Environmental shocks and internal migration in Tanzania

Shock ambientali e migrazioni interne in Tanzania

Maria Francesca Marino, Alessandra Petrucci, and Elena Pirani

Abstract There is substantial evidence that climate has been changed in recent decades, and this trend will intensify in the near future. Climate changes are expected to exert additional pressure on areas which mainly rely on agricultural activities and, as result, act as push factors for population movements. By implementing a gravity-type model, in this study we aim at examining migration flows across Tanzanian districts. We analyze the potential impact of climate changes on inter-district migration in Tanzania, while controlling for socio-economic and geo-physical features for both the district of origin and destination.

Abstract Numerosi studi dimostrano come il clima sia cambiato negli ultimi decenni e come questa tendenza si intensificherà nel prossimo futuro. Si prevede che tali cambiamenti avranno ripercussioni soprattutto su aree la cui attività economica è prevalentemente basata sull'agricoltura e, in tali aree, agiranno come fattore di spinta per le migrazioni. Implementando un modello gravitazionale, in questo studio si propone un'analisi delle migrazioni interne in Tanzania. In particolare, si analizza l'effetto di shock ambientali sui flussi migratori inter-distrettuali in Tanzania, tenendo conto delle diverse caratteristiche socio-economiche e geofisiche delle aree interessate dal processo migratorio, ovvero quella di partenza e quella di arrivo.

Key words: Climate change, gravity-type models, population flows, socio-economic conditions.

Maria Francesca Marino
Dipartimento di Statistica, Informatica, Applicazioni, Università degli Studi di Firenze, e-mail: mariafrancesca.marino@unifi.it

Alessandra Petrucci
Dipartimento di Statistica, Informatica, Applicazioni, Università degli Studi di Firenze, e-mail: alessandra.petrucci@unifi.it

Elena Pirani
Dipartimento di Statistica, Informatica, Applicazioni, Università degli Studi di Firenze, e-mail: elena.pirani@unifi.it

1 Introduction

Recent studies have clearly shown that climate has been changed in recent decades and that this trend will not only persist but will also intensify in the near future. Climate changes are expected to have an impact especially on those areas whose economy mainly relies on agricultural activities; here, they are expected to exert additional pressure on the economy and, as a result, act as push factors for population movements. Despite the growing acknowledgement that the environment is becoming increasingly important in explaining large-scale movements of migrants, some of whom have been described as environmental refugees [5, 6, 8], understanding the association between migrations flows and changes in climate conditions remains a challenging task, especially in developing countries. Most of the previous studies are based on a micro approach [10, 7]. That is, they focus on the individual migrating unit (person, group or household) and study those factors that influence the decision of the potential migrant to remain in the current location or to move. In contrast, when considering a macro perspective, the interest is in analyzing aggregated migration flows, and understanding the relation between migration and objectively-determined macro variables, such as population sizes, economic descriptors of the context, and/or environmental conditions.

In this study, we follow the macro approach and focus on migration flows across Tanzania districts. This country represents an interesting case study due to its diversity in terms of agro-climatic conditions and ecological zones, and for the richness of available data on climate conditions. In this respect, our aim is that of analyzing the potential impact of climate changes on inter-districts migration in Tanzania via a gravity-type model. This also allows us to properly control for socio-economic and geo-physical features of both the district of origin and destination.

2 Tanzania National Panel Survey

Tanzania National Panel Survey (TNZPS) is a longitudinal survey carried out in Tanzania between October 2008 and January 2016. The survey collects information on a wide range of topics, including agricultural production, non-farm income generating activities, consumption expenditures, and other socio-economic features, with the aim of monitoring poverty dynamics in the country. The survey is made up by three different questionnaires: (*i*) the household questionnaire, which collects information on households economic conditions (expenditures, loans and credits, crimes and justice, etc...) and on education, labor, and health conditions of household members; (*ii*) the agriculture questionnaire, which collects information on household's agricultural activities in terms of production and sales; (*iii*) the community questionnaire, which provides information on physical and economic infrastructure and events in surveyed communities. The first wave of the survey (October 2008 - September 2009) provides information on $3,265$ households; the second (October 2010 - November 2011) and the third (October 2012 - November 2013) provide information on $3,924$ and $5,010$ households, respectively. At the last wave (October 2014 - January 2016), $3,352$ households were involved in the study. In the following, we will focus on data from the household questionnaire coming from the third wave of the survey, that is, that providing the highest sample size.

