Public Disclosure Authorized

Public Disclosure Authorized

WPS6429

Policy Research Working Paper 6429

Risk Sharing and Internal Migration

Joachim De Weerdt
Kalle Hirvonen

The World Bank
Development Research Group
Poverty and Inequality Team
April 2013

Policy Research Working Paper 6429

Abstract

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 migrants unilaterally insure their extended family members at home. This finding contradicts risk-sharing models based on reciprocity, but is consistent with assistance driven by social norms. Migrants sacrifice 3 to 7 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.

This paper is a product of the Poverty and Inequality Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at j.deweerdt@edi-africa.com and k.v.hirvonen@sussex.ac.uk.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

Risk Sharing and Internal Migration

Joachim De Weerdt [*]

Kalle Hirvonen [**]

JEL codes: O12, O15, O17, R23

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

Sector board: POV (Poverty Reduction)

* EDI, Tanzania, ** University of Sussex, UK

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, 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.v.hirvonen@sussex.ac.uk).

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 – 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) put it (p. 1230) "[...] *a complete*

[1] We miss 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.

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".

Geographical mobility in rural Tanzania is associated with large income gains. 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.

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 migrants' consumption responds to shocks experienced by others in their extended family network. Much of the migration literature has a very strong focus on dealing with the selectivity of the migration decision. Interestingly, in this paper, the endogeneity of migration turns out to be irrelevant for our most important contribution: the documentation of the long-run dynamics of risk-sharing arrangements among extended family members in a context of large internal migration flows. Whether or not migration is causally responsible for any of our findings is an interesting, but secondary question, which we will nevertheless attempt to answer in Section 8.

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.

The observed divergence between migrants and non-migrants in these data also persists within extended family networks (Beegle, De Weerdt, and Dercon 2011), which violates the full risk-sharing hypothesis (Townsend 1994), and does not support the notion that migration is the result of a household level maximization strategy (Stark and Bloom 1985; Rosenzweig and Stark 1989; Grimard 1997). It could, however, be consistent with other reciprocity-based models (e.g. limited commitment, moral hazard, or hidden income) that permit the co-existence of divergent consumption growth and risk sharing. In our empirical analysis, we find that migrants are affected by shocks to others in the network whereas non-migrants are not. Such unilateral insurance leads us to reject the reciprocity-based risk-sharing models.

One explanation for this observed lack of reciprocity could be that migrants insure non-migrants in exchange for other benefits (Lucas and Stark 1985; Hoddinott 1994). These benefits could accrue to the migrant later in life and outside the purview of our survey data. We consider, but reject a number of such longer-run transactional motives for the observed unilateral insurance. The results are, however, 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.

Our results speak 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 7.7 percentage points out of a total growth of 108 percent to insure their relatives. This estimate is equivalent to a 'tax' of 3 to 7 percent. We regard this tax rate as too trivial to exert any constraining effect on migrants.

After describing the model, the data and the estimation strategy in the next three sections, we discuss the results in Section 5. Section 6 tackles the issue of longer-run transactional motives. In Section 7 we calculate the cost of this insurance provision. Section 8 discusses endogeneity of migration and Section 9 contains some further robustness checks. 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:

[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 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 ($v' > 0$ & $v'' < 0$).

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

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

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

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. Furthermore, assuming that there are no frictions between the households in the extended family, the model predicts that all idiosyncratic shocks experienced by households are completely smoothed through the extended family. These two predictions form the basis for our rejection of the full risk-sharing model. First, the descriptive statistics in Section 3 confirm highly unequal consumption growth between migrants and non-migrants. Second, in Section 5 we show that after controlling for extended family fixed effects, household consumption growth remains responsive to idiosyncratic income shocks.

The rejection of full risk sharing 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

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

full commitment) to the full risk-sharing model. We will discuss each of these constraints in turn, but point out that 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 and are therefore no longer constant over time. This could allow the share of some members (migrants in our case) to increase over time.

