Inter-regional migration in Tanzania:

The role of socio-demographic and environmental factors

Cristina Ocello [1], A. Petrucci [1], M.Rita Testa [2]

[1] *Department of Statistics, University of Florence*
[2] *Vienna Institute of Demography, Austrian Academy of Sciences*

ocello@ds.unifi.it

Abstract

By using data from the first wave of the Tanzania National Panel survey (TZNPS) conducted in 2008/2009, we aim to investigate the relationship between environment and internal migration flows in Tanzania. In the TZNPS survey 16709 individuals nested in 3265 households (and 126 districts) were asked to report the three heaviest environmental shocks experienced in the last five years. We model the inter-regional migration flows occurred in the five years before the time of the survey by assessing the relative importance of socio-demographic and environmental determinants. The main hypothesis is that inter-regional migrations are influenced by environmental shocks at the origin and favorable environmental and socio-economic conditions at the destination. First, we run logistic models with the aim to figure out the most relevant individual predictors of migration. Second, we run Poisson regression models to identify the most important directions and the distance of the migration flows as well as the determinants of these spatial patterns in terms of push and pull factors. Results show that environmental shocks are an important driving force of inter-regional migration flows. However, other factors related to the individual's socio-demographic characteristics as well as to the socio-demographic conditions at the origin and destination region do also play an important role.

Introduction

Over the last few years there has been an upsurge of interest in the potential impact of climate change on human migration. Climate change is not the real direct cause of human migration but it produces environmental events such as cyclones, hurricanes, tsunamis, drought, floods that make it difficult, if not impossible, for people to survive in those places affected by these events. Even the gradual environmental changes, such as desertification, coastal and soil erosion, may have a great impact on the movement of people in the future.

Regions all over the world are likely to experience some adverse effects of climate change. However, developing countries are especially vulnerable because a large share of their economies depends on climate-sensitive sectors. In this vein, this study aims at exploring the complex relationship between climate change and migration in Tanzania, where the impact of environmental factors on migrations is under-investigated.

Data

The data are from the Tanzania National Panel Survey carried out between October 2008 and October 2009 in Tanzania. It is the first part of an ongoing project on Living Standards Measurement Study-Integrated Surveys on Agriculture (LMS-ISA), conducted by the World Bank. This project involves seven countries of Sub-Saharan Africa including Tanzania. One of the objectives of the project is to improve the understanding of the links between agriculture, socioeconomic status, and non-farm income activities. To gather all this information three different questionnaires were administered: 1) a multi-topic household questionnaires, 2) a community questionnaires and 3) a questionnaire focused on agriculture. From these questionnaires, individual-level variables and community-level factors were used to evaluate the impact of environmental events in driving migration flows both at a micro and macro level.

Measures

Environmental shock. In the household questionnaire, there is a specific section to gather information just about environmental shocks. The question is worded as follows: "*Over the past five years, was your household severely affected negatively by any of the following events?"* Different shocks are listed in the response options: shocks that can be considered related to environmental events (as drought, floods), to family (as "Death of a member of household") or economic reasons (as "Loss of salaried employment or non-payment of salary). We considered events like drought, floods, crop disease, loss of land, severe water shortage, livestock died related to environmental shocks.
The year in which the shock occurred is also asked.

Migration. We defined the migrant a person aged 15 or above who moved in the last five years before the interview. Consistently with previous studies (Henry et al.2002 and 2004), we made this age section because it can be reasonably assumed that at age 15 a person begins to made decisions independently from the family. The 5-year period is either 2004-2008 or 2005-2009 depending on the year in which the interview was made. We use the information on migration, present in the household questionnaires, which compare people's address at the time of the interview and during the last five years: this decision is due to the fact that information about environmental events has been recorded just for the past five years before the interview. With the aim to study the effects of environmental changes on migrations, it was necessary to have the right temporal sequence between the two events. Thus, it was considered the people affected by an environmental shock before the migration.

Sample size. The analytical sample includes 9,523 individuals aged 15 years or above. Out of this selected sample 1,078 (11.3%) migrated in the 5-year period before the survey and 246 (2.6%) moved after having experienced an environmental shock.

