



HAL
open science

The effect of weather-induced internal migration on local labor markets. Evidence from Uganda

Eric Strobl, Marie-Anne Valfort

► **To cite this version:**

Eric Strobl, Marie-Anne Valfort. The effect of weather-induced internal migration on local labor markets. Evidence from Uganda. World Bank Economic Review, 2015, 29 (2), pp.385-412. 10.1093/wber/lht029 . halshs-00977051

HAL Id: halshs-00977051

<https://shs.hal.science/halshs-00977051>

Submitted on 28 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The effect of weather-induced internal migration on local labor markets. Evidence from Uganda.

Eric Strobl*

Marie-Anne Valfort[†]

June 13, 2013

Abstract

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. Consistent with the prediction of a simple theoretical model, our results reveal a larger negative impact than the one documented for developed countries. They further show that this negative impact is significantly stronger in Ugandan 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 10 percentage points.

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

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

*Ecole Polytechnique Paris and IPAG Research Lab. Department of Economics. Route de Saclay. 91128 Palaiseau Cedex. E-mail: eric.strobl@polytechnique.edu.

[†]Corresponding author. Paris School of Economics - Paris 1 Panthéon Sorbonne University. 106-112, Boulevard de l'Hôpital. 75013 Paris. France. E-mail: marie-anne.valfort@univ-paris1.fr. Phone: 33(0)6 40 12 84 07.

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. In the same vein, Ruist and Bigsten (2013) demonstrate that international capital adjustment substantially diminishes the negative impact of immigration from developing countries on native wages in developed countries.

Surprisingly, little attention has been paid to the impact of migration on labor market outcomes in developing countries.² 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.³ 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

¹See Hunt (1992), Friedberg and Hunt (1995), Carrington and de Lima (1996), Pischke and Velling (1997), Friedberg (2001), Suen (2000), Card (2005), Longhi, Nijkamp and Poot (2005), McIntosh (2008), Hanson (2009), Boustan, Fishback and Kantor (2010).

²Berker (2011) is an exception.

³The 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.

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 in Uganda.⁴

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. We construct this variable for each destination region such that it depends on the weather shocks affecting the source regions.

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. Munshi (2003) was the first to find a strong negative impact of rainfall in the Mexican origin-communities (where rain-fed agriculture is the major occupation) on migration from these communities to the United States. Since then, a vast literature has confirmed that extreme weather conditions impose considerable strains on populations that depend on rain-fed agriculture (see Miguel, Satyanath and Sergenti (2004), Barrios, Bertinelli and Strobl (2006), Yang and Choi (2007), Gray and Mueller (2012), Beegle, De Weerd and Dercon (2011), Miguel and Satyanath (2011), Marchiori, Maystadt and Schumacher (2012)).⁵ 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 de-

⁴According to the World Bank, Uganda was ranked 190th of 215 countries in 2010 in terms of GNI/capita (PPP).

⁵Note 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).

pendence 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 destination region as an instrumental variable is that this variable is expected to be orthogonal to the unobserved correlates of the employment probability of the non-migrants in that region. The weather predicted value of the net in-migration rate indeed depends on the weather shocks that affect the *source* regions, not on the weather shocks that affect the destination 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.

Consistent with the prediction of a simple theoretical model, our results reveal a 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 7.8 percentage points. Our results further show that this negative impact is significantly stronger in Ugandan 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 10 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 economies 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 economies

with lower road density.⁶

We consider an 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, meaning that the labor supply increases with w . As a consequence, the employment probability also increases with w . We suppose that firms have to incur a cost to finance capital that we denote by $c(K^a, Q)$, where K^a stands for the aggregate capital in the economy and $Q \in [0, \bar{Q}]$ captures road density in the economy.

We suppose that $\frac{\partial c}{\partial Q} < 0$, meaning that c decreases with road density. Transportation infrastructure in general, and road infrastructure in particular, have indeed been shown to trigger factor mobility and therefore the productivity of firms, thanks to a decrease in transportation costs. There is widespread evidence of this phenomenon in developed countries (see Michaels (2008) for evidence for the US or Holl (2004a and 2004b) and Cieslik (2005) for evidence for Europe). But this phenomenon has been widely documented for developing countries too, where poor transportation infrastructure is presented as a key constraint for industrial development (see Bloom and Sachs (1998) and Tybout (2000)). In South Africa, for instance, McPherson (1995) finds that micro and small enterprises located along the road and close to commercial centers have a better chance of survival. Similarly, Renkow, Hallstrom and Karanja (2004) find that transaction costs in rural Kenya increase with remoteness of villages, which ultimately constrains farmers' market participation. More recently, Rijkers, Söderbom and Loening (2010), Fafchamps and Söderbom (2013) and Shiferaw, Söderbom, Siba and Alemu (2013) have provided compelling evidence of the positive impact of road

