Risk Sharing and Internal Migration

Joachim De Weerdt
EDI, Tanzania

Kalle Hirvonen
University of Sussex
International Food Policy Research Institute

March 2014

Over the past two decades, more than half the population in rural Tanzania migrated within the country, profoundly changing the nature of traditional institutions such as informal risk sharing. Mass internal migration has created geographically disperse networks, on which the authors collected detailed panel data. By quantifying how shocks and consumption co-vary across linked households, they show how a previously reciprocal risk sharing relationship becomes unilateral after migrants move, with migrants insuring their extended family members at home, but not vice versa. This finding contradicts risk sharing models based on reciprocity, but is consistent with assistance driven by social norms. Migrants sacrifice 3 to 5 percent of their very substantial consumption growth to provide this insurance, which seems too trivial to have any stifling effect on their growth through migration.

JEL codes: O12, O15, O17, R23

Keywords: internal migration, risk, insurance, institutions, Africa, tracking data

Acknowledgements: The fieldwork was primarily funded by the Rockwool Foundation and the World Bank, with additional funds provided by AFD, IRD and AIRD through the "Health Risks and Migration" grant of the William and Flora Hewlett Foundation. Kalle Hirvonen gratefully acknowledges the financial support from the Economic and Social Research Council [grant number ES/I900934/1], the Finnish Cultural Foundation and Yrjö Jahnsson Foundation. Stefan Dercon was instrumental in conceptualising this paper and alerted us to the beauty of contrasting full and partial insurance through Equations (1) and (4). We further thank Kathleen Beegle, Marcel Fafchamps, Garance Genicot, Markus Goldstein, Flore Gubert, Cynthia Kinnan, Andy McKay, Imran Rasul, Barry Reilly and seminar and conference participants at BREAD, CSAE, LICOS, NEUDC, Sussex University, Paris School of Economics, FUNDP and UNU-WIDER for useful comments. The usual disclaimer applies. For more information, contact Joachim De Weerdt (j.deweerdt@edi-africa.com) or Kalle Hirvonen (k.hirvonen@cgiar.org).

1. Introduction

If, in the next decades, Africa catches up with the rest of the world, then that will almost certainly coincide with intergenerational mobility out of rural into urban areas and out of agriculture into non-agricultural activities (Lewis 1954; Harris and Todaro 1970). Historically, in both rich developed countries and fast-growing developing countries, this type of migration has moved in lockstep with development and poverty reduction (Collier and Dercon 2009). Recently, China's urban population officially surpassed its rural one: of China's 1.35 billion people, 51 percent lived in urban areas at the end of 2011, rising from less than 20 percent in 1980 (UN, 2012). Furthermore, UNDP (2009) reports that of the one billion migrants worldwide, three-quarters are internal migrants. With international migration open to only very few Africans, we should expect massive internal migration to form a core part of the development process.

The scale of this demographic process is captured in the data that form the basis of this paper, further motivating our focus on internal migration. These data are part of an exceptional panel data set from the Kagera region in Tanzania, spanning nearly two decades of migration and development. The 2010 follow-up survey attempted to trace all 6,353 individuals listed on the baseline 1991/94 household rosters and re-interview them irrespective of their location. Once we exclude the 1,275 individuals who had died by 2010, we are left with 4,996 baseline individuals whose 2010 locations are known.[1] Of those, 45 percent were found residing in the baseline village, 53 percent had migrated within the country, 2 percent to another East African country (primarily Uganda) and 0.3 percent had moved outside of East Africa. This region –

[1] We lack location information on 82 individuals. Because this is after multiple attempts through various sources it is unlikely that these individuals have moved outside of East Africa. Information on such an important, low-occurrence event is unlikely to be hidden.

not atypical of remote rural Africa – is clearly on the move, with internal migration dwarfing international migration.

We attempt to understand how this powerful current of internal migration, which is part and parcel of the modernization process, interacts with a traditional institution like informal risk sharing to shape economic mobility and vulnerability. This is a key question because, as Munshi and Rosenzweig (2006, p. 1230) put it

> [...] a complete understanding of the development process must not only take account of the initial conditions and the role of existing institutions in shaping the response to modernization and globalization, but must also consider how these traditional institutions are shaped in turn by the forces of change.

Our analysis departs from a number of other studies in the migration literature by focusing on consumption instead of transfers. This choice of the outcome variable is motivated by the fact that risk sharing and other economic exchange could happen through a multitude of different mechanisms, of which transfers is just one. Other mechanisms could include looking for a job for someone, employing them directly, providing them with tips, advice or a network link, or providing migration opportunities (Munshi 2003). By analyzing consumption we focus on the joint and final effect of all such mechanisms.

Geographical mobility in rural Tanzania is associated with large income gains (Beegle, De Weerdt, and Dercon 2011). Our data show that despite only minor welfare differences during the 1991-94 baseline survey, those who moved out of the region to other parts of Tanzania have grown roughly twice as rich as those who did not by the time we interviewed them again nearly two decades later. As we are measuring consumption and not income, it is clear that the main beneficiaries of this migration-led growth were the migrants themselves and certainly not their relatives who remained at home. This large divergence is also at odds with the notion that migration is the result of a household level maximization strategy (Stark and Bloom 1985; Rosenzweig and Stark 1989; Grimard 1997).

But did these migrants simply leave and never look back, or did they maintain links with the home community? We investigate this question by exploiting the fact that the 3,314 households interviewed in 2010 are grouped in 817 geographically disperse extended family networks. Using techniques from the risk sharing literature, we quantify how household consumption responds to shocks experienced by other households in the extended family network. We find that while everyone suffers from own shocks, only migrants are affected by shocks to others in the network whereas non-migrants are not. This leads us to reject risk sharing models that are based on reciprocity, including the full risk sharing and limited commitment models. After considering and rejecting a number of longer-run transactional motives (Lucas and Stark 1985; Hoddinott 1994), we argue that the results are very much in line with findings from the diverse literature on social norms (Platteau 2000; Cox and Fafchamps 2007; Burke and Young 2011), where those who move ahead remain obligated to their extended family in the home community.

Household division and migration could be correlated with unobserved characteristics that are themselves associated with risk sharing. We will dedicate Section 8 of the paper to measuring the likely extent of this type of endogeneity and what it means for the interpretation of our empirical tests.

Our analysis speaks to an emerging literature that worries about home communities imposing a stifling 'kin-tax' on the upwardly mobile. Baland, Guirkinger, and Mali (2011) show how people take out costly loans in order to conceal their income, while Platteau (2012) sees migration as a means to escape the implied prying eyes and incessant demands of the kinship group. The kinship poverty trap model of Hoff and Sen (2006) predicts possible resistance from the home communities as they feel threatened by productive forces leaving and severing links with home to escape taxing demands for assistance. Anticipating this, the home community may set up subtle exit barriers, which could lead to below-optimal levels of migration. Jakiela and Ozier (2012) report laboratory evidence from Kenya that women feel obliged to share 4 to 8 percent of the income gains realized in the experiment. In our sample, Tanzanian migrants sacrifice 2.9 to 5.0 percentage points out of a total growth of 108 percent to insure their relatives. We regard this amount as too trivial to exert any constraining effect on migrants.

Section 2 starts with an overview of the full risk sharing model and then discusses what happens when we add participation, incentive compatibility or truth-telling constraints to it. It goes on to explain how social norms could weaken these constraints to the extent that they never bind and allow for the existence of sustained, unreciprocated assistance. Section 3 describes the data and Section 4 distills formal empirical tests from the theories presented in Section 2. These tests are taken to the data in Section 5, where we conclude that households do indeed share risk, but the relationship is not always reciprocal. In Section 6 we then run some additional tests to convince that we are not confusing an apparent lack of reciprocity with pay-offs that will happen beyond the survey period, such a future inheritance. In Section 7 we calculate the cost of this insurance provision for the migrant. Section 8 discusses the endogeneity of household division and migration, Section 9 contains some further robustness checks, and Section 10 provides a concluding discussion.

2. Risk sharing in theory

The full risk sharing hypothesis is based on the idea that the network acts as if it was a single household that maximized utility subject to a joint budget constraint. The model predicts that incomes are completely pooled (according to predetermined weights) and all idiosyncratic income shocks are smoothed through the network (e.g. Altonji, Hayashi, and Kotlikoff 1992; Townsend 1994).

In a simple two-household extended family network, both households derive utility from consumption: $v(c)$. Insurance and credit markets are missing and income (y_s) is uncertain and depends on the state of the world (s). [2] We assume that households live infinitely. [3]

Assuming that households maximize a well behaving utility function [4], the standard utility maximization problem yields a following first order condition:

$$\frac{u'[c_1(y)]}{u'[c_2(y)]} = \lambda = \frac{\omega_2}{\omega_1}, \tag{1}$$

where λ, the Lagrange multiplier, is the marginal utility of income. According to Equation (1), households equate their marginal utilities of consumption in all states of the world. The allocation depends on the Pareto weights ω_1 and ω_2 that are determined by the extended family.

If utility functions follow a constant relative risk aversion function: $u(c) = \frac{c^{1-\psi}}{1-\psi}$, where ψ is a measure of risk aversion [5], the first order conditions for household i at time t become:

[2] To simplify notation, we abstract away savings. This does not affect the main predictions of the model (see, for example, Ligon, 1998 for a characterisation of the full risk sharing model with savings). However, the ability to save may exacerbate the efficiency problems if the key assumptions listed below do not hold (see Ligon, 1998; Chandrasekhar, Kinnan and Larreguy, 2012).

[3] If the time frame is finite, in the absence of altruism, households would not have any incentive for risk sharing in the final period, and as result in T-1, T-2, etc. The assumption of an infinite time frame holds if the new household head inherits from the previous head and maintains the risk sharing contract with same households. See Fafchamps (1992) for an alternative justification for this assumption.

[4] The utility function is inter-temporally separable, strictly increasing but concave ($v' > 0$ & $v'' < 0$).

[5] We assume that the risk preferences within the networks are identical. The implications of this assumption are discussed in Section 4.

$\omega_i c_{it}(y)^{-\psi} - \lambda = 0$. Equating these conditions for the two households, taking logarithms and re-arranging yields:

$$\Delta \ln c_1(y) = \Delta \ln c_2(y). \qquad (2)$$

Equation (2) implies that if full risk sharing takes place, we should not expect to see households within the same extended family growing at different rates. The descriptive statistics in Section 3 document substantial divergence of growth rates between family members who migrated and those who stayed at home. This finding contradicts Equation (2), rejecting full risk sharing.

Furthermore, the model predicts that fluctuations in the individual incomes of members of the extended family network do not influence their individual consumption, once total network resources are controlled for. The consumption of all households in the network goes up, or down, in lockstep with idiosyncratic shocks being perfectly absorbed by everyone in the network. Section 4 will present a formal test for this – Equation (5) in that section – which, once brought to the data in Section 5, will also clearly reject full risk sharing.