Methods

We performed analyses at micro and macro levels. At the individual level we used a logistic model where the response variable is measured on a binary scale, migrated/no migrated. At the aggregated level, we used a Poisson regression and a negative Binomial regression, that can be regarded as a generalized Poisson model (Flowerdew and Lovett, 1989). Poisson regression is particularly suited to analysis of migration flows data (Flowerdew 2010). It was introduced in the context of migration analysis by Flowerdew and Aitkin (1982): starting from the gravity models, based on the idea of the Newton's Law of gravity and usually used for migration data, they suggest an alternative method to analyze migration flows. They proposed to fit a gravity model with an assumed Poisson distributed response. Furthermore, Flowerdew in 1991 provided an update account of Poisson models of migration.

In this analysis, at the beginning, a series of Poisson regression models were run to determine the model with the best fit to the data, to identify the most significant flows, the regions that seem most or less attractive for the migrants, always evaluating the influence of the environmental shocks on movements. The basic model, the unconstrained model, has the form:

$$M_{ij=}exp\left(\beta_0 + \beta_1 \ln P_i + \beta_2 \ln P_j + \beta_3 \ln d_{ij} + \beta_4 C_{ij} + \beta_5 X_i\right) + \varepsilon_i$$

where M_{ij} is the predicted number of migrants moving from the region *i* to the region *j*. P_i, P_j and d_{ij} are the independent variables, respectively the origin and destination population size and the distance between *i* and *j*, with the corresponding regression coefficients β_1, β_2 and β_3; β_0 is the intercept term and ε_i the error term. According to the gravity law, it's expected that migration is to be positively related to the population in the origin and destination and negatively related to the distance between the two places.

Thus, the dependent variable was defined as the flows of migrants, aged 15 years and above, from the region *i* to the region *j*. As to calculate the origin and destination population size was used the sample of the TZNPS aged 15 years and over.

The distance was computed between centroids derived for each of the 26 regions of Tanzania with the Euclidean metric, measured in kilometers. This distance measure may not capture the reality of the situation, most of the migration between i and j may be shorter distances. Thus, to improve the fit of the model, a contiguity dummy variable between regions, C_{ij}, was included in the model: for each pair of regions which border on each other the value of the variable is 1, otherwise 0. Furthermore, to evaluate the effects of the environmental shocks on migrations, it was added a new variable, X_i, that represents the number of environmental shocks occurred in each region of origin flows over the last five years before the interview.

Many migrants don't move independently: they can move with the family, friends, partners and this has the effect of increasing the variance to be greater than the mean. This phenomena takes the name of overdispersion. The most common approach to this issue is to try negative Binomial regression. The negative Binomial distribution, like the Poisson, can only have a non negative integers as its values, but, unlike the Poisson, it has two parameters: one is the mean and the other, alpha, is an indication of the spread of the data around the Poisson mean, which can takes values greater than 1. The Negative Binomial model can be regarded as a generalized Poisson model (Flowerdew and Lovett, 1989), with mean μ_i and the variance $(\mu_i + \alpha\mu_i^2)$. It follows that when the overdispersion is zero the Poisson model is obtained. Looking at the data it's possible to observe that the value of the mean is much lower than the variance, 0.37 and 2.5 respectively, so the data are affected by overdisperison. In this way the Poisson assumption of equal mean and variance is violated and thus the use of a Poisson model may not be appropriate. Thus, to overcome this problem of overdispersion in the data, a negative binomial model was implemented.

The whole number of flows analyzed in the models is 650 corresponding to all possible 26*25 combinations between the regions of the country. All models were fitted using STATA Version 11.