In the presence of enforcement problems, the better-off households have an incentive to leave the arrangement and live in autarky. The limited commitment model (e.g. Coate and Ravallion 1993; Attanasio and Ríos-Rull 2000; Ligon, Thomas, and Worrall 2002; Kinnan 2012) appends the full risk-sharing model with participation constraints (one for each household):

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

where u_A is the expected utility received in autarky. Solving the augmented maximization problem yields a following first-order condition:

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

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. In the context of migration, a growth premium has to be granted to the migrant whose autarky options have improved. As a result, risk sharing is no longer efficient: the impact of idiosyncratic income shocks is not equally shared within the extended family network.

The other frictions have similar analytical consequences. 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 are motivated to exert effort by rewarding them with higher consumption. This comes

with an efficiency cost: idiosyncratic shocks are not completely smoothed within the network. 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, migrants are allowed to enjoy a larger share of the consumption cake. As a consequence, Pareto-efficient risk sharing is again sacrificed.

These frictions can have important implications for the degree of risk sharing. [6] Distinguishing which of the three models of constrained insurance explains our data best is beyond the scope of this paper. [7] One common feature, however, is that despite friction, reciprocity remains intact: households engage in reciprocal risk sharing but the degree of its efficiency varies. In Section 5, we study the existence of such reciprocal but partial risk-sharing arrangements by testing whether households are responsive to income shocks faced by other households in the same extended family network.

In this paper, we contrast these reciprocity-based models with models that take into account social norms. 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 also buy social prestige, political power or serve to perpetuate subordination (Platteau 2012; Platteau and Sekeris 2010). 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 modeled as subjective satisfaction that individuals receive from participation. [8] The satisfaction can stem from the

[6] For example, Chandrasekhar, Kinnan and Larreguy (2011, 2012), using field experiments from Southern India find that limited commitment reduces transfers by 10 percent and hidden income by 40 percent.
[7] See Kinnan (2012) for such exercise with data from rural Thailand.
[8] 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).

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.

A recent empirical literature relying on experimental design 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. Furthermore, 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. [9] Below we will find evidence of such unilateral relations and argue that this is consistent with risk sharing motivated by social norms.

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. At the time of the last national census in 2002, Kagera had a population of roughly two million people.

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

[9] Schechter and Yuskavage (2011) empirically document unreciprocated relations in Paraguay.
[10] See World Bank (2004).

alive in 2004. This effectively meant that the original household panel survey turned into a panel of individuals. A full 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 (17 households and 27 respondents), 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.

Figure 1 provides an overview of migration patterns. By 2010 nearly 45 percent of the households were still residing in their original baseline community. We define migrants as households that in 2010 are not located in the original village but are found from a nearby village, elsewhere within Kagera or outside Kagera. [13]

[Figure 1 here]

[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.

Remittances offer one medium for risk sharing between households. Table 1 provides a summary of the average remittance flows over the past 12 months in 2010 between the migrant households and households living in or near their baseline villages. While non-migrant households were net receivers of remittances, Table 1 shows that transfers flow both ways. This could lead one to think – mistakenly as the analysis below reveals – that these are relationships of reciprocal risk sharing. The data in Table 1 are self-reported and it is interesting to note that migrants claim to send more home than non-migrants acknowledge. A similar discrepancy does not exist in migrant-migrant or stayer-stayer dyads [14].

[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.

[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 questionnaire included in the survey. Consumption is expressed in annual per capita terms using 2010 Tanzanian shillings. [15]

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:

[14] By dyad we refer to a pair of households.
[15] Using adult equivalent units as the denominator instead of household size produces almost identical results across all specifications.

nearly all respondents who left the region managed to escape poverty, while poverty reduction among non-migrants was more modest. These descriptive statistics reinforce the results reported in Beegle, De Weerdt, and Dercon (2011): individuals who moved did considerably better than those who decided to stay.

