DISCUSSION PAPER SERIES

IZA DP No. 6923

The Effect of Weather-Induced Internal Migration on Local Labor Markets: Evidence from Uganda

Eric Strobl
Marie-Anne Valfort

October 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Effect of Weather-Induced Internal Migration on Local Labor Markets: Evidence from Uganda

Eric Strobl
Ecole Polytechnique Paris, IPAG Research Lab and IZA

Marie-Anne Valfort
PSE - Paris 1 Panthéon Sorbonne University

Discussion Paper No. 6923
October 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 6923
October 2012

ABSTRACT

The Effect of Weather-Induced Internal Migration on Local Labor Markets: Evidence from Uganda*

Relying on census data collected in 2002 and historical weather data for Uganda, we estimate the impact of weather-induced internal migration on the probability for non-migrants living in the destination regions to be employed. Our results reveal a significant negative impact. Consistent with the prediction of a simple theoretical model, they further show that this negative impact is significantly stronger in regions with lower road density and therefore less conducive to capital mobility: a 10 percentage points increase in the net in-migration rate in these areas decreases the probability of being employed of non-migrants by more than 20 percentage points.

JEL Classification: E24, J21, J61, Q54, R23

Keywords: weather shocks, internal migration, labor market, Sub-Saharan Africa

Corresponding author:

Eric Strobl
Département d'Économie
Ecole Polytechnique
91128 Palaiseau Cedex
France
E-mail: eric.strobl@polytechnique.edu

* This research was funded by the Agence Nationale de Recherche, Grant ANR-JCJC-0127-01. We thank Milo Bianchi, Pierre Cahuc, Marie Gachassin, Mathilde Maurel, Katrin Millock, and participants at the DIAL Development Conference "Shocks in developing countries" (Paris, 2011) for their generous guidance and helpful comments.

1 Introduction

There is widespread evidence from developed countries that migration has relatively benign effects on the employment outcomes of non-migrants in the destination regions. Card (1990) was the first to show, based on the study of the Mariel Boatlift, that even a sudden large inflow of migrants virtually has no effect on native wages and employment probability. This research was followed by a plethora of studies which all drew similar conclusions.[1] The adjustment process that is typically advocated to explain why researchers find no labor market effects of immigration is capital mobility: capital inflows are expected to mitigate the negative impact of immigration on native employment outcomes. For instance, Angrist and Kugler (2003) show that the negative effect of immigration is much stronger in countries with high business entry costs than in countries with more flexible markets.[2]

Surprisingly, little attention has been paid to the impact of migration on labor market outcomes in developing countries.[3] Yet, it is particularly developing countries that are subjected to large migration flows, although these concern mostly internal rather than international migration. For instance, Barrios, Bertinelli and Strobl (2006) note that rural-urban migration has accounted for roughly half of Africa's spectacular urban growth between the 1960s and 1990s.[4] Moreover, one expects the negative effect of migration on labor market outcomes to be much more pronounced in developing than in developed countries. In developing countries (especially those located in Africa), road infrastructure indeed tends to be poor (see Yepes, Pierce and Foster (2009)), and therefore capital mobility low, thus undermining the potential for wages and hence job opportunities to return to their pre-migration levels.

The objective of this paper is therefore to investigate the impact of internal migration on local labor markets in a developing country. More precisely, we estimate the impact of the internal net in-migration rate on the employment probability of non-migrants within regions

[1] See Hunt (1992), Friedberg and Hunt (1995), Carrington and de Lima (1996), Friedberg (2001), Suen (2000), Card (2005), Longhi, Nijkamp and Poot (2005), McIntosh (2008), Hanson (2009), Boustan, Fishback and Kantor (2010) and Glitz (2012).

[2] See Aghion, Algan, Cahuc and Shleifer (2010) and Aghion, Algan and Cahuc (2011) for a discussion of the differences of market regulation in OECD countries.

[3] Berker (2011) is an exception.

[4] The authors report that Africa's growth rate of urbanization, defined as the share of urban to total population, has been extraordinary by international standards, averaging 140 percent between the 1960s and the 1990s – which is a rate of ten times that of OECD countries.

in Uganda.[5]

Of course, a simple regression analysis of the correlation between these two variables will provide only a biased estimate of the impact due to unobserved factors (e.g.: work opportunities at the regional level) likely to influence both the net in-migration rate and the employment probability of the non-migrants. To solve this endogeneity problem we rely on an instrumental variable approach. More specifically, following the empirical strategy developed by Boustan, Fishback and Kantor (2010) in their study of the impact of internal migration on local labor markets during the Great Depression in the US, we use the weather-predicted value of the net in-migration rate as an instrumental variable. To do so we construct this variable for each region such that it depends on: (i) the weather shocks affecting the other regions; (ii) the geographic distance between these other regions and the region under scrutiny.

The first advantage of relying on the weather predicted value of the net in-migration rate as an instrumental variable is that this variable is expected to be highly correlated with the actual net in-migration rate in a country dependent on rain-fed agriculture like Uganda. As a matter of fact, extreme weather conditions have been shown to impose considerable strains on populations living in such countries (see Miguel, Satyanath and Sergenti (2004), Barrios, Bertinelli and Strobl (2006), Gray and Mueller (2012), Beegle, De Weerdt and Dercon (2011), Miguel and Satyanath (2011), Marchiori, Maystadt and Schumacher (2012)).[6] Uganda is no exception and arguably constitutes a particularly good case study. According to FAOSTAT (2007), Uganda is among the countries in Sub-Saharan Africa showing the lowest share of irrigated cropland (less than 1%). Moreover, according to UN data (2005), a large majority of the Ugandan people (68.7%) make their living on rain-fed agriculture. Due to its heavy dependence on this sector, Uganda is widely considered as one of the most vulnerable countries in Sub-Saharan Africa to climate shocks.

The second advantage of relying on the weather predicted value of the net in-migration rate in a given region as an instrumental variable is that this variable is expected, by con-

[5]According to the World Bank, Uganda was ranked 190th of 215 countries in 2010 in terms of GNI/capita (PPP).

[6]Note that extreme weather conditions can affect individuals' conditions of living and therefore their decision to migrate in developed countries as well, as shown by Deschênes and Moretti (2009) and by Boustan, Fishback and Kantor (2010).

struction, to be orthogonal to the unobserved correlates of the employment probability of the non-migrants in that region. Put differently, the weather-predicted value of the net in-migration rate can be deemed as a good instrument to the extent that it is not only correlated with the endogenous explanatory variable, but also satisfies the exclusion restriction.

Our results confirm a much larger negative impact of migration on local labor outcomes than the one documented for developed countries: we find that a 10 percentage points increase in the net in-migration rate decreases the employment probability of non-migrants in the destination region by 7.5 percentage points. Consistent with the prediction of a simple theoretical model, our results further reveal that this negative impact is significantly stronger in regions less conducive to capital mobility (i.e., showing below-median road density): a 10 percentage points increase in the net in-migration rate in these areas decreases the probability of being employed of non-migrants by more than 20 percentage points.

The paper proceeds as follows. In Section 2, we develop a simple theoretical model that shows that the impact of an influx of migrants on the employment probability of the non-migrants is more negative in regions less conducive to capital mobility. In Section 3, we present our data. Section 4 describes our empirical strategy. Section 5 presents our results. Section 6 provides robustness checks. In Section 7, we summarize our conclusions and their policy implications.

2 Theoretical model

The purpose of this simple theoretical model is to show that the impact of an influx of migrants on the employment probability of the non-migrants is more negative in regions with lower road density.

We consider a regional economy with two goods: a good produced, consumed and used as capital, and labor. A representative competitive firm produces the good in quantity Y with capital K and labor L thanks to the following production function:

$$Y = F(K, L),$$

that is increasing with respect to its arguments, concave, and homogeneous of degree 1.

