Internal migration and vulnerability to poverty in Tanzania

Rebecca Pietrelli* and Pasquale Scaramozzino†

September 5, 2016

Abstract

This paper investigates internal migration and vulnerability to poverty in Tanzania. It examines whether migration reduced household vulnerability to poverty for a panel of households from the Kagera region over the period 2004-2010. The dataset allows the analysis of two samples of households: those with the same head in the periods considered and an enlarged network of split-off households. The potential endogeneity of migration is controlled by both matching methods and an exogenous variation. A severe drought in 2008-09 affected the areas of the country with a bimodal rain season, but not those with a unimodal rain season. It is thus possible to study the heterogeneity of migrants with respect to an unanticipated shock in the region of destination. The evidence shows that migration reduced vulnerability to basic needs and to food consumption poverty for families which experienced migration to unimodal regions. The results are consistent with migration as a risk management strategy by households.

JEL codes: C21, I32, O15, O55.
Key words: Migration, vulnerability, matching methods, exogenous variation, Tanzania.
Acknowledgements: We are grateful to Tiziano Arduini, Richard Black, Ciaran Driver, Kalle Hirvonen, Majlinda Joxhe, Pierluigi Montalbano, Donato Romano, Furio C. Rosati, Antonia Schwarz and L. Alan Winters for very helpful suggestions.

*Agricultural Development Economics Division, The Food and Agriculture Organization of the United Nations; rebecca.pietrelli@fao.org.

†School of Finance and Management, SOAS, University of London and Dipartimento di Economia e Finanza, Università degli Studi di Roma Tor Vergata; ps6@soas.ac.uk.

1 Introduction

According to the traditional approach, migration is motivated by the attempt by individuals to benefit from a positive expected wage gap in the region of destination relative to the region of origin. The original analysis by Harris and Todaro (1970) is based on a dual model of the labour market, with a traditional (/agricultural/rural) sector which pays a subsistence wage and a modern (/industrial/urban) sector where a higher marginal product of labour is achieved. Migration is the mechanism through which the labour force moves towards the more productive sectors of the economy, thereby ensuring a more efficient allocation of labour in the economy.

The traditional analysis is based on two key assumptions. First, migration is an individual decision. There is therefore no explicit consideration of the possibility that migration decisions are made at the level of the household, with some members remaining in the region of origin and some others migrating and then pooling resources with their family. Second, the decision to migrate is solely based on the expected value of the migrants' wage in the region of destination, without consideration of the risks associated with their decision.

The New Economics of Labour Migration (NELM: Stark and Bloom (1985); Stark and Levhari (1982); Stark (1991)) addresses the first issue by focussing on the household as the unit where the migration decision is made. Some family members might migrate in order to remit back to the original family a share of their labour income from the destination region. More generally, migration can act as an insurance mechanism, which enables the household to reduce its vulnerability to adverse shocks in the absence of fully developed insurance markets (Ray (1998), chapter 15; Bardhan and Udry (1999), chapter 8).

The role of uncertainty in informing migration decisions has been studied by Burda (1995) in the context of real options theory (Dixit and Pindyck (1994)). Income in the region of destination is risky, and migration involves a sunk cost which cannot be recovered if the decision is reversed at a later date. Under these conditions, the option to migrate should only be exercised when favourable further delay would not be optimal. A reduction in income uncertainty in the region of destination might lead the household to bring forward the migration decision.[1] In this framework, migration can be seen as an element of the risk management strategy. If income is pooled within the family, migration by some household members can be an effective method to diversify the household's income sources across different regions, thereby reducing its overall risk.

We adopt a stochastic outcome approach to the measurement of household vul-

[1] See also Khwaja (2002), Anam et al. (2008), Moretto and Vergalli (2008) and Vergalli (2011) for applications of the real options approach to migration.

nerability. Specifically, we follow Chaudhuri (2000) and Chaudhuri et al. (2002) and consider Vulnerability to Expected Poverty (VEP). The approach to household risk in terms of its vulnerability to poverty is intrinsically dynamic and forward-looking: vulnerability is defined as the probability that the household might fall below a critical poverty line, as a result of unfavourable shocks. This approach is consistent with the World Bank's "Social Risk Management" framework which considers the ability of a community to manage risks as the main source of vulnerability (WB (2005)).

In spite of its potential relevance, empirical analysis of the risk mitigating aspects of migration decisions is still very limited because of the difficulty of controlling for the endogeneity of the migration decision, and for the lack of experimental data. In this paper we are able to identify the risk reducing effects of migration thanks to a unique meteorological feature of Tanzania. The country exhibits two rainfall regimes: unimodal (which covers the areas in the South, Centre and West) and bimodal (in the North, the Northern Coast and the North-West). The unimodal regions experience only one long rainy season in the course of the agricultural year, whereas the bimodal regions have two short rainy periods. During the 2008-09 season, the bimodal regions experienced an extreme drought. This shock affected eastern African countries and was described as "one of the worst in living memory" (IDRC (2010)). Unimodal regions were, however, not affected. We look at the extreme meteorological events which occurred in Tanzania in 2009 to examine how the vulnerability to expected poverty by households in the region of origin was affected by the migration of household members. In particular, we analyse if migration had a different effect depending on whether it had been directed towards a unimodal or a bimodal region.

We use a comprehensive data set from the region of Kagera in Tanzania. The data are obtained from surveys carried out over the period 2004-2010. Individual household members are traced over time, and the area of destination of migrants is recorded alongside information about both the original family and the new family in the destination regions. It is therefore possible to measure changes in the vulnerability of the original household in a fully dynamic setting.

Specifically, we examine whether household whose members had migrated to a unimodal region of the country experienced a reduction in their vulnerability to expected poverty relative to households whose members had migrated to a bimodal region which was affected by the drought. An ex ante unanticipated shock would have resulted in changes in the ex post vulnerability of the household. The main empirical results are consistent with migration as an insurance mechanism for the household. We first adopt a matching approach to examine the differential changes in the vulnerability to basic needs and food insecurity by households with and without migrants, and show that migration by some family members significantly reduced the vulnerability of the household of origin. We then exploit the "natural experiment" of the drought in one of the two meteorological regions in Tanzania to control for time-varying unobserved heterogeneity, and observe that migration to the drought-free unimodal zones resulted in a significant decline of vulnerability by the household of origin. These novel empirical results show that migration can be an effective strategy for households to mitigate their risks.

Our results are consistent with Hirvonen and Lilleør (2015) and De Weerdt and Hirvonen (2016), who used the same data set for Kagera. Hirvonen and Lilleør (2015) establish the existence of links between migrants and their home communities, both during the migration spell and following return migration. De Weerdt and Hirvonen (2016) find evidence of obligations of migrants towards family members who remain at home, consistent with social norms associated with kinship (Lévi-Strauss (1969)).

The structure of the paper is as follows. Section 2 provides background information on the Kagera region and on the severity of the drought of 2009. Section 3 describes household sample data used in the analysis and how the latter was obtained and motivated our choice of measure of vulnerability. Section 4 illustrates the matching approach adopted in the paper and discusses the empirical results of the analysis. Section 5 explains our methodology for the exogenous variation and shows that migration to unimodal zones resulted in a significant decline in vulnerability for the households of origin. Section 6 concludes.

2 The Kagera region and the drought of 2009

Kagera is the remote north-western region of Tanzania, bordering Lake Victoria, Rwanda, Burundi and Uganda. The region covers 40,838 km^2 of land surface and 11,885 km^2 of water surface and is overwhelmingly rural. According to the latest Population and Housing Census (2012), the population residing in the region is about 2.5 million. The main ethnic groups of the region are Haya and Nyambo in the north and Subi, Sukuma, Zinza and Hangaza in the south.

The agricultural sector is dominant in Kagera economy. The sector accounts for about 50 percent of the regional GDP and employs about 90 percent of the economically active population in the production of food and cash crops. Bananas, beans, maize and cassava are the main food crops while coffee, tea and cotton are the main cash crops. Livestock is the second most important economic activity in the region. Recently, fishing in Lake Victoria has provided an alternative source of income. The industrial base in the region is limited and mainly comprises agro-industrial operations (URT (2012)). The region is relatively remote and is the farthest from the political and commercial capital, Dar es Salaam.

Tanzania presents two rainfall regimes. The country is accordingly divided in the unimodal zone (covering south, central and west) and the bimodal zone (north, northern coast and north western), as represented in figure 1. They have different rainy seasons and consequently harvesting periods. The unimodal zone experiences only one long rainy season from December to April. Sowing takes place in November and harvesting in June and July. The bimodal zone has short rainy periods from October to December and a long rainy one from March to May. It has a short harvesting in January/February and a long one in July/August (WFP, 2013).