[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. [16] This difference is statistically significant at the 1 percent level. The severing of the most basic 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

[16] The mean consumption growth among those who maintained contact was 394,679 TZS and among those who severed links 286,991 TZS.

reformatted from the individual to the household level. [17] 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 respondents moved to their 2010 location, allowing us to include only shocks that occurred one year after this move. [18]

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.

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.

[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:

$$\Delta \ln c_{ij} = \beta s_{ij} + x'_{ij}\gamma + \alpha_j + \varepsilon_i \quad (5)$$

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

[18] 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.

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 [19] 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 relation to the 2010 household head and their marital status. [20] The term α_j represents the network fixed effect and ε_i is the error term. The inclusion of the network fixed effects means that we compare the impact of shocks between 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 (Schulhofer-Wohl 2011; Chiappori et al. 2011; Mazzocco and Saini 2012). [21] 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. We also include (logged) per capita consumption at the baseline ($\ln c_{j,1991}$). The network shock variable, z_{ij}, measures the number of

[19] Previous household member refers to a person interviewed at the baseline in 1991/94.

[20] 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 specifications.

[21] 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 an upward bias in β in Equation (5). The standard full risk sharing test is then biased against the null-hypothesis of full-risk sharing.

households affected by an income shock. The household's own shocks are excluded from this variable. The partial risk-sharing specification is formulated as:

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

A negative and statistically significant δ would imply that some risk sharing takes place within the extended families.

We will assess the impact of these network shocks separately for migrant and non-migrant households and fully acknowledge that selection into migration is unlikely to be random. The differences in the observed level of risk sharing may be caused by migration or by some unobserved characteristics that differ between migrant and non-migrant 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 home-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). Our primary contribution lies in documenting the fact that migrants provide unilateral insurance to non-migrants, while at the same time shooting ahead of them in consumption terms. Nonetheless, we will dedicate Section 9 to shedding light on how selection fits into this and will utilize information from additional survey rounds to exclude some possible types of endogeneities.

Finally, the baseline per capita consumption variable in Equation (6) raises a concern about endogeneity. The error term ε_i could be correlated, for example due to measurement error, with the lagged consumption variable. 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. [22] 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.

[22] 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 them 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 later in the analysis on 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) implies 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. 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 reports the results for the migrant households and the second

column for the non-migrant households. For migrants, the network shock coefficient 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 is nearly zero and insignificant. 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, 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 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. 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]

A final 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. This does not seem consistent with the findings of Table 1, where we noted that migrants are net senders of transfers.

7. Is there a kin tax?

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). [23] Put another way, migrants share about 2.7 percent of their very substantial growth by insuring family members at their original location. [24]

This is a lower-bound estimate because we cannot exclude the possibility that we are only measuring a subset of relevant shocks: if shocks are self-reported then respondents may fail to mention those that were effectively insured. Fortunately, the survey provides an alternative shock measure, which is not self-reported. We have historical rainfall data from the Tanzanian Meteorological Agency for gauges in 8 weather stations across the Kagera Region. Each baseline village can be linked to its closest rainfall station as both databases record GPS location. The mean distance to a rainfall station is 19 km and the median is 15 km. For each village 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 two columns in Table 10 show that rainfall shocks are important in determining consumption growth, with every standard deviation decrease in (negative) rainfall deviation causing consumption growth to decline by 20 to 32 percentage points, depending on the specification.

[Table 10 here]

Knowing that rainfall shocks drive the incomes of the stayer households, we can use them as an alternative shock indicator. We replace the network shock variable with the baseline village rainfall shock variable in Equation (6). This rainfall shock is constructed as the most negative rainfall deviation in the baseline village after the migrant left. With baseline village fixed effects, the specification exploits the variation arising from the fact that migrants leave at different times. [25] Column 3 repeats the results from Table 8 for comparison purposes and column 4 replaces the network shocks in non-migrant households with the baseline village rainfall shock.