The rejection of full risk sharing hypothesis is neither novel nor surprising and emerged as an empirically established stylized fact early on within this strand of literature, being valid across a variety of different contexts (e.g. Altonji, Hayashi, and Kotlikoff 1992; Townsend 1994; Grimard 1997). Most studies, however, find that at least some degree of insurance takes place and explain this theoretically by adding additional constraints (relating to the failure of assumptions regarding perfect information and full commitment) to the full risk sharing model. An important common feature across all these augmented models is that, if the risk sharing contract survives, the ratios of marginal utilities become state contingent, which could allow the share of some members (migrants in our case) to increase over time. This constitutes an inefficiency because agents would be better off were they able to make a credible contract *ex-ante* that equalizes the ratio of marginal utilities over time, taking into account any future changes in income.

Take, for example, the limited commitment model (e.g. Coate and Ravallion 1993; Attanasio and Ríos-Rull 2000; Ligon, Thomas, and Worrall 2002; Kinnan 2012), which augments the full risk sharing model with participation constraints (one for each household):

$$(3) \qquad \sum_{t=1}^{\infty} \beta^t \sum_{s=1}^{S} \pi(y_s)\{v_1[c_{1t}(y_s)]\} \geq \mathrm{u}_A ,$$

where u_A is the expected utility received in autarky, β is the discount rate and π is the probability attached to the state of the world s. Solving the augmented maximization problem yields a following first-order condition:

$$(4) \qquad \frac{u'[c_1(y)]}{u'[c_2(y)]} = \frac{\omega_2 + \sum_{s=1}^{S} \mu_2(y_s)}{\omega_1 + \sum_{s=1}^{S} \mu_1(y_s)} .$$

where μ_1 and μ_2 are the Lagrange multipliers attached to the participation constraints. Now, as can be seen from Equation (4), if the participation constraints bind, the ratio of marginal utilities becomes state contingent. This model provides a framework for thinking of risk sharing and migration. In particular, it could explain why migrants become so much richer than their extended family members at home (see Section 3). The migrant's rising income creates a threat of reneging, so her share of total resources needs to be increased until she is exactly indifferent between leaving and staying. Indeed, another way to give meaning to Equation (4) is that Pareto weights are re-established post-migration.

Other frictions have analytical consequences that are similar to those of limited commitment.[6,7] If households cannot monitor other network members, the problem of free riding emerges. In moral hazard models (Lim and Townsend 1998; Kinnan 2012), the full risk sharing model is augmented with incentive-compatibility constraints. The *ex-ante* information asymmetry leaves the extended family to balance effort and insurance; migrants, whose incomes rise on average, are motivated to exert effort by rewarding them with higher consumption. Finally, if there is imperfect information about the realized incomes, households may have an incentive to misreport their incomes to avoid payments or even claim transfers from other households. In hidden income models (Townsend 1982; Fafchamps 1992; Kinnan 2012), the maximization problem is augmented with truth-telling constraints that require that households will not gain from misreporting. To encourage truthful reporting, those whose incomes rise over time (migrants in our study) are allowed to enjoy a larger share of the consumption cake.

Our results are not, however, consistent with all predictions of these basic models of limited commitment and other frictions. A further empirical consequence of these models is that, controlling for aggregate resources, agents' consumption co-moves with their own individual shocks, as well as with shocks of their network members: an idiosyncratic income shock will get partly absorbed by the network and partly by the individual. Furthermore, there should be no subgroups of households – delineated along exogenous or endogenous characteristics – that are completely unresponsive to the shocks of others, as that would violate the reciprocity assumption that underlie these models. We formulate this as a testable hypothesis in Section 4 with regression Equation (6). In Section 5 we will show that our data contradict this

[6] Distinguishing which of the three models of constrained insurance explains our data best is beyond the scope of this paper. See Kinnan (2012) for such an exercise with data from rural Thailand.

[7] These frictions have important implications for the degree of risk sharing as highlighted by experimental studies, such as Barr and Genicot (2008), Chandrasekhar, Kinnan and Larreguy (2011, 2012) and Charness and Genicot (2009).

prediction: migrants respond to stayers' shocks, but stayers do not respond to migrants' shocks. This finding leads us to the reject basic risk sharing models with friction.[8]

One way to understand this lack of reciprocity in the data is by introducing social norms in the model. Redistributive values may have been instilled since childhood and carefully nurtured through oral transmission, rituals and ceremonies in which the importance of the kinship group is strongly emphasized (Lévi-Strauss 1969). Remittances and other forms of assistance may buy social prestige, political power or serve to perpetuate subordination (Platteau and Sekeris 2010; Platteau 2012). In the risk sharing literature, social norms have been seen as the glue that keeps the risk sharing contract from breaking apart by alleviating enforcement and information problems (Stark and Lucas 1988; Fafchamps 1999; Foster and Rosenzweig 2001). Theoretically this can be modelled as subjective satisfaction that individuals receive from participation. [9] The satisfaction can stem from the fulfillment of obligations and the avoidance of social sanctions, such as guilt, shame or ridicule, or fear of witchcraft. It can also include altruism, which we do not attempt to distinguish from social norms. Social norms could weaken the constraints to risk sharing to the extent that they never bind and allow for the existence of sustained, unreciprocated transfers, as documented, for example, for Paraguay by Schechter and Yuskavage (2011) and for Tanzania by De Weerdt and Fafchamps (2011). Below we will find evidence of such unilateral relations and argue that this is consistent with risk sharing motivated by social norms.

[8] By referring to basic models of limited commitment, we mean the simple version of the model presented in this paper, which assumes no saving, full observability of income and reciprocity only possible by means of insuring shocks. More sophisticated versions of these models may have different predictions. We devote Section 6 to a discussion of alternative forms of reciprocity, not incorporated in these basic models.

[9] In the context of limited commitment, we can re-write the right-hand side of Equation (3) as $u_A - A$, where A captures such satisfaction (Fafchamps 1999; Foster and Rosenzweig 2001; De Weerdt and Fafchamps 2011).

A recent empirical literature relying on experimental design also highlights the importance of these forces. Chandrasekhar, Kinnan, and Larreguy (2011, 2012) find that in the presence of hidden income and limited commitment, social proximity between the risk sharing partners increases the amounts transferred. The field experiments of Leider et al. (2009) and Ligon and Schechter (2012) show that altruism is more important than repeated interaction in determining the size of the transfer. Attanasio et al. (2012) conclude that high pay-off assortative grouping on risk-attitudes only occurs among close friends and relatives in their risk pooling experiment in Colombia.

3. Data and descriptive analysis

Kagera is a region in the north-western part of Tanzania. A large part of Lake Victoria is contained within this region and it shares a border with Burundi, Rwanda, and Uganda. The region is overwhelmingly rural and agricultural production is the most important source of income, with more than 80 percent of the region's economically active population engaged in it (URT 2012). Bananas, beans, maize, and cassava comprise the main food crops while coffee, tea, and cotton are important cash crops. Recent years have seen a rise in improved banana varieties and sugar for use as cash crops. According to the 2012 census, the region has a population of roughly 2.5 million people (URT 2013).

The Kagera Health and Development Survey (KHDS) was originally designed and implemented by the World Bank and the Muhimbili University College of Health Sciences. It consisted of 915 households from 51 villages that were interviewed up to four times from autumn 1991 to January 1994.[10] The KHDS-2004 survey aimed to re-interview all individuals that were ever interviewed in the baseline survey and were alive in 2004. This effectively meant that the original household panel survey turned into a panel of individuals. A full

[10] See World Bank (2004).

household questionnaire was administered in a household where a panel respondent was found residing. Due to household dynamics, the sample size increased to more than 2,700 households.[11] The second KHDS follow-up was administered in 2010 with this time more than 3,300 households interviewed.[12]

Although KHDS is a panel of individuals and the definition of a household loses meaning after 10-19 years, it is common in panel surveys to consider re-contact rates in terms of households. Excluding households for which all previous members were deceased the KHDS 2004 field team managed to re-contact 93 percent of the baseline households. In 2010, 92 percent of the initial households were re-contacted. Taking into account the long, 10 or 16 year periods between surveys, the attrition rates in KHDS-2004 and KHDS-2010 are extremely low by the standards for such panels (Alderman et al. 2001).

This paper exploits the fact that the survey includes all tracked split-offs from the original household and contains particularly rich information on the current links between them. The 2010 sample contains 3,314 households, originating from 816 initial households. The average baseline household spawned 4.1 households by 2010, out of which 1.8 were non-migrant and 2.3 were migrant households. Approximately 3 percent of the initial households (99 households) did not have any split-offs. In what follows we will refer to these networks as extended family networks.

In this paper we will define a migrant as anyone who has moved out of the baseline village. [13] By this definition 55 percent of the household sample is considered migrant. Details on where they moved to are given in Figure 1.

[11] See Beegle, De Weerdt and Dercon (2006).
[12] See De Weerdt et al (2012).
[13] Our results are robust to alternative migrant definitions, such as also defining households that moved to a nearby village as non-migrant households.

[Figure 1 here]

These internal migration flows described above are associated with structural transformation.[14] Table 1 shows that out of the 1,850 migrant households, only one-third reported agriculture as their main income generating activity. For the 1,460 non-migrant households this is 65 per cent. More than 25 per cent of the migrant households engage in informal or formal wage employment and 11 per cent are self-employed (non-agriculture). Furthermore, migrants who move farther from the baseline village are less likely to engage in agriculture and more likely to be in wage employment.

[Table 1 here]

Table 2 provides an overview of the reasons for leaving the baseline village. More than one-third of the female respondents but none of the male respondents cited marriage as the reason for migrating, which is what one would expect in a culture with patrilocal marriages. Less than 15 percent of the female respondents reported that they left because of work. In contrast, almost 45 percent of the male migrants reported to have moved because they had found work or went looking for work.[15]

[Table 2 here]

The consumption data originate from extensive food and non-food consumption modules in the survey, carefully designed to maintain comparability across survey rounds and controlling for seasonality. The aggregates are temporally and spatially deflated using data from a price

[14] This is also documented by Christiaensen, De Weerdt, and Todo (2013) who use the same data to study the role of urbanization and diversification in poverty reduction.
[15] Despite these differences in migration motives across the two gender groups, we do not find any statistically significant differences in risk sharing provision between male and female migrants. Results are not reported but available upon request.

questionnaire included in the survey. Consumption is expressed in annual per capita terms using 2010 Tanzanian shillings. [16]

Table 3 provides the summary of the consumption and poverty developments of the panel respondents with respect to their 2010 location. On average, consumption levels in the sample almost doubled over 19 years. Individuals who stayed in their community saw their consumption increase by more than 40 percent. Consumption growth for migrants was much higher: those who left Kagera saw their consumption nearly triple over the same two decades. The poverty statistics tell the same story: nearly all respondents who left the region managed to escape poverty, while poverty reduction among non-migrants was more modest. These descriptive statistics, which reinforce the results reported in Beegle, De Weerdt, and Dercon (2011), also contradict the full risk sharing model, but are consistent with, for example, the limited commitment model, as characterized by Equation (4): as migrants' income goes up they are allocated a larger share of the total network resources to ensure they do not renege on the contract. Section 5 will show, however, that the reciprocity prediction of the basic limited commitment model is violated, leading us, ultimately, to reject this model.