The good produced by the firm is the numeraire and the real wage is denoted by w. All markets are perfectly competitive. Labor supply or, equivalently, the employment probability increases with w. We suppose that firms have to incur a borrowing/opportunity cost to finance capital. This cost is decreasing with road density in the region. This assumption is in congruence with the well-known stylized fact according to which rural firms have lower access to credit than their urban counterparts in developing countries (see Rijkers, Söderbom and Loening (2010)). In the same vein, we assume that low road density can induce congestion effects: it becomes more costly to borrow capital when aggregate capital increases, such congestion effects becoming more pervasive with road scarcity. More precisely, we denote the cost of capital by $c(K^a, Q) = 1 + r(K^a, Q)$ where r stands for the interest rate, K^a for the aggregate capital in the region and $Q \in [0, \overline{Q}]$ captures road density in the region. We assume that $\frac{\partial r}{\partial Q} < 0$, $\frac{\partial r}{\partial K^a} \geq 0$ and that $\frac{\partial^2 r}{\partial K^a \partial Q} < 0$. More precisely, we suppose that $\frac{\partial r}{\partial K^a} = 0$ when $Q = \overline{Q}$: road density allows perfect mobility of capital between the region and the rest of the world. As soon as $Q < \overline{Q}$, however, firms have lower access to credit markets, thereby leading to congestion effects that increase with the isolation of the region.

The maximization program of the representative firm is defined by:

$$\max_{K,L} F(K, L) - wL - cK.$$

Since c is considered as given by the firm, the solutions K and L of the maximization program are determined by the following two first order conditions:

$$\begin{aligned} c &= F_K(K, L); \\ w &= F_L(K, L). \end{aligned}$$

Moreover, we have $K^a = K$ at equilibrium.

Let us denote by $\mathrm{d}L > 0$ an increase in labor supply subsequent to an influx of migrants in the economy. How is w impacted by $\mathrm{d}L$? To address this question, we compute the elasticity of w with respect to L that we denote by:

$$\epsilon_{w,L} \equiv \frac{\mathrm{d}w}{\mathrm{d}L}\frac{L}{w}.$$

Following standard calculus detailed in the Appendix, we obtain:

$$\epsilon_{w,L} \equiv (1 - \frac{1}{1 - \epsilon_{c,K}\epsilon_{k,c}})\epsilon_{w,L}|_{K=cst},$$

where $\epsilon_{c,K}$ expresses the elasticity of c with respect to K, $\epsilon_{k,c}$ stands for the elasticity of the ratio of capital per worker ($k = \frac{K}{L}$) with respect to c and $\epsilon_{w,L}|_{K=cst}$ denotes the elasticity of w with respect to L when K is fixed.

This expression of $\epsilon_{w,L}$ clearly shows that the elasticity of w with respect to L is a decreasing function of the elasticity of c with respect to K (since $\epsilon_{k,c} < 0$ and $\epsilon_{w,L}|_{K=cst} < 0$). More precisely, in the limit case of perfect mobility of capital ($Q = \overline{Q}$), $\epsilon_{c,K} = 0$ and therefore $\epsilon_{w,L} = 0$. In this setting, as shown in Figure 1, the line $w(L)$ that describes the variation of w with respect to L is flat: an influx of migrants has no impact on w. Consequently, the labor supply and hence employment probability of the non-migrants is itself unaffected. We indeed observe that, in equilibrium, the increase in employment is equal to the number of arrivals of migrants.

When capital is not perfectly mobile ($Q < \overline{Q}$), however, the expression of $\epsilon_{w,L}$ shows that the arrival of migrants decreases the equilibrium wage. In this case, $\epsilon_{c,K}$ is strictly positive and $\epsilon_{w,L}$ strictly negative, the absolute values of these two elasticities increasing at a rate that itself increases with road scarcity. Put differently, the lower Q, the more negative the impact of an influx of migrants on w will be. In the limit case where $Q = 0$ (the road density allows no mobility of capital between the region and the rest of the world), then $\epsilon_{c,K} \to +\infty$ and therefore $\epsilon_{w,L} = \epsilon_{w,L}|_{K=cst}$: the decrease in w is maximal. Figure 2 illustrates such mechanisms. The slope of the line $w(L)$ is negative (assuming that its absolute value increases with road scarcity). We observe that an influx of migrants decreases w and hence the labor supply and employment probability of the non-migrants: the increase in employment is indeed smaller than the number of arrivals of migrants in equilibrium.

This simple theoretical model allows us to derive the following proposition:

The impact of an influx of migrants on the employment probability of the non-migrants is more negative in regions with lower road density.

Our objective in the following is to empirically test this proposition. It is important to note that road density may not only influence capital mobility, but also the easiness for residents in a given region to respond to the wage impact of immigration by moving to other regions. This phenomenon is another potential adjustment typically advocated to explain why researchers find no local wage effects of immigration (Borjas, Freeman and Katz (1997), Card (2001), Borjas (2003), Borjas (2006), Federman, Harrington and Krynski (2006), Boustan, Fishback and Kantor (2010)). In other words, if we find that the negative impact of an influx of migrants is lower in regions with higher road density, we will have to ensure that this result is driven by the mechanism uncovered by the previous theoretical model, and not by an outflow of specific residents to other regions leading to a selection bias (i.e., in regions with higher road density residents who are more likely to be negatively affected by population inflows also have a greater chance to move elsewhere than in regions with lower road density).

3 Data

In this section, we first present the census data collected in Uganda in 2002. Besides the socioeconomic characteristics of the respondent, these data allow us to exploit two critical pieces of information: the employment probability of non-migrants as well as the net in-migration rate in each region. We then describe the weather data that help us construct the instrument for this net in-migration rate and present additional region-specific controls. Finally, we comment on the descriptive statistics related to each of these variables.

3.1 Census data

The population universe of the 2002 Uganda Census is composed of all persons living in the national territory. Respondents are the head of household or compound, or the person who has authority on the compound or the household. The dataset comprises 4,045,909 households (compared to a total of 24,442,084 inhabitants).

3.1.1 The employment probability

The employment probability derives from a question that asks respondents to indicate their employment status during the week preceding the census. Respondents can report to be employed, unemployed, or inactive. The employed population consists of persons working for pay for an employer, self-employed persons, unpaid (usually family) workers engaged in the production of economic goods, and persons who have a job but were temporarily absent for some reason. Unemployed persons are those who report to actively seek work. The inactive population encompasses persons not actively seeking work, persons unable to work (disabled), houseworkers, students, and retired people.

There are 647,983 individuals whose employment and migration status is known. 92% (532,454 individuals) are non-migrants. Among these non-migrants, 58% report to be employed, 2% to be unemployed, and 41% to be inactive.

We create an "employment" variable that stands for the employment status of the respondent. This "employment" variable is binary and takes the value 1 if the respondent reports to be employed and 0 if she reports to be unemployed or inactive.

3.1.2 The net in-migration rate

The 2002 census was the first census in Uganda to gather systematic information on internal mobility. Each interviewee was asked to report: (i) the number of years she had been living in the region where the census was being conducted (the answer ranges from "less than 1 year" to "more than 95 years") and (ii) the region in which she was living before. The regional breakdown in the data set is the district, where districts are the major administrative division in the country. In 2002 Uganda was composed of 56 districts, with an average area of 4,215 km^2 (approximately 65 km*65 km) each. We depict these in Figure 3. In the following we refer to districts as "regions".

We calculate a one-year-net-in-migration rate at the regional level. More precisely, for each region j, we first calculate the number of migrants arriving in and the number of migrants leaving region j between 2001 and 2002 as a share of the population of region j in 2001. We then compute the difference between these two ratios in order to obtain the net in-migration rate in region j. One should note that we focus on one-year net in-migration rates in order to minimize missing the number of migrant flows that occur within our time periods: with one-year migration flows we are simply missing intra-annual population movements.

3.1.3 Socioeconomic characteristics

The census data also inform us on a set of socioeconomic characteristics of the respondent, notably her gender, age, education, and whether she lives in a rural or urban area. We define the variable "male" as a dummy that takes the value 1 if the respondent is a male and 0 otherwise. The variable "age" is constructed as an ordinal variable that captures the four-year age interval to which the respondent belongs. It ranges from 1 to 19, where 1 stands for the interval "5 to 9 year old" and 19 stands for "more than 80 year old". We create the variable "education" as an ordinal variable that ranges from 1 to 4 where 1 stands for "less than primary completed", 2 for "primary completed", 3 for "secondary completed" and 4 for "university completed". Finally, we define the variable "urban" as a dummy that takes the value 1 if the respondent lives in a urban area and 0 if she lives in a rural area.

