global Economy*. Berkeley: University of California Press.
- Basu, A.K., and N. Chau (2003) Targeting Child Labor in Debt Bondage: Evidence Theory and Policy Prescriptions. *World Bank Economic Review*, 17(2): 255-281.
- Bauer, T., and K. Zimmermann (1999). Assessment of Possible Migration Pressure and Its Labor Market Impact Following EU Enlargement to Central and Eastern Europe. IZA Research Reports 3, Institute for the Study of Labor (IZA).
- Bauer, T., G. Epstein, and I.N. Gang (2007). The Influence of Stocks and Flows on Migrants' Location Choices. *Research in Labor Economics*, 26: 199-229.
- Beine, M., F. Docquier and Ç. Özden (forthcoming). Diasporas. *Journal of Development Economics*.
- Bhabha, J. (2005). Trafficking, Smuggling, and Human Rights. *Migration Information Source: Special Issue on Migration and Human Rights*. Migration Policy Institute, Washington, D.C.
- Blum, J. (1957). The Rise of Serfdom in Eastern Europe. *American Historical Review*, 62(4): 807-836.
- Borjas, G.J. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review*, 77(4): 531-53.
- Bruckert, C., and C. Parent (2002). Trafficking in Human Beings and Organized Crime: A Literature Review. Royal Canadian Mounted Police. Mimeo.
- Carrington, W.J., E. Detragiache, and T. Vishwanath (1996). Migration with Endogenous Moving Costs. *American Economic Review*, 86(4): 909-30.
- Chand, S. (2008). Economics of Human Trafficking: A Survey. The Australian National University. Mimeo.
- Chiquiar, D. and G.H. Hanson (2005). International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States. *Journal of Political Economy*, 113(2): 239-281.
- Clemens, M., C. Montenegro, and L. Pritchett (2008). The Place Premium: Wage Differences for Identical Workers Across the US Border. World Bank Policy Research Working Paper 4671.

- Collier, P., and A. Hoeffler (2004). Greed and Grievance in Civil War. *Oxford Economic Papers*, 56(4): 563-595.
- Coomaraswamy, R. (2001). United Nations Addendum Report to the Human Rights Commission regarding Mission to Bangladesh, Nepal and India on the Issue of Trafficking of Women and Children.
- Cranford, C.J. (2005). Networks of Exploitation: Immigrant Labor and the Restructuring of the Los Angeles Janitorial Industry. *Social Problems*, 52(3): 379-397.
- Dessy, S., and S. Pallage (2006). Some Surprising Effects of Better Law Enforcement Against Child Trafficking. *Journal of African Development*, 8(1): 115-132.
- Domar, E. (1970). The Causes of Slavery or Serfdom: A Hypothesis. *Economic History Review*, 30(1): 18-32.
- Fearon, J.D., and D.D. Laitin (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review*, 97(1): 75-90.
- Fiorentini, G., and S. Peltzman (1995). *The Economics of Organized Crime*. Cambridge: Cambridge University Press.
- Freeman, R.J. (1999). The Economics of Crime. In: Orley Ashenfelter and David Card (Eds.), *Handbook of Labor Economics, Volume 3C*. Amsterdam: North Holland.
- Freeman, S., J. Grogger and J. Sonstelie (1996). The Spatial Concentration of Crime. *Journal of Urban Economics*, 40(2): 216-231.
- Friebel, G., and S. Guriev (2006). Smuggling Humans: A Theory of Dept-Financed Migration. *Journal of the European Economic Association*, 4(6): 1085-1111.
- Gathmann, C. (2008). Effects of Enforcement on Illegal Markets: Evidence from Migrant Smuggling Along the South-Western Border. *Journal of Public Economics*, 92(10-11): 1926-1941.
- Glaeser, E.L., and B. Sacerdote (1999). Why Is There More Crime in Cities? *Journal of Political Economy*, 107(6): 225-229.
- Gozdziak, E.E., and E.A. Collett (2005). Research on Human Trafficking in North America: A Review of Literature. *International Migration*, 43(1-2): 99-128.
- Görlich, D., and C. Trebesch (2008). Seasonal Migration and Networks – Evidence on Moldova’s Labor Exodus. *Review of World Economics*, 144(1): 107-133.
- Harff, B. (2003). No Lessons Learned from the Holocaust: Assessing Risks of Genocide and Political Mass Murder Since 1955. *American Political Science Review*, 97(1): 57-74.
- Hausmann, R., L. Pritchett and D. Rodrik (2005). Growth Accelerations. *Journal of Economic Growth*, 10(4): 303-329.
- ILO (2005). A Global Alliance Against Forced Labor. ILO, Geneva.
- IOM (2006). Human Trafficking Survey: Belarus, Bulgaria, Moldova, Romania, and Ukraine. December 2006. IOM Kyiv, Ukraine.
- King, G., and L. Zeng (2001). Logistic Regression in Rare Events Data. *Political Analysis*, 9(2): 137-63.
- Kraler, A., and K. Iglicka (2002). Labor Migration in Central and Eastern European Countries. In F. Laczko, I. Stacher, A. Klekowski von Koppenfels (Eds.), *New Challenges for Migration Policy in Central and Eastern Europe*. The Hague: Asser Press.