[Table 3 here]

After moving, migrants remain linked to extended family members at home: 90 percent of the migrants report that they communicated with a non-migrant network member in the 12 months preceding the survey. Migrants who maintained some form of communication experienced an average consumption growth of 110 percent, while those who did not grew by 88 percent. [17] This difference is statistically significant at the 1 percent level. The severing of the most basic

[16] Using adult equivalent units as the denominator instead of household size produces almost identical results across all specifications.
[17] The mean consumption growth among those who maintained contact was 394,679 TZS and among those who severed links 286,991 TZS.

links does not seem to be associated with higher consumption growth; if anything, the reverse is true.

We use data from shock modules administered in 2004 and 2010. During both of these rounds, the panel respondents were asked to consider each year between the survey rounds and indicate whether a particular year was, in economic terms, 'Very good', 'Good', 'Normal', 'Bad', 'Very bad'. For each 'Very bad' response, the respondents were asked to provide the main reason for the hardship. We consider each 'Very bad' response as an economic shock. More than 60 percent of the panel respondents reported experiencing at least one such shock between 1994 and 2009.

Table 4 provides an overview of the shocks experienced. Most frequently reported economic shocks were death of a family member, serious illness and poor harvest due to bad weather.

[Table 4 here]

The shock data were collected at the individual level – in particular for each person on the 2010 roster who also appears on the original 1991/94 rosters. Since our focus is to examine the role of shocks on household consumption, the data had to be reformatted from the individual to the household level. [18] If at least one individual in the household reported to have experienced a shock, we interpret it as a household level shock. We should also exclude shocks that occurred before the households split. Fortunately, we know the year in which the

[18] We repeated the complete analysis of the following sections using individual level data and find it does not change the conclusions.

respondents moved to their 2010 location, allowing us to include only shocks that occurred at least one year after this move. [19]

Furthermore, some of the shock categories are problematic to our network analysis. Mortality shocks may trigger inheritance flows within extended families. As such, a negative shock in one household may actually be a positive income shock in another household. A similar problem arises with the loss of remittance shocks, if these capture the loss of transfers from a household within the same extended family. We therefore exclude these two shock categories from our final shock variable.

Another worry is that because of the self-reported nature of the shock variable uninsured shocks may go unreported or reports may differ along unobserved characteristics. To address this, in Section 7, we extend our analysis with an alternative shock measure based on historical rainfall data collected from 212 weather stations in Kagera and the migration destinations of our panel respondents.

Finally, there are 439 households that belong to a network that contains only non-migrants or only migrants. As our interest lies in the role of migration in risk sharing, we cannot use these households for empirical identification for risk sharing between migrant and non-migrant households. These households are therefore dropped from the final sample. Table 5 presents the summary statistics for the final sample of 2,349 households by 2010 migration status. Two facts from this table are worth noting here because we will refer to them later on in the manuscript. First, transfers from migrants to non-migrants are 55 per cent higher than those

[19] This means that for households that remained in the baseline village we consider shocks that took place between 1994 and 2009. An alternative strategy would be to only use shocks that occurred after these household lived with *any* other network household member. Applying this strategy does not, however, change the conclusions.

from non-migrants to migrants. Second, travel costs related to the migrant's move amount to roughly 3 per cent of the migrants annual consumption in 2010.

[Table 5 here]

4. Econometric strategy

We begin the econometric analysis by testing the full risk sharing hypothesis for those extended family networks that contain both migrant and non-migrant households. The difference in logged per capita consumption between 2010 and the baseline ($\Delta \ln c_{ij}$) for household i in extended family j is formally modeled as:

$$\Delta \ln c_{ij} = \beta s_{ij} + x'_{ij}\gamma + \alpha_j + \varepsilon_{ij} \,, \tag{5}$$

where s_{ij} has a value 1 if the household experienced a shock in 1994-2009, or if a migrant household, after migrating to its current location. The term x_{ij} is a vector of household characteristics in 2010 capturing the characteristics of the previous household members [20] such as the number of previous household members in the 2010 household, the age of the oldest and the education (in years) of the most educated previous household members in the household. We also include dummies capturing their relationship to the 2010 household head and their marital status. [21] The term α_j represents the network fixed effect that, among other things, controls for the change in aggregate network resources and ε_{ij} is the error term. The inclusion of the network fixed effects means that we compare the impact of shocks between

[20] Previous household member refers to a person interviewed at the baseline in 1991/94.
[21] To address concerns about some of these 2010 household characteristics variables being potentially endogenous, we run all main regressions again, but drop each of these control variables in turn. We find the shock and network shock coefficients remain stable across all such re-specifications.

the households originating from the same initial household. As such, the full risk sharing model presented earlier requires that $\beta = 0$.

The rejection of the full risk sharing model using Equation (5) implies either that the risk sharing arrangement is not efficient – or that the network does not engage in risk sharing at all. The rejection may also stem from the violation of the assumption that the risk preferences are identical within the network (Chiappori et al. 2011; Schulhofer-Wohl 2011; Mazzocco and Saini 2012). [22] To explore the existence of reciprocal risk sharing, we assess whether household per capita consumption growth is responsive to shocks experienced by other households in the same extended family. This test builds on Equation (5). We drop the network fixed effects and replace them with baseline village fixed effects (θ_v) and network characteristics (w_j) comprising the number of migrant and non-migrant households in the network and variables capturing characteristics of the initial household, such as its demographic composition, the household head's characteristics, including education, gender, age and the quadratic of age. Since growth rates are likely to be related to the initial levels of income (e.g. Dercon et al. 2009), we also include (logged) per capita consumption at the baseline ($\ln c_{j,1991}$). The network shock variable, z_{ij}, measures the number of households affected by an income shock. The household's own shocks are excluded from this variable. The partial risk sharing specification is formulated as:

$$(6) \qquad \Delta \ln c_{ij} = \beta s_{ij} + \delta z_{ij} + x'_{ij}\gamma + w'_j\vartheta + \gamma \ln c_{j,1991} + \theta_v + \varepsilon_{ij}.$$

[22] In a context of heterogenous risk preferences, Pareto-efficient contract allocates more aggregate risk to less risk-averse households. As demonstrated by Schulhofer-Wohl (2011), Chiappori et al (2011) and Mazzocco and Saini (2012) this would lead to a downward bias in β in Equation (5). The standard full risk sharing test is then biased against the null-hypothesis of full-risk sharing.

Equation (4) implies that $\beta < 0$ and $\delta < 0$: individual consumption is negatively affected by own income shocks (idiosyncratic shocks are not completely smoothed) and negatively affected by income shocks to others the network (some of the individual shock gets absorbed by the extended family).

Furthermore, the basic partial risk sharing models discussed in Section 2 assume reciprocity, which can be tested by interacting all right hand side variables of Equation (6) by an indicator variable[23] G_{ij}, representing some exogenous or endogenous grouping of the households within the network:

$$\begin{aligned}(7)\qquad \Delta \ln c_{ij} &= \beta_1 s_{ij} + \delta_1 z_{ij} + x_{ij}'\gamma_1 + w_j'\vartheta_1 + \gamma_1 \ln c_{j,1991} + \theta_v \\ &+ G_{ij}(\beta_2 s_{ij} + \delta_2 z_{ij} + x_{ij}'\gamma_2 + w_j'\vartheta_2 + \gamma_2 \ln c_{j,1991} + \theta_v) + \varepsilon_{ij}.\end{aligned}$$

With reciprocity across the groups we would expect $\delta_1 < 0$ and $\delta_1 + \delta_2 < 0$.[24] Alternatively, we can conduct two separate regressions using Equation (6), one for each value of G_{ij}, and test that $\delta_n < 0$ for each n group. The difference in the magnitude of these two coefficients is represented by δ_2 in the pooled regression.

Our interest in this paper is to test for reciprocity between migrant and non-migrant households, which is an endogenous grouping of households. As noted in the theory section, whether or not groups are delineated along endogenous or exogenous characteristics does not matter for the validity of our test of the partial insurance model. It does, however, bear on how we interpret the results. Because selection into migration is unlikely to be random and any differences in reciprocity as measured through Equations (6) and (7) may be caused by migration or by some unobserved characteristics that differ between migrant and non-migrant

[23] Or vector in case there are more than two groups.

[24] Note that the level effect of migration (G_{ij}) is absorbed in the migrant specific village fixed effect component ($G_{ij}\theta_v$).

households, or by some combination of both. As a result, these regressions do not allow us to say whether migration is causally responsible for the migrant taking on the role of insuring sedentary extended family network members, or whether the effect is driven by unobservables. In particular, we cannot make any statements about what would have happened if migrants had stayed home or the stayers had migrated. It is possible that in this parallel universe roles would have switched (migration is causally responsible) or not (it is driven by the unobserved differences between migrants and non-migrants). We will dedicate Section 8 to a discussion of the extent to which selection is at play and what that means for the interpretation of our results. This same section will also address a similar concern about the existence of unobserved characteristics that simultaneously influence the propensity for households to divide and the level of risk sharing post-division.

Finally, the baseline per capita consumption variable in Equation (6) raises a concern about endogeneity. The error term ε_{ij} could be correlated, for example due to measurement error, with the lagged consumption variable on the right-hand side of Equation (6). This would then bias the estimate measuring the impact of the lagged consumption but it may also affect other coefficients. Fortunately, we can think of a credible instrument that allows us to assess this possibility. Rainfall is one of the main inputs in agricultural production in Kagera and poor rainfall (i.e. droughts) can have serious consequences for incomes. Excess rains are less of a problem due to the focus of the production on tree crops and also because the terrain is relatively undulating. The region has two rainy seasons, a long rainy season usually between March and May and a short rainy season usually between October and December. The agricultural production takes place during these seasons. Therefore, we employ average monthly z-score deviations of rainfall during the two rainy seasons preceding the interview and truncate the positive rainfall deviations to zero. [25] Rainfall during the agricultural production is expected to influence consumption through income fluctuations but is unlikely to be correlated with the potential measurement error in the per capita consumption variable. The baseline village fixed effects (θ_v) in Equation (6) wipe out the level effects of rainfall in the first stage regression. Therefore, exploiting the fact that rainfall shocks will affect different types of households in different ways, we interact the rainfall variable with head's gender, age and education yielding a total of three instruments.

5. Results

We begin by testing the full risk sharing model described above. Column 1 in Table 6 provides the results for the base specification of Equation (5) with network fixed effects (NFE). The control variables capture the characteristics of the previous household members, including their position within the 2010 household. The signs of the control variables are *a priori* correct. For example, education has a positive impact on consumption growth, while households with widowed or divorced previous household members experience lower consumption growth than others within the same extended family network.

The statistical significance of the shock coefficient, despite the inclusion of NFE, reveals that shocks are not insured within extended families. Households that experienced a shock had 14 percentage points lower consumption growth, on average and *ceteris paribus*, than households from the same extended family who did not experience a shock. The emergence of this wedge in the face of a shock implies a clear rejection of the full risk sharing model in the extended family networks in this study.

[25] Beegle, De Weerdt and Dercon (2008) employ a similar instrumental variable approach for their lagged consumption variable in assessing the long-term impact of adult deaths on consumption growth in Kagera.