3.2 Weather data

We obtain our weather data by computing a Standardized Precipitation Index (SPI) for each region. To do so, we first rely on the Inter-Governmental Panel on Climate Change (IPCC) dataset[7] that provides measures of monthly precipitations at the 0.5 degree level over the entire 20th century. We calculate monthly regional precipitation by placing the grids within regions. We then fit these rainfall data to estimate a gamma distribution. For each year in each region, the SPI is subsequently computed as the standard deviation of rainfall, i.e., as the variation of rainfall around its regional historical mean, as predicted by the gamma distribution. As such a SPI greater than 2 (1) indicates an extremely (moderately) wet event. Conversely, a SPI lower than -2 (-1) indicates an extremely (moderately) dry event (Hayes, Svoboda, Wilhite and Vanyarkho (1999)).

In terms of using SPI to capture the appropriate weather shocks that may affect migration, the choice of time frame is important. To the best of our knowledge, Dercon (2004) is the first to examine the long-term economic impact of extreme weather conditions in developing

[7] This dataset is available at http://www.cru.uea.ac.uk/cru/data/hrg/.

countries. He finds in the context of Ethiopia that the loss in food consumption persists five years after a drought has occurred. Relying on Brazilian data, Mueller and Osgood (2009) also point to a five year persistence effect: they show that droughts can cause wages in rural municipalities to be lower than their peers for five years after the event. Following this evidence, we therefore assume that the impact of extreme weather conditions can affect individuals' decision to migrate up to five years after their surge. This means that we compute the mean of the yearly average values of the SPI in each region during the five years preceding the year of the census, i.e., between 1997 and 2002. One may want to note that in the absence of panel data (and therefore of controls for regional fixed effects), the SPI is valuable. Given that SPI is defined relative to each region's own rainfall distribution, any cross-regional variations in it are indeed truly capturing regional differences in shocks rather than regional differences in mean historical rainfall.

3.3 Region-specific variables

We create two region-specific variables: one which allows us to distinguish between regions with higher and lower road density, the other which controls for the level of regional economic development at the beginning of the migration period (i.e., in 2001).

3.3.1 Road density

Road density at the regional level is captured by the number of kilometers of road per square kilometer in each region. This information stems from the USGS (US Geological Survey) Global GIS (Geographic Information System) database that was released in 2002.[8] We categorize Uganda regions into two groups: regions with below-median road density and regions with above-median road density. The road network in Uganda is shown in Figure 4. As can be seen, while the road network covers most of Uganda, there does appear to be a higher concentration of roads in the southeast of Uganda, near the capital city of Kampala and in the northwestern part of the Lake Victoria.

[8]This database is available at http://www.agiweb.org/pubs/globalgis/description.html.

3.3.2 Initial regional economic development

Regional economic development at the beginning of the migration period is proxied by the average intensity of nightlights in each region in 2001. Satellite imagery of nightlights are provided by the United States Airforce Defense Meterological Satellite Program (DMSP) since 1992 and measure the intensity of lights at night around the globe at the approximately 1 squared kilometer grid cell level. More specifically, each satellite of the DMSP observes every location on the globe at some point in time at night, between 8:30 and 10:30 pm. These images are then processed to remove intensity due to moonlight, late lighting during summer months, auroral activity and forest fires. The remaining light intensity, arguably due to human activity, is then averaged on an annual basis and normalized to scale of integers ranging from 0 (no light) and 63. One should note that these nightlights data have been shown to constitute good proxies for GDP and its growth (Henderson, Storeyguard and Weil (2012)), especially in African countries where national income figures are widely thought to be unreliable (Behrman and Rosenzweig (1994), Heston (1994)). Moreover, they have been argued to serve as an alternative measure of local income where disaggregated figures are not available on a comprehensive basis, as is the case for Uganda. We depict the nightlights distribution in 2001 at the grid cell level for the Ugandan regions in Figure 5. In contrast to the road network, nightlights in 2001 display a much higher concentration. In particular, large parts of Uganda, as is the case for most of the African continent, are completely dark (i.e., with a normalized nightlights value of zero). Only a few pockets of agglomerated brightness can be observed around the larger cities. For instance, the largest area of light is centered around the capital city of Kampala. We use the average per square kilometer intensity per region in 2001 as our measure of initial regional economic development.

3.4 Summary statistics

Table 1 presents summary statistics for each of the variables used in our analysis. Individual level variables (employment probability and socioeconomic characteristics) are presented as regional averages over the population of non-migrants whose employment and migration status is known ($N = 532,454$). The employment probability among non-migrants amounts to 57.7%. The sample of non-migrants is well balanced across gender. On average non-migrants tend to be in their forties, have not completed primary school, and live in rural areas (only 25.4% live in urban areas).

Table 2 reports a difference of means analysis that compares individual level variables of migrants and non-migrants. It reveals that migrants are significantly less likely to be employed than non-migrants. This result suggests that it takes some time for migrants to find a new job upon arrival to their destination region. Moreover, migrants are more likely to be male, which is consistent with preliminary findings on the characteristics of internal migrants (see Lucas (1997)). This result suggests that the bulk of internal migration cannot be accounted for by the prevalence of patrilocality in Uganda, whereby females move out of the paternal location at the time of marriage. Migrants are also younger than non-migrants, more educated, and more likely to choose an urban area as their new place of residence, a set of results also consistent with those reported by Lucas (1997).

With regard to regional level variables, we observe in Table 1 that the average net in-migration rate is close to 0%. This should come as no surprise since regional in- and out-flows tend to compensate each other. The mean of the variable "SPI" is equal to 0.326. According to Hayes, Svoboda, Wilhite and Vanyarkho (1999), this stands for a "near normal" level,[9] meaning that rainfall is, on average, close to its historical values. The analysis of the minimum and maximum values show that variations outside of this range of "near normal" levels are driven by "moderately wet" events. Put differently, rainfall departures from their historical mean in Uganda are due to unusually wet, not dry events between 1997 and 2002. Finally, region-specific variables confirm the low-income country status of Uganda: each square kilometer is endowed with an average of 58 meters of roads only, while the mean of the nightlights intensity is low (equal to 3.616) as compared to the range of values (from 0 to 63) it could theoretically take. By contrast, in the UK, which has roughly the same geographical area as Uganda, the average road density amounts to 105 meters per square kilometer while the average nightlights intensity is equal to 15.208.

[9]The "near normal" category concerns SPI values ranging from -0.99 to 0.99, i.e., rainfall shocks classified below "moderate".

4 Empirical strategy: constructing the weather-predicted value of the net in-migration rate

The objective of this paper is to estimate, in the context of a developing country, the impact of the net in-migration rate in region j on the employment probability of the non-migrants in that region. To do so, following the empirical strategy developed by Boustan, Fishback and Kantor (2010), we instrument the net in-migration rate in region j over the 2001-2002 period by its weather-predicted value denoted $n^{WP}_{j,01-02}$. This variable stands for the difference between the weather-predicted value of the in-migration rate in region j ($i^{WP}_{j,01-02}$ hereafter) and the weather-predicted value of the out-migration rate from region j ($o^{WP}_{j,01-02}$ hereafter) over this period. In the following, we describe how we compute $i^{WP}_{j,01-02}$ and $o^{WP}_{j,01-02}$, respectively.