Combining information on the household's district of residence with the answers provided by the household's head to items “For how many years have you lived in this community?” and “From which district did you move?”, we derive information on: (*i*) the migration status of the household (the household migrated if the household's head moved from one district to another in the 5 years prior to the interview); (*ii*) the district of origin and of destination of the household. By aggregating these information at the district level, we are able to build a migration variable, M_{ij}, counting the number of households migrating between district i and j during the five years prior to the interview, with $i = 1, \dots, 129$, $j = 1, \dots, 129$, and $j \neq i$.

3 A gravity-type model for migration flows across Tanzanian districts

The use of gravity models represents a well-established practice in the economic literature. Recently, thanks to the enhanced availability of migration data, this class of models has attracted researcher's interest also for the analysis of migration flows. The traditional gravity model (e.g. [11] , so-called by analogy with Newtons gravity law) is defined as

$$M_{ij} = \frac{P_i P_j}{d_{ij}},$$

where, as before, M_{ij} denotes the number of recorded migrants between area i and j, P_i and P_j the corresponding population size, and d_{ij} denotes the distance between the centroids of i and j. This equation has the nice property that, taking the logarithm of both sides, we end up with a multiple linear regression model, which is simply to fit and to interpret. However, the restrictive assumptions upon which the log-normal model is built have led researchers to consider more elaborated gravity-type models. Two main extensions have been explored: (*i*) migration flows are assumed to follow a Poisson distribution, rather than a log-normal one; (*ii*) together with the origin and destination populations and the corresponding distance, additional environmental, demographic, and socio-economic factors are included in the model with the aim of providing a better description of the phenomenon under investigation. See e.g. [4, 2, 3]. In this framework, migration flows are modeled according to the following regression model:

$$\log M_{ij} = \beta_0 + \beta_1 P_i + \beta_2 P_j + \beta_3 d_{ij} + x_i' \gamma_1 + x_j \gamma_2 + w_i' \phi_1 + w_j' \phi_2,$$

where x_i and x_j denote the vector of environmental covariates for the area of origin and destination, respectively, and γ_1 and γ_2 the corresponding vectors of parameters. Here we focus on three environmental shock variables: for each district, we consider the percentage of households in the survey indicating to have been affected by (*i*) drought or floods, (*ii*) crop disease or pests, (*iii*) severe water shortage, on the total number of households. These events have been cited as negative important events for a relevant part of the Tanzanian population in the period 2007/08 – 2012/13, with a substantial variation across districts (Table 1).

On the other hand, w_i and w_j denote the vector of socio-economic covariates, with ϕ_1 and ϕ_2 summarizing the corresponding effects on the (transformed) response M_{ij}. Given the scarcity of information on contextual socio-economic characteristics at district level from external databases, we decided to derive such information directly from the TNZSP survey. Based on previous theoretical and empirical literature [9], we built some indicators describing the demographic structure, the level of social and economic development and urbanization, and some geo-physical features of each district. In preliminary analyses, we tested a large set of indicators. At end, the variables included in the model are those listed and described in Table 1, together with their distribution across Tanzanian districts.

Table 1: Gravity-type model for migration flows: variables description

Variable	**Description**	**Mean**	**Std dev.**	**Min**	**Max**
Environmental variables					
Drought or floods	n. of household indicating to have been affected by drought or floods in the last 5 years	22.5	16.5	0	78.1
Crop disease or pests	n. of household indicating to have been affected by crop disease or crop pests in the last 5 years	13.0	14.4	0	87.9
Severe water shortage	n. of household indicating to have been affected by severe water shortage in the last 5 years	20.7	11.7	0	63.6
Demographic and socio-economic variables					
Cultivated area	% under agriculture in the area	27.4	15.9	0.8	84
Mean age	Mean age of the population	23.3	3.1	17	33.5
Urbanization	% of persons living in urbanized areas	22.5	28.4	0	100
Education	% of people with secondary of higher education	16.6	15.9	0	74.9
Tenure status	% of households of property	74.5	17.2	30.6	100