- Laczko, F. (2005). Data and Research on Human Trafficking. *International Migration*, 43(1): 5-16.
- Laczko, F., and M.A. Gramegna. (2003). Developing Better Indicators of Human Trafficking. *Brown Journal of World Affairs*, 10(1): 179-94.
- Langberg, L. (2005). A Review of Recent OAS Research on Human Trafficking in the Latin American and Caribbean Region. *International Migration*, 43(1-2): 129-139.
- Leblang, D., and S. Satyanath (2008). Politically Generated Uncertainty and Currency Crises: Theory, Tests, and Forecasts. *Journal of International Money and Finance*, 27(3): 480-497.
- Leman, J., and S. Janssens (2007). Travel Agencies as a Linking Element for Human Smuggling and Trafficking from Eastern Europe. *Migration Studies*, 44(166): 443-459.
- LexisNexis (2008). LexisNexis Global Survey of Human Trafficking Laws. July 2008. Available at http://www.lexisnexis.com/documents/pdf/20081218105838_large.pdf
- Lücke, M., T. Omar Mahmoud and A. Steinmayr (2009). Labor Migration and Remittances in Moldova: Is the Boom Over? IOM, Chisinau.
- Mansoor, A., and B. Quillin (2006). *Migration and Remittances: Eastern Europe and the Former Soviet Union*. The World Bank, Washington, D.C.
- Martin, N., S. Morales and N. Theodore (2007). Migrant Worker Centers: Contending with Downgrading in the Low-Wage Labor Market. *GeoJournal*, 68(2-3): 155-165.
- McKenzie, D., and H. Rapoport (forthcoming). Self-Selection Patterns in Mexico-U.S. Migration: the Role of Migration Networks. *Review of Economics and Statistics*.
- Mora, J., and J.E. Taylor (2005). Determinants of Migration, Destination, and Sector Choice: Disentangling Individual, Household, and Community Effects. In Ç. Özden and M. Schiff (Eds.) *International Migration, Remittances, and the Brain Drain*. Washington D.C.: World Bank/Palgrave Macmillan.
- Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *Quarterly Journal of Economics*, 118(2): 549-599.
- Naim, M. (2005). *Illicit: How Smugglers, Traffickers and Copycats Are Hijacking the Global Economy*. New York, NY: Doubleday.
- Obuah, E. (2006). Combating Global Trafficking in Persons: The Role of the United States Post-September 2001. *International Politics*, 43(2): 241-265.
- Omar Mahmoud, T., and C. Trebesch (2009). Emigration and Skills Shortages in Moldova. Background paper for Country Economic Memorandum, The World Bank, Washington, D.C.
- Orrenius, P. (1999). The Role of Family Networks, Coyote Prices and the Rural Economy in Migration from Western Mexico: 1965-1994. Working Paper 99-10, Federal Reserve Bank of Dallas.
- Orrenius, P., and M. Zavodny (2005). Self-Selection Among Undocumented Immigrants from Mexico. *Journal of Development Economics*, 78(1): 215-240.
- Piper, N. (2005). A Problem by a Different Name? A Review of the Research on Trafficking in South-East Asia and Oceania. In F. Laczko and E. Gozdzik (Eds), *Data and Research On Human Trafficking: a Global Survey*. IOM, Geneva: 202-233.
- Ruggiero, V. (1997). Trafficking in Human Beings: Slaves in Contemporary Europe. *International Journal of the Sociology of Law*, 25(3): 231-244.