4.1 Computing the weather-predicted value of the in-migration rate

We proceed in three steps. We first regress the out-migration flow from source region k between 2001 and 2002, denoted $O_{k,01-02}$, on the mean of the yearly values of the SPI in region k between 1997 and 2002:

$$O_{k,01-02} = \alpha + \beta.\text{SPI}_{k,97-02} + u_k. \tag{1}$$

For each source region k, we then regress the share of people leaving source region k who settle in destination region j on a function that is quadratic in the distance between regions k and j:

$$P_{kj,01-02} = \delta_k + \theta_k dist_{kj} + \eta_k (dist_{kj})^2 + \mu_k. \tag{2}$$

The weather-predicted in-migration flow to destination region j is then the sum over all source regions k ($k \neq j$) of the predicted number of migrants leaving source region k who are expected to settle in destination region j:

$$I^{WP}_{j,01-02} = \sum_{k=1,\dots,n(k\neq j)} \widehat{O_{k,01-02}} * \widehat{P_{kj,01-02}}.$$

We finally obtain the weather predicted in-migration rate to destination region j by dividing $I^{WP}_{j,01-02}$ by the population of destination region j in 2001:

$$i^{WP}_{j,01-02} = \frac{I^{WP}_{j,01-02}}{pop_{j,01}}.$$

4.2 Computing the weather-predicted value of the out-migration rate

We proceed in an analogous fashion to calculate the out-migration rate, i.e., in three steps. We first regress the in-migration flow in destination region k between 2001 and 2002 denoted $I_{k,01-02}$ on the mean of the yearly values of the SPI in region k between 1997 and 2002:

$$I_{k,01-02} = \alpha + \beta.\text{SPI}_{k,97-02} + u_k. \tag{3}$$

For each destination region k, we then regress the share of people leaving source region j who settle in destination region k on a function that is quadratic in the distance between regions j and k:

$$P_{jk,01-02} = \delta_k + \theta_k dist_{jk} + \eta_k (dist_{jk})^2 + \mu_k. \tag{4}$$

The weather-predicted out-migration flow from source region j is then the sum over all destination regions k ($k \neq j$) of the predicted number of migrants settling in destination region k who are expected to come from source region j:

$$O^{PW}_{j,01-02} = \sum_{k=1,\dots,n(k\neq j)} \widehat{I_{k,01-02}} * \widehat{P_{jk,01-02}}.$$

We finally obtain the weather predicted out-migration rate from source region j by dividing $O^{PW}_{j,01-02}$ by the population of source region j in 2001:

$$o^{WP}_{j,01-02} = \frac{O^{WP}_{j,01-02}}{pop_{j,01}}.$$

Two sets of regressions allow us to construct $n^{WP}_{j,01-02}$, i.e., the weather-predicted value of the net in-migration rate in region j over the 2001-2002 period. The first set concerns Equation (1) and Equation (3) which regress the out-migration flow (resp. in-migration flow) from source region k (resp. in destination region k) between 2001 and 2002 on the mean of the yearly values of the SPI in region k between 1997 and 2002. The second set concerns Equation (2) and Equation (4) which regress, for each source region k (resp. destination region k) the share of people leaving source region k (resp. source region j) who settle in destination region j (resp. destination region k) on a function that is quadratic in the distance between those regions. In the following we show that the results from these ancillary regressions are intuitive and show statistical significance.

Table 3 reports OLS estimates for the first set of regressions. The relationship between the SPI and out-migration flows (Equation (1)), as well as in-migration flows (Equation (3)) is estimated in column (1) and column (2), respectively. These results confirm that weather shocks have a statistically significant effect on migration flows in a country dependent on rain-fed agriculture like Uganda. Yet, it is *a priori* difficult to predict the sign of the impact of the SPI on these flows. Uganda indeed encompasses both traditionally wet and dry districts. In other words, a high SPI can be a blessing (in traditionally dry districts) or a curse (in already traditionally wet districts). Estimates in column (1) and in column (2) confirm this ambiguity: a high SPI generates both higher out- and in-migration flows (on net, we observe that a higher SPI increases in-migration flows).

As for the second set of regressions (Equation (2) and Equation (4)), due to their large numbers (55*2=110) not reported here, we find that for all but 9 of these either the negative coefficient on the linear distance term is significantly different from zero at the 10% confidence level or the Fisher test rejects the null hypothesis that both the coefficient on the linear distance and quadratic distance terms are jointly equal to zero (again at the 10% confidence level). Put differently, we find confirmation that the share of people leaving source region k (resp. source region j) who settle in destination region j (resp. destination region k) is negatively and significantly impacted by the distance between these regions.

5 Results

In the following, we first follow a naïve probit approach, likely to underestimate the negative impact of the net in-migration rate on the employment probability of the non-migrants. We then turn to an IV probit approach that allows us to address this potential bias. Finally, we test the prediction of our theoretical model according to which non-migrants living in regions less conducive to capital mobility (i.e., showing lower road density) should be more negatively impacted by an influx of migrants.

5.1 A naïve probit approach

A naïve probit approach consists of computing the probit estimates of Equation (5):

$$P(E_{ij,02} = 1) = G(a + b.n_{j,01-02} + c.D_{j,01} + \mathbf{X_{ij,02}}'.\mathbf{e} + \epsilon_{ij,02}), \tag{5}$$

where G is the cumulative distribution function for a standard normal density. The dummy $E_{ij,02}$ stands for the employment status of the non-migrant i who lives in region j in 2002. The variable $n_{j,01-02}$ is the net in-migration rate in region j between 2001 and 2002. Ideally, we would have liked to use, as the dependent variable, the change in the employment probability of the non-migrant i between 2001 and 2002. However, we do not have information on the employment status of the non-migrant i in 2001. Instead, we include a proxy for the level of economic development of region j in 2001. This proxy is the variable $D_{j,01}$ which represents the average nightlights intensity of region j in 2001. Finally, $\mathbf{X_{ij,02}}$ is a vector of socio-economic characteristics of the non-migrant i who lives in region j in 2002. This vector contains information on the gender, age, and education of the non-migrant i, as well as on whether she lives in a urban area.

Table 4 reports the marginal effects of the probit estimation of Equation (5), where robust standard errors are bootstrapped at the regional level.[10] We observe that an increase

[10] An alternative would be to simply cluster standard errors at the regional level. However, as shown by Cameron, Gelbach and Miller (2008), bootstrapping is preferable to clustering when the number of clusters is relatively small since it limits the tendency of clustering to over-reject the null hypothesis. Another reason for relying on bootstrapping is unbalanced cluster size. As we observe both of these characteristics in our clusters – regions that are both small in number (56) and unbalanced in size (the least populated encompasses 320 individuals for whom the employment and migration status is known while the most populated hosts

in the net in-migration rate by 10 percentage points is associated with a decrease in the employment probability of the non-migrants by roughly 3 percentage points (significant at the 1% confidence level). This correlation is of low magnitude. This is possibly due to an unobserved factor (e.g.: work opportunities at the regional level) that influences both the net in-migration rate and the employment probability of the non migrants, thereby leading to underestimate the negative impact of the net in-migration rate on the employment probability of the non migrants. To solve this potential endogeneity problem, an instrumental variable approach is needed.

5.2 The IV probit approach

In this section, we estimate the impact of the net in-migration rate in region j on the employment probability of the non-migrants in that region after having instrumented this net in-migration rate by its weather-predicted value.

5.2.1 First- and second- stage results

The first stage of the IV probit approach consists of computing the OLS estimates of Equation (6):

$$n_{j,01-02} = a + b.n^{WP}_{j,01-02} + c.D_{j,01} + \mathbf{X_{ij,02}}'.\mathbf{e} + \epsilon_{ij,02}. \tag{6}$$

These estimates are reported in Table 5. Since $n^{WP}_{j,01-02}$ is generated through statistical estimation, standardly derived standard errors are no longer correct (see Wooldridge (2002): 139-141). This constitutes a third justification (in addition to the small number of clusters and unbalanced cluster size) for generating robust standard errors bootstrapped at the regional level. As expected, our results show a strongly significant and positive correlation between the net in-migration rate and its weather-predicted value (the z-statistic related to the coefficient of $n^{WP}_{j,01-02}$ is equal to 1106.42).

The second stage of the IV probit approach entails computing the probit estimates of

99,025) – we consider bootstrapping at the regional level to be a more conservative approach than clustering.

Equation (7):

$$P(E_{ij,02}=1)=G(a+b.\widehat{n_{j,01-02}}+c.D_{j,01}+\mathbf{X_{ij,02}}'.\mathbf{e}+\epsilon_{ij,02}), \tag{7}$$

where $\widehat{n_{j,01-02}}$ is the predicted value of $n_{j,01-02}$, as derived from the first stage (Equation (6)).

Table 6 reports the marginal effects of the probit estimation of Equation (7), where robust standard errors are bootstrapped at the regional level. We observe that an increase in the net in-migration rate by 10 percentage points leads to a decrease in the employment probability of non-migrants by roughly 7.5 percentage points. This impact is both substantial in magnitude and strongly significant. It confirms that the measure of this negative impact stemming from the naïve probit approach was underestimated.