[23] We use geometric (rather than arithmetic) means to calculate the average annual growth rates.
[24] The 95%-confidence interval for the annual growth penalty is [0.0002, 0.1515] and for the 'kin-tax' [0.01, 5.40].
[25] Replacing the baseline village fixed effects with district fixed effects yields nearly identical results.

As expected, once we use rainfall shocks the kin tax goes up to 7.7 percentage points, with the 95 percent confidence interval ranging from 0.03 to 0.13. We consider the parsimonious rainfall specification from Table 10 as the upper bound effect, binding the estimate between 2.7 and 7.1 percent.

It is interesting, if slightly misleading (see below), to point out that our implicit tax rate of 2 to 7 percent is of the same order of magnitude as that found in two other studies. Jakiela and Ozier (2012) estimate that women in a laboratory setting in Kenya acted as if they were expecting to be pressured to share 4 percent of their experiment winnings with relatives that were not present at the experiment. Ambler (2012) reports that El Salvadorian migrants living around Washington DC remit 5 percent more of a windfall income if they know the potential recipients at home will be informed about it. However, we consider these striking similarities in magnitude to be slightly misleading: the above studies focus on the effect, on sharing, of providing full information on a windfall income, while our figures reflect the effect of shocks on sharing within real-world belief sets.

8. Endogeneity of migration

Much of the research on migration is concerned with establishing its causal effects. The primary goal of this paper, however, is quite different. Here, we aim to document what happens to a traditional institution, like informal insurance, in a society that modernizes and is characterized by massive internal migration. The strength of the paper lies in its ability to describe this process within linked extended family networks, over a long two decade time period and within all the richness of the real world. Many readers, however, will wonder to what extent migration is causally responsible for the empirical patterns described in the previous sections. 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. Fortunately, we are able to exploit more survey rounds in order to speak to the causality issue and exclude that certain forms of unobserved heterogeneity explain the results. The purpose of this section is to be very specific about which remaining types of endogeneity could compete with causality to explain the results.

Thus far the paper has used waves 1 and 6 of the survey. This section will exploit the four waves that lie in between. In what follows it is useful to bear in mind that consumption was measured comparably within waves 1, 5 and 6 and within waves 2, 3 and 4 – but not between these two groups of waves.[26]

As our left hand side variable is consumption growth, a first set of concerns relate to whether migrants start out from similar consumption positions at baseline (compared to non-migrants) and whether, prior to migration, they are following a similar growth path. We can test this by regressing baseline consumption on future migration status ($M_{i,2010}$) and use baseline village fixed effects (θ_v) and control for the initial household characteristics (w_j) used in Equation 6 (and listed in Table 5):

$$Y_{i,} = \propto M_{i,2010} + w_j'\vartheta + \theta_v + \varepsilon_i. \quad (7)$$

In the first column of Table 11, Y_i is the natural logarithm of the consumption level at baseline. The insignificance of the future migration coefficient shows that our extensive controls manage to capture all heterogeneity with respect to baseline start-off levels: migrants started off at the same position as non-migrants, *ceteris paribus*.

We then append the right-hand side of Equation (7) with (logged) consumption at the baseline. The second column of Table 11 exploits information on consumption growth between rounds 2, 3 and 4 (1992-1994) to ask whether growth across those three rounds can be explained by future migration status. We find that it cannot: migrants and non-migrants were on the same growth curve prior to migration, given our controls.

[Table 11 here]

Finally, we exclude that there are any time invariant traits of migrants and non-migrants, such as risk preferences, that are jointly determining their migration decision and the respective roles they take on in the insurance arrangement. This test exploits the fact that some migrants migrated after we observe them in 2004. We

[26] The primary concern is the difference in recall period. Beegle et al. (2012) look at how differential recall periods affect consumption aggregates, using data from a survey experiment conducted in Tanzania.

restrict the sample to 1,146 households that had not migrated by 2004 and re-estimate Equation (6) using growth in (logged) consumption from 1991 to 2004 as the dependent variable. [27] We also include a dummy that captures those 151 households that will migrate between 2004 and 2010.