[Table 6 here]

In column 2 we drop the NFE and replace with network characteristics, such as the number of migrant and non-migrant network members (which together control for network size and composition) and the wealth and demographics of the baseline household from which the network is formed. We also include baseline village fixed effects. The size of the shock coefficient is nearly identical to the one obtained with NFE, giving confidence in the network level controls we use in Table 7 below for the test of reciprocal risk sharing.

Finally, column 3 provides the Two-Stage Least Squares results that address the potential endogeneity problem arising from the inclusion of the initial logged per capita consumption variable. The first stage regression results and the standard IV-diagnostic tests are presented in Table A1 in the Appendix. The included instruments show how households headed by older and more educated males enjoy higher baseline consumption. The excluded instruments are zero-truncated negative z-score deviations of rainfall interacted with the household head's age, education and gender. They show that the positive level effects of each of these three household head characteristics are attenuated with the inclusion of negative rainfall shocks. The Cragg and Donald (1993) test yields 19.1 indicating that our instruments are relevant. Comparison with the critical values provided in Stock and Yogo (2005) suggests that the bias of our IV-estimate is less than 5 percent of the OLS estimate. The Hansen (1982) J-test provides a p-value of 0.570. Thus, the null hypothesis of zero correlation between the instrument and the error term is upheld at conventional levels. The shock coefficient and the standard error from the 2SLS estimates are almost identical to those from OLS, indicating that the potential endogeneity of the logged per capita baseline consumption has a negligible influence on the shock variable. [26] In the light of this, we use the more efficient OLS method to make inferences in the remainder of the text.

Next we test whether any risk sharing takes place in these networks. As discussed earlier, we replace the NFE with network characteristics and baseline village fixed effects and augment the specification with the network shock variable. The first column in Table 7 estimates Equation (6) on a pooled sample and shows $\beta < 0$ and $\delta < 0$, consistent with partial risk sharing.

We now turn to the issue of reciprocity. The second and third columns in Table 7 estimate Equation (6) on the subsamples of migrants and non-migrants, respectively. The own shock coefficient, β, remains negative and is statistically different from zero in both columns. For migrants, the network shock coefficient, δ, is negative and highly significant. These network shocks have a sizeable impact on migrant households' consumption: on average, a shock in one household in the network resulted in a drop of 5 percentage points in consumption growth. As shocks are not correlated within the extended family networks (the intra-class correlation coefficient equals 0.017 with a standard error of 0.016), this finding reveals that migrants insure other households in their extended families. Non-migrant households, on the other hand, do not appear to be affected by the network shocks. The point estimate of δ is nearly zero and insignificant. Finally, column (4) in this table presents the estimates of the interacted coefficients of Equation (7) – with G_{ij} set to our migration dummy – on a pooled sample, providing a formal test of the difference between the coefficients of column (2) and (3). Of

[26] A C-test (see e.g., Hayashi 2000, p. 220-221) with one degree of freedom yields $\chi 2$-test statistic of 0.104 (p=0.748). Thus, we cannot reject the null hypothesis that the lagged consumption variable is exogenous.

note in this column is that δ differs significantly by migration status. These results suggest that the risk sharing arrangement is not reciprocal.

[Table 7 here]

In order to investigate this further, we decompose the network shock variable into shocks in non-migrant and migrant households. The first variable measures the number of non-migrant households that experienced a shock in the extended family. The second network shock variable measures the number of migrant households affected by shocks. As before, the household's own shocks have been excluded from these variables. Table 8 presents the regression results. Migrants are susceptible to shocks affecting other migrant and non-migrant households within their extended family network, while non-migrants are sensitive to neither. On average, a shock in one non-migrant household in the network leads to a drop of 5.5 percentage points in migrant household's consumption growth. Shocks in other migrant households have a negative effect of similar magnitude on a migrant's consumption than shocks experienced in stayer households but this coefficient is not statistically significant at a conventional level (p=0.127).

[Table 8 here]

We conclude that migrant households are partially and unilaterally insuring households that stay behind. This lack of reciprocity violates the predictions of the reciprocity-based models (without a social norms term). Because, on average, migrants are nearly twice as rich as those who remained at home, these findings are consistent with reciprocity-based models augmented with a social norms term, which attenuates the participation, truth-telling or incentive compatibility constraints.

6. Other transactional insurance motives

An alternative explanation to the observed lack of reciprocity could be that migrants insure non-migrants in exchange for other benefits. By concentrating on consumption differences we have considered only current pay-offs from any risk sharing arrangement. It is quite possible that the benefits are still to accrue to the migrant in the more distant future. Lucas and Stark (1985) mention that there could be exchange motives for insurance provision relating to the desire for non-migrants to look after local assets, the intention to return home and the aspiration to inherit. In a context that lacks technology to allow future income to be consumed now, we could confuse unilateral insurance with postponed reciprocity. Fortunately, the KHDS questionnaire is particularly rich and we are thus able to explore some of these issues.

The questionnaire asks each migrant about asset holdings in the baseline village. As our outcome variable is consumption growth we cannot use these asset holdings as explanatory variables: current wealth is surely endogenous to growth in wealth. We attempt to circumvent this problem by looking at the share of assets in the current portfolio that are located in the village. While it remains possible that portfolio composition is endogenous to consumption growth, we believe the results are informative enough to report.

About 28 percent of migrants have assets in the baseline village and 25 percent of migrants own land in the baseline village. For land we have exact area measurements, but not monetary values. If migrants engage in risk sharing with those who remain at home for the purpose of maintaining land and ensuring their continued entitlement to the land (which is important in a country with few formal land deeds), then we would expect more responsiveness to network shocks from people with a larger share of their land holdings in the baseline village. The first column of Table 9 explores this. As before, the dependent variable is logged per capita consumption growth. We interact the non-migrant network shocks with a variable measuring the share of the land in the baseline village. The coefficient on this interacted variable turns out insignificant implying that the share of land in the baseline village neither increases nor decreases the insurance provision.

In the second column in Table 9 we interact the non-migrant network shock variable with the length of the migration spell. Following Dustmann and Mestres (2010), we argue this to be a measure of the permanence and success of the move and an inverse measure of the return likelihood. We find that the duration of the migration spell does not have any impact on migrant's insurance provision. This also holds when we use non-linear versions of the migration duration in the form of a piecewise linear spline.

The third column in Table 9 investigates whether the expectation to inherit is a plausible motive for unilateral insurance. Since traditional law excludes women from inheriting land, a household can only inherit land if it has male members whose parents own clan land in the baseline village. [27] We construct a variable that captures these households and also control for cases where the parents no longer live in the baseline village in order to isolate the inheritance story from the effect of having parents in the baseline village. Nearly 42 percent of the migrant households have parental clan land holdings waiting for them in the baseline village. By interacting the non-migrant network shocks with a parental clan land holdings dummy, we find that that these households are no more (or less) engaged in insurance provision than households that do not expect to inherit land.

[Table 9 here]

Another transactional motive that could be consistent with the regression results is that non-migrants pay insurance premiums to migrants in return for their continued insurance provision;

[27] Land in this context normally belongs to the clan and the purpose of the traditional law is to keep in the clan (see De Weerdt 2010).

the rich, in effect, sell insurance to the poor (Fafchamps 1999; Genicot 2006) and regressive transfers result. This does not, however, fit with the descriptive statistics from Table 5 that show non-migrants are net recipients of transfers.

Finally, the empirical patterns we describe could occur if migrants receive loans for their move from those who remain at home and then pay back these loans state contingently post-migration. We would still need to see evidence of some future benefit in order to make that repayment promise credible in the absence of social norms.

7. The cost of insurance provision

Does the migrant incur a significant cost for providing this unilateral insurance? From Table 8 we observe that for each shock in the extended family network at home there is a drop of 5.5 percentage points in the migrant's consumption, which appears to be a permanent deviation from the growth curve. The average migrant has 0.53 network shocks of non-migrants, resulting in an implied overall consumption growth penalty of 2.9 percentage points, on average, over the 19-year period. Over this same period, the average consumption growth among migrants was 108 percent, implying that insurance constituted an average annual growth penalty of around 0.077 of one percentage point (reducing average annual growth roughly from 3.93 percent to 3.86 percent). [28] Put another way, migrants share about 2.7 percent of their very substantial growth by insuring people at their original location. [29]

This is a lower-bound estimate because we cannot exclude the possibility that we are only measuring a subset of relevant shocks. First, if shocks are self-reported then respondents may fail to mention those that were effectively insured. Second, the extended family network in the home community may extend beyond the networks as defined in our data. Fortunately, the

[28] We use geometric (rather than arithmetic) means to calculate the average annual growth rates.
[29] The 95%-confidence interval is [0.01, 5.40].

survey provides an alternative shock measure, which is community-wide and not self-reported. We have historical rainfall data from the Tanzanian Meteorological Agency for gauges in 212 weather stations in Kagera and the migration destinations in our sample. In a first step, each household is linked to all rainfall stations within a 100 km radius. Next, a monthly rainfall figure is calculated, for each household, by weighing each monthly rainfall reading with the inverse of the distance of the rainfall station where it was recorded to the household in question.

The mean distance to the nearest rainfall station is 10 km and the median is 23 km. For each household we can calculate average monthly z-score deviations of rainfall during the two rainy seasons, in relation to the 30 year average (1980-2010) for that village. Rainfall shocks are then constructed by truncating the positive yearly average rainfall deviations to zero. We calculate a non-migrant household's own shock as the most negative shock in the 1994-2009 period. The first row in Table 10 shows that rainfall shocks are important in determining consumption growth, with every standard deviation decrease in (negative) rainfall deviation causing consumption growth to decline by 12 percentage points for migrants and 18 percentage points for non-migrants.[30]

[Table 10 here]

Knowing that rainfall shocks drive the incomes of both stayer and migrant households, we can use them as an alternative network shock indicator. We replace the network shock variable with the baseline village rainfall shock variable in Equation (6). For migrant households, this rainfall shock is constructed as the most negative rainfall deviation in the baseline village after the migrant left. For stayer households, we take the most negative rainfall deviation among

[30] Out of the 1,270 migrant households 77 per cent report to derive at least some income from agricultural production and 73 per cent own farm land. The results reported in Columns 1 and 2 of Table 10 are robust to constraining the migrant sample to these households.

the migrant household locations, after the migrant arrived at the current location. Due to the covariate nature of the rainfall shocks we cannot use the baseline village fixed effects in the stayer household specification. Similarly, there is little variation in the baseline village rainfall shock variable among migrants that originate from the same village. We therefore replace the baseline village fixed effects with baseline district fixed effects. [31] Column 1 reports the results for the migrant households. We see that after the migrants leave their consumption remains responsive to rainfall shocks at the baseline village, amounting to 5.0 percentage points out of their growth, on average. [32] With this as the upper bound effect, we conclude that migrants share between 2.7 and 4.6 percent of their growth with home communities through insurance provision.

Column 2 reports the corresponding results for the non-migrant households. Consistent with the results presented in Section 5, we see that non-migrants are not affected by rainfall shocks that take place in migrant households.