Results

Descriptive results

Table 1 shows the number of immigrants, emigrants, net migrations and immigration-emigration rates for the 26 regions of Tanzania in the last 5 years. The regions are grouped into eight zones: North, South, East, West, Centre, Lake, South of Highlands and Islands. Dar es Salaam has the highest immigration rate (7.71%), followed by Kusini Unguja (4.44%) and Mjini/Magharibi Unguja (4.18%). Immigration and emigration rates are also described in the geographical maps shown in ***Figures 1 and 2*** respectively**.** Dar es Salaam is in the central part of Tanzania while Kusini Unguja and Mjini/Magharibi Unguja are the southern and urban-western part of the island of Unguja, known by all as Zanzibar. As mentioned above, Dar es Salaam and Mjini/Magharibi Ungujia are the two hotspots of the country presenting the highest values of net migrations. The five highest emigration rates concerned the regions of Singida (9.52%), Kilimanjaro (7.03%), Pwani (6.88%), Mara (6.58%) and Morogoro (6.53%). In ***Figure 2*** we also show the distribution of the environmental shocks occurred in each region of the country before the migration. The regions more affected by enviromental shocks are Tanga in the north part, Shinyaga in the western, Matwara in the southern and Dar es Salaam in the Eastern.

Table 1 – Number of migrants, net migration and immigration-emigration rates for zones in the last 5 years before the interview, Tanzania

Region	Sample of residents aged 15 and above	Proportion of residents aged 15 years and above	Immigrants	Emigrants	Net migration	Immigration rate (%)	Emigration rate (%)
Northern							
Arusha	268	2.81	5	7	-2	1.87	2.61
Manyara	219	2.30	3	1	2	1.37	0.46
Kilimanjaro	313	3.29	4	22	-18	1.28	**7.03**
Tanga	318	3.34	0	13	-13	0.00	4.09
Central							
Dodoma	220	2.31	2	12	-10	0.91	5.45
Singida	168	1.76	2	16	-14	1.19	**9.52**
Eastern							
Morogoro	291	3.06	4	19	-15	1.37	**6.53**
Pwani	189	1.98	3	13	-10	1.59	**6.88**
Dar es salaam	1660	17.43	128	11	**117**	**7.71**	0.66
Western							
Tabora	396	4.16	8	6	2	2.02	1.52
Kigoma	305	3.20	0	6	-6	0.00	1.97
Shinyanga	493	5.18	12	14	-2	2.43	2.84
Southern							
Lindi	377	3.96	1	12	-11	0.27	3.18
Mtwara	497	5.22	5	6	-1	1.01	1.21
Ruvuma	369	3.87	0	3	-3	0.00	0.81
Southern Highlands							
Iringa	357	3.75	8	16	-8	2.24	4.48
Mbeya	368	3.86	4	5	-1	1.09	1.36
Rukwa	236	2.48	0	3	-3	0.00	1.27
Lake		0.00			0		
Kagera	311	3.27	9	7	2	2.89	2.25

Mwanza	448	4.70	9	14	-5	2.01	3.13
Mara	152	1.60	1	10	-9	0.66	**6.58**
Islands							
KASKAZINI UNGUJA (Zanzibar North)	208	2.18	4	9	-5	1.92	4.33
KUSINI UNGUJA (Zanzibar South)	90	0.95	4	1	3	**4.44**	1.11
MJINI/MAGHARIBI UNGUJA (Zanzibar Urban and west)	717	7.53	30	5	**25**	**4.18**	0.70
KASKAZINI PEMBA (North of Pemba)	261	2.74	0	4	-4	0.00	1.53
KUSINI PEMBA (South of Pemba)	292	3.07	0	7	-7	0.00	2.40
Burundi				3			
Mozambique				1			
Total	**9523**	**100.00**	**246**	**246**			