In column 1 of Table 12, we see that the future migration status is not significant indicating that, *ceteris paribus*, migrants and non-migrants were on the same long-run growth path. Households are negatively affected by own shocks and network shocks, with the latter slightly more imprecisely estimated at p=0.116. In the 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. [28] 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 shows that the sample of 1,146 non-migrant households was sharing risk in the period prior to their migration. In columns 2 and 4 we interact the network shock variable with 2010 migration status. This interaction is not significant, irrespective of using cluster or district fixed effects. This shows that migrant and non-migrant households are both responsive to each other's shocks prior to the move and the insurance relationship becomes unilateral only after the move. Akin to a difference-in-difference estimator, this excludes any time invariant characteristics of either party to be driving the results.

[Table 12 here]

Taken together, the results in Table 11 and Table 12 tell us that there is no unobserved heterogeneity between migrants and non-migrants with respect to the starting position and slopes of their short-run (1992-1994) and long-run (1991-2004) growth paths, prior to migration. Results from Table 12 further show that the insurance contract was characterized by reciprocity before the move and only became unilateral after the move. 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

[27] The term x_{ij} now refers to household characteristics in 2004 with respect to its previous household members.
[28] Households group into 51 baseline village and 6 baseline districts.

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 take on the special role in the insurance network.

9. Robustness

We conducted an array of robustness checks to verify our findings. [29] 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, changing the way we isolate the shocks that occurred before the households split does not change our results either.

Fifth, we also checked whether the potential endogeneity of some our control variables is driving our results. 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.

10. Conclusions

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

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

with a traditional institution like informal risk sharing to shape economic mobility and vulnerability.

We find that internal migrants provide unilateral insurance to those who remain at home, which seems to be driven by social norms rather than exchange motives. The total, final, long-run effect of this insurance provision on the migrant's growth amounts to a 3 to 7 percent sacrifice in consumption (2.9 to 7.1 percentage points off the 108 % total growth realized by the migrant).

While our study cannot conclusively say where migrants would be without their extended family networks back home, a tax rate of 3 to 7 percent seems too low to be an important brake on a migrant's growth. We do know that migrants who have severed links with home perform worse than other migrants and one should not overlook the fact that, while starting out from similar baseline welfare levels, migrants realize a total consumption growth which is three times higher than that of non-migrants.

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 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.

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

———. 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 Alberto Bisin Jess Benhabib and O. Jackson Matthew, 311-338. North-Holland.

Chandrasekhar, Arun G, Cynthia Kinnan, and Horacio Larreguy. 2011. Information, networks and informal insurance: evidence from a lab experiment in the field. MIT and Northwestern working paper.

———. 2012. Informal Insurance, Social Networks, and Savings Access: Evidence from a lab experiment in the field. MIT Working Paper.

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.

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, 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.

Dustmann, Christian, and Josep Mestres. 2010. "Remittances and temporary migration." *Journal of Development Economics* no. 92 (1):62-70. doi: DOI 10.1016/j.jdeveco.2008.12.002.

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., 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.

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.

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.

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.

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.