Insert figure 1 here

In 2009, a severe drought hit many countries of East Africa, leading to crop failures and mortality of livestock (Goldman and Riosmena (2013)). This event has been described as "one of the worst in living memory" (IDRC (2010)). The northern regions of Tanzania were extremely damaged by the drought of 2009 (Goldman and Riosmena (2013)): indeed, Kagera was one of the worst hit regions.

Insert figure 2 here

The figure 2 underlines that Kagera region as well as all the regions under the bimodal rainfall regimes received a significantly lesser amount of rain in 2009 compared to the 2004-08 average than the unimodal regions of Tanzania.

3 Data

3.1 Sample of households

The Kagera Health and Development Survey (KHDS) was designed and implemented by the World Bank and the Muhimbili University College of Health Sciences. It is a survey of households, living in Kagera region, originally interviewed in 4 rounds from 1991 to 1994 (KHDS I). Resurveys were then administrated in 2004 (KHDS II) and 2010 (KHDS III).

The baseline KHDS I sample is composed of 915 households and was drawn in two stages. The stratification was based on geography in the first stage and on mortality risk in both stages. The sampling strategy of KHDS II and KHDS III consisted of re-interviewing all individuals who were interviewed in any round of the KHDS 1991-1994, the so called Previous Household Members (PHHMs). In 2004, the field team managed to re-contact 832 households of the baseline sample (KHDS 1991-1994). Re-contact means that at least one PHHM was re-interviewed in the KHDS 2004.[2] Because since 13 years a number of people had moved out of

[2] A few households (20) had all previous members deceased while 63 households remained untraced.

their original households, the new sample consisted of 2,774 households. In 2010, 818 households from KHDS 1991-1994 sample were re-contacted.[3] After 18 years, several of the original households had split up into many more households compared to KHDS II. In fact the new sample of households included 3,313 households compared to the original 915 (De Weerdt et al.; 2012).

The analysis in the present paper mainly employs KHDS II and III. Thus, the sample covers the period between 2004 and 2010. Nevertheless, the very low attrition rate[4] in both waves preserves the representativeness of the survey. Since KHDS I was a household survey which was turned into an individual longitudinal survey in 2004 and 2010, the definitions of households and migrants need to be carefully addressed.

For the purposes of the present analysis, the 2004 sample of households is restricted to those residing in the original community (1,235). This choice is constrained by the data, since the household localization in each wave of the survey is reported with respect to KHDS I.[5] Consequently focusing on the total sample of KHDS II would not allow the analysis to disentangle whether they have any member who moved between 2004 and 2010. Furthermore, this allows to rule out the effect of previous migration (between 1991 and 2004), which may confound the effect of migration between 2004 and 2010. Therefore the analysis focuses only on recent migration following the sample of households that did not move between 1991 and 2004. The summary statistics for the 2004 sample of households are reported in table 1.

Insert table 1 here

To define the migrants this paper looks at the household members' localization in 2010 compared to 2004 (which, according to the selected sample of households, corresponds to the original localization in KHDS I). The paper adopts the same definition of migration proposed by Beegle et al. (2011) and De Weerdt and Hirvonen (2016). An internal migrant is an household member found to not reside in the baseline community in 2010. By contrast, the migrant can be localized in nearby village, elsewhere in the same region, or elsewhere in Tanzania. Therefore, an internal migrant is anyone who has moved out of the baseline village between 2004 and 2010 and resides elsewhere inside the country. The sample of individuals

[3] A total of 26 households had all previous household members deceased and 71 were untraced.

[4] In KHDS II the rate of re-contacted households is 93 percent while in KHDS III it is equal to 92 percent of the original KHDS I sample.

[5] The question on household localization in KHDS I and II is the following: "Is the household living in (i) original village; (ii) nearby village; (iii) elsewhere in Kagera; (iv) elsewhere in Tanzania; (v) neighbouring country; (vi) other country."

(specifically PHHMs) who did not leave the original village between 1991 and 2004 but moved internally[6] during the considered period is composed of 197 people.

The use of this survey for conducting a household level analysis presents two main implications. First, by sampling design, some individuals are not traced between 2004 and 2010 (if they are not PHHMs, namely they are new members in 2004, or they are not found to reside with a PHHMs at 2010). This dataset does not allow to control for their location in 2010 if they do not reside with a PHHMs. For that reason, the number of internal migrants (197 people) is likely underestimated (according to De Weerdt and Hirvonen (2016), the percentage of internal migrants is 37% in 2004 and 48% in 2010 with respect to KHDS I). Consequently, the effect of migration found in this paper can be considered as a lower bound of the actual effect.

Second, the presence of household split-offs between 2004 and 2010 needs to be taken into consideration. The present analysis has been conducted on two samples of households: the extended sample of (1,235) households including all the splits at 2010 as well as on the restricted sample of (881) household with same head in the two periods, both samples of households with residence in Kagera in 2010. This implies testing the impact of migration on the entire network of the households. The latter is in the same spirit of Angelucci et al. (2010). Indeed the inter-generational family ties between new households of siblings (living together in the previous wave) are considered to form the extended sample. The summary statistics are reported in table 2 for the two samples and by typology of households, with and without migrants.

Insert table 2 here

Finally, this analysis uses two additional datasets provided by the Economic Development Initiatives (EDI) Group.[7] First, the total, food and non-food consumption data for all the waves have been used in the analysis. Consumption data are expressed in annual per capita terms. The data are aggregate controlling for the seasonality of food consumption and for the slight changes between rounds. Also all the consumption aggregates are deflated using information from the KHDS price questionnaire. Second, we control for migration distance by means of a distance matrix, elaborated by Jose Funes and Jean-Francois Maystadt. The matrix is composed by all distances between all households interviewed in KHDS III. The distances are expressed in Euclidean distances (km).

[6]Due to the fact that the analysis focuses on internal migration, 62 international migrants are not considered in the analysis.

[7]Additional information on the EDI Group and specifically on KHDS data can be found here: http://edi-global.com/publications/.

3.2 Outcome variable

This paper follows the Vulnerability as Expected Poverty (VEP) approach proposed by Chaudhuri et al. (2002) and extended by Günther and Harttgen (2009) to take into account idiosyncratic and covariate shocks. They define vulnerability as "the ex-ante risk that a household will, if currently non-poor, fall below the poverty line, or, if currently poor, will remain in poverty". According to this definition, a household is classified as vulnerable if it has a high probability of being poor in the future. Conversely, a household is classified as not vulnerable, if it is likely to be non-poor in the future. Formally, the vulnerability level of a household h at time t is defined as the probability that the household will find itself consumption poor at time $t+1$ (Chaudhuri et al. (2002)).

The VEP measure presents two main attractive features. First, it estimates vulnerability using a single round of cross-sectional data. Second, it is easily interpretable since the results are expressed in terms of expected values of Foster et al. (1984) measure of poverty.

The VEP measure has been estimated in 2004 and 2010 for all households which did not move relative to their initial location reported in KHDS I. Tables from A1 to A4 in the Appendix show the regressions results[8] of both total and food consumption. All coefficients display the expected signs. For example, the age of household head has a concave impact on food consumption. The proportion of adults of the household has a positive impact on consumption, while the number of children has a negative effect.

As shown in table 3, the vulnerability rates in 2010 are greater than the poverty rates, for both samples and measures. Beside the vulnerability rates, table 3 presents the estimated expected means and variances of total and food consumption, as well as the vulnerability means. The mean estimates are similar in magnitude for the two groups of households with and without migrants (respectively 0.41 and 0.43 for basic needs poverty and 0.34 and 0.36 for food poverty) and the hypothesis of their equality in both cases cannot be rejected by a t-test of equality of vulnerability means. The same pattern is confirmed in the case of food and basic needs poverty.

Insert table 3 here

Finally, table 4 reports the vulnerability variation between 2004 and 2010 for (both extended and same head) households with and without migrants. In the period considered, households with migrants perform better in terms of vulnerability reduction than households without them, both for basic needs and food

[8]The reported results of consumption and weighted consumption regressions are respectively the first and last steps estimations of the FGLS procedure proposed by Chaudhuri et al. (2002) to estimate vulnerability to poverty.

poverty. For samples, the vulnerability to food poverty decreases in 2010 relative to 2004, while for households without migrants it increases. The vulnerability to basic needs poverty decreases in the extended sample for households without and with migrants, but for the latter the decline is larger. The difference of the vulnerability variation between household with and without migrants is statistically significant for basic needs poverty in the extended household sample and for food poverty in the sample of households with the same head.

Insert table 4 here

4 Matching approach

4.1 Methodology

It is well-known in the development literature[9] that if one is interested in estimating the impact of migration on outcomes of origin households (or migrant's outcomes), one needs to deal with the endogeneity issue.[10] Specifically, the outcomes of households with and without migrants are not simply comparable each other because households self-select into migration. In fact, both observed and unobserved attributes of households are likely to be correlated with the decision to send one member away as well as with the outcome of interest, namely household's vulnerability to poverty.