There are two related studies that have looked at similar issues in a more controlled setting. Jakiela and Ozier (2012) find that women in a laboratory setting in Kenya purposefully reduced their income in order to keep it hidden. They acted as if they were expecting to be taxed any observable winnings to be taxed at around 4 to 8 percent. Ambler (2012) reports that El Salvadorian migrants living around Washington DC remit 5 percent more of a windfall income if they are told that the organizers of the experiment will inform potential recipients at home about it. One important difference between these experiments and our observational data is that they look at the short-run reactions to windfall incomes, while we study the long-run

[31] Households group into 51 baseline village and 6 baseline districts. For migrants, the baseline fixed effects exploit the variation arising from the fact that migrants leave at different times. This yields almost an identical coefficient but due to the limited variation remaining in the data, the coefficient is not significant at a conventional level (p=0.155).

[32] The 95 percent confidence interval ranges between 0.7 and 10.7 percentage points.

consequences of reactions to actual income shocks. Another difference is that they look at how people change remittance behavior when going from actual belief sets to full information, or how much they would be willing to sacrifice to avoid giving full information. We look at the effect of shocks within real-world belief sets. The advantage of the experimental approach of the two studies, however, is to be able to deal with endogeneity much better. It is this issue we turn to next.

8. Endogeneity of household division and migration

It is very plausible that the nature of household division and migration is related to unobserved variables associated with risk pooling.[33] For example, households that are more unequal or uncooperative may be more likely to split or uncooperative individuals more likely to migrate.

Our data are not experimental and their real-world richness comes at the cost of not being able to provide iron clad proof of causality. The primary goal of this paper is to document what happens to a traditional institution, like informal insurance, in a society that modernizes and is characterized by massive internal migration. While the endogeneity of household division and migration does not diminish the validity of tests of the risk sharing theory, nor of the stylized facts, it does come to bear importantly on how we give meaning to the results. It also relates to their external validity. For example, if migration is *not* causally responsible for the empirical patterns described in the previous sections then that should increase one's degree of skepticism about these same patterns applying in other settings where the selection process into migration is different.

Fortunately, we are able to exploit other survey rounds in order to speak to the causality issue and exclude that certain forms of unobserved heterogeneity explain the results. The purpose

[33] By household division we refer to an event where a household splits into two or more households. Migration is then one, special, form of household division.

of this section is to be very specific about which remaining types of endogeneity could compete with causality to explain the results.

We start off by looking at the correlates of household division, and migration. Household division or out-migration could be related to high inequality within the household (Foster 1993; Foster and Rosenzweig 2002), which may then be correlated with the risk sharing arrangement after the household splits – or produces a migrant. The decision to split or to migrate will then be related to the level of risk sharing provided in the baseline household, complicating the interpretation our results. We explore this possibility by looking at whether unequal households spawn more splits or migrants, than other households. Following Dercon and Krishnan (2000) we use the Body Mass Index (BMI) as our measure of intra-household inequality and risk sharing. One can think of the head of the household, or a planner, allocating consumption shares (the Pareto weights) to its household members. Unequal division of the consumption cake would then result in an unequal distribution of the BMI within the household. [34]

In Equation 8, the outcome variable is the number of split-off households, or migrant households, the baseline household had generated by 2010. Building on the specification in Foster and Rosenzweig (2002), we regress this on the mean (B_i^{mean}), variance (B_i^{var}) and maximum (B_i^{max}) of BMI z-score [35] while controlling for a host of other initial household level variables (X_i') and village fixed effects (θ_v):

[34] The advantage of using BMI, rather than consumption, is that it is measured at the individual level and can also be considered as durable good (Dercon and Krishnan 2000).

[35] The BMI z-score provides a more comparable measure of the body-mass across household members of different age and sex (see e.g., Cole et al. 2007). We use the US 2000 NCHS/CDC as the reference population (Kuczmarski et al. 2002). To further control for the age and sex specific developments in the BMI *z-score* we include variables capturing the number of males and females in the following age groups: 0-5 years, 6-15 years, 16-60 years and 61+ years. Together these variables also control for the household size.

(8) $$N_{i,2010} = \xi_1 B_i^{mean} + \xi_2 B_i^{var} + \xi_3 B_i^{max} + X_i'\vartheta + \theta_v + \varepsilon_i.$$

Table 11 presents the OLS estimates regressing $N_{i,2010}$: the number of split-off (column 1) or migrant households (column 2) that the baseline household has generated by 2010 on the right-hand side of Equation (8). [36] The coefficient on the variance of the BMI z-score appears insignificant in both columns suggesting that the initial inequality in the allocation of the Pareto weights does not predict household division, or out-migration.

[Table 11 here]

However, the concern about endogeneity is not primarily a concern about heterogeneity across non-related households, as addressed to some degree by Table 11; it is a concern about which of the households within the network decide to migrate. For migration we can investigate this issue further by looking at a restricted sample of 1,146 households that had not migrated by the time we observe them in the 2004 round. Out of these, 151 will migrate between 2004 and 2010. This allows us to compare risk sharing arrangements before-after and with-without the migration event (difference-in-differences, in other words) to control for all unobserved time invariant traits of migrants and non-migrants, such as risk preferences, that are jointly determining their decision to migrate and the respective roles they take on in the insurance arrangement. [37]

[36] It is worth noting that the outcome of interest, the number of split off or migrant households produced by the baseline household, takes only non-negative integer values. Given the resulting departure from normality and the count nature of the data, the use of OLS may provide biased estimates. However, our results are robust to using a non-linear count modelling approach, such as the Poisson model.

[37] While we can unambiguously, and at the individual level, establish who migrates and who does not, it is conceptually problematic to do the same for household division. In this two-decade panel an average household had 5.7 members at baseline and sees it surviving members divide into 4.1 new households by 2010. There is no meaningful way to define who forms a split and who the original household. Household division is not an individual, but a household

We re-estimate Equation 6 on the 2004 sample of non-migrant households.[38] In column 1 of Table 12, we see that households are negatively affected by own shocks and network shocks, with the latter slightly more imprecisely estimated at p=0.112. In this reduced sample baseline village fixed effects take a high toll on the degrees of freedom, with only an average of 3 migrant households per village, compared to 25 migrant households per village in the main regressions. In column 3, we replace the baseline village fixed effects with baseline district fixed effects. The coefficient on the migration dummy remains insignificant and both own and network shocks yield a significant and negative effect on consumption growth. Taken together this reveals that the 2004 sample of 1,146 non-migrant households was sharing risk in the period prior to their migration.

In columns 2 and 4 of Table 12 we interact the own shock and the network shock variable with 2010 migration status. These two interactions are not significant, irrespective of using baseline village or district fixed effects. The insignificance of the coefficient on the interaction with the network shock variable shows that the risk sharing relationship between (future) migrant and non-migrant households is reciprocal prior to the migrant's move and becomes unilateral only after the move. This excludes the possibility that time invariant characteristics of either party are driving the results.

[Table 12 here]

Of further note is that the future migration dummy is not significant in any of the specifications, which shows that migrants and non-migrants were on similar growth paths

trait, the correlates of which are explored in Table 11. By consequence, we cannot run any analogous difference-in-differences regressions for household division.

[38] The term x_{ij} now refers to household characteristics in 2004 with respect to its previous household members.

prior to migration. The last column in Table 12 shows that this result is robust to running these regressions with NFE as in Equation 5.

We can therefore be confident that the effect of migration on informal insurance is either causal, or it is driven by the occurrence of a time variant event (like a shock pulling or pushing someone into migration), or change in individual characteristic (like coming of age, achieving higher levels of education or winning a lottery), which causes one to both migrate and assume the special role in the insurance network.

9. Robustness

We conducted an array of robustness checks to validate our findings. [39] First, we find that the results are robust to an alternative migrant definition where also households that moved to a nearby village are defined as non-migrants.

Second, the results are not driven by the configuration of the data. The shock data were initially defined at individual level while our outcome variable is measured at household level. Conducting the empirical analysis at individual level does not affect our main findings.

Third, defining household consumption per adult equivalent instead of per household member yields close to identical results in all specifications.

Fourth, we also checked whether the potential endogeneity of some of our control variables is driving our results. As discussed in Section 5, instrumenting the lagged consumption variable does not affect the shock coefficient. In addition, when the 2010 household level control variables are omitted one-by-one, the estimated shock coefficients remain stable across all specifications.

[39] The results of these robustness checks are available upon request.

Finally, estimating Equation 7 in levels does not change our conclusions either. More specifically, replacing the outcome variable in Equation 7, the change in log consumption ($\Delta \ln c_{ij}$), with the logged consumption measured in 2010 and dropping the baseline consumption variable from the right-hand side yields almost identical coefficients in Tables 7, 8 and 9.

10. Conclusions

In the next few decades internal migration is likely to drive the development process in Africa. This demographic process is visible in the data that form the basis of this paper. Starting from the household rosters of a representative household survey conducted nearly two decades ago in Kagera, we find that over half of the original household members had moved internally, while very few moved internationally. We document how this powerful current of migration, an integral part of development, interacts with a traditional institution like informal risk sharing to shape economic mobility and vulnerability.

Prior to their move the (future) migrants were as rich as stayers and shared risk reciprocally with them. After moving migrants realize a total consumption growth that is three times higher than that of non-migrants and the reciprocity in the risk sharing contract disappears. Post migration, migrants provide unilateral insurance to those who remain at home, which seems to be driven by social norms rather than exchange motives. Migrants share 2.7 to 4.6 percent of their substantial growth to provide this insurance. While our study cannot conclusively say where migrants would be without their extended family networks back home, this seems too little to be an important brake on their growth.

References

Alderman, Harold, Jere Behrman, Hans-Peter Kohler, John Maluccio, A., and Susan Watkins. 2001. "Attrition in Longitudinal Household Survey Data." *Demographic Research* no. 5 (4):79-124.

Altonji, Joseph G., Fumio Hayashi, and Laurence J. Kotlikoff. 1992. "Is the Extended Family Altruistically Linked - Direct Tests Using Micro Data." *American Economic Review* no. 82 (5):1177-1198.

Ambler, Catherine. 2012. Don't Tell on Me: Experimental Evidence of Asymmetric Information in Transnational Households. Mimeo, University of Michigan.

Attanasio, Orazio, Abigail Barr, Juan Camilo Cardenas, Garance Genicot, and Costas Meghir. 2012. "Risk Pooling, Risk Preferences, and Social Networks." *American Economic Journal: Applied Economics* no. 4 (2):134-167.

Attanasio, Orazio P., and José-Víctor Ríos-Rull. 2000. "Consumption smoothing in island economies: Can public insurance reduce welfare?" *European Economic Review* no. 44 (7):1225-1258.

Baland, Jean-Marie, Cathrine Guirkinger, and Charlotte Mali. 2011. "Pretending to Be Poor: Borrowing to Escape Forced Solidarity in Cameroon." *Economic Development and Cultural Change* no. 60 (1):1-16.

Barr, Abigail, and Garance Genicot. 2008. "Risk sharing, commitment, and information: an experimental analysis." *Journal of the European Economic Association* no. 6 (6):1151-1185.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2006. Kagera Health and Development Survey, 2004: Basic Information Document. Mimeo, The World Bank.