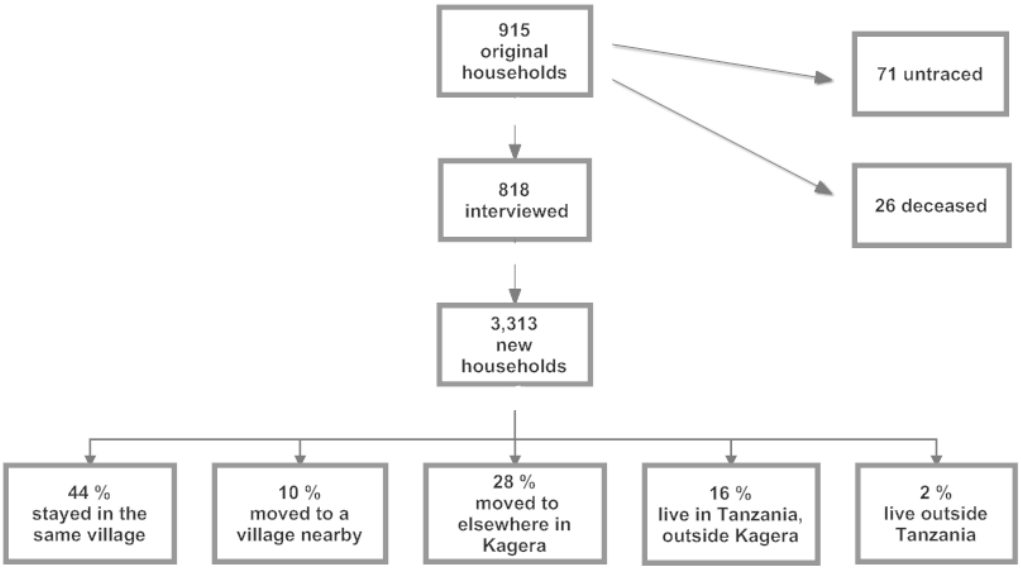
United Nations, Department of Economic and Social Affairs, Population Division 2012. World Urbanization Prospects: The 2009 Revision.CD-ROM Edition, 2010. Accessed at http://esa.un.org/unpd/wup/index.htm on 19 January 2012.

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).

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

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



Tables

Table 1: Reported remittances in and out between migrants and non-migrants, as reported by the first half of the dyad with respect to transfers to (first column) or from (second column) the second half of the dyad. [30]

dyad	gifts out	gifts in	net in
stayer-migrant	8,920	13,577	4,656
migrant-stayer	19,044	10,567	-8,477
stayer-stayer	8,310	10,476	2,167
migrant-migrant	17,096	14,816	-2,280
Total	**13,208**	**12,328**	**-880**

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**

[30] This table is based on self-reported remittance flows in households in the past 12 months in 2010.

Table 3: Consumption and poverty movements of the panel respondents in 1991-2010 by 2010 location[31]

	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: Significance of the difference in means using a paired 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%**

[31] All consumption values are in annual per capita terms and expressed in 2010 Tanzanian shillings.

Table 5: Descriptive statistics

	Migrant households		Non-migrant households	
	mean	**sd**	**mean**	**sd**
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
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	

Table 6: The effect of shocks on consumption growth

Dependent variable: (logged) per capita consumption growth	1	2	3
	OLS, NFE	**OLS**	**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)

Table 6: The effect of shocks on consumption growth

Dependent variable: (logged) per capita consumption growth	**1**	**2**	**3**
	OLS, NFE	**OLS**	**2SLS**
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)
Number of males 0-5 years in the household		0.005	0.016
		(0.019)	(0.025)
Number of males 6-15 years in the hh		0.049***	0.056***
		(0.012)	(0.016)
Number of males 16-60 years in the hh		0.006	0.002
		(0.016)	(0.017)
Number of males 61+ years in the hh		0.164***	0.199***
		(0.049)	(0.070)
Number of females 0-5 years in the hh		0.012	0.018
		(0.017)	(0.020)
Number of females 6-15 years in the hh		0.022*	0.027*
		(0.012)	(0.015)
Number 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