Two examples are useful to clarify the issue. First, the amount of assets owned by a farmer household constitutes an example of an observable characteristic, which can be correlated both with migration and with vulnerability to poverty. Indeed the household may decide to sell part of its assets to cover the migration costs for its member. On the other side, the household may decide to sell assets after the occurrence of a shock as part of a strategy of consumption smoothing, which will in turn affect the probability of the household experiencing a consumption below the poverty line.

Second, household attributes unobserved to the researcher, such as loss aversion, entrepreneurial ability or families' ties, may play an important role for migration. For example, a loss averse household may be reluctant to encourage migration of its members, when this is perceived as a risky strategy, as well as crop diversification, which may result in a useful risk-management strategy for agricultural production risks. Further, a household engaged in crop diversification may reduce

[9]Following McKenzie (2015), while the development literature aims at estimating the impact of migration on welfare of migrants and origin families, the immigration literature focus on the impact of migration on outcomes of natives in destination countries.

[10]See McKenzie and Yang (2010) for a very exhaustive explanation.

the agricultural product risk, thus reducing the gains from migration. To assess the impact of migration on vulnerability to poverty of origin households, firstly the following model is explored:

$$ATT = [Y_t - Y_{t'}|X_{t'}, D = 1] - [Y_t - Y_{t'}|X_{t'}, D = 0] \tag{1}$$

The treatment variable (dummy D) is equal to 1 if the household had migrant members between the baseline period ($t' = 2004$) and the following interview ($t = 2010$). The outcome of interest is the variation of household vulnerability to poverty between the first (2004) and the second period (2010). The latter is measured as VEP variation, both in terms of basic needs and food poverty. The main advantage of that method is that it allows for "temporally invariant differences in outcomes between households with and without migrants" (Smith and Todd (2005)). In other words, it allows to control for selection into having migrants based on time-invariant unobservable characteristics at household level. The model is estimated using many matching estimators, as the difference-in-difference with propensity score matching (D-i-D PSM) proposed by Smith and Todd (2005). The D-i-D PSM estimator requires that:

$$E(Y_{0,t} - Y_{0,t'}|P, D = 1) = E(Y_{0,t} - Y_{0,t'}|P, D = 0) \tag{2}$$

where t and t' are respectively time periods after and before migration and $P = Pr(D = 1|Z)$ is the conditional probability of having migrants. Furthermore, this estimator requires that the support condition holds in both periods, t and t'. For all the observable conditioning variables Z, there is a positive probability of having ($D = 1$) or not ($D = 0$) migrants. In order to obtain the propensity score we estimate a logit model[11] that links the probability of having a migrant in t to household characteristics in t'. The model is defined as in equation:

$$P(D_{i,j,t} = 1) = F(X_{i,j,t'}, Z_{j,t'}) \tag{3}$$

The dependent variable is the probability that household i in village j has a migrant member in period t. The binary dummy $D_(i, j, t)$ equals one if household i has at least one migrant in period t and zero otherwise. The probability of having a migrant member is a function of household characteristics measured at period t'.

Second, the bias-adjusted estimator proposed by Abadie and Imbens (2006) has been employed to estimate the Average Treatment on Treated (ATT). Its main advantage is that it allows for matching on multiple covariates without imposing parametric assumptions. According to Gibson et al. (2011) and Gibson

[11]We obtain similar results by using a probit model. The results are presented in table A5 of the Appendix.

et al. (2013), the bias-adjusted matching estimator of Abadie and Imbens (2006) performs next best after the instrumental variable (IV) estimator among the non-experimental methods. IV estimation has been avoided in this analysis because of the nature of the dependent variable, which makes the exclusion restriction unlikely to hold.[12]

Third, the Coarsened Exact Matching (CEM) estimator, recently developed by Iacus et al. (2011) has been employed. "After pre-processing, some type of estimator must be applied to the data to make causal inferences. For example, if one-to-one exact matching is used, then a simple differences-in-means between Y in the treated and control groups provide a fully non-parametric estimator of the causal effect. Under a matching method that produces a one-to-one match, any analysis method that might have been appropriate without matching can alternatively be used on the matched data set with the benefit of having a lower risk of model dependence" (Iacus et al. (2011)). After pre-processing the data the vulnerability variation (both basic needs and food poverty) has been regressed on a migration dummy (equal to one if the household has at least one member who migrated between 2004 and 2010, and zero otherwise).

The key element of this strategy is to match households with migrants to those without them with similar pre-treatment characteristics measured in 2004. Thus the attributes used to match households are aimed to control for the propensity of having migrants in the next years (household size; number of males; number of members with primary education; number of members with secondary education, a dummy for whether the origin household is living in Bukoba urban; number of cattle owned by the household; number of sheep owned by the household; acre of plots owned by the household; dummy for having a bank account, dummy for having experienced a shock; dummy for participating in informal organizations). We specifically avoid matching on characteristics in 2010 because they may be endogenous to migration.[13]

4.2 Results

As described above matching occurs over pre-treatment covariates, measured in 2004. They are related to household demographic characteristics affecting the propensity of having migrants and the risk management strategy. The inclusion of a relatively large number of covariates is motivated by the need to satisfy the Conditional Independence Assumption, that there are no other observable factors influencing migration and the potential outcomes that would be realized in the

[12] McKenzie and Yang (2010) warn against the use of IV to estimate the impact of migration.

[13] This approach followed by Esquivel and Huerta-Pineda (2007) to assess the effect of remittances on poverty of Mexican households is reasonably criticized by McKenzie and Yang (2010).

absence of migration. The balance test is satisfied for all the used covariates for the two samples of households. After matching, baseline variables for the treatment group are well balanced in both samples (see p-values for matched and unmatched households in tables A6 and A7 of the Appendix).

Before looking at the results, figure A1 in the Appendix confirm the common support condition for propensity score matching. The evidence suggests that the large majority of observations are found to be on common support, hence they can be compared in a meaningful way.

Insert table 5 here

Table 5 reports the results from all the matching methods employed: propensity score (matching to the nearest neighbour, with the caliper of 0.2% and kernel), matching on multiple covariates (with three different distances metrics: Mahalanobis, Euclidean, and the inverse variance), and coarsened exact matching. The results are presented for both samples of households, in column (a) for vulnerability to basic needs poverty while in column (b) for vulnerability to food poverty (this format will also be followed in the following table 6).

The estimated gain from having migrants is a reduction in vulnerability, ranging from 6 to 15 percentage points for basic needs poverty and from 11 to 26 for food poverty, depending on the method employed.

An interesting finding is that the results are specular for the two samples of households. While for extended households the migration effect is significant for vulnerability to basic needs poverty and not for food poverty, for households with the same head in the two periods a significant effect is only detected for food poverty. This result may be driven by the fact that households with the same head (which may be composed by older household members since the younger members may have left the original household) are likely to be poorer than households resulting from split-offs, and may face barrier to migration. Consequently, whenever they can invest in migration, they will have a stronger return in vulnerability to food poverty. In other words, while the return of migration for the origin household can be interpreted in terms of substantial needs, the return for the extended network is more for consumer goods.

The robustness of the results has been checked by running the same estimates using KHDS I. The period considered is now 1992-1994. Since KHDS I is a household survey, split-offs cannot be considered and there is only one sample of households. The results confirming a negative effect are reported in table A8 of the Appendix (tables A9 and A10 report the consumption regressions for estimating vulnerability to poverty in 1992 and 1994).

5 Exogenous variation

5.1 Methodology

Time-varying unobserved heterogeneity may still affect the previous results. In order to control for it, our analysis takes advantage of a natural event. As discussed in the seminal work by Rosenzweig and Stark (1989), the covariance of shocks between the area of destination, where migrants live, and the place of origin, where their households reside, has a crucial relevance for risk-management. Indeed they find that migration contributes to the reduction in the variability of original household consumption, for given variability in household income from crop production. Furthermore, they find that "households exposed to higher income risk are more likely to invest in longer-distance migration-marriage arrangements" (Rosenzweig and Stark (1989)).

The crucial assumption tested in our work is that migration in unimodal zones of the country acted as an insurance mechanism for the origin households against the effects of the extreme drought. Migrants who moved from Kagera to unimodal areas of Tanzania, are expected to be less damaged by the drought than their household of origin. They will therefore be able to send remittances back to Kagera region. Thus, in the subsample of households with migrants, those with members who migrated to the unimodal areas of the country should be expected to be less vulnerable than households with migrants in the bimodal areas. This exogenous variation (the extreme drought) negatively affected, ceteris paribus, the households living in Kagera. If households with migrants in unimodal areas are found to perform better in their vulnerability variation than households with migrants in bimodal areas, this would be evidence in support of risk management through migration.