5.2.2 Testing the theoretical prediction: does internal migration impact regions with low and high road density differently?

Our results so far show a much larger negative impact of the net in-migration rate on the employment probability of the non-migrants in Uganda than the one documented for developed countries. This finding is consistent with the low income status of Uganda and more precisely its limited access to capital. To strengthen this interpretation, we test the prediction of our theoretical model by examining whether this negative impact of the net in-migration rate on the employment probability of the non-migrants is significantly stronger in regions characterized by lower road density (and therefore lower prospects of capital mobility).

To this end, we estimate Equation (7) on two sub-samples: the sub-sample of regions characterized by a below-median road density, and the sub-sample of regions characterized by an above-median road density. Results are reported in Table 7. They confirm that the negative impact of the net in-migration rate in a specific region on the employment probability of the non-migrants in that region is significantly stronger in regions showing lower road density.[11] While an increase in the net in-migration rate by 10 percentage points leads to a decrease in the employment probability of the non-migrants by only 4 percentage points in

[11] A test of the equality of the coefficients related to $\widehat{n_{j,01-02}}$ in both equations shows that these coefficients significantly differ at the 1% confidence level.

regions showing above-median road density, the employment probability decreases by more than 20 percentage points in regions showing below-median road density, i.e., arguably a very large impact.

To be sure, road density is not only a proxy for capital mobility but also for the easiness for residents in a given region to respond to the wage impact of immigration on a local labor market by moving to other regions. Put differently, the results reported in Table 7 might be merely driven by a selection bias (i.e., in regions with higher road density residents who are more likely to be negatively affected by population inflows also have a greater chance to move elsewhere than in regions with lower road density). To rule out this possibility, we compute the probit estimates of Equation (8):

$$P(M_{ij,01-02} = 1) = G(a + b.R_j + c.D_{j,01} + \mathbf{X_{i,02}}'.\mathbf{d} + \epsilon_{ij,02}), \tag{8}$$

where G is the cumulative distribution function for a standard normal density. The dummy $M_{ij,01-02}$ stands for the decision of individual i to leave region j between 2001 and 2002. The variable R_j is the road density in region j. The variable $D_{j,01}$ represents the average night-lights intensity of region j in 2001. Finally, $\mathbf{X_{i,02}}$ is a vector of socio-economic characteristics of individual i in 2002. This vector contains information on the gender, age, and education of individual i.

Table 8 reports the marginal effects of the probit estimation of Equation (8), where robust standard errors are bootstrapped at the regional level. We observe that road density in region j is *negatively*, not positively correlated with individuals' decision to leave region j. This finding is clearly consistent with the proposition derived from our simple theoretical model according to which higher road density allows to mitigate the negative impact of an influx of migrants on the probability of non-migrants to be employed. More precisely, this finding shows that this mitigating effect of higher road density dominates the fact that higher road density also eases the possibility for residents in a given region to move to other regions as a response to an influx of immigrants. Put differently, this finding makes us confident that the results reported in Table 7 are not driven by a selection of non-migrants.

6 Robustness checks

Two factors could jeopardize the validity of our instrument. First, the weather-predicted value of the net in-migration rate in region j depends not only on the SPI in the other regions but also negatively on the distance between these other regions and region j. In other words, our instrument is particularly correlated with the SPI in the regions located closer to region j. This would imply that our instrument might be correlated with the employment probability of the non-migrants in region j not only through its impact on the net in-migration rate in region j, but also through is correlation with the SPI in region j. If so, the exclusion restriction would be violated.

Second, it may be that weather shocks in regions other than region j have an impact on the economic conditions in region j (and therefore on the employment probability of the non-migrants in this region) that does not transmit through internal migration, but through other channels, as for instance, when region j and these other regions are trade partners (or trade competitors). If so, the exclusion restriction would again be violated.

To ensure that the exclusion restriction is satisfied, we therefore test whether the IV-probit results reported in Tables 6 and 7 are robust to controlling for the SPI in region j as well as for the economic growth rate in region j between 1997 and 2002. The latter is proxied by the growth rate in the nightlights intensity between these two dates. Our results are reported in Table 9. They show that the negative impact of the net in-migration rate on the employment probability of the non-migrants is, if anything, reinforced by these controls.

7 Conclusion

This paper investigates the impact of weather-induced internal net in-migration rates on the employment probability of non-migrants within regions in Uganda. Our results reveal a much larger negative impact of migration on local labor outcomes than the one documented for developed countries: we find that a 10 percentage points increase in the net in-migration rate decreases the employment probability of non-migrants in the destination region by 7.5 percentage points. Consistent with the prediction of a simple theoretical model, our results further show that this negative impact is significantly stronger in regions less conducive to capital mobility (i.e., showing below-median road density): a 10 percentage points increase in the net in-migration rate in these areas decreases the probability of being employed of non-migrants by more than 20 percentage points.

Our findings suggest that the development of road infrastructure which ranks high on the World Bank's agenda (see World Bank (2009)) could considerably mitigate the negative spill-over effects of weather shocks on local labor markets in countries dependent on rain-fed agriculture (the bulk of countries in Sub-Saharan Africa). Note that road infrastructure development may also help circumvent the expected negative impact of internal migration induced by other types of shocks in source regions, such as conflict outbreaks that have plagued Sub-Saharan African countries over the last decades. Estimating such impacts, as well as defining the conditions of an efficient road infrastructure policy (prioritizing construction or maintenance; focusing on international, national or rural road networks... etc) constitute important avenues for future research.

References

[1] Aghion, Philippe, Yann Algan, Pierre Cahuc and Andrei Shleifer. 2010. "Regulation and distrust." *Quarterly Journal of Economics* 125(3): 1015-1049.

[2] Aghion, Philippe, Yann Algan and Pierre Cahuc. 2011. "Civil society and the state: the interplay between cooperation and minimum wage regulation." *Journal of the European Economic Association* 9(1): 3-42.

[3] Angrist, Joshua D. and Adriana Kugler. 2003. "Protective or counter-productive? Labour market institutions and the effect of immigration on EU natives." *Economic Journal* 113(488): F302-F331.

[4] Barrios, Salvador, Luisito Bertinelli and Eric Strobl. 2006, "Climatic change and rural-urban migration: the case of Sub-Saharan Africa." *Journal of Urban Economics* 60(3): 357-371.

[5] Beegle, Kathleen, Joachim De Weerdt and Stefan Dercon. 2011. "Migration and economic mobility in Tanzania: evidence from a tracking survey." *Review of Economics and Statistics* 93(3): 1010-1033.

[6] Behrman, Jere R. and Mark R. Rosenzweig. 1994. "Caveat emptor: cross-country data on education and the labor force." *Journal of Development Economics* 44(1): 147-171.

[7] Berker, Ali. 2011. "Labor-Market Consequences of Internal Migration in Turkey." *Economic Development and Cultural Change* 60(1): 197-239.

[8] Boustan, Leah Platt, Price V. Fishback and Shawn Kantor. 2010. "The effect of internal migration on local labor markets: American cities during the great depression." *Journal of Labor Economics* 28(4): 719-746.

[9] Borjas, George J., Richard B. Freeman and Lawrence F. Katz. 1997. "How much do immigration and trade affect labor market outcomes?" *Brookings Papers on Economic Activity* 28(1): 1-67.

[10] Borjas, George J. 2003. "The labor demandcurve is downward sloping: reexamining the impact of immigration on the labor market." *Quarterly Journal of Economics* 118(4): 1335-1174.

[11] Borjas, George J. 2006. "Native internal migration and the labor market impact of immigration." *Journal of Human Resources* 56(2): 221-258.

[12] Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *Review of Economics and Statistics*, 90(3): 414-427.

[13] Card, David. 1990. "The impact of the Mariel Boatlift on the Miami labor market." *Industrial and Labor Relations Review* 43(2): 245-257.

[14] Card, David. 2001. "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration." *Journal of Labor Economics* 19(1): 22-64.

[15] Card, David. 2005. "Is the new immigration really so bad?" *Economic Journal* 115(507): F300-F323.

[16] Carrington, William and Pedrode de Lima. 1996. "The impact of 1970s repatriates from Africa on the Portugese labor market." *Industrial and Labor Relations Review* 49(1): 330-347.