4 Model results

To evaluate the role of environmental shocks in internal migration, we estimated a set of models. First, we considered the basic gravity model (Model 1) including only the population size of the district of origin and of destination and the distance between them. Second, we added to Model 1 the variables associated to perceived environmental shocks in the district of origin and of destination (Model 2). Finally, we included also those variables describing the socio-economic context of the districts involved in the migration (Model 3). Model parameter estimates are reported in Table 2. In the last lines of the table, we also report the value of the log-likelihood function, the deviance, and the AIC index [1] of all three models under investigation to be used for model comparison.

Table 2: Estimates, standard errors, and p-values of model parameters under different model specifications

	Mod 1			Mod 2			Mod 3		
	Est.	Std Er.	p-value	Est.	Std Er.	p-value	Est.	Std Er.	p-value
Intercept	0.554	0.046	0.000	0.673	0.061	0.000	0.642	0.394	0.103
$\text{Pop}_i/1000$	0.002	0.000	0.000	-0.001	0.001	0.046	-0.002	0.001	0.025
$\text{Pop}_j/1000$	0.002	0.000	0.000	0.003	0.001	0.000	0.002	0.001	0.006
d_{ij}	-0.244	0.016	0.000	-0.238	0.016	0.000	-0.235	0.016	0.000
Crop disease or pests$_i$				-0.017	0.005	0.001	0.020	0.005	0.000
Crop disease or pests$_j$				-0.013	0.005	0.009	0.015	0.006	0.006
Drought or floods$_i$				0.020	0.005	0.000	-0.010	0.006	0.092
Drought or floods$_j$				0.003	0.005	0.605	-0.010	0.006	0.095
Water shortage$_i$				0.024	0.004	0.000	0.020	0.005	0.000
Water shortage$_j$				0.005	0.004	0.222	0.004	0.005	0.478
Urbanization$_i$							0.002	0.001	0.097
Urbanization$_j$							0.006	0.001	0.000
Cultivated area$_i$							-0.005	0.002	0.002
Cultivated area$_j$							0.002	0.002	0.266
Education$_i$							0.000	0.002	0.908
Education$_j$							0.007	0.002	0.001
Tenure status$_i$							0.000	0.002	0.845
Tenure status$_j$							0.009	0.002	0.000
Mean Age$_i$							-0.007	0.010	0.484
Mean Age$_j$							-0.036	0.012	0.003
Deviance	975.774			931.291			856.831		
AIC	3049.133			3016.651			2962.190		

By looking at the last three lines of Table 2, we may observe that the optimal specification includes in the linear predictor both environmental and socio-economic factors. In particular, by comparing the deviance of Model 2 with that of Model 1, we obtain a Likelihood Ratio Test (LRT) equal to 44.48 with 6 degrees of freedom; the low p-value (< 0.001) associated to such a statistic leads us to prefer the more complex model specification (i.e. Model 2). Similarly, when comparing Model 3 with Model 2, we obtain a LRT equal to 74.46 with 10 degrees of freedom. Once again, the low p-value (< 0.001) of such a test statistic leads us to retain the more complex model specification which includes in the linear predictor the population size of the district of origin and destination, the distance between then, the environmental variables, as well as the socio-economic ones. Identical conclusions can be drawn when looking at the AIC values reported in the last line of Table 2. Based on these findings, we will discuss in the following only results obtained under Model 3.

By looking at the estimates derived under the optimal model specification, we may firstly observe that the largest the population of destination, the biggest the migration flow, whereas, the population of origin seems to be negatively linked to migratory movements (significance level at 5%). In agreement with the standard theory behind gravity-type models, longer distances discourages movements (coefficient equal to -0.23); that is, when Tanzanian inhabitants migrate, they generally tend to prefer districts which are not that far from the district of origin.