- Salt, J., and J. Stein (1997). Migration as Business: The Case of Trafficking. *International Migration*, 35(4): 467-494.
- Schloenhardt, A. (1999). Organized Crime and the Business of Migrant Trafficking. *Crime, Law and Social Change*, 32(3): 203-233.
- Stalker, P. (2000). *Workers without Frontiers: The Impact of Globalisation on International Migration*. ILO, Geneva.
- Surtees, R. (2005). Second Annual Report on Victims of Trafficking in South-Eastern Europe. IOM, Geneva.
- Tamura, Y. (forthcoming). Migrant Smuggling. *Journal of Public Economics*.
- Tyldum, G., and A. Brunovskis (2005). Describing the Unobserved: Methodological Challenges in Empirical Studies on Human Trafficking. *International Migration*, 43(1-2): 17-34.
- Tyuryukanova, E. (2005). *Forced Labor in the Russian Federation Today*. Geneva: ILO.
- UN (2000a). Protocol to Prevent, Suppress and Punish Trafficking in Persons, Especially Women and Children, Supplementing the United Nations Convention Against Transnational Organized Crime. United Nations.
- UN (2000b). Protocol against the Smuggling of Migrants by Land, Sea and Air, Supplementing the United Nations Convention Against Transnational Organized Crime. United Nations.
- UNODC (2006). *Trafficking in Persons: Global Patterns*. United Nations Office on Drugs and Crime.
- US State Department (2008). *Trafficking in Persons Report*. Office to Monitor and Combat Trafficking in Persons. Washington, D.C.
- Väyrynen, R. (2005). Illegal Immigration, Human Trafficking and Organized Crime. In: G. Borjas and J. Crisp (eds.), *Poverty, International Migration and Asylum*. New York, NY: Palgrave Macmillan: 143-170.
- Winters, P., A. de Janvry, and E. Sadoulet (2001). Family and Community Networks in Mexico-U.S. Migration. *Journal of Human Resources*, 36(1): 159-184.
- World Bank (2005). *Global Economic Prospects 2006: Economic Implications of Remittances and Migration*. Washington, D.C.
- Zenou, Y. (2003). The Spatial Aspects of Crime. *Journal of the European Economic Association*, 1(2-3): 459-467.
- Zhang, S., and K.-L. Chin (2002). Enter the Dragon: Inside Chinese Human Smuggling Organizations. *Criminology*, 40(4): 737-767.