Dependent variable: (logged) per capita consumption growth	**Migrant households** **1** **OLS**	**Non-migrant households** **2** **OLS**
Number of households that experienced a shock in the network	-0.050***	0.008
	(0.018)	(0.015)
Own shock	-0.094**	-0.060*
	(0.043)	(0.032)
Number of split-off hhs stayed	-0.031*	-0.032**
	(0.016)	(0.013)
Number of split-off hhs moved	-0.008	-0.026**
	(0.013)	(0.013)
Age of oldest PHHM in the 2010 hh	0.002	-0.000
	(0.002)	(0.001)
A PHHM is head of this 2010 hh	0.160***	0.220***
	(0.056)	(0.054)
A PHHM is spouse of this 2010 hh's head	-0.042	0.185***
	(0.059)	(0.057)
A PHHM is child of this 2010 hh's head	-0.356***	-0.008
	(0.099)	(0.044)
A Divorced PHHM in 2010 hh	-0.363***	-0.165*
	(0.111)	(0.088)
A widowed PHHM in 2010 hh	-0.412***	-0.124**
	(0.125)	(0.052)
A married PHHM in 2010 hh	-0.431***	-0.246***
	(0.052)	(0.051)
Max years of education of PHHM in this 2010 hh	0.072***	0.030***
	(0.007)	(0.007)
Number of PHHMs in this 2010 hh	0.050	-0.063***
	(0.047)	(0.023)
(logged) hh per capita consumption in 1991	-0.978***	-0.848***
	(0.049)	(0.053)
Number of observations	1,270	1,079
R^2	0.462	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 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

Dependent variable: (logged) per capita consumption growth	**Migrant households**	**Non-migrant households**
	1	**2**
	OLS	**OLS**
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.*

Table 10: Re-calculating the kin-tax through rainfall data

Dependent variable: (logged) per capita consumption growth		Non-migrant HHs		Migrant HHs	
	mean	**1**	**2**	**3**	**4**
max rain shock 1994-2009	-0.743	0.319***	0.204***		
		(0.101)	(0.011)		
own shock (self-reported)	0.210			-0.093**	-0.073
				(0.043)	(0.046)
number of stayer hhs in NW affected by shock	0.531			-0.055**	
				(0.028)	
Most negative rainfall deviation in baseline village *after* migrant left	-0.435				0.176***
					(0.058)
Number of observations		1,079	1,079	1,270	1,270
R^2		0.353	0.389	0.462	0.463
Adjusted R^2		0.336	0.374	0.450	0.451
baseline village fixed effects?		no	no	yes	yes
baseline district fixed effects?		no	yes	no	no

*note: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors by baseline village (columns 1, 3 and 4) or district (column 2) 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 future migration on baseline consumption levels and short-run growth (1992-94)

Dependent variable:	**ln (conspc w1)**	**Δ ln (conspc)**
	1	**2**
migrant in 2010	-0.017	0.017
	(0.038)	(0.045)
(logged) household per capita consumption in 1991		-0.064
		(0.041)
Number of observations	803	782
R^2	0.198	0.032
Adjusted R^2	0.183	0.012

*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 (θ_v) and control for the initial household characteristics (w_j) used in Equation 6 (and listed in Table 5).*

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**
migrant in 2010	0.058	0.065	0.047	0.059
	(0.051)	(0.066)	(0.043)	(0.062)
own shock	-0.064*	-0.064*	-0.069*	-0.069*
	(0.037)	(0.037)	(0.037)	(0.037)
number of (other) stayer households affected by shock	-0.047	-0.046	-0.047***	-0.045**
	(0.029)	(0.029)	(0.017)	(0.021)
--- Interacted with:				
* (migrant in 2010)		-0.009		-0.017
		(0.049)		(0.053)
Number of observations	1,146	1,146	1,146	1,146
R^2	0.405	0.405	0.412	0.412
Adjusted R^2	0.390	0.389	0.397	0.396
baseline village fixed effects?	yes	yes	no	no
baseline district fixed effects?	no	no	yes	yes

*note: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors by baseline village (columns 1-2) or district (columns 3-4) are in parenthesis. Regressions include baseline village fixed effects, 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)

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

Males 6-15 years in 1991	-0.045***
	(0.015)
Males 16-60 years in 1991	0.022
	(0.018)
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).*