As for environmental factors, we found that being affected by crop diseases or crop pest, or by a severe water shortage are important factors in pushing internal migrations (both coefficients equal to 0.020): as expected, the higher is the number of households experiencing one of such environmental shocks, the higher is the likelihood for them to move. On the other hand, experiencing droughts or floods does not seem to have a significant effect on migration flows. Also crop-related shocks associated to the district of destination are positively associated to migration flows, even if the magnitude of this effect is lower (coefficient equal to 0.015). This result may be possibly related to the negative sign of the parameter associated to the distance between districts. Tanzanian households tend to prefer closer destinations, which are more likely to be affected by similar problems in term of crop disease or pests with respect to the area of origin. Overall, estimated parameters suggest that environmental variables mainly act as as pushing factors for migration. Similarly, we found that the percentage of cultivated area of the district of origin – a characteristic which is related to the environment, although not representing a shock – is negatively related to inter-districts migrations: the average number of migrants is lower for those districts characterized by a low percentage of cultivated areas.

As for the variables describing the demographic, the social and the economic characteristics of the district, we found that migrants prefer destinations where the population is younger (probably the more dynamic areas). The variables referring to the percentage of population living in urban areas and the percentage of people who are owner of their home can be interpreted as proxies of the richness and development of an area. In this sense, the higher these indicators are for the districts of destination, the higher is the incoming migration flow they attract. Finally, in line with the literature, the attractiveness of an area increases when the corresponding educational level is higher. Generally, when considering demographic and socio-economic factors, it seems that they are relevant especially for the district of destination; in other words, such factors appear to be important as pull factors rather than as pushing ones.

We tested the effect of other objective climatic variables, coming from external databases, such as temperature and rainfall, but none of them were statistically significant. In the analysis, we also considered other demographic and socio-economic variables (such as the percentage of household headed by a woman; the illiteracy rate; the proportion of male population), but once again they were not significant.

5 Conclusions

The results of this analysis highlight that, as expected, demographic and socio-economic characteristics of Tanzanian districts are correlated to migration patterns, and that environmental variables are also significant. In particular, the contribution of environmental variables is mainly important in explaining the departure from a given district, whereas demographic and socio-economic characteristics are especially relevant in explaining the destination of the migration process. It is also worthwhile noticing that the distance remains the most important factor in determining the district of destination (the coefficient, equals to -0.23, has the biggest magnitude). In conclusion, Tanzanian inter-districts migrations are especially pushed by crop-related shocks and severe water shortage. At the same time, migrants generally prefer closer destinations, where they can expect to improve their social and economic situation (i.e., more developed and higher educated areas).

In the next steps of the analysis, we aim at adding a contiguity variable in the model specification, which is expected to better characterize migration flows (previous literature stated that areas that share a boundary tend to have more migration between them because this migration will include a proportion of short-distance moves from one side of the boundary to the other). Moreover, we aim at improving the description of the origin and destination districts, testing the effect of other environmental and socio-economic variables.

References

1. AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory* (1973), B. N. Petrov and F. Csaki, Eds., Akadémiai Kiado, pp. 267–281.
2. FLOWERDEW, R. Poisson regression modelling of migration. *Migration models: Macro and micro approaches* (1991), 92–112.
3. FLOWERDEW, R. Modelling migration with poisson regression. *Technologies for migration and commuting analysis: Spatial interaction data applications* (2010), 261–279.
4. FLOWERDEW, R., AND AITKIN, M. A method of fitting the gravity model based on the poisson distribution. *Journal of regional science 22* (1982), 191–202.
5. MYERS, N. Environmental refugees. *Population and environment 19* (1997), 167–182.
6. MYERS, N., KENT, J., ET AL. *Environmental exodus: an emergent crisis in the global arena*. Climate Institute, 1995.
7. OCELLO, C., PETRUCCI, A., TESTA, M. R., AND VIGNOLI, D. Environmental aspects of internal migration in tanzania. *Population and Environment 37* (2015), 99–108.
8. RAMLOGAN, R. Environmental refugees: a review. *Environmental conservation 23* (1996), 81–88.
9. STILLWELL, J. Inter-regional migration modelling-a review and assessment. Tech. rep., European Regional Science Association, 2005.
10. WAJDI, N., MULDER, C. H., AND ADIOETOMO, S. M. Inter-regional migration in indonesia: a micro approach. *Journal of Population Research 34* (2017), 253–277.
11. ZIPF, G. K. The P_1P_2/D hypothesis: on the intercity movement of persons. *American sociological review 11* (1946), 677–686.