In order to exploit this exogenous variation, the analysis is limited to the subsample of households having at least one migrant in 2010. Thus the focus is the impact of migrants' location on the change of vulnerability to poverty of origin household before (2004) and after (2010) the drought. Formally, the following equation is estimated:

$$\Delta Vul_i = \beta_0 + \beta_1 Dist_i + \beta_2 Unim_i + \beta_3 Dist * Unim_i + \beta_4 \Delta X_i + \epsilon_i \qquad (4)$$

The focus on changes (2010-2004) in household vulnerability to poverty (ΔVul_i) and attributes (ΔX_i) allows to purge the estimates of household time-invariant heterogeneity. $Dist_i$ measures the (Km) distance between the migrant and his family of origin i; the dummy variable $Unim_i$ is equal to one if the migrant from household i is located in a unimodal zone and zero otherwise; $Dist_i * Unim_i$ is an interaction term between the distance and the unimodal dummy. The idea is that migrants' location in unimodal zone may have a different effect according to the distance which separates migrant and origin household.

If the household has more than one migrant, the $Dist_i$ variable is replaced with the mean distance over all migrants. In this case, the unimodal dummy is equal to one if at least one migrant moved to unimodal areas of the country. Table A11 in the Appendix reports the summary statistics of variables used.

5.2 Results

Table 6 reports the impact of migrant location on the change in household VEP, before (2004) and after (2010) the occurrence of the drought, for the sub-sample of households with migrants.

Insert table 6 here

The coefficient on the unimodal dummy is significant and highly negative: it implies a reduction of about 42 percentage points for vulnerability to basic needs and 55 percentage points for vulnerability to food. Its effect declines with the distance from the area of origin, as indicated by the positive coefficient on the interaction terms, which is only significant for vulnerability to food poverty. The unimodal and the interaction coefficients are however jointly significant in both cases: the p-value for the F test is 0.0880 for vulnerability to basic needs poverty (a) and 0.0476 for food poverty (b).

We can have an estimate of the average effect of the unimodal location of migrants on vulnerability to poverty change by considering the mean value of the Distance variable in the sample is employed, which is 249 Km. Thus, at the mean distance, the effect is: $0.42 + 0.0002 \cdot (249) = 0.37$ for vulnerability to basic needs poverty and $0.55 + 0.0006 \cdot (249) = 0.40$ for vulnerability to food poverty. Migration of family members to unimodal zones of Tanzania is associated, on average, with a differential decrease in vulnerability for the household of origin of about 40 percentage points of the pre-shock level.

6 Conclusions

Migration can be an important strategy to reduce the risks faced by households, but empirical evidence on the risk mitigating aspects of migration is still limited. This paper investigates how vulnerability to expected poverty was influenced by migration from the Kagera region of Tanzania. The paper uses matching methods and exploits an exogenous variation due to an unanticipated drought which only affected regions with a bimodal rainy pattern. The empirical findings show that vulnerability to expected poverty was significantly reduced for families whose members migrated to unimodal regions, relative to those with migration to bimodal regions. The evidence, therefore, supports that migration can be an effective risk management strategy for households.

References

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects, *Econometrica* **74**(1): 235–267.

Anam, M., Chiang, S.-H. and Hua, L. (2008). Uncertainty and international migration: an option cum portfolio model, *Journal of Labor Research* **29**(3): 236–250.

Angelucci, M., De Giorgi, G., Rangel, M. A. and Rasul, I. (2010). Family networks and school enrolment: Evidence from a randomized social experiment, *Journal of Public Economics* **94**(3): 197–221.

Bardhan, P. and Udry, C. (1999). *Development microeconomics*, Oxford University Press, USA.

Beegle, K., De Weerdt, J. and Dercon, S. (2011). Migration and economic mobility in Tanzania: Evidence from a tracking survey, *Review of Economics and Statistics* **93**(3): 1010–1033.

Burda, M. C. (1995). Migration and the option value of waiting, *Economic and Social Review* **27**: 1–19.

Chaudhuri, S. (2000). Empirical models for assessing household vulnerability to poverty, *Technical report*, Mimeo, Columbia University.

Chaudhuri, S., Jalan, J. and Suryahadi, A. (2002). Assessing household vulnerability to poverty from cross-sectional data: A methodology and estimates from Indonesia.

De Weerdt, J., Beegle, K., Lilleør, H. B., Dercon, S., Hirvonen, K., Kirchberger, M. and Krutikova, S. (2012). Kagera health and development survey 2010: Basic information document.

De Weerdt, J. and Hirvonen, K. (2016). Risk sharing and internal migration, *Economic Development and Cultural Change* .

Dixit, A. K. and Pindyck, R. S. (1994). *Investment under uncertainty*, Princeton University Press.

Esquivel, G. and Huerta-Pineda, A. (2007). Remittances and poverty in Mexico: a propensity score matching approach, *Integration and Trade Journal* **27**(11): 45–57.

Foster, J., Greer, J. and Thorbecke, E. (1984). A class of decomposable poverty measures, *Econometrica* **52**(3): 761–766.

Gibson, J., McKenzie, D. and Stillman, S. (2011). The impacts of international migration on remaining household members: omnibus results from a migration lottery program, *Review of Economics and Statistics* **93**(4): 1297–1318.

Gibson, J., McKenzie, D. and Stillman, S. (2013). Accounting for selectivity and duration-dependent heterogeneity when estimating the impact of emigration on incomes and poverty in sending areas, *Economic Development and Cultural Change* **61**(2): 247–280.

Goldman, M. J. and Riosmena, F. (2013). Adaptive capacity in Tanzanian Maasailand: Changing strategies to cope with drought in fragmented landscapes, *Global Environmental Change* **23**(3): 588–597.

Günther, I. and Harttgen, K. (2009). Estimating households vulnerability to idiosyncratic and covariate shocks: a novel method applied in madagascar, *World Development* **37**(7): 1222–1234.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis, *The American Economic Review* **60**(1): 126–142.

Hirvonen, K. and Lilleør, H. B. (2015). Going back home: internal return migration in rural Tanzania, *World Development* **70**: 186–202.

Iacus, S. M., King, G. and Porro, G. (2011). Causal inference without balance checking: Coarsened exact matching, *Political Analysis* **20**: 1–24.

IDRC (2010). Reducing vulnerability among pastoralists in northern Kenya.

Khwaja, Y. (2002). Should I stay or should I go? Migration under uncertainty: a real options approach.

Lévi-Strauss, C. (1969). *The elementary structures of kinship*, Boston, Beacon Press.

McKenzie, D. (2015). Learning about migration through experiments, *Migration: Economic Change, Social Challenge* p. 68.

McKenzie, D. J. and Yang, D. (2010). *Experimental approaches in migration studies*, Handbook of Research Method in Migration.

Moretto, M. and Vergalli, S. (2008). Migration dynamics, *Journal of Economics* **93**(3): 223–265.

Ray, D. (1998). *Development economics*, Princeton University Press.

Rosenzweig, M. R. and Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural India, *The Journal of Political Economy* **97**(4): 905–926.

Smith, J. A. and Todd, P. E. (2005). Does matching overcome Lalonde's critique of nonexperimental estimators?, *Journal of Econometrics* **125**(1): 305–353.

Stark, O. (1991). *The Migration of Labor*, Oxford, Basil Blackwell.

Stark, O. and Bloom, D. E. (1985). The new economics of labor migration, *The American Economic Review* **75**(2): 173–178.

Stark, O. and Levhari, D. (1982). On migration and risk in LDCs, *Economic Development and Cultural Change* **31**(1): 191–196.

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

Vergalli, S. (2011). Entry and exit strategies in migration dynamics, *Journal of Labor Research* **32**(4): 362–389.

WB (2005). *Afghanistan. Poverty, Vulnerability and Social Protection: An Initial Assessment*, Human Development Unit, South Asia Region, Report No. 29694-AF, Washington DC.

WFP (2013). *Comprehensive Food Security and Vulnerability Analysis: Tanzania 2012*, World Food Programme report.