Appendix

TABLE A7.1: SUMMARY STATISTICS FOR HOUSEHOLDS WITH AND WITHOUT TRAFFICKED MEMBERS

| | All households (n=5,513) | | | | Migrant households only (n=1,560) | | | |
|---|--------------------------|-----------|------------------------|-----------|-----------------------------------|-----------|------------------------|-----------|
| | No trafficked member | | With trafficked member | | No trafficked member | | With trafficked member | |
| | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| <i>Regional characteristics</i> | | | | | | | | |
| rural locality | 0.48 | 0.50 | 0.56 | 0.50 | 0.55 | 0.50 | 0.61 | 0.49 |
| capital district | 0.18 | 0.38 | 0.21 | 0.41 | 0.14 | 0.34 | 0.10 | 0.30 |
| share of rural population | 0.38 | 0.22 | 0.51 | 0.26 | 0.48 | 0.24 | 0.58 | 0.23 |
| infant mortality | 10.93 | 3.75 | 11.08 | 2.57 | 11.87 | 3.72 | 11.39 | 2.56 |
| density of physicians (per 10,000 inh.) | 35.40 | 16.24 | 28.84 | 12.51 | 29.41 | 15.43 | 27.09 | 12.25 |
| crime rate (times per 10,000 inh.) | 160.92 | 107.93 | 115.89 | 77.72 | 122.28 | 80.46 | 104.41 | 65.21 |
| migration prevalence | 0.28 | 0.17 | 0.42 | 0.19 | 0.37 | 0.18 | 0.46 | 0.20 |
| share of illegal migrants | 0.28 | 0.17 | 0.31 | 0.17 | 0.30 | 0.17 | 0.35 | 0.17 |
| share of stayers | 0.51 | 0.12 | 0.48 | 0.11 | 0.49 | 0.12 | 0.46 | 0.11 |
| awareness of trafficking | 0.79 | 0.14 | 0.75 | 0.18 | 0.79 | 0.15 | 0.70 | 0.19 |
| share willing to do any job abroad | 0.05 | 0.05 | 0.07 | 0.05 | 0.06 | 0.05 | 0.08 | 0.05 |
| rule of law (perception-based) | 0.27 | 0.14 | 0.24 | 0.12 | 0.27 | 0.15 | 0.24 | 0.12 |
| corruption (perception-based) | 0.20 | 0.14 | 0.22 | 0.13 | 0.23 | 0.14 | 0.21 | 0.13 |
| <i>Household characteristics</i> | | | | | | | | |
| family network (other migrants in hh) | 0.25 | 0.92 | 0.86 | 1.85 | 0.90 | 1.58 | 1.30 | 2.14 |
| illegal migration experience | 0.09 | 0.29 | 0.40 | 0.49 | 0.34 | 0.47 | 0.60 | 0.49 |
| number of children (0-16) | 0.46 | 0.75 | 0.63 | 0.84 | 0.55 | 0.83 | 0.76 | 0.91 |
| TV use | 0.83 | 0.38 | 0.78 | 0.42 | 0.85 | 0.35 | 0.79 | 0.41 |
| poor financial status | 0.41 | 0.49 | 0.48 | 0.50 | 0.45 | 0.50 | 0.48 | 0.50 |
| employed hh member (dummy) | 0.46 | 0.50 | 0.49 | 0.50 | 0.41 | 0.49 | 0.50 | 0.50 |
| <i>Respondent's characteristics</i> | | | | | | | | |
| age | 43.86 | 17.71 | 41.50 | 16.04 | 42.32 | 17.14 | 39.99 | 16.38 |
| male | 0.45 | 0.50 | 0.47 | 0.50 | 0.44 | 0.50 | 0.49 | 0.50 |
| no or primary education | 0.29 | 0.45 | 0.27 | 0.45 | 0.29 | 0.46 | 0.31 | 0.46 |
| secondary education | 0.43 | 0.50 | 0.46 | 0.50 | 0.41 | 0.49 | 0.46 | 0.50 |
| higher education | 0.28 | 0.45 | 0.27 | 0.45 | 0.30 | 0.46 | 0.24 | 0.43 |
| married | 0.60 | 0.49 | 0.67 | 0.47 | 0.63 | 0.48 | 0.67 | 0.47 |

Chapter 8

Conclusion

Rural life in developing and transition countries is plagued by risks. This dissertation has explored how rural households cope with shocks and manage risks. A special focus has been put on the welfare implications of household behavior.

The first part has analyzed the socio-economic consequences of AIDS-related mortality. Chapter 2 has shown that households in rural Zambia that experience the death of a prime-age member are able to stabilize their per-capita incomes in the short to medium run. Adjustments of income-generating activities as well as household size and composition are likely to explain this finding. Yet, risk-sharing arrangements spread the burden beyond directly afflicted households. Households that have not lost a member but support afflicted households by fostering orphans incur considerable reductions in per-capita incomes. These spillover effects suggest that the targeting of AIDS-related interventions should not be limited to households in which the death occurred. Overall, however, the impact of AIDS-related mortality on monetary welfare appears to be limited. Nevertheless, this finding does not render public support unnecessary if coping with the death of a household member adversely affects other dimensions of welfare. Chapter 3 has demonstrated for North-Western Tanzania that the age at parental bereavement has important consequences for children's long-term capital accumulation in terms of both health and education. The effects, however, depend on the gender of the deceased parent. Preferences of the surviving parent partly protect same-sex children from the detrimental effects of orphanhood, suggesting that risks are not equally shared within households. The results provide helpful information for designing policy interventions with an age- and gender-specific approach.