———. 2008. "Adult mortality and consumption growth in the age of HIV/AIDS." *Economic Development and Cultural Change* no. 56 (2):299-326.

———. 2011. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics* no. 93 (3):1010-1033.

Burke, Mary A., and H. Peyton Young. 2011. "Chapter 8 - Social Norms." In *Handbook of Social Economics*, edited by Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 311-338. San Diego, CA: North-Holland.

Chandrasekhar, Arun G, Cynthia Kinnan, and Horacio Larreguy. 2011. Information, networks and informal insurance: evidence from a lab experiment in the field. Mimeo, Massachusetts Institute of Technology and Northwestern University.

———. 2012. Informal Insurance, Social Networks, and Savings Access: Evidence from a lab experiment in the field. Mimeo, Massachusetts Institute of Technology and Northwestern University.

Charness, Gary, and Garance Genicot. 2009. "Informal Risk Sharing in an Infinite-Horizon Experiment." *The Economic Journal* no. 119 (537):796-825.

Chiappori, Pierre-André, Krislert Samphantharak, Sam Schulhofer-Wohl, and Robert M Townsend. 2011. "Heterogeneity and risk sharing in village economies." *NBER Working Paper* no. 16696.

Coate, Stephen, and Martin Ravallion. 1993. "Reciprocity without commitment: Characterization and performance of informal insurance arrangements." *Journal of Development Economics* no. 40 (1):1-24.

Cole, Tim J, Katherine M Flegal, Dasha Nicholls, and Alan A Jackson. 2007. "Body mass index cut offs to define thinness in children and adolescents: international survey." *Bmj* no. 335 (7612):194.

Collier, Paul, and Stefan Dercon. 2009. African Agriculture in 50 Years: Smallholders in A Rapidly Changing World? Paper presented at the 2009 Food and Agriculture

Organization Expert Meeting on ‘How to Feed the World in 2050. Available at ftp://ftp.fao.org/docrep/fao/012/ak983e/ak983e00.pdf.
Cox, Donald, and Marcel Fafchamps. 2007. "Extended family and kinship networks: economic insights and evolutionary directions." *Handbook of Development Economics* no. 4:3711-3784.
Cragg, John G., and Stephen G. Donald. 1993. "Testing identifiability and specification in instrumental variable models." *Econometric Theory* no. 9:222-222.
De Weerdt, J. 2010. "Moving out of Poverty in Tanzania: Evidence from Kagera." *Journal of Development Studies* no. 46 (2):331-349. doi: 10.1080/00220380902974393.
De Weerdt, Joachim, Kathleen Beegle, Helene Bie Lilleør, Stefan Dercon, Kalle Hirvonen, Martina Kirchberger, and Sofya Krutikova. 2012. "Kagera Health and Development Survey 2010: Basic Information Document." *Rockwool Foundation Working Paper Series* no. 46.
De Weerdt, Joachim, and Marcel Fafchamps. 2011. "Social Identity and the Formation of Health Insurance Networks." *Journal of Development Studies* no. 47 (8):1152-1177. doi: Doi 10.1080/00220388.2010.527952.
Dercon, Stefan, Daniel O Gilligan, John Hoddinott, and Tassew Woldehanna. 2009. "The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages." *American Journal of Agricultural Economics* no. 91 (4):1007-1021.
Dercon, Stefan, and Pramila Krishnan. 2000. "In sickness and in health: Risk sharing within households in rural Ethiopia." *Journal of Political Economy* no. 108 (4):688-727.
Dustmann, Christian, and Josep Mestres. 2010. "Remittances and temporary migration." *Journal of Development Economics* no. 92 (1):62-70.
Fafchamps, Marcel. 1992. "Solidarity networks in preindustrial societies: Rational peasants with a moral economy." *Economic Development and Cultural Change* no. 41 (1):147-174.
———. 1999. "Risk sharing and quasi-credit." *Journal of International Trade & Economic Development* no. 8 (3):257-278.
Foster, Andrew D. 1993. "Household partition in rural Bangladesh." *Popul Stud (Camb)* no. 47 (1):97-114.
Foster, Andrew D, and Mark R Rosenzweig. 2002. "Household division and rural economic growth." *The Review of Economic Studies* no. 69 (4):839-869.
Foster, Andrew D., and Mark R. Rosenzweig. 2001. "Imperfect commitment, altruism, and the family: Evidence from transfer behavior in low-income rural areas." *Review of Economics and Statistics* no. 83 (3):389-407.
Genicot, Garance. 2006. "Does wealth inequality help informal insurance." *Unpublished, Unpublished Manuscript*.
Grimard, Franque. 1997. "Household consumption smoothing through ethnic ties: evidence from Cote d'Ivoire." *Journal of Development Economics* no. 53 (2):391-422.
Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized-Method of Moments Estimators." *Econometrica* no. 50 (4):1029-1054.
Harris, John R., and Michael P. Todaro. 1970. "Migration, Unemployment and Development - A Two-Sector Analysis." *American Economic Review* no. 60 (1):126-142.
Hayashi, Fumio. 2000. *Econometrics*. Princeton: Princeton University Press.
Hoddinott, John. 1994. "A Model of Migration and Remittances Applied to Western Kenya." *Oxford Economic Papers-New Series* no. 46 (3):459-476.
Hoff, Karla, and Arjit Sen. 2006. "The Kin System as a Poverty Trap?" In *Poverty Traps*, edited by Samuel Bowles, Steven Durlauf and Karla Hoff, 91-115. Princeton University Press.

Jakiela, Pamela, and Owen Ozier. 2012. "Does africa need a rotten kin theorem? experimental evidence from village economies." *World Bank Policy Research Working Paper* (6085).

Kinnan, Cynthia. 2012. Distinguishing barriers to insurance in Thai villages. Mimeo, Northwestern University.

Kuczmarski, R. J., C. L. Ogden, S. S. Guo, L. M. Grummer-Strawn, K. M. Flegal, Z. Mei, R. Wei, L. R. Curtin, A. F. Roche, and C. L. Johnson. 2002. "2000 CDC Growth Charts for the United States: methods and development." *Vital and Health Statistics, Series 11, Data from the national health survey* (246):1-190.

Leider, Stephen, Markus M. Möbius, Tanya Rosenblat, and Quoc-Anh Do. 2009. "Directed Altruism and Enforced Reciprocity in Social Networks." *Quarterly Journal of Economics* no. 124 (4):1815-1851.

Lévi-Strauss, Claude. 1969. *Elementary structures of kinship*. London: Social Science Paperbacks (Revised edition).

Lewis, W. Arthur. 1954. "Economic Development with Unlimited Supplies of Labour." *The Manchester School* no. 22 (2):139-191.

Ligon, Ethan. 1998. "Risk sharing and information in village economies." *Review of Economic Studies* no. 65 (4):847-864.

Ligon, Ethan, and Laura Schechter. 2012. "Motives for sharing in social networks." *Journal of Development Economics* no. 99 (1):13-26.

Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall. 2002. "Informal insurance arrangements with limited commitment: Theory and evidence from village economies." *Review of Economic Studies* no. 69 (1):209-244.

Lim, Youngjae, and Robert M Townsend. 1998. "General equilibrium models of financial systems: Theory and measurement in village economies." *Review of Economic Dynamics* no. 1 (1):59-118.

Lucas, Robert EB, and Oded Stark. 1985. "Motivations to remit: Evidence from Botswana." *The Journal of Political Economy*:901-918.

Mazzocco, Maurizio, and Shiv Saini. 2012. "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences." *The American Economic Review* no. 102 (1):428-468.

Munshi, Kaivan. 2003. "Networks in the modern economy: Mexican migrants in the US labor market." *The Quarterly Journal of Economics* no. 118 (2):549-599.

Munshi, Kaivan, and Mark Rosenzweig. 2006. "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy." *The American Economic Review*:1225-1252.

Platteau, Jean-Philippe. 2000. *Institutions, Social Norms, and Economic Development*: Harwood Academic Publishers.

———. 2012. "Redistributive pressures in Sub-Saharan Africa: Causes, consequences, and coping strategies." *The Historical Roots of Africa's Underdevelopment. Cambridge University Press, Cambridge*.

Platteau, Jean-Philippe, and Petros G Sekeris. 2010. "On the feasibility of power and status ranking in traditional setups." *Journal of Comparative Economics* no. 38 (3):267-282.

Rosenzweig, Mark R, and Oded Stark. 1989. "Consumption Smoothing, Migration, and Marriage - Evidence from Rural India." *Journal of Political Economy* no. 97 (4):905-926.

Schechter, Laura, and Alex Yuskavage. 2011. Reciprocated versus unreciprocated sharing in social networks. Mimeo, University of Wisconsin and U.S. Treasury.

Schulhofer-Wohl, Sam. 2011. "Heterogeneity and Tests of Risk Sharing." *Journal of Political Economy* no. 119 (5):925-958.

Stark, Oded, and David E Bloom. 1985. "The New Economics of Labor Migration." *American Economic Review* no. 75 (2):173-178.

Stark, Oded, and Robert EB Lucas. 1988. "Migration, remittances, and the family." *Economic Development and Cultural Change*:465-481.

Stock, James, and Motohiro Yogo. 2005. "Testing for weak instruments in linear IV regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by Donald Andrews and James Stock. Cambridge University Press.

Townsend, Robert M. 1982. "Optimal multiperiod contracts and the gain from enduring relationships under private information." *The Journal of Political Economy*:1166-1186.

———. 1994. "Risk and Insurance in Village India." *Econometrica* no. 62 (3):539-591.

United Nations, Department of Economic and Social Affairs, Population Division 2012. World Urbanization Prospects: The 2009 Revision.CD-ROM Edition, 2010. Accessed at http://esa.un.org/unpd/wup/index.htm on 19 January 2012.

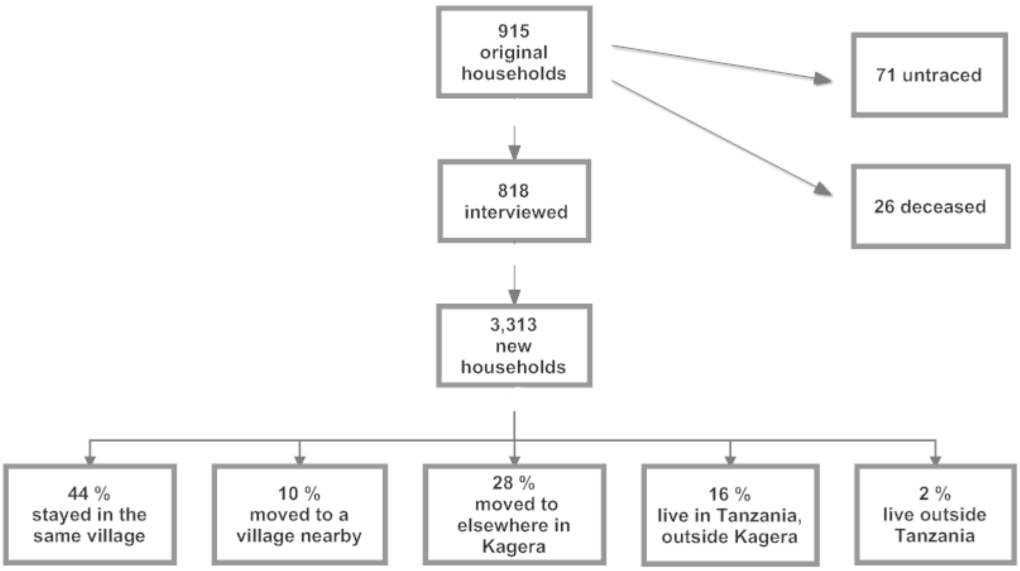
United Nations Development Programme. 2009. Human Development Report 2009. Overcoming barriers: Human mobility and development. United Nations, New York.