Tables and figures

Figure 1: Map of rainfall regimes in Tanzania
Source: WFP (2013)

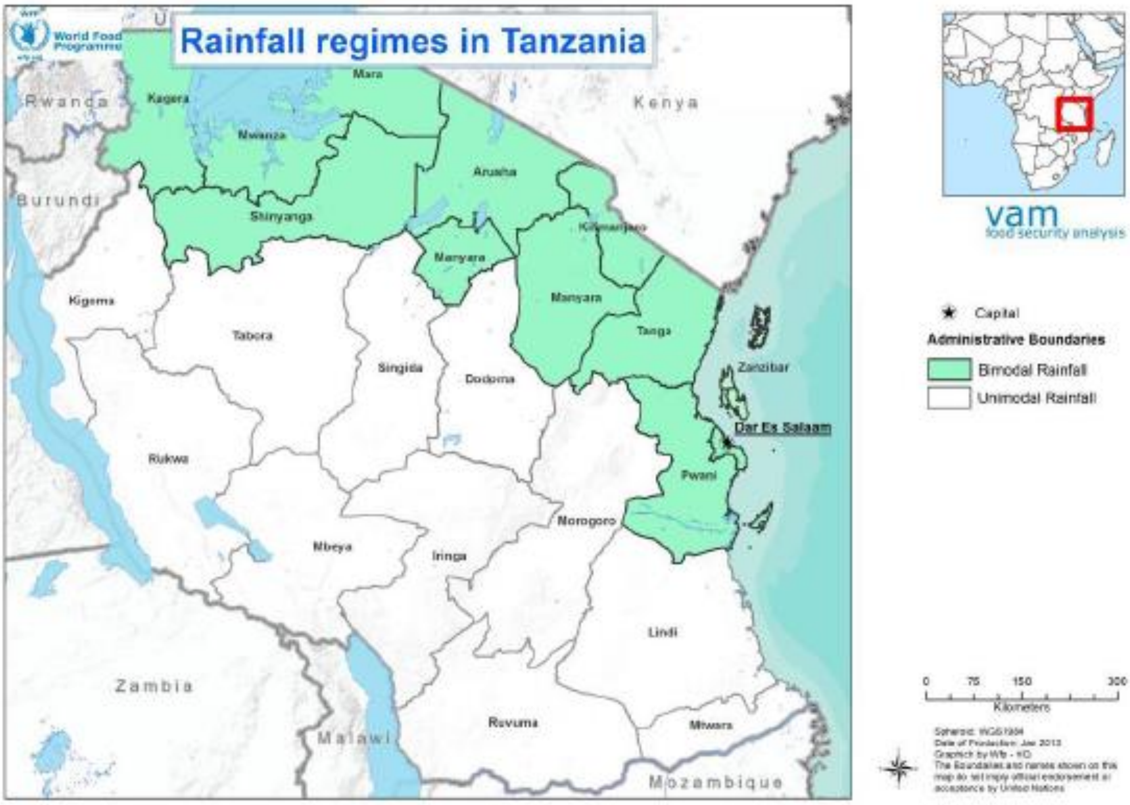


Figure 2: Rainfall variability (difference 2009 mm of rainfall - 2004/08 average) by region and rainfall regime
Source: Authors' elaboration from FAO GIEWS

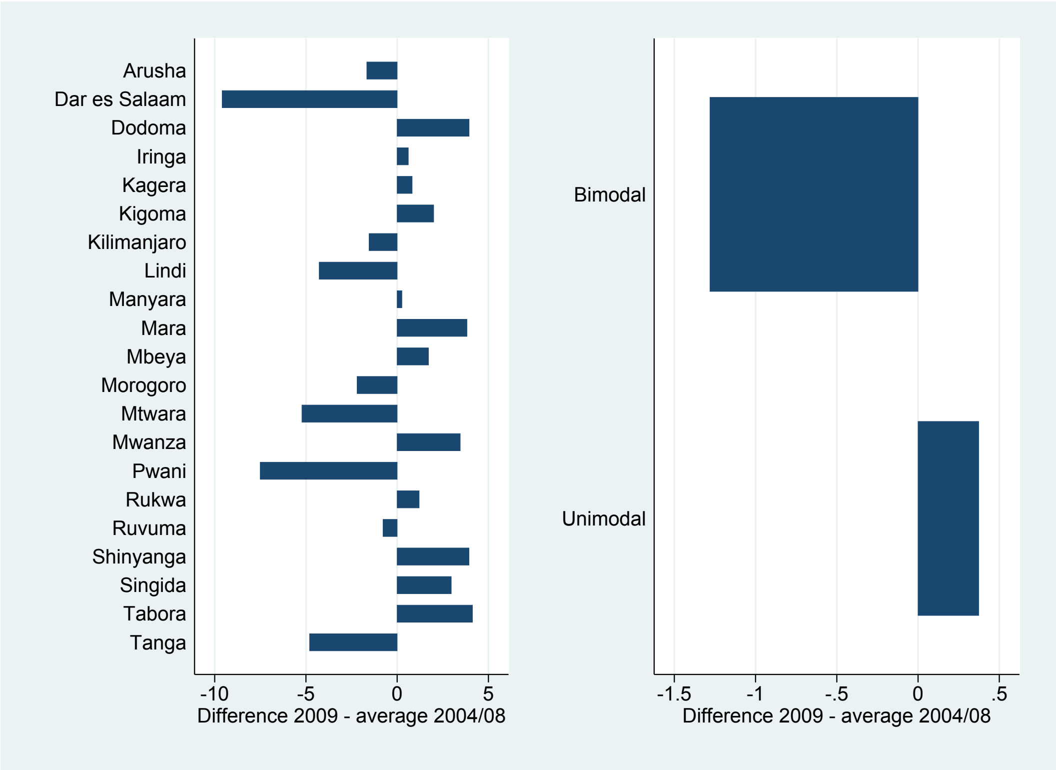


Table 1: Summary statistics of 2004 variables

	Mean	Std. Dev.	Min	Max
Age (HH head)	46.23	17.32	10	99
Male (HH head)	0.75	0.43	0	1
Married (HH head)	0.69	0.46	0	1
Employee (HH head)	0.34	0.48	0	1
Farmer in own field (HH head)	0.90	0.30	0	1
Farmer not in own field (HH head)	0.34	0.48	0	1
HH size	5.32	2.61	1	21
Number of children ($< 6 years$)	1.06	1.02	0	5
Number of adult ($> 18 years$)	2.48	1.17	0	10
Number of educated members[a]	2.90	1.81	0	12
Number of employees	0.71	0.87	0	5
Number of farmers in HH field	3.26	1.94	0	16
Number of farmers not in HH field	0.64	0.80	0	5
Acre of owned plots[b]	3.37	3.93	0	63
Number of cattle[c]	6.53	10.89	0	45
Number of owned enterprises[d]	0.6	0.70	0	2
Toilet[e]	0.93	0.25	0	1
Electricity[f]	0.05	0.22	0	1
Karagwe	0.14	0.34	0	1
Bukoba	0.52	0.50	0	1
Muleba	0.15	0.36	0	1
Biharamulo	0.06	0.24	0	1
Ngara	0.13	0.34	0	1
Total consumption per capita (in TSh)	392,444	249,268	39,456	2,162,022
Food consumption per capita (in TSh)	275,699	179,095	28,475	1,160,340
Observations	1,235			

[a] Number of household members with a primary (from 1 to 8 years of schooling) or secondary education (from 9 to 19 years).
[b] Acre of total plots owned by the household: they include cultivated, rented out or left fallow plots.
[c] Number of total cattle, of all ages, owned by the household at the time of the interview.
[d] Number of all non-farm enterprises owned by household members. They include, for example, selling fruits, car washing, driving car, hair dressing, bicycle transportation and mechanics etc.
[e] Dummy equal to one if the household uses a flush toilet and zero otherwise.
[f] Dummy equal to one is the main source of lighting for the household dwelling is electricity.

Table 2: Summary statistics of 2010 variables by typology of households

	Extended HH				**Same HH head**			
	Without migrants		**With migrants**		**Without migrants**		**With migrants**	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Age (HH head)	45.31	17.09	50.49	18.56	47.96	16.53	59.25	15.06
Sq. Age (HH head)	2345.04	1827.88	2892.94	1995.89	2573.65	1833.05	3736.75	1843.94
Male (HH head)	0.78	0.41	0.66	0.47	0.81	0.38	0.63	0.48
Married (HH head)	0.71	0.45	0.61	0.48	0.74	0.43	0.63	0.48
Employee (HH head)	0.40	0.49	0.28	0.45	0.38	0.48	0.18	0.38
Farmer own field (HH head)	0.89	0.31	0.90	0.28	0.89	0.31	0.92	0.26
Farmer not own field (HH head)	0.41	0.49	0.44	0.49	0.40	0.49	0.41	0.49
HH size	4.80	2.23	4.65	2.44	5.15	2.28	5.20	2.68
N. children ($< 6 years$)	1.02	1.00	0.79	0.97	1.05	1.03	0.64	0.87
N. adults ($> 18 years$)	1.66	1.48	1.55	1.56	1.90	1.48	2.02	1.65
N. educated members	2.89	1.74	2.94	1.79	3.15	1.76	3.44	1.91
N. employees	0.66	0.76	0.54	0.73	0.66	0.79	0.49	0.73
N. farmers in HH field	2.91	1.65	3.20	1.88	3.11	1.71	3.74	1.89
N. farmers not in HH field	0.71	0.80	0.74	0.83	0.71	0.81	0.77	0.82
Acre of owned plots	2.78	2.75	3.41	4.46	3.05	2.94	4.46	5.39
N. cattle	0.61	2.62	0.91	4.34	0.73	2.93	1.56	5.93
N. enterprises	0.71	0.81	0.78	0.92	0.70	0.81	0.81	0.92
Toilet	0.94	0.22	0.95	0.21	0.95	0.20	0.95	0.20
Electricity	0.05	0.23	0.05	0.22	0.04	0.21	0.05	0.22
Bank account	0.12	0.33	0.12	0.33	0.13	0.33	0.16	0.37
Karagwe	0.14	0.35	0.12	0.33	0.14	0.35	0.13	0.33
Bukoba rural	0.32	0.46	0.35	0.48	0.33	0.47	0.35	0.47
Muleba	0.13	0.34	0.19	0.40	0.12	0.33	0.23	0.42
Biharamulo	0.07	0.26	0.08	0.27	0.07	0.25	0.05	0.23
Ngara	0.14	0.35	0.07	0.25	0.14	0.35	0.08	0.28
Bukoba urban	0.18	0.38	0.16	0.37	0.17	0.38	0.14	0.34
Total cons. pc (in TSh)	477,112.90	365,293.20	505,627.00	360,985.80	464,439.00	349,791.80	495,579.40	347,482.70
Food cons. pc (in TSh)	303,919.70	232,828.20	310,091.70	193,605.30	297,578.80	213,831.20	302,675.50	159,426.20
Observations	910		325		686		195	