[17] Dercon, Stefan. 2004. "Growth and shocks: evidence from rural Ethiopia." *Journal of Development Economics* 74(2): 304-329.

[18] Deschenes, Olivier and Enrico Moretti. 2009. "Extreme weather events, mortality and migration." *Review of Economics and Statistics* 91(4): 659-681.

[19] Federman, Maya N., David E. Harrington and Kathy J. Krynski. 2006. "Vietnamese manicurists: are immigrants displacing natives or finding new nails to polish?" *Industrial and Labor Relations Review* 59(2): 302-318.

[20] Friedberg, Rachel M. 2001. "The impact of mass migration on the Israeli labor market." *Quarterly Journal of Economics* 116(4): 1373-1408.

[21] Friedberg, Rachel M. and Jennifer Hunt. 1995. "The impact of immigrants on host country wages, employment and growth." *Journal of Economic Perspectives* 9(2): 23-44.

[22] Glitz, Albrecht. 2012. "The labor market impact of immigration: a quasi-experiment exploiting immigrant location uules in Germany." *Journal of Labor Economics* 30(1): 175-213.

[23] Gray, Clark and Valerie Mueller. 2012. "Drought and population mobility in rural Ethiopia." *World Development* 40(1): 13-145.

[24] Hayes, Michael J., Mark D. Svoboda, Donald A. Wilhite and Olga V. Vanyarkho. 1999. "Monitoring the 1996 drought using the standardized precipitation index." *Bulletin of the American Meteorological Society* 80(3): 429-438.

[25] Hanson, Gordon H. 2009. "The economic consequences of the international migration of labor." *Annual Review of Economics* 1(1): 179-208.

[26] Henderson, J. Vernon, Adam Storeygard and David N. Weil. 2012. "Measuring economic growth from outer space." *American Economic Review* 102(2): 994-1028.

[27] Heston, Alan. 1994. "A brief review of some problems in using national accounts data in level of output comparisons and growth studies." *Journal of Development Economics* 44(1): 29-52.

[28] Hunt, Jennifer. 1992. "The impact of the 1962 repatriates from Algeria on the French labor market." *Industrial and Labor Relations Review* 45(3): 556-572.

[29] Longhi, Simonetta, Peter Nijkamp and Jacques Poot. 2005. "A meta-analytic assessment of the effect of immigration on wages." *Journal of Economic Surveys* 19(3): 451-477.

[30] Lucas, Robert E.B. 1997. "Internal migration in developing countries." *Handbook of Population and Family Economics* 13: 721-798.

[31] Marchiori, Luca, Jean-François Maystadt and Ingmar Schumacher. 2012. "The impact of weather anomalies on migration in Sub-Saharan Africa." *Journal of Environmental Economics and Management* 63(3): 355-374.

[32] McIntosh, Molly Fifer. 2008. "Measuring the labor market impacts of Hurricane Katrina migration: evidence from Houston, Texas." *American Economic Review* 98(2): 54-57.

[33] Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004. "Economic shocks and civil conflict: an instrumental variables approach." *Journal of Political Economy* 112(4), 725-753.

[34] Miguel, Edward and Shanker Satyanath. 2011. "Re-examining economics shocks and civil conflict." *American Economic Journal: Applied Economics* 3(4), 228-232.

[35] Mueller, Valérie A. and Daniel Osgood. 2009. "Long-term impacts of droughts on labour markets in developing countries: evidence from Brazil." *Journal of Development Studies* 45(10): 1651-1662.

[36] Rijkers, Bob, Mans Söderbom and Josef L. Loening. 2010. "A rura-urban comparison of manufacturing enterprise performance in Ethiopia." *World Development* 38(9): 1278-1296.

[37] Suen, Wing. 2000. "Estimating the effects of immigration in one city." *Journal of Population Economics* 13(1): 99-112.

[38] Wooldridge, Jeffrey. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

[39] World Bank. 2009. *World Development Report 2009: Reshaping economic geography.* The World Bank: Washington D.C.

[40] Yepes, Tito, Justin Pierce and Vivien Foster. 2009. "Making sense of Africa's infrastructure endowment. A benchmark approach." World Bank Policy Research Working Paper 4912.

Tables and Figures

Figure 1: The impact of an influx of migrants on the employment probability of the non-migrants: the case of perfect capital mobility.

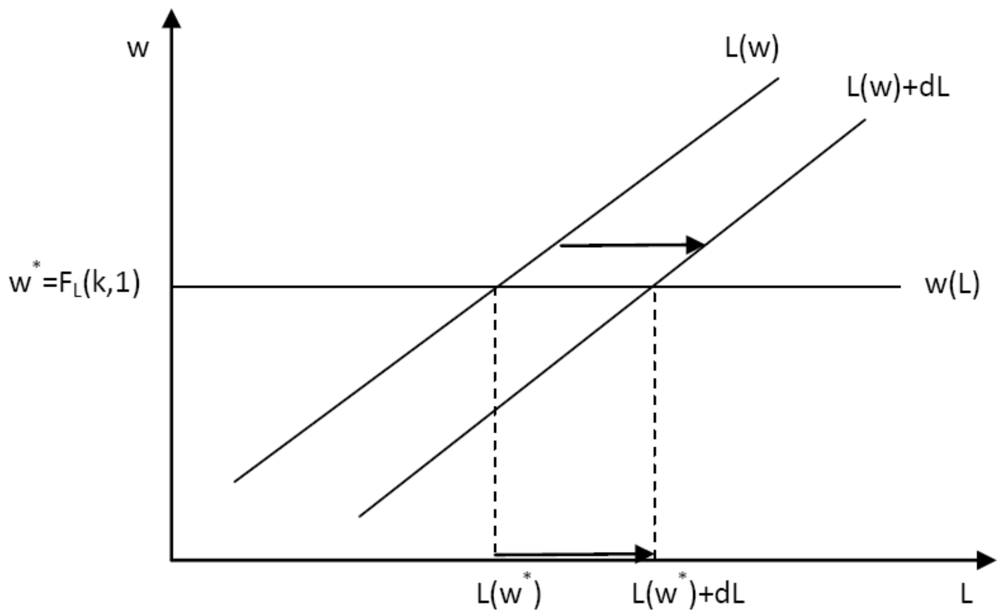

Figure 2: The impact of an influx of migrants on the employment probability of the non-migrants: the case of imperfect capital mobility.

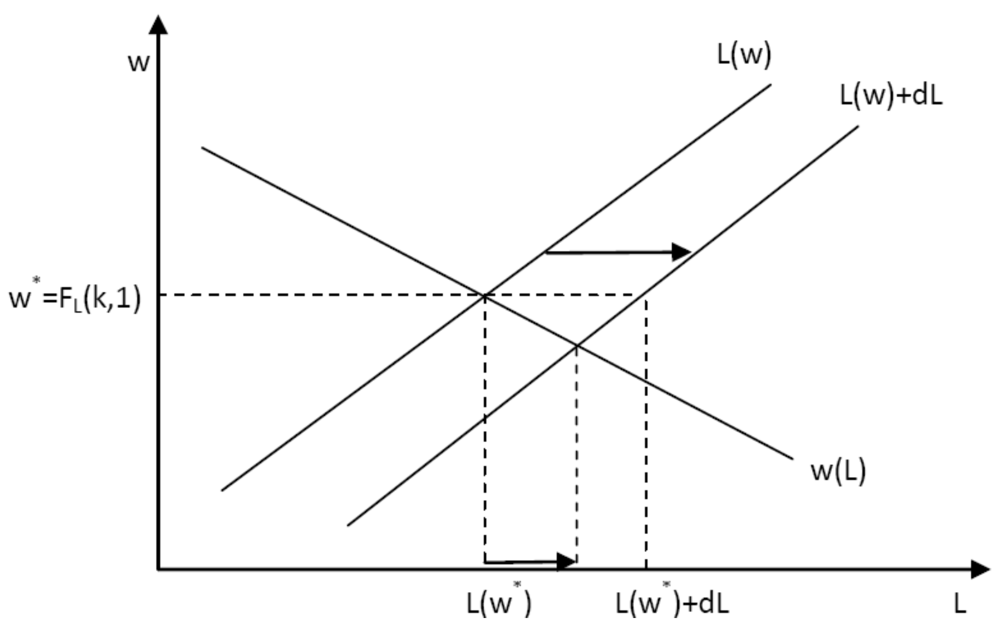

Figure 3: The 56 Ugandan districts (2002).