URT. 2012. National Sample Census Of Agriculture 2007/2008: Regional Report – Kagera Region (volume V) The United Republic of Tanzania (URT).

———. 2013. Population and Housing Census: Population Distribution by Administrative Units; Key Findings. National Bureau of Statistics (NBS) and Office of Chief Government Statistician (OCGS), Zanzibar. United Republic of Tanzania (URT).

World Bank. 2004. User’s Guide to the Kagera Health and Development Survey Datasets. Mimeo, The World Bank.

Figure 1: KHDS-2010 – Re-contacting after 16+ years



Tables

Table 1: Main income generating activity by migrant status

	non-migrant HHs	migrant HHs			
		all	nearby village	elsewhere in Kagera	outside Kagera
	%	%	%	%	%
agriculture	64.9	33.0	51.3	41.9	8.5
wage employment	6.2	26.8	12.0	20.1	45.8
self-employed	8.8	11.2	10.5	9.5	14.4
trading	11.7	17.2	17.2	15.1	20.7
casual labour	5.5	7.6	6.7	9.5	5.3
fishing	1.8	1.7	0.6	2.6	0.9
transfers & savings	1.2	2.5	1.8	1.4	4.6
number of HHs	**1,460**	**1,850**	**343**	**917**	**590**

Note: Agriculture category includes farming and livestock keeping, trading includes agriculture and non-agricultural trading. Wage employment can be either in formal or informal employment. Transfers include pensions, remittances and rental income. Self-employed category only considers self-employment outside agriculture. The information is missing for 2 non-migrant and 5 migrant households.

Table 2: Reasons for leaving the baseline village

Reason	males (%)	females (%)
To look for work	29.8	7.5
Own schooling	16.0	10.3
Found work	15.1	6.7
To live in a healthier environment	10.4	11.7
Marriage	0.0	38.9
Other reason	28.8	24.9
Total	**100.0**	**100.0**

Table 3: Consumption and poverty movements of the panel respondents in 1991-2010 by 2010 location

	mean 91	mean 2010	difference in means	N
Consumption per capita (TZS) by 2010 location				
Within community	343,718	492,398	148,680***	2,224
Nearby community	364,099	569,438	205,339***	382
Elsewhere in Kagera	357,930	695,951	338,021***	1,007
Out of Kagera	389,379	1,110,827	721,449***	658
Full Sample	355,926	642,558	286,632***	4,271
Consumption Poverty Head Count (%) by 2010 location				
Within community	31	19	-13***	2,224
Nearby community	30	20	-10***	382
Elsewhere in Kagera	31	16	-15***	1,007
Out of Kagera	23	3	-21***	658
Full Sample	30	16	-14***	4,271

*Note: All consumption values are in annual per capita terms and expressed in 2010 Tanzanian shillings. Significance of the difference in means using a t-test; *** p<0.01, ** p<0.05, * p<0.1.*

Table 4: Shocks reported by the panel respondents 1994-2009

Type of shock	Freq.	Percentage
Death of family member	797	26%
Poor harvest due to adverse weather	638	21%
Serious illness	577	19%
Loss in wage employment	219	7%
Loss of assets	205	7%
Eviction/resettlement	99	3%
Poor harvest due to pests or crop diseases	98	3%
Low crop prices	85	3%
Loss in off-farm employment	78	3%
Low income due to lower remittances	43	1%
Loss of livestock	6	0.2%
Loss of gifts and support by organizations	4	0.13%
Other reasons	172	6%
Total	**3,021**	**100%**

Table 5: Descriptive statistics

	Migrant households		Non-migrant households	
	mean	**std dev**	**mean**	**std dev**
1991 household per capita consumption	355,038	193,321	344,095	188,122
2010 household per capita consumption	739,033	634,925	488,830	358,197
per capita consumption growth in 1991-2010	383,995	643,235	144,736	352,247
natural log of per capita consumption growth in 1991-2010	0.5754	0.793	0.2944	0.618
own shock	0.2102	0.408	0.5320	0.499
# of hhs that reported a shock in the network	0.7937	1.120	1.4467	1.322
transfers in from migrant households	10,452	38,917	12,491	32,175
transfers in from non-migrant households	8,084	31,005	11,856	35,553
travel expenses	19,522	38,791	n/a	n/a
2010 household characteristics:				
age of oldest PHHM in the 2010 hh	31.376	11.132	44.405	18.288
a PHHM is head of this 2010 hh	0.4094	0.492	0.8100	0.392
a PHHM is spouse of this 2010 hh's head	0.4528	0.498	0.3058	0.461
a PHHM is child of this 2010 hh's head	0.0559	0.230	0.2067	0.405
divorced PHHM in 2010 hh	0.0433	0.204	0.0556	0.229
a widowed PHHM in 2010 hh	0.0417	0.200	0.1881	0.391
a married PHHM in 2010 hh	0.6614	0.473	0.6942	0.461
max yrs edu of PHHM in this 2010 hh	6.7811	3.083	6.1696	2.945
number of PHHMs in this 2010 hh	1.1024	0.419	1.5681	1.021
household size in 2010 hh	4.4732	2.450	4.8406	2.322
hh size in aeu in 2010 hh	3.5236	1.939	3.8314	1.878
Initial household characteristics:				
natural log value of assets	13.7058	1.100	13.7210	1.061
Educ of hh head	4.4189	3.139	4.2132	2.979
head was male	0.7646	0.424	0.7850	0.411
Age of hh head	48.9764	15.697	48.9296	15.599
age of head squared	2,645	1,596	2,637	1,575
Males 0-5 years	0.7622	0.896	0.7090	0.875
Males 6-15 years	1.3283	1.188	1.3040	1.123
Males 16-60 years	1.3756	1.022	1.4365	1.059
Males 61+ years	0.1913	0.394	0.2048	0.404
Females 0-5 years	0.8386	0.959	0.7609	0.877
Females 6-15 years	1.4591	1.340	1.3661	1.246
Females 16-60 years	1.8929	1.320	1.7822	1.186
Females 61+ years	0.2236	0.446	0.1937	0.407
hh had a non-earth floor in 1991	0.1811	0.385	0.1455	0.353
Observations	1,270		1,079	

Note: all monetary values in this table are expressed in 2010 Tanzanian shillings. PHHM refers to previous household member (i.e. person interviewed at the baseline).

Table 6: The effect of shocks on consumption growth

Dependent variable: (logged) per capita consumption growth	1 OLS, NFE	2 OLS	3 2SLS
Own shock	-0.141***	-0.144***	-0.140***
	(0.026)	(0.024)	(0.024)
2010 household characteristics:			
Age of oldest PHHM in the household	-0.001	0.000	-0.000
	(0.001)	(0.001)	(0.001)
A PHHM is head of the household	0.164***	0.152***	0.151***
	(0.047)	(0.034)	(0.034)
A PHHM is spouse of the household head	0.110**	0.077*	0.071*
	(0.048)	(0.039)	(0.042)
A PHHM is child of the household head	-0.196***	-0.184***	-0.182***
	(0.051)	(0.037)	(0.036)
A divorced PHHM in the household	-0.342***	-0.306***	-0.300***
	(0.068)	(0.076)	(0.073)
A widowed PHHM in the household	-0.333***	-0.332***	-0.328***
	(0.056)	(0.052)	(0.051)
A married PHHM in the household	-0.483***	-0.454***	-0.449***
	(0.040)	(0.038)	(0.039)
Max years of education of PHHM in the hh	0.058***	0.063***	0.060***
	(0.006)	(0.005)	(0.006)
Number of PHHMs in the household	-0.008	-0.018	-0.013
	(0.023)	(0.020)	(0.019)
Network characteristics:			
Number of split-off households stayed		-0.056***	-0.060***
		(0.009)	(0.010)
Number of split-off households moved		-0.008	-0.011
		(0.010)	(0.011)
Household characteristics at the baseline:			
Natural log value of assets in 1991		0.006	-0.009
		(0.017)	(0.027)
Education of 1991 household head		0.003	0.001
		(0.006)	(0.007)
Head was male in 1991		-0.064*	-0.082
		(0.038)	(0.053)
Age of household head in 1991		0.010**	0.009*
		(0.005)	(0.005)
Age of head squared		-0.000**	-0.000**
		(0.000)	(0.000)

Table 6: The effect of shocks on consumption growth

Dependent variable: (logged) per capita consumption growth	**1** **OLS, NFE**	**2** **OLS**	**3** **2SLS**
Num of males 0-5 years in the hh		0.005	0.016
		(0.019)	(0.025)
Num of males 6-15 years in the hh		0.049***	0.056***
		(0.012)	(0.016)
Num of males 16-60 years in the hh		0.006	0.002
		(0.016)	(0.017)
Num of males 61+ years in the hh		0.164***	0.199***
		(0.049)	(0.070)
Num of females 0-5 years in the hh		0.012	0.018
		(0.017)	(0.020)
Num of females 6-15 years in the hh		0.022*	0.027*
		(0.012)	(0.015)
Num of females 16-60 years in the hh		0.008	0.011
		(0.013)	(0.013)
Number of females 61+ years in the hh		0.026	0.033
		(0.033)	(0.036)
Household had a non-earth floor in 1991		-0.008	-0.054
		(0.047)	(0.095)
(logged) hh per capita consumption in 1991		-0.911***	-0.732**
		(0.042)	(0.290)
Number of observations	2,349	2,349	2,349
R^2	0.202	0.421	0.412
Adjusted R^2	0.199	0.414	0.392

*note: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors by baseline village are in parenthesis. Regressions in column 1 includes NFE, regressions in columns 2 and 3 include baseline village fixed effects. PHHM refers to previous household member (i.e. person interviewed at the baseline).*

Table 7: The effect of network shocks on consumption growth (OLS estimates)