Table 3: Vulnerability to basic needs and food poverty at 2010

	HHs with migrants	HHs without migrants	t-test[a]
Basic needs			
Poverty rate	0.38	0.38	
Poverty line	12.6	12.6	
Vulnerability mean	0.41	0.43	0.170
	(0.401)	(0.384)	
(estimated) Mean	13.17	13.04	
	(4.288)	(3.731)	
(estimated) Variance	1.01	1.01	
	(0.307)	(0.198)	
Vulnerability rate	0.47	0.53	
Food			
Poverty rate	0.18	0.18	
Poverty line	11.9	11.9	
Vulnerability mean	0.34	0.36	0.240
	(0.415)	(0.385)	
(estimated) Mean	13.20	12.96	
	(6.738)	(12.345)	
(estimated) Variance	1.02	0.99	
	(0.360)	(0.270)	
Vulnerability rate	0.48	0.52	
Sample size	325	910	

[a] p-value of the t-test of equality of vulnerability rates between households with and without migrants.
Standard deviation in parenthesis.

Table 4: Vulnerability variation (2010-2004) by categories of households

	Basic needs		Food	
	Extended HH	Same HH head	Extended HH	Same HH head
With migrants	-0.094	-0.060	-0.002	-0.103
	(0.527)	(0.615)	(0.587)	(0.489)
Without migrants	-0.015	0.001	0.033	0.048
	(0.501)	(0.525)	(0.507)	(0.411)
t-test[a]	0.025	0.431	0.368	0.000
Sample size	1,235	881	1,235	881

[a] p-value of the t-test of equality of vulnerability variation between households with and without migrants.
Standard deviation in parenthesis.

Table 5: The effect of migration (ATT) on VEP variation 2010-2004

	Extended household		**Same household head**	
	(a)	(b)	(a)	(b)
	Basic needs	Food	Basic needs	Food
Propensity score matching				
Nearest-neighbour	-0.153**	-0.0616	-0.0135	-0.155**
	(0.0618)	(0.0659)	(0.0509)	(0.0605)
Caliper	-0.153**	-0.0616	-0.0135	-0.155**
	(0.0618)	(0.0659)	(0.0509)	(0.0605)
Kernel	-0.0878**	-0.0236	-0.0616*	-0.154***
	(0.0352)	(0.0363)	(0.037)	(0.0379)
Matching on multiple variables				
Mahalanobis distance	-0.106**	-0.0351	-0.0132	-0.119**
	(0.0472)	(0.0543)	(0.0513)	(0.0571)
Euclidean distance	-0.0563	-0.0252	-0.0185	-0.168***
	(0.0452)	(0.0559)	(0.0539)	(0.0595)
Ivariance distance	-0.102**	-0.0654	-0.0117	-0.150**
	(0.0483)	(0.0588)	(0.0549)	(0.0597)
Coarsened exact matching				
	-0.0753	-0.0291	-0.0698**	-0.262***
	(0.0731)	(0.059)	(0.0297)	(0.0565)
Observations	1,235		881	
Treated	325		195	
Control	910		686	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Matching and bias-adjustment variables: household size; number of males; number of members with primary education; number of members with secondary education, a dummy for whether the origin household is living in Bukoba urban; number of cattle owned by the household; number of sheep owned by the household; acre of plots owned by the household; dummy for having a bank account, dummy for having experienced a shock; dummy for participating in informal organizations.

Table 6: Impact of migrant location on origin household VEP variation

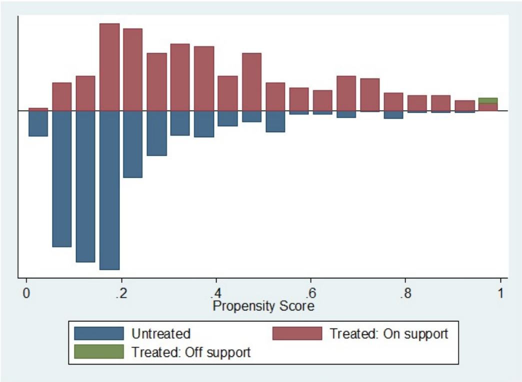
	VEP variation	
	(a)	(b)
	Basic needs	Food
Distance	0.0000	-0.0000
	(0.0001)	(0.0001)
Unimodal	-0.4185*	-0.5514**
	(0.2499)	(0.2219)
Distance * Unimodal	0.0002	0.0006**
	(0.0004)	(0.0003)
Diff Number of cattles	-0.0016	-0.0016
	(0.0027)	(0.0024)
Diff Acre of owned plots	-0.0117*	0.0144**
	(0.0063)	(0.0057)
Diff Number of enterprises	-0.0303	0.0958***
	(0.0404)	(0.0359)
Diff Toilet	-0.2181	-0.3169**
	(0.1441)	(0.1278)
Diff Electricity	0.5160**	0.2309
	(0.2274)	(0.2017)
Constant	-0.1167***	-0.1624***
	(0.0446)	(0.0398)
Observations	325	325
R-squared	0.077	0.113

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix

Figure A1: Common support

(a) Extended HHs



(b) Same HH head

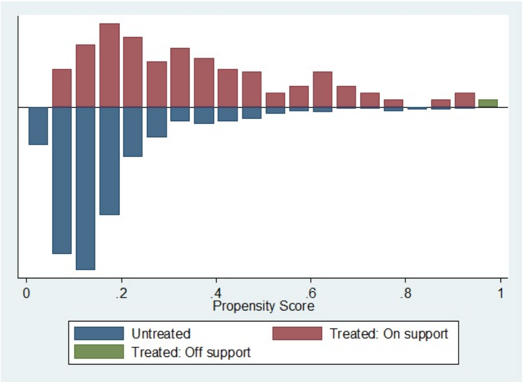


Table A1: Regression results of 2004 (per capita) total consumption

	Total Consumption	**Weighted Total Consumption**
Age (HH head)	0.0032	0.0054
	(0.0051)	(0.0050)
Squared age (HH head)	-0.0000	-0.0001
	(0.0000)	(0.0000)
Male (HH head)	0.1877***	0.1726***
	(0.0532)	(0.0533)
Married (HH Head)	-0.0732	-0.0527
	(0.0526)	(0.0530)
Employee (HH Head)	0.0680*	0.0673
	(0.0399)	(0.0418)
Farmer in own field (HH Head)	-0.0996**	-0.1104***
	(0.0432)	(0.0400)
Farmer not in own field (HH Head)	0.0770*	0.0947**
	(0.0397)	(0.0412)
HH size	-0.1674***	-0.1763***
	(0.0184)	(0.0190)
Squared HH size	0.0062***	0.0066***
	(0.0010)	(0.0011)
Number of children (<6 years)	-0.0567***	-0.0522**
	(0.0203)	(0.0213)
Number of adults (>18 years)	0.0566***	0.0637***
	(0.0177)	(0.0180)
Number of educated members	0.0476***	0.0535***
	(0.0135)	(0.0140)
Number of employees	-0.0318	-0.0260
	(0.0212)	(0.0228)
Acre of owned plots	0.0084**	0.0034
	(0.0040)	(0.0034)
Number of cattle	0.0089***	0.0087***
	(0.0015)	(0.0014)
Number of owned enterprises	0.0751***	0.0637**
	(0.0261)	(0.0272)
Toilet in dwelling	0.1688***	0.1751***
	(0.0542)	(0.0534)
Electricity	0.3996***	0.3744***
	(0.0583)	(0.0612)
Karagwe	0.4035***	0.3951***
	(0.0529)	(0.0513)
Bukoba	0.3698***	0.3491***
	(0.0437)	(0.0423)
Muleba	0.3283***	0.3202***
	(0.0522)	(0.0508)
Biharamulo	-0.1566**	-0.1639***
	(0.0622)	(0.0599)
Constant	12.5189***	12.4932***
	(0.1437)	(0.1420)
Observations	1,235	1,235
R-squared	0.359	0.999