Figure 1 – **Immigration rate for each region of Tanzania**

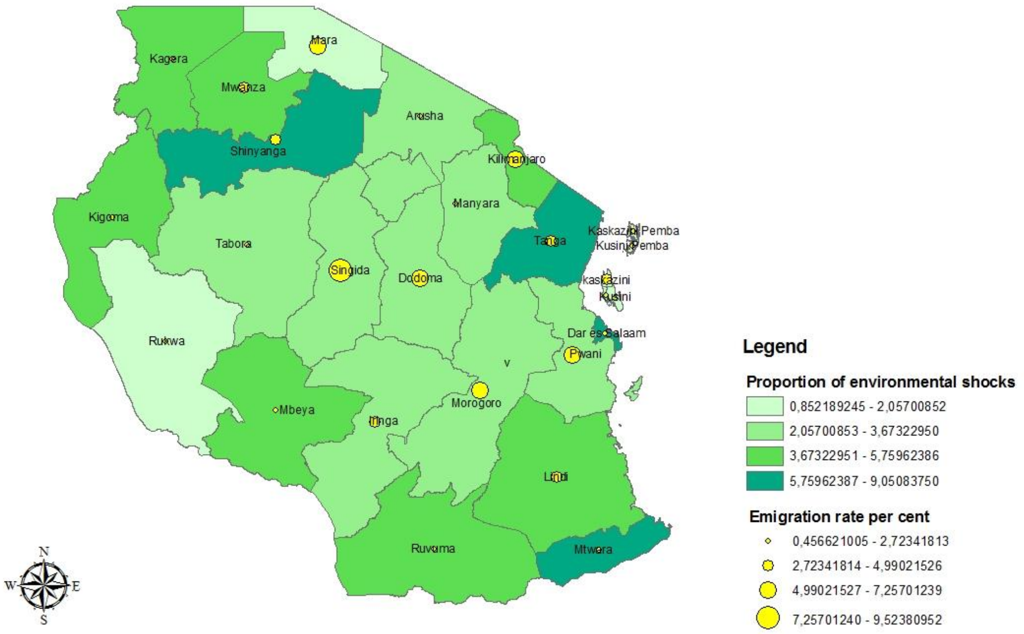

Figure 2 – **Proportion of number of environmental shocks and rate of emigration for each region of Tanzania**

Table 2 – Odds ratio for inter-regional migration in Tanzania, standard errors and p-value of the complete model. Years 2004-2008, 2005-2009.

Variable	**Odds ratio**	**Standard errors**	**P > |z|**
[30-60) years	0.353	0.097	0.000
60 years and over	0.688	0.449	0.567
Sex - female	1.845	0.414	0.006
No education	0.018	0.006	0.000
High education (secondary or university)	0.663	0.183	0.136
Environmental shock - dummy variable	3.630	1.029	0.000

Table 3 - Negative Binomial model parameter estimates and standard error for inter-regional migration in Tanzania. Years 2004-2008, 2005-2009.

	Negative Binomial model				
Variable	Parameter (β)	SE	exp(β)	se	P > |z|
ln (origin population)	0.172	0.30	1.188	0.35	0.561
ln (destination population)	2.506	0.20	12.257	2.51	0.000
ln (distance in km)	-1.365	0.23	0.255	0.06	0.000
Contiguity	1.311	0.32	3.708	1.29	0.000
ln (shock_E)	0.680	0.32	1.973	0.63	0.032
Constant	-13.179	2.36			0.000
ln (alpha)	0.654		1.923	0.45	
No. Of observations	650				

Likelihood-ratio test of alpha= 0: chibar2(01) = 148.29 Prob ≥ chibar2 = 0.000

Further steps

The results presented here clearly indicate that there is a strong association between environmental events and migration flows in Tanzania. This result holds at micro and macro level.

The next step will be to develop a multi-level model including a spatial component. This innovative approach will allows us to circumvent the problems associated with the ecological and atomistic fallacy while at the same time will provide us a model for the different spatial distribution of migration flows that overcomes the problems of multilevel analyses, (i.e., a predefined hierarchy of spatial units).

References

Anselin L. (1998), *Spatial Econometrics: Methods and Models.* Kluwer Academic, Dordrecht.

Black R, Kniveton D, Skeldon R, Coppard D, Murata A, et al.(2008), Demographics and climate change: future trends and their policy implications for migration. Working Paper, Brighton, Development Centre on Migration, Globalisation and Poverty, University of Sussex; T-27.

Flowerdew R. (1982), Fitting the Lognormal Gravity Model to Heteroscedastic Data, *Geographical Analysis* **14** (3): 263-267.

Flowerdew R. (1991), Poisson Regression Modelling Of Migration. In J. Stillwell and P. Congdon (Eds.), *Migration Models: Macro and Micro Approaches*., Chapter 6, pp. 92-113. London, England: Belhaven Press.