Dependent variable: (logged) per capita consumption growth	**Eq. (6) pooled sample 1**	**Eq. (6) Migrant hhs 2**	**Eq (6) Non-migrant hhs 3**	**Eq (7), interaction terms only 4**
Number of households that experienced a shock in the network	-0.043***	-0.050***	0.008	-0.058***
	(0.013)	(0.018)	(0.015)	(0.021)
Own shock	-0.126***	-0.094**	-0.060*	-0.034
	(0.024)	(0.043)	(0.032)	(0.054)
Number of split-off hhs stayed	-0.041***	-0.031*	-0.032**	0.001
	(0.011)	(0.016)	(0.013)	(0.018)
Number of split-off hhs moved	-0.003	-0.008	-0.026**	0.018
	(0.010)	(0.013)	(0.013)	(0.018)
Age of oldest PHHM in the 2010 hh	0.000	0.002	-0.000	0.002
	(0.001)	(0.002)	(0.001)	(0.002)
A PHHM is head of this 2010 hh	0.155***	0.160***	0.220***	-0.059
	(0.034)	(0.056)	(0.054)	(0.092)
A PHHM is spouse of this 2010 hh's head	0.072*	-0.042	0.185***	-0.227***
	(0.039)	(0.059)	(0.057)	(0.087)
A PHHM is child of this 2010 hh's head	-0.186***	-0.356***	-0.008	-0.348***
	(0.036)	(0.099)	(0.044)	(0.123)
A Divorced PHHM in 2010 hh	-0.303***	-0.363***	-0.165*	-0.198
	(0.076)	(0.111)	(0.088)	(0.138)
A widowed PHHM in 2010 hh	-0.335***	-0.412***	-0.124**	-0.289**
	(0.051)	(0.125)	(0.052)	(0.139)
A married PHHM in 2010 hh	-0.439***	-0.431***	-0.246***	-0.185**
	(0.038)	(0.052)	(0.051)	(0.073)
Max years of education of PHHM in this 2010 hh	0.062***	0.072***	0.030***	0.042***
	(0.005)	(0.007)	(0.007)	(0.009)
Number of PHHMs in this 2010 hh	-0.021	0.050	-0.063***	0.113**
	(0.019)	(0.047)	(0.023)	(0.052)
(logged) hh per capita consumption in 1991	-0.915***	-0.978***	-0.848***	-0.131**
	(0.043)	(0.049)	(0.053)	(0.062)
Number of observations	2,349	1,270	1,079	2,349
R^2	0.424	0.462	0.390	0.439
Adjusted R^2	0.417	0.450	0.374	0.425

Column 4: estimates of Equation (7), a fully interacted model, showing only the interaction terms with the migration dummy variable. These represent a formal test for the difference between column (2) and (3).

*note: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors by baseline village are in parenthesis. Regressions include baseline village fixed effects and variables controlling for household characteristics at the baseline. PHHM refers to previous household member (i.e. person interviewed at the baseline).*

Table 8: Network shocks in migrant and non-migrant households (OLS estimates)

Dependent variable: (logged) per capita consumption growth	**Migrant households**	**Non-migrant households**
	1	**2**
Number of non-migrant hhs that experienced a shock in the network	-0.055**	0.013
	(0.028)	(0.022)
Number of migrant hhs that experienced a shock in the network	-0.043	0.002
	(0.028)	(0.023)
Own shock	-0.093**	-0.059*
	(0.043)	(0.032)
Number of split-off hhs stayed	-0.030*	-0.034**
	(0.018)	(0.015)
Number of split-off hhs moved	-0.009	-0.024*
	(0.013)	(0.014)
Number of observations	1,270	1,079
R^2	0.463	0.390
Adjusted R^2	0.450	0.374

*note: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors by baseline village are in parenthesis. Regressions include baseline village fixed effects, 2010 household level variables capturing characteristics of the previous household members and variables controlling for household characteristics at the baseline.*

Table 9: Other transactional insurance motives

	Migrant households		
Dependent variable: (logged) per capita consumption growth	**1**	**2**	**3**
Number of non-migrant hhs that experienced a shock in the network	-0.042	-0.050	-0.048
	(0.028)	(0.063)	(0.045)
--- Interacted with:			
* Share of land in BLV in total land portfolio	-0.055		
	(0.054)		
* Number of years since the last PHHM migrated into this hh		-0.001	
		(0.005)	
* Hh has inheritable land in the baseline village			0.019
			(0.055)
* Hh member's parent lives in BLV			-0.032
			(0.063)
Own shock	-0.085**	-0.097**	-0.090**
	(0.043)	(0.045)	(0.044)
Share of land in BLV in total land portfolio	0.300***		
	(0.064)		
Household does not own land	0.262***		0.171***
	(0.047)		(0.043)
Number of years since the last PHHM migrated into this hh		-0.001	
		(0.004)	
Hh has inheritable land in the baseline village			-0.016
			(0.066)
Hh member's parent lives in BLV			0.054
			(0.063)
Number of split-off hhs stayed	-0.021	-0.028	-0.026
	(0.017)	(0.018)	(0.018)
Number of split-off hhs moved	-0.015	-0.014	-0.015
	(0.013)	(0.013)	(0.013)
Number of observations	1,270	1,270	1,270
R^2	0.487	0.462	0.470
Adjusted R^2	0.474	0.449	0.456

*note: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors by baseline village are in parenthesis. Regressions include baseline village fixed effects, 2010 household level variables capturing characteristics of the previous household members and variables controlling for household characteristics at the baseline. PHHM refers to previous household member (i.e. person interviewed at the baseline).BLV refers to baseline village.*

Table 10: Re-calculating the cost of insurance through rainfall data

	Migrant households		**Non-migrant households**	
	mean	**1**	**mean**	**2**
max rain shock in own location [a]	-0.70 [0.52]	0.115** (0.045)	-1.17 [0.30]	0.179** (0.076)
max rain shock in deviation in baseline village [b]	-0.63 [0.56]	0.079* (0.046)		
max rain shock in deviation in migrant locations [b]			-1.21 [0.29]	-0.014 (0.086)
Number of observations	1,270		1,079	
R^2	n/a	0.471	n/a	0.400
Adjusted R^2	n/a	0.458	n/a	0.382
baseline district FE?	n/a	yes	n/a	yes

*note: *** p<0.01, ** p<0.05, * p<0.1.*
[a] For migrants this is after they moved to their 2010 location, for non-migrants this refers to 1994-2009.
[b] After the migrant moved to their 2010 location.
Standard deviations in brackets.
Cluster-robust standard errors by baseline village are in parenthesis.
Regressions include 2010 household level variables capturing characteristics of the previous household members and variables controlling for household characteristics at the baseline.

Table 11: The effect of initial within-household inequality on future household division and migration

Dependent variable:	**mean**	**number of split-off hhs in 2010** **1**	**number of migrant hhs in 2010** **2**
Mean BMI z-score	-0.704	0.148	0.085
	[0.711]	(0.152)	(0.181)
Variance of BMI z-score	1.039	0.068	0.106
	[0.914]	(0.069)	(0.079)
Maximum BMI z-score	0.408	-0.138	-0.104
	[0.852]	(0.142)	(0.165)
Number of observations	787	787	787
R^2	n/a	0.599	0.396
Adjusted R^2	n/a	0.590	0.382

*note: *** p<0.01, ** p<0.05, * p<0.1. Unit of observation is the baseline household in wave 1 (1991-92). Standard deviations are in brackets and cluster-robust standard errors by baseline village in parenthesis. Regressions include baseline village fixed effects and variables controlling for household characteristics at the baseline (wave 1): number of female and male children of the head, land size, head's characteristics (sex, age and education), highest level of education in the household and variables capturing household size and demographics.*

Table 12: The effect of future migration on long-run growth and pre-migration insurance contract type

Dependent variable: (logged) per capita consumption growth 1991-2004	**1**	**2**	**3**	**4**	**5**
migrant in 2010	0.058	0.117	0.047	0.116	0.077
	(0.051)	(0.095)	(0.053)	(0.092)	(0.062)
own shock	-0.064*	-0.050	-0.069**	-0.054*	
	(0.037)	(0.035)	(0.033)	(0.032)	
--- Interacted with:					
* (migrant in 2010)		-0.094		-0.102	
		(0.121)		(0.115)	
number of (other) stayer households affected by shock	-0.047	-0.045	-0.047*	-0.043*	
	(0.029)	(0.029)	(0.027)	(0.026)	
--- Interacted with:					
* (migrant in 2010)		-0.014		-0.022	
		(0.049)		(0.047)	
Number of observations	1,146	1,146	1,146	1,146	1,146
R^2	0.405	0.406	0.424	0.424	0.033
Adjusted R^2	0.390	0.389	0.406	0.406	0.024
Baseline village FE?	yes	yes	no	no	no
Baseline district FE?	no	no	yes	yes	no
Network FE (NFE)?	no	no	no	no	yes

*note: *** p<0.01, ** p<0.05, * p<0.1. Unit of observation is 2004 household. Cluster-robust standard errors by baseline village are in parenthesis. Regressions include 2004 household level variables capturing characteristics of the previous household members and variables controlling for household characteristics at the baseline.*

Table A1: First-stage regression results of Column 3 in Table 6

Dependent variable: (logged) hh per capita consumption in 1991	
Included instruments:	
Own shock	-0.019
	(0.018)
Number of split-off hhs stayed	0.018
	(0.019)
Number of split-off hhs moved	0.019
	(0.014)
Age of oldest PHHM in the 2010 hh	0.001**
	(0.001)
A PHHM is head of this 2010 hh	0.005
	(0.023)
A PHHM is spouse of this 2010 hh's head	0.036
	(0.025)
A PHHM is child of this 2010 hh's head	-0.008
	(0.034)
A divorced PHHM in 2010 hh	-0.037
	(0.038)
A widowed PHHM in 2010 hh	-0.015
	(0.033)
A married PHHM in 2010 hh	-0.032
	(0.023)
Max years of education of PHHM in this 2010 hh	0.015***
	(0.004)
Number of PHHMs in this 2010 hh	-0.024*
	(0.012)
Natural log value of assets in 1991	0.081***
	(0.027)
Education of hh head in 1991	0.024**
	(0.011)
Head was male in 1991	0.212***
	(0.075)
Age of hh head in 1991	0.006
	(0.006)
Age of head squared	-0.000
	(0.000)
Males 0-5 years in 1991	-0.064***
	(0.023)
Males 6-15 years in 1991	-0.045***
	(0.015)
Males 16-60 years in 1991	0.022
	(0.018)

Table A1: First-stage regression results of Column 3 in Table 6

Males 61+ years in 1991	-0.176**
	(0.079)
Females 0-5 years in 1991	-0.035
	(0.025)
Females 6-15 years in 1991	-0.034**
	(0.017)
Females 16-60 years in 1991	-0.017
	(0.015)
Females 61+ years in 1991	-0.030
	(0.042)
Hh had a non-earth floor in 1991	0.268***
	(0.059)
Excluded instruments:	
(Negative rainfall deviation) * (Age of hh head in 1991)	0.002
	(0.004)
(Negative rainfall deviation) * (Education of hh head in 1991)	0.048*
	(0.027)
(Negative rainfall deviation) * (Head was male in 1991)	0.355**
	(0.159)
Number of observations	2,349
R^2	0.228
Adjusted R^2	0.201
Under-identification test:	
Kleibergen-Paap rk LM statistic	8.780
p-value	0.032
Weak identification tests:	
Cragg-Donald Wald F Statistic	19.12
Kleibergen-Paap rk Wald F statistic	5.419
Over-identification test:	
Hansen-J statistic	1.124
p-value	0.570

*note: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors by baseline village are in parenthesis. Regression includes baseline village fixed effects. PHHM refers to previous household member (i.e. person interviewed at the baseline).*