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2: Regression results of 2004 (per capita) food consumption

	Food Consumption	**Weighted Food Consumption**
Age (HH head)	0.0091	0.0098*
	(0.0059)	(0.0057)
Squared age (HH head)	-0.0001*	-0.0001**
	(0.0001)	(0.0001)
Male (HH head)	0.2090***	0.1810***
	(0.0614)	(0.0603)
Married (HH Head)	-0.0751	-0.0549
	(0.0606)	(0.0603)
Employee (HH Head)	0.0400	0.0428
	(0.0460)	(0.0495)
Farmer in own field (HH Head)	-0.0966*	-0.1140**
	(0.0498)	(0.0462)
Farmer not in own field (HH Head)	0.0950**	0.1192**
	(0.0457)	(0.0489)
HH size	-0.1184***	-0.1197***
	(0.0212)	(0.0228)
Squared HH size	0.0045***	0.0047***
	(0.0011)	(0.0012)
Number of children (<6 years)	-0.0655***	-0.0823***
	(0.0234)	(0.0251)
Number of adults (>18 years)	0.0568***	0.0648***
	(0.0204)	(0.0223)
Number of educated members	0.0282*	0.0292*
	(0.0156)	(0.0171)
Number of employees	-0.0219	-0.0205
	(0.0245)	(0.0282)
Acre of owned plots	0.0080*	0.0137**
	(0.0047)	(0.0060)
Number of cattle	0.0081***	0.0070***
	(0.0017)	(0.0016)
Number of owned enterprises	0.0777***	0.0644**
	(0.0300)	(0.0319)
Toilet in dwelling	0.1013	0.1074*
	(0.0625)	(0.0614)
Electricity	0.3253***	0.3332***
	(0.0672)	(0.0744)
Karagwe	0.4024***	0.3895***
	(0.0610)	(0.0638)
Bukoba	0.3504***	0.3572***
	(0.0504)	(0.0504)
Muleba	0.2612***	0.2792***
	(0.0602)	(0.0586)
Biharamulo	-0.2229***	-0.2095***
	(0.0717)	(0.0717)
Constant	11.9410***	11.9382***
	(0.1657)	(0.1643)
Observations	1,235	1,235
R-squared	0.258	0.998

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3: Regression results of 2010 (per capita) total consumption

	Total Consumption	**Weighted Total Consumption**
Age (HH head)	-0.0019	-0.0009
	(0.0049)	(0.0049)
Squared age (HH head)	-0.0000	-0.0000
	(0.0000)	(0.0000)
Male (HH head)	0.1907***	0.2119***
	(0.0537)	(0.0516)
Married (HH Head)	-0.1132**	-0.1331***
	(0.0527)	(0.0497)
Employee (HH Head)	0.0272	0.0300
	(0.0352)	(0.0363)
Farmer in own field (HH Head)	-0.2549***	-0.2654***
	(0.0415)	(0.0402)
Farmer not in own field (HH Head)	0.0914**	0.0995***
	(0.0355)	(0.0360)
HH size	-0.2131***	-0.2177***
	(0.0200)	(0.0197)
Squared HH size	0.0051***	0.0052***
	(0.0012)	(0.0011)
Number of children (<6 years)	-0.0134	-0.0175
	(0.0207)	(0.0220)
Number of adults (>18 years)	0.0917***	0.0877***
	(0.0186)	(0.0196)
Number of educated members	0.1112***	0.1140***
	(0.0148)	(0.0158)
Number of employees	-0.0360*	-0.0376*
	(0.0208)	(0.0224)
Acre of owned plots	0.0204***	0.0211***
	(0.0043)	(0.0044)
Number of cattle	0.0043***	0.0036***
	(0.0011)	(0.0010)
Number of owned enterprises	0.0102	0.0097
	(0.0203)	(0.0205)
Toilet in dwelling	0.0922	0.1012
	(0.0618)	(0.0621)
Electricity	0.5968***	0.5777***
	(0.0547)	(0.0601)
Karagwe	0.2855***	0.2948***
	(0.0497)	(0.0507)
Bukoba	0.0698	0.0623
	(0.0426)	(0.0451)
Muleba	0.0415	0.0318
	(0.0496)	(0.0515)
Biharamulo	-0.1415**	-0.1407**
	(0.0613)	(0.0633)
Constant	13.2530***	13.2654***
	(0.1473)	(0.1446)
Observations	1,235	1,235
R-squared	0.421	0.999

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A4: Regression results of 2010 (per capita) food consumption

	Food Consumption	**Weighted Food Consumption**
Age (HH head)	0.0022	0.0041
	(0.0054)	(0.0051)
Squared age (HH head)	-0.0000	-0.0001
	(0.0001)	(0.0000)
Male (HH head)	0.3318***	0.3650***
	(0.0592)	(0.0549)
Married (HH Head)	-0.1594***	-0.1719***
	(0.0580)	(0.0540)
Employee (HH Head)	0.0179	0.0192
	(0.0388)	(0.0408)
Farmer in own field (HH Head)	-0.2618***	-0.2755***
	(0.0458)	(0.0443)
Farmer not in own field (HH Head)	0.1095***	0.1106***
	(0.0392)	(0.0411)
HH size	-0.1600***	-0.1685***
	(0.0220)	(0.0249)
Squared HH size	0.0037***	0.0048***
	(0.0014)	(0.0017)
Number of children (<6 years)	-0.0115	-0.0193
	(0.0229)	(0.0247)
Number of adults (>18 years)	0.0582***	0.0552**
	(0.0206)	(0.0224)
Number of educated members	0.0840***	0.0812***
	(0.0164)	(0.0180)
Number of employees	-0.0120	-0.0176
	(0.0229)	(0.0257)
Acre of owned plots	0.0152***	0.0173***
	(0.0048)	(0.0057)
Number of cattle	0.0021*	0.0006
	(0.0012)	(0.0009)
Number of owned enterprises	0.0095	0.0123
	(0.0223)	(0.0248)
Toilet in dwelling	0.0822	0.0904
	(0.0682)	(0.0779)
Electricity	0.4825***	0.4666***
	(0.0603)	(0.0660)
Karagwe	0.1433***	0.1451**
	(0.0548)	(0.0593)
Bukoba	0.0194	0.0193
	(0.0469)	(0.0484)
Muleba	-0.0425	-0.0444
	(0.0547)	(0.0562)
Biharamulo	-0.1578**	-0.1693***
	(0.0675)	(0.0654)
Constant	12.6002***	12.5864***
	(0.1624)	(0.1618)
Observations	1,235	1,235
R-squared	0.288	0.998

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A5: Probit and Logit models of having migrants on 2004 HH characteristics by typology of households

	Extended HH		Same HH head	
	(1)	(2)	(3)	(4)
	Probit	Logit	Probit	Logit
HH size	0.111***	0.186***	0.114***	0.190***
	(0.0305)	(0.0529)	(0.0385)	(0.0675)
N. males	-0.107***	-0.180**	-0.150***	-0.258***
	(0.0409)	(0.0705)	(0.0525)	(0.0923)
N. members with primary education	0.240***	0.408***	0.255***	0.446***
	(0.0373)	(0.0653)	(0.0480)	(0.0860)
N. members with secondary education	0.441***	0.766***	0.443***	0.777***
	(0.0877)	(0.159)	(0.111)	(0.205)
Acre of plots	0.000641	-0.000440	0.0161	0.0256
	(0.0120)	(0.0206)	(0.0158)	(0.0278)
Cattle	0.0217	0.0375	0.0170	0.0321
	(0.0142)	(0.0231)	(0.0179)	(0.0294)
Sheep	0.00225	0.00449	-0.00108	-0.00171
	(0.00928)	(0.0166)	(0.00798)	(0.0131)
Bukoba urban	-0.214*	-0.364	-0.254	-0.464*
	(0.128)	(0.224)	(0.156)	(0.280)
Informal organizations	0.0447	0.0964	0.130	0.246
	(0.102)	(0.178)	(0.124)	(0.224)
Bank account	-0.498***	-0.925***	-0.522**	-0.981**
	(0.176)	(0.323)	(0.228)	(0.431)
Shock at 2003	0.172	0.280	0.244*	0.396
	(0.124)	(0.213)	(0.147)	(0.259)
Constant	-1.757***	-2.949***	-1.884***	-3.185***
	(0.135)	(0.244)	(0.162)	(0.298)
Prob > chi2	0.000	0.000	0.000	0.000
Pseudo R2	0.156	0.154	0.158	0.156
Log likelihood	-604.015	-605.078	-391.8044	-392.685
Observations	1,235	1,235	881	881