The second part of this dissertation has been concerned with income diversification patterns of farm households in sub-Saharan Africa. For Western Kenya, the analysis in chapter 4 has confirmed the dichotomy of the non-agricultural sector. As expected, only rich households are able to overcome the entry barriers into high-return activities. Low-return activities, however, are not concentrated among the poor. They are pursued by households across the entire income distribution, possibly reflecting the high risks associated with high-return activities. The chapter also provides evidence that engagement in high-return, but not low-return non-agricultural activities is associated with increased agricultural productivity. Apparently, high-return activities play an important role in triggering positive interactions between different income strategies and may hence contribute to sustainable livelihood success. Chapter 5 has examined the determinants of diversification in Burkina Faso between 1994 and 2003. Diversification into non-agricultural activities is mainly motivated by insurance motives. During the severe drought in 1997/1998, households earned increased incomes from migration and local non-agricultural activities. Many households were also forced to sell livestock to ensure survival. Yet, the patterns of

diversification also reflect structural change as non-agricultural activities tend to become more opportunity-led.

In the last part, this dissertation has looked at international migration in Eastern Europe. Chapter 6 has examined the reasons for reduced labor supply of migrant-sending households in Moldova. The findings do not support the common view that decreased labor market activity is the result of remittances-driven leisure consumption. Instead, the departure of a migrant appears to raise remaining members' productivity in home production. In addition, young adults in migrant families are substantially more likely to pursue higher education. This effect may be due to remittances relieving credit constraints or migration-induced incentives for additional investments in human capital. Chapter 7 has illustrated that the individual risk of falling victim to human trafficking increases with the size of regional migration flows. The reasons are lower recruitment costs for traffickers and, to a lesser extent, more negative self-selection into migration. Accordingly, counter-trafficking efforts should be targeted at areas where emigration rates are high or on the rise and not at regions with no or little emigration. Our results also indicate that awareness campaigns and other measures designed to reduce information asymmetries in the migration process may be a promising way to fight the crime.

Taken together, the findings of this dissertation illustrate that rural households in developing and transition countries have adapted their livelihoods in various ways to cope with shocks or reduce their exposure to risks. These strategies help households to smooth income and consumption in the presence of dysfunctional insurance and credit markets. In the longer run, however, they are likely to perpetuate poverty and reduce economic growth.

Coping strategies such as the depletion of productive assets or the withdrawal of children from school may have detrimental and persistent effects on future welfare and vulnerability to shocks. Households might even be trapped in poverty if a shock pushes their asset base below a critical level that is needed for entry into profitable activities. Risk-management strategies may come at the cost of welfare, too. Diversification into low-risk activities typically not only lowers the variance, but also the level of (expected) income. This situation creates efficiency losses, as households forego potential gains from specialization or engagement in additional high-return activities, hampering the accumulation of productive capital or savings.

Consequently, strategies that protect households from risks should play a prominent role in poverty-reduction and growth policies. Better access to microfinance including index-based insurance and the expansion of public safety nets could be promising ways to improve rural livelihoods.

Danksagung

Zuallererst möchte ich Rainer Thiele und Jann Lay für die engagierte Betreuung danken. Mit ihren konstruktiven Kommentaren und den mir eingeräumten Freiheiten haben sie mein Denken und diese Arbeit nachhaltig geprägt. Ein ganz besonderer Dank gilt auch meinen Koautoren Dennis Görlich, Jens Hagen, Jann Lay, George Michuki M'Mukaria, Ulf Narloch, Rainer Thiele, Natalia Trofimenko und allen voran Christoph Trebesch für die großartigen Diskussionen und die bereichernde Zusammenarbeit. Sie haben in großem Maße zum Gelingen dieser Arbeit beigetragen.

Desweiteren bedanke ich mich bei meinem externen Doktorvater Awudu Abdulai für die hilfreichen Kommentare in den Brown-Bag-Seminaren und die Übernahme des Erstgutachtens. Ebenso danke ich Christian Henning für die Übernahme des Zweitgutachtens.