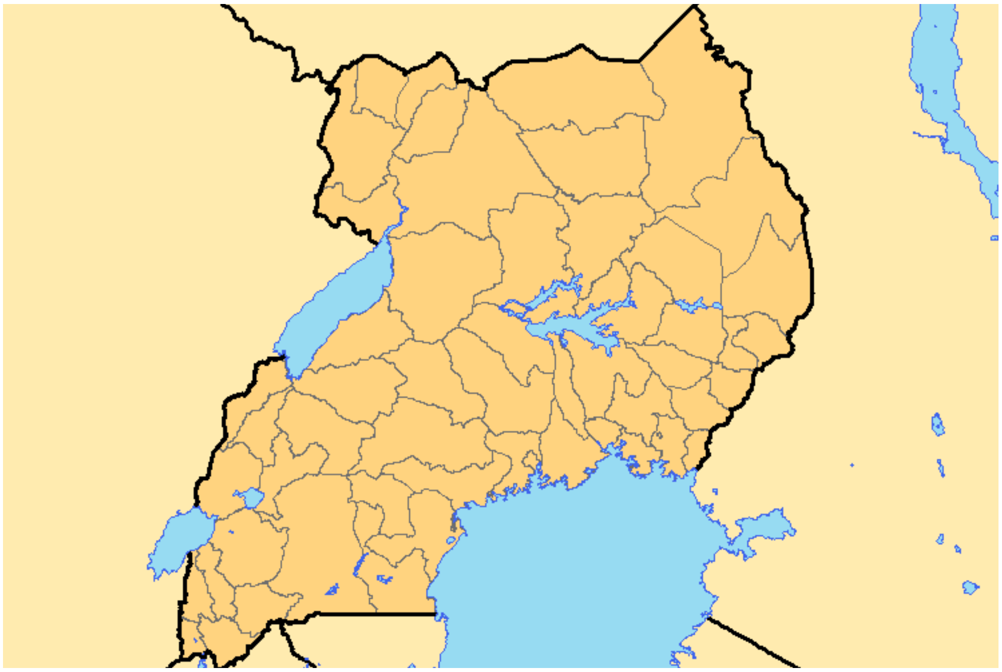

Figure 4: Road density in Uganda (2002).

Figure 5: Nightlights in Uganda (2001).

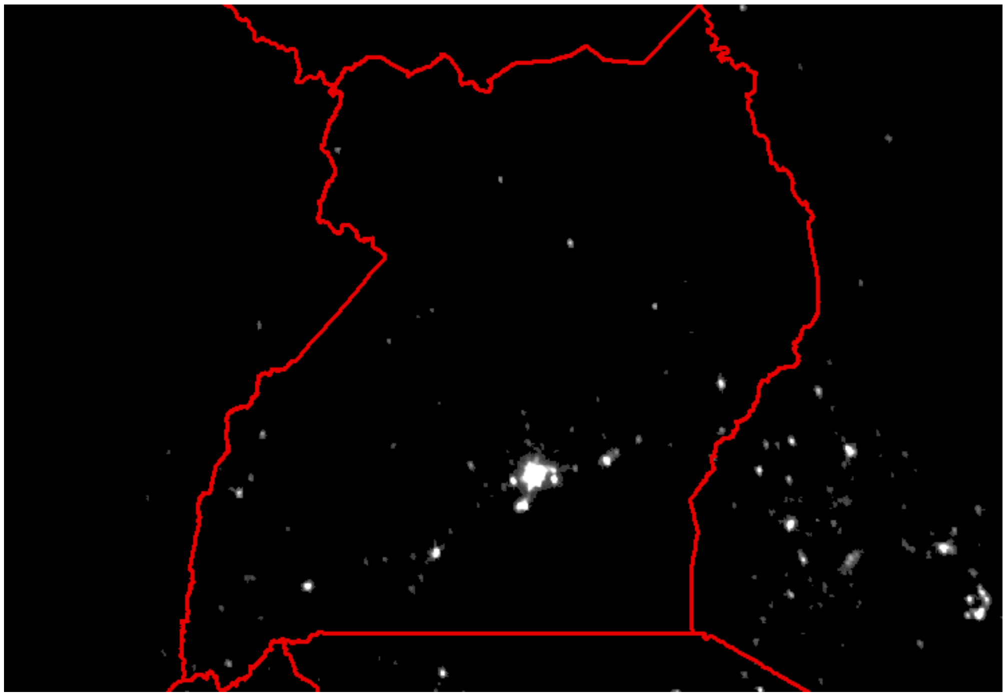

Table 1: Summary statistics.

Variables	Mean	Std. Dev.	Min	Max
A. Employment probability	0.577	0.075	0.145	0.743
B. Socioeconomic characteristics				
Male	0.480	0.035	0.340	0.667
Age	8.328	0.746	6.528	11.523
Education	1.454	0.253	1.179	1.961
Urban	0.254	0.328	0.003	1.000
C. Net in-migration rate (2001-2002)	0.003	0.054	-0.090	0.309
D. SPI (1997-2002)	0.326	0.412	-0.324	1.329
E. Region-specific variables				
Road density (2002)	0.058	0.035	0.007	0.175
Nightlights intensity (2001)	3.616	0.463	3.063	4.842

Notes: The table reports summary statistics at the regional level. Individual level variables (employment probability and socioeconomic characteristics) are presented as regional averages over the population of non-migrants whose employment status is known ($N = 532,454$).

Table 2: Socioeconomic characteristics of migrants and non-migrants. Difference of means analysis.

Variable	Non-migrants (a)	Migrants (b)	Difference (b-a)
Employed	0.58 (N=532,454)	0.48 (N=115,529)	-0.10 p=0.00
Male	0.48 (N=532,454)	0.49 (N=115,529)	-0.01 p=0.00
Age	8.33 (N=532,319)	5.87 (N=115,501)	-2.46 p=0.00
Education	1.45 (N=530,959)	1.53 (N=114,952)	+0.08 p=0.00
Urban	0.25 (N=532,454)	0.35 (N=115,529)	+0.10 p=0.00

Notes: The table reports arithmetic means for the sub-samples of migrants and non-migrants whose employment status is known, and two-tailed t-tests assuming unequal variances.

Table 3: The relationship between the out- and in-migration flows and the SPI. OLS analysis.

	Out-migration flow	In-migration flow
SPI (yearly average between 1997 and 2002)	2737.011**	4558.948**
	(1069.263)	(1719.766)
R^2	0.205	0.295
Observations	56	56

Notes: The table reports OLS estimates. The unit of observation is the region. Standard errors are robust. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Table 4: The relationship between the employment probability of the non-migrants and the net in-migration rate. Probit analysis.

	Employment probability
Net in-migration rate	-0.295***
	(0.017)
Economic development in 2001	-0.100***
	(0.002)
Male	0.231***
	(0.001)
Age	0.204***
	(0.001)
Age^2	-0.009***
	(0.000)
Education	-0.005***
	(0.001)
Urban area	-0.076***
	(0.002)
Pseudo-R^2	0.237
Observations	530,827

Notes: The table reports marginal effects, based on a probit estimation. The unit of observation is the non-migrant. Robust standard errors are bootstrapped at the regional level. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Table 5: The relationship between the net in-migration rate and its weather-predicted value. OLS analysis.

	Net in-migration rate
Weather-predicted value of the net in-migration rate	0.431***
	(0.001)
Economic development in 2001	-0.009***
	(0.000)
Male	0.001***
	(0.000)
Age	0.000**
	(0.000)
Age^2	-0.000***
	(0.000)
Education	0.010***
	(0.000)
Urban area	0.039***
	(0.000)
R^2	0.331
Observations	530,827

Notes: The table reports OLS estimates. The unit of observation is the non-migrant. Robust standard errors are bootstrapped at the regional level. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Table 6: The impact of the net in-migration rate on the employment probability of the non-migrants. IV probit analysis.