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A6: Variables used for matching households (extended sample)

Variable	Matched/Unmatched	Mean treated	Mean Control	Bias	Reduction	t	p-value
HH size	U	6.755	4.728	79.8		12.59	0.000
	M	6.656	6.756	-3.9	95.1	-0.45	0.651
Acre of owned land	U	4.370	3.127	29.3		4.82	0.000
	M	4.303	4.107	4.6	84.3	0.5	0.614
Number of cattle	U	1.035	0.436	16.6		2.9	0.004
	M	0.997	0.820	4.9	70.4	0.54	0.587
Living in Bukoba urban	U	0.152	0.172	-5.4		-0.78	0.433
	M	0.153	0.150	0.7	86.9	0.09	0.931
Participating in informal organizations	U	0.766	0.676	20.1		2.88	0.004
	M	0.767	0.743	5.5	72.7	0.68	0.495
N. of males	U	3.166	2.361	49.2		7.88	0.000
	M	3.115	3.131	-1	98	-0.11	0.91
N. of members with primary educ.	U	3.717	2.211	87.1		13.71	0.000
	M	3.677	3.689	-0.7	99.2	-0.08	0.94
N. of members with primary educ.	U	0.359	0.122	37.2		6.41	0.000
	M	0.347	0.304	6.8	81.8	0.68	0.498
N. of sheep	U	2.759	1.739	13.2		2.47	0.014
	M	2.674	2.064	7.9	40.2	0.94	0.347
Having a bank account	U	0.100	0.085	5.3		0.79	0.427
	M	0.097	0.107	-3.5	34.3	-0.4	0.690
Shock	U	0.148	0.131	4.9		0.74	0.462
	M	0.146	0.162	-4.6	6.1	-0.54	0.593
Observations	1,235						

Table A7: Variables used for matching households (same head households sample)

Variable	Matched/Unmatched	Mean treated	Mean Control	Bias	Reduction	t	p-value
HH size	U	6.506	4.546	76		9.88	0.000
	M	6.341	6.437	-3.7	95.1	-0.34	0.732
Acre of owned land	U	4.505	2.844	36.2		5.28	0.000
	M	4.380	3.721	14.4	60.3	1.31	0.189
Number of cattle	U	1.163	0.448	20.5		2.87	0.004
	M	1.108	0.904	5.9	71.4	0.49	0.623
Living in Bukoba urban	U	0.129	0.172	-11.8		-1.36	0.175
	M	0.131	0.142	-3.2	73.4	-0.31	0.759
Participating in informal organizations	U	0.781	0.665	26.2		2.98	0.003
	M	0.778	0.756	5.1	80.4	0.5	0.615
N. of males	U	3.000	2.266	44.3		5.95	0.000
	M	2.909	2.892	1	97.7	0.1	0.923
N. of members with primary educ.	U	3.466	2.077	84.2		10.75	0.000
	M	3.403	3.475	-4.3	94.9	-0.35	0.723
N. of members with primary educ.	U	0.360	0.110	37.8		5.49	0.000
	M	0.330	0.232	14.8	61	1.25	0.212
N. of sheep	U	3.169	1.810	14.5		2.36	0.019
	M	3.034	2.340	7.4	48.9	0.68	0.495
Having a bank account	U	0.112	0.074	13.1		1.64	0.101
	M	0.102	0.095	2.6	80.6	0.23	0.816
Shock	U	0.163	0.130	9.4		1.14	0.256
	M	0.153	0.190	-10.3	-10.4	-0.91	0.365
Observations	881						

Table A8: Impact of migration on household VEP variation (1994-92)

	ATT
Propensity score matching (Caliper)	-0.0652**
	(0.0324)
Matching on multiple variable (Mahalanobis metric distance)	-0.0592**
	(0.0241)
Observations	737

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Matching and bias-adjustment variables: household size; number of males; number of self-employed farmers; number of children ($< 6years$); a dummy for whether the origin household is living in Bukoba urban; number of cattle owned by the household; acre of plots owned by the household.

Table A9: Regression results of 1992 (per capita) total consumption

	Total consumption	**Weighted total consumption**
Age (HH head)	0.006	0.005
	(-0.007)	(-0.007)
Squared Age (HH head)	0.000	0.000
	(0.000)	(0.000)
Male (HH head)	0.017	0.024
	(-0.066)	(-0.066)
Married (HH head)	0.180***	0.173***
	(-0.063)	(-0.063)
Employee (HH head)	0.130***	0.126***
	(-0.047)	(-0.047)
Farmer in own field (HH head)	-0.205***	-0.197***
	(-0.064)	(-0.061)
Self-employed not in agriculture (HH head)	0.166***	0.163***
	(-0.048)	(-0.048)
HH size	-0.115***	-0.117***
	(-0.023)	(-0.023)
Squared HH size	0.00413***	0.00424***
	(-0.001)	(-0.001)
Number of children (< 6 years)	-0.0773***	-0.0796***
	(-0.024)	(-0.024)
Number of adults (> 18 years)	0.0505**	0.0472**
	(-0.020)	(-0.020)
Number of educated HH members	0.029	0.031
	(-0.020)	(-0.020)
Number of cattle	0.0170***	0.0189***
	(-0.003)	(-0.004)
Acre of own plots	0.00108***	0.00110***
	(0.000)	(0.000)
Karagwe	0.595***	10.94***
	(-0.096)	(-0.216)
Bukoba rural	0.414***	10.76***
	(-0.084)	(-0.203)
Muleba	0.529***	10.88***
	(-0.091)	(-0.205)
Ngara	0.221**	10.34***
	(-0.097)	(-0.217)
Bukoba urban	0.799***	10.56***
	(-0.089)	(-0.212)
Month of interview	-0.0540***	11.15***
	(-0.012)	(-0.211)
Constant	10.34***	-0.0528***
	(-0.219)	(-0.012)
Observations	737	737
R-squared	0.330	0.998

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A10: Regression results of 1994 (per capita) total consumption

	Total consumption	**Weighted total consumption**
	Total consumption	Weighted total consumption
Age (HH head)	0.000	0.000
	(-0.006)	(-0.006)
Squared Age (HH head)	0.000	0.000
	(0.000)	(0.000)
Male (HH head)	0.008	0.049
	(-0.059)	(-0.057)
Married (HH head)	0.142**	0.124**
	(-0.056)	(-0.054)
Employee (HH head)	0.0871**	0.0909**
	(-0.043)	(-0.041)
Farmer in own field (HH head)	-0.157**	-0.144**
	(-0.063)	(-0.067)
Self-employed not in agriculture (HH head)	0.178***	0.195***
	(-0.039)	(-0.040)
HH size	-0.106***	-0.120***
	(-0.019)	(-0.022)
Squared HH size	0.00376***	0.00456***
	(-0.001)	(-0.001)
Number of children (< 6 years)	-0.0833***	-0.0857***
	(-0.022)	(-0.023)
Number of adults (>18 years)	0.013	0.001
	(-0.016)	(-0.018)
Number of educated HH members	0.0363**	0.0451***
	(-0.016)	(-0.017)
Number of cattle	0.0136***	0.0124***
	(-0.002)	(-0.001)
Acre of own plots	0.0137***	0.00888***
	(-0.002)	(-0.001)
Karagwe	0.601***	10.79***
	(-0.083)	(-0.201)
Bukoba rural	0.407***	10.60***
	(-0.073)	(-0.189)
Muleba	0.522***	10.72***
	(-0.080)	(-0.192)
Ngara	0.190**	10.26***
	(-0.084)	(-0.203)
Bukoba urban	0.671***	10.33***
	(-0.077)	(-0.192)
Month of interview	-0.0273***	10.88***
	(-0.010)	(-0.193)
Constant	10.17***	-0.0263***
	(-0.197)	(-0.010)
Observations	737	737
R-squared	0.347	0.998

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A11: Summary statistics of variables used for semi-experimental approach

	Mean	Std. Dev.	Min	Max
Distance (km)	279.738	459.533	0.0102	1578.88
Unimodal	0.0843	0.2786	0	1
Distance (km) * Unimodal	54.1165	214.3794	0	1323.13
Diff acre of owned plot	0.1608	6.7044	-55.2	36.5
Diff N. owned cattle	0.4551	3.4000	-16	32
Diff toilet	-0.0281	0.2237	-1	1
Diff electricity	0.0112	0.1499	-1	1
Diff N. of enterprises	0.1404	1.0237	-3	5
Observations	325			