Einen großen Beitrag zu dieser Arbeit hat auch die stimulierende und angenehme Atmosphäre im Forschungsbereich Armutsminderung und Entwicklung geleistet. Ein großer Dank geht an Matthias Lücke, der mein starkes Interesse am Thema Migration geweckt hat und mit dem ich viele erkenntnisreiche Wochen in Moldawien verbracht habe. Desgleichen möchte ich mich sehr bei Christiane Gebühr für die gewissenhafte Durchsicht und Korrektur der einzelnen Aufsätze bedanken. Dem Institut für Weltwirtschaft danke ich für die sehr guten Promotionsbedingungen und die großzügige Förderung dieser Arbeit.

Außerdem danke ich von ganzem Herzen meinen Eltern Carola und Faidi Omar Mahmoud für die stetige Unterstützung, mit der sie mir so viele wichtige Erfahrungen möglich gemacht haben.

Zusammenfassung

Haushalte in ländlichen Regionen von Entwicklungs- und Transformationsländern sind vielen Risiken ausgesetzt. Unvorhergesehene Ereignisse wie Dürren, Preisfluktuationen oder Krankheiten stellen eine erhebliche Belastungsprobe für ländliche Haushalte dar und führen nicht selten zu Verarmung. Aufgrund des Versagens formaler Versicherungs- und Kreditmärkte im ländlichen Raum haben Haushalte informelle Strategien entwickelt, um Risiken zu streuen und Einkommenseinbußen auszugleichen. Diese Dissertation untersucht inwieweit es ländlichen Haushalten dadurch gelingt, ihre Risiken zu managen und im Falle des Eintritts von unvorhergesehenen Ereignissen ihre Wohlfahrtsverluste zu minimieren.

Der erste Teil der Arbeit befasst sich mit den sozio-ökonomischen Folgen der durch AIDS gestiegenen Erwachsenensterblichkeit in Sub-Sahara-Afrika. In Kapitel 2 wird gezeigt, dass Haushalte im ländlichen Sambia in der Lage sind, ihr Pro-Kopf-Einkommen nach dem Tod eines erwachsenen Familienmitglieds zu stabilisieren. Dieses Ergebnis lässt sich vermutlich auf die Anpassung von Einkommensaktivitäten und der demografischen Zusammensetzung seitens betroffener Haushalte zurückführen. Allerdings gehen diese demografischen Anpassungen zu Lasten der Wohlfahrt anderer Haushalte. Für Haushalte, die Waisenkinder adoptieren, können erhebliche Rückgänge im Pro-Kopf-Einkommen festgestellt werden. Insgesamt halten sich die Einkommenseinbußen aber in Grenzen. Allerdings können Todesfälle die Lebensqualität verbliebener Familienmitglieder auch in nicht-monetären Dimensionen verringern. Wie in Kapitel 3 gezeigt wird, hat das Alter zum Todeszeitpunkt der Eltern langfristige Folgen für die Gesundheit und Bildung von Kindern. Diese Effekte hängen jedoch vom Geschlecht des verstorbenen Elternteils ab. Kinder, deren gleichgeschlechtliches Elternteil überlebt, sind weniger stark von den negativen Folgen für Gesundheit und Bildung betroffen. Dieses Ergebnis deutet darauf hin, dass Risiken im Haushalt nicht im gleichen Maße geteilt werden.

Der zweite Teil der Dissertation beschäftigt sich mit dem Einkommensdiversifizierungsverhalten bäuerlicher Kleinbetriebe in Sub-Sahara-Afrika. In Kapitel 4 wird der dichotome Charakter des nicht-landwirtschaftlichen Sektors in Westkenia untersucht. Die Analyse legt nahe, dass nur reiche Haushalte in der Lage sind, die Eintrittsbarrieren in nicht-landwirtschaftliche Beschäftigung mit hohen Erträgen zu überwinden. Allerdings gehen sowohl arme als auch reiche Haushalte nicht-landwirtschaftlichen Tätigkeiten mit niedrigen Erträgen nach. Dies deutet darauf hin, dass Tätigkeiten, die potentiell höhere Erträge erwirtschaften, möglicherweise auch mit höheren Risiken gepaart sind. Darüber hinaus sind derartige Tätigkeiten aber auch mit höherer landwirtschaftlicher Produktivität verbunden und spielen somit möglicherweise eine wichtige Rolle bei der nachhaltigen Verbesserung der Lebensgrundlagen im ländlichen Raum. In Kapitel 5 werden die Determinanten der Einkommensdiversifizierung in Burkina Faso über den Zeitraum 1994-2003 analysiert. Das Versicherungsmotiv scheint bei der Diversifizierungsentscheidung von großer Bedeutung zu sein. Während der schweren Dürre 1997/1998 bezogen Haushalte mehr