	Employment probability
Predicted value of the net in-migration rate	-0.757***
	(0.033)
Economic development in 2001	-0.108***
	(0.002)
Male	0.232***
	(0.001)
Age	0.204***
	(0.001)
Age^2	-0.009***
	(0.000)
Education	0.000
	(0.001)
Urban area	-0.067***
	(0.002)
Observations	530,827

Notes: The table reports marginal effects, based on an IV probit estimation. The unit of observation is the non-migrant. Robust standard errors are bootstrapped at the regional level. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Table 7: The impact of the net in-migration rate on the employment probability of the non-migrants in regions with low and high road density. IV probit analysis.

	Employment probability	
	Below-median road density	Above-median road density
Predicted value of the net in-migration rate	-2.096***	-0.412***
	(0.106)	(0.030)
Economic development in 2001	-0.127***	-0.060***
	(0.002)	(0.003)
Male	0.222***	0.240***
	(0.002)	(0.002)
Age	0.198***	0.209***
	(0.001)	(0.001)
Age^2	-0.009***	-0.009***
	(0.000)	(0.000)
Education	-0.003	-0.003
	(0.002)	(0.002)
Urban area	-0.079***	-0.067***
	(0.004)	(0.002)
Observations	248,560	282,267

Notes: The table reports marginal effects, based on an IV probit estimation. The unit of observation is the non-migrant. Robust standard errors are bootstrapped at the regional level. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Table 8: The relationship between the probability to migrate out of region j and the road density of region j. Probit analysis.

	Out-migration probability
Road density	-0.025***
	(0.001)
Economic development in 2001	-0.049***
	(0.001)
Male	0.010***
	(0.001)
Age	-0.013***
	(0.000)
Age^2	-0.000***
	(0.000)
Education	0.035***
	(0.001)
Pseudo-R^2	0.064
Observations	637,853

Notes: The table reports marginal effects, based on a probit estimation. The unit of observation is the population of migrants and non-migrants whose employment status is known. Robust standard errors are bootstrapped at the regional level. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Table 9: The impact of the net in-migration rate on the employment probability of the non-migrants in region j when one controls for the SPI of region j as well as for its economic growth rate between 1997 and 2002. IV probit analysis.

	Employment probability		
	All	Below-median road density	Above-median road density
Predicted value of the net in-migration rate	-0.897***	-2.328***	-0.509***
	(0.032)	(0.100)	(0.034)
Economic development in 2001	-0.146***	-0.146***	-0.112***
	(0.002)	(0.003)	(0.006)
SPI (yearly average between 1997 and 2002)	0.027***	-0.043***	0.039
	(0.002)	(0.003)	(0.004)
Economic growth between 1997 and 2002	0.251***	0.053***	0.256***
	(0.010)	(0.014)	(0.025)
Male	0.231***	0.220***	0.239***
	(0.001)	(0.002)	(0.002)
Age	0.204***	0.197***	0.209***
	(0.001)	(0.001)	(0.001)
Age^2	-0.009***	-0.009***	-0.009***
	(0.000)	(0.000)	(0.000)
Education	0.001	-0.000	-0.002
	(0.001)	(0.002)	(0.002)
Urban area	-0.067***	-0.072***	-0.073***
	(0.002)	(0.004)	(0.003)
Observations	2,789,561	1,388,705	1,400,856

Notes: The table reports marginal effects, based on an IV probit estimation. The unit of observation is the non-migrant. Robust standard errors are bootstrapped at the regional level. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Appendix

The purpose of this Appendix is to show that

$$\epsilon_{w,L} \equiv \frac{\mathrm{d}w}{\mathrm{d}L}\frac{L}{w} = (1 - \frac{1}{1-\epsilon_{c,K}\epsilon_{k,c}})\epsilon_{w,L}|_{K=cst}.$$

We know that $w = F_L(K,L)$ at the optimum. Differentiating this equation with respect to K and L yields:

$$\mathrm{d}w = [F_{KL}(K,L)\frac{\mathrm{d}K}{\mathrm{d}L} + F_{LL}(K,L)]\mathrm{d}L,$$

which can be rewritten as follows:

$$\epsilon_{w,L} \equiv \frac{\mathrm{d}w}{\mathrm{d}L}\frac{L}{w} = [\frac{F_{KL}(K,L)}{F_L(K,L)}\frac{\mathrm{d}K}{\mathrm{d}L}L + \frac{F_{LL}(K,L)L}{F_L(K,L)}].$$

Since $F(K,L)$ is homogeneous of degree 1, $F_L(K,L)$ is homogeneous of degree 0 and therefore, according to Euler's formula:

$$F_{KL}(K,L) = -\frac{F_{LL}(K,L)L}{K}.$$

By plugging this equation into the expression of $\epsilon_{w,L}$, we obtain:

$$\epsilon_{w,L} \equiv \frac{F_{LL}(K,L)L}{F_L(K,L)}[-\frac{\mathrm{d}K}{\mathrm{d}L}\frac{L}{K} + 1] = (1-\epsilon_{K,L})\epsilon_{w,L}|_{K=cst},$$

where $\epsilon_{K,L}$ denotes the elasticity of K with respect to L and $\epsilon_{w,L}|_{K=cst}$ stands for the elasticity of w with respect to L when K is fixed. As a matter of fact, differentiating $w = F_L(K,L)$ with respect to L when K is fixed leads to: $\mathrm{d}w = F_{LL}(K,L)\mathrm{d}L$, which can be rewritten as follows: $\epsilon_{w,L}|_{K=cst} \equiv \frac{\mathrm{d}w}{\mathrm{d}L}\frac{L}{w}|_{K=cst} = \frac{F_{LL}(K,L)L}{F_L(K,L)}$.

We now proceed to further arrangements to demonstrate that

$$\epsilon_{K,L} \equiv \frac{1}{1-\epsilon_{c,K}\epsilon_{k,c}}.$$

We know that $c = F_K(K,L)$ at the optimum. Differentiating this equation with respect to K and L yields:

$$\mathrm{d}K[c' - F_{KK}(K,L)] = \mathrm{d}LF_{KL}(K,L),$$

which can be rewritten as follows:

$$\epsilon_{K,L} \equiv \frac{\mathrm{d}K}{\mathrm{d}L}\frac{L}{K} = \frac{1}{\frac{Kc'}{LF_{KL}(K,L)} - \frac{KF_{KK}(K,L)}{LF_{KL}(K,L)}}.$$

Since $F(K,L)$ is homogeneous of degree 1, $F_K(K,L)$ is homogeneous of degree 0 and therefore, according to Euler's formula:

$$KF_{KK}(K,L) + LF_{KL}(K,L) = 0.$$

Relying on this equation, the expression of $\epsilon_{K,L}$ can be rewritten as follows:

$$\epsilon_{K,L} \equiv \frac{1}{1 - \frac{Kc'}{KF_{KK}(K,L)}}. \tag{1}$$

Let us define the elasticity of c with respect to K as $\epsilon_{c,K} \equiv \frac{\mathrm{d}c}{\mathrm{d}K}\frac{K}{c}$. Therefore:

$$Kc' \equiv c\epsilon_{c,K}. \tag{2}$$

Moreover, we know that $c = F_K(K,L)$ at the optimum. Since $F_K(K,L)$ is homogeneous of degree 0, this equation can be rewritten as follows:

$$c = F_K(k,1),$$

where k denotes the ratio of capital per worker ($k = \frac{K}{L}$).

After differentiating this equation with respect to c and k, we get:

$$\mathrm{d}c = \mathrm{d}k F_{KK}(k,1),$$

and therefore:

$$\epsilon_{k,c} \equiv \frac{\mathrm{d}k}{\mathrm{d}c}\frac{c}{k} = \frac{c}{kF_{KK}(k,1)} = \frac{cL}{KF_{KK}(k,1)}.$$

Yet, since $F(K,L)$ is homogeneous of degree 1, $F_{KK}(K,L)$ is homogeneous of degree -1 and therefore:

$$F_{KK}(k,1) = LF_{KK}(K,L).$$

Hence:

$$\epsilon_{k,c} \equiv \frac{c}{KF_{KK}(K,L)}. \tag{3}$$

As a consequence, using Equations (1), (2) and 3), the expression of $\epsilon_{K,L}$ can be rewritten as follows:

$$\epsilon_{K,L} \equiv \frac{1}{1 - \epsilon_{c,K}\epsilon_{k,c}}.$$