Flowerdew R. (2010), Modeling Migration with Poisson Regression. In J. Stillwell, O. Duke-Williams and A. Dennett (Eds.), *Technologies for Migration and Commuting Analysis: Spatial Interaction Data Applications,* Chapter 14, pp. 261-279. IGI Global.

Flowerdew R. and M. Aitkin (1982). A Method Of Fitting The Gravity Model Based On The Poisson Distribution, *Journal of Regional Science* **22** (2): 191-202.

Flowerdew R. and A. Lovett (1988). Fitting Constrained Poisson Regression Models To Interurban Migration Flows, *Geographical Analysis* **20** (4): 297-307.

Frank Laczko and Christine Aghazarm (2009). *Migration, Environment and Climate change: assessing the evidence*. International Organization for Migration (IOM), Geneva.

Franziska Kroll, Dagmar Haase (2009), Does demographic change affect land use patterns? A case of study from Germany, *Land Use Policy,* **27**: 726 -737.

Goldstein H. (2003), *Multilevel Statistical Models*. Arnold, London.

Grothmann T, Patt A. (2005), Adaptive capacity and human cognition: the process of individual adaptation to climate change. *Glob Environ Change,* **15**:199–213.

Haining Robert P., Kerry R. , Oliver Margaret A. (2010), Geography, Spatial Data Analysis, and Geostatistics: An Overview, *Geographical Analysis***, 42**: 7-31.

Haining R. (2003), *Spatial Data Analysis: Theory and Practice*. Cambridge University Press.

Henry S., Schoumaker B., Beauchemin C. (2004), The Impact of Rainfall on the First Out-Migration: a Multilevel Event-History Analysis in Burkina Faso, *Population and Environment* **25** (5): 423-60.

Henry S., Boyle P., Lambin Eric F. (2003), Modelling inter-provincial migration in Burkina Faso, West Africa: the role of socio-demographic and environmental factors, *Applied Geography* **23**: 115–136.

Hox, J. (2002), *Multilevel Analysis: Techniques and Applications.* Lawrence Erlbaum Associates , Hillsdale, NJ.

Perch-Nielsen Sabine L., Bättig Michèle B., Imboden D. (2008), Exploring the link between climate change and migration, *Climatic Change*, **91**: 375 -393.

Petrucci A., Bocci C., Borgoni R., Civardi M., Salvati N., Salvini S. and Vignoli D. (2009), Indagine sulla georeferenziazione dei dati nella statistica ufficiale, Rapporto di Indagine, Commissione per la Garanzia dell'Informazione Statistica, Presidenza del Consiglio dei Ministri, Roma.

Piguet E. (2010), Linking climate change, environmental degradation, and migration: a methodological overview, *Climate Change*, **1:** 517-524.

Reibel M. (2007), Geographic Information Systems and Spatial Data Processing in Demography: a Review, *Population Research and Policy Review*, **26**: 601-608.

Saldaña-Zorrilla Sergio O., Sandberg Krister (2009), Spatial econometric model of natural disaster impacts on human migration in vulnerable regions of Mexico, *Disasters,* **33**(4): 591−607.

Snijders T. A. B., Bosker, R. J. (1999), *Multilevel Analysis: an introduction to basic and advanced multilevel modelling*. SAGE Publications , Thousand Oaks, CA.

Stevens Candice, *Measuring Sustainable Development.* Statistics Brief, OECD September 2005. No 10.

UNDP (2008), A Survey of Composite Indices Measuring Country Performance: 2008 Update, A UNDP/ODS Working Paper, Office of Development Studies United Nations Development Programme, New York.

Warner K, Ehrhart C, de Sherbinin A, Adamo S, Chai-Onn T (2009), In Search of Shelter - Mapping the Effects of Climate Change on Human Migration and Displacement. CARE/CIESIN/UNHCR/UNU-EHS/World Bank.

Zolnik EJ. (2009), Context in human geography: a multilevel approach to study human–environment interactions, *Prof Geogr* , **61**:336–349.