Einkommen aus Rücküberweisungen von Migranten und lokalen nicht-landwirtschaftlichen Aktivitäten. Am stärksten betroffen waren die ärmsten Haushalte, die sich oft gezwungen sahen, ihren Viehbestand zu verkaufen. Allerdings spiegelt das Diversifikationsverhalten auch strukturellen Wandel wider, der vielen Haushalten bessere nicht-landwirtschaftliche Beschäftigungsmöglichkeiten verschafft hat.

Der letzte Teil der Dissertation befasst sich mit den Wohlfahrtseffekten internationaler Migration auf zurückbleibende Haushaltsmitglieder, aber auch auf die Migranten selbst. Der geografische Schwerpunkt liegt dabei auf Osteuropa. In Kapitel 6 wird untersucht, warum moldawische Haushalte mit einem im Ausland lebenden Migranten ihr Arbeitsangebot auf dem heimischen Arbeitsmarkt reduzieren. Die Ergebnisse deuten darauf hin, dass die gesunkene Arbeitsmarktaktivität nicht auf erhöhten Freizeitkonsum zurückzuführen ist, der durch Rücküberweisungen aus dem Ausland finanziert wird. Stattdessen scheint der Weggang des Migranten die Produktivität verbliebener Haushaltsmitglieder in häuslicher Produktion zu steigern. Außerdem gehen junge Erwachsene aus Migrantenhaushalten verstärkt einer Ausbildung oder einem Studium nach. In Kapitel 7 wird berücksichtigt, dass Migration an sich eine mit enormen Risiken behaftete Aktivität ist. Hier werden die ökonomischen Mechanismen, die den Menschenhandel treiben, herausgearbeitet. Menschenhandel (*human trafficking*) bezeichnet eine Situation, in der ein Migrant unter falschen Versprechen angeworben und dann mit Zwangsmitteln genötigt wird, seine Arbeitskraft umsonst oder für wenig Geld zur Verfügung zu stellen. Basierend auf einer speziellen Haushaltsbefragung in Bulgarien, Moldawien, Rumänien, der Ukraine und Weißrussland stellt sich als grundlegendes Ergebnis heraus, dass das individuelle Risiko in die Hände von Menschenhändlern zu geraten mit der Größe regionaler Emigrationsströme wächst. Die Gründe dafür liegen in den geringeren Anwerbekosten für Menschenhändler und in geringerem Maße in negativer Selbstselektion von Migranten.

Zusammen betrachtet zeigen die Ergebnisse dieser Dissertation, dass sich ländliche Haushalte in Entwicklungs- und Transformationsländern in vielfältiger Weise auf ihre risikobehafteten Lebensbedingungen einstellen. Zwar helfen die ergriffenen Strategien den Haushalten dabei, ihre Einkommens- und Konsumniveaus vorübergehend zu glätten. Auf längere Sicht können sie jedoch Armut festigen und das Wirtschaftswachstum mindern. Deswegen sollten Politikmaßnahmen, die Haushalte vor Risiken schützen, einen prominenten Stellenwert in ländlichen Entwicklungsstrategien einnehmen.

Lebenslauf

Toman Omar Mahmoud

Persönliche Daten

Geboren am 1. März 1981 in Dannenberg an der Elbe

Deutsche Staatsangehörigkeit

Schule

1987-1991 Grundschule Loschgeschule, Erlangen

1991-2000 Gymnasium Fridericianum, Erlangen

Universität

2000-2003 B.A. European Economic Studies,
Otto-Friedrich-Universität Bamberg

2004-2005 M.Sc. Economics,
University of Warwick, Großbritannien

2006-2010 Externer Doktorand am
Institut für Ernährungswirtschaft und Verbrauchslehre,
Christian-Albrechts-Universität zu Kiel

seit 2006 Wissenschaftlicher Mitarbeiter im
Forschungsbereich Armutsminderung und Entwicklung,
Institut für Weltwirtschaft an der Universität Kiel