⁶The predictions of our model remain unchanged if, instead of focusing on road density, we focus on the density of transportation infrastructure generally speaking. We choose to focus on road density because roads play the dominant role worldwide as a percentage of freight. For instance, roads carry over 60% of freight in the US (US Bureau of Transportation Statistics (2006)) and 75% of freight in Africa (World Bank (2011)).

infrastructure on Ethiopian firms' performance.⁷

Moreover, we suppose that $\frac{\partial c}{\partial K^a} \geq 0$, meaning that c increases weakly with aggregate capital. Indeed, the higher the aggregate capital, the lower the ability to convey capital if congestion effects emerge, and therefore the higher c . More precisely, when road density is maximal ($Q = \bar{Q}$), we assume no congestion effects: capital is perfectly mobile between the economy and the rest of the world and therefore $\frac{\partial c}{\partial K^a} = 0$. However, as soon as road density is lower than its maximal value ($Q < \bar{Q}$), congestion effects emerge, that decrease with road density. In this case, $\frac{\partial c}{\partial K^a} > 0$ and $\frac{\partial^2 c}{\partial K^a \partial Q} < 0$.

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

$$\epsilon_{w,L} \equiv \frac{dw}{dL} \frac{L}{w}.$$

Following standard calculus detailed in the Appendix, we obtain:

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

⁷See Limão and Venables (2001), Buys, Deichmann and Wheeler (2010), Escribano, Guash and Pena (2010), Jedwab and Moradi (2012) for additional evidence from Africa. See also Chen (1996), Wei, Liu, Parker and Vaidya (1998), Banerjee, Duflo and Qian (2012), Datta (2012), Rothenberg (2012) and Donaldson (2013) for evidence from Asia.

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 ($Q = \bar{Q}$), $\epsilon_{c,K} = 0$ and therefore $\epsilon_{w,L} = 0$. Figure A1 in the Appendix illustrates such mechanisms. 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, since the labor supply is a function of w , the labor supply does not change either. Therefore, the employment probability of the non-migrants is itself unaffected. In equilibrium, the increase in employment is equal to the number of arrivals of migrants.

When capital is not perfectly mobile ($Q < \bar{Q}$), 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 of no mobility of capital ($Q = 0$), $\epsilon_{c,K} \rightarrow +\infty$ and therefore $\epsilon_{w,L} = \epsilon_{w,L}|_{K=cst}$: the negative impact of an influx of migrants on w is maximal. Figure A2 in the Appendix 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 . Since the labor supply increases with w , an influx of migrants also decreases the employment probability of the non-migrants. In equilibrium, the increase in employment is smaller than the number of arrivals of migrants.

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

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

This proposition in turn leads to the two following hypotheses:

*Hypothesis 1: The impact of an influx of migrants on the employment probability of the non-migrants is more negative in developing than in developed countries.*⁸

⁸As already emphasized, developing countries are characterized by poorer road infrastructure.

Hypothesis 2: Within a developing country, 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 these two hypotheses. Recall that empirical evidence from developed countries has shown no impact of immigration on the employment probability of natives (see Card (1990 and 2005) and McIntosh (2008) for the US, Hunt (1992) for France, Pischke and Velling (1997) for Germany, Carrington and de Lima (1996) for Portugal, Friedberg (2001) for Israel). Therefore, we test Hypothesis 1 by estimating whether the (negative) impact of an influx of migrants on the employment probability of non-migrants in Uganda is significantly different from 0. We test Hypothesis 2 by examining whether this negative impact is significantly stronger in Ugandan regions characterized by lower road density (and therefore lower prospects of capital mobility).

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, a negative impact of an inflow of migrants that is lower in regions with higher road density (Hypothesis 2) could be accounted for by the fact that, in these regions, the inflow of migrants is accompanied by a larger outflow of residents. To rule out this possibility, we focus in the following on the impact of the *net* in-migration rate (i.e., on the impact of the in-migration rate once the impact of the out-migration rate has been netted out).

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 status 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 (i.e., non-employed and actively searching for a job), 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 non-employed and actively searching for a job or inactive.

Note that unemployment is typically defined in the developed world in terms of being non-employed and actively searching for a job (in the formal sector). Yet, the ILO (1982) has recognized that the job search criteria may not be meaningful in developing countries where labour markets are dominated by the informal sector. The unsuitability of the standard

definition of unemployment for developing countries is clearly apparent in our data: if we consider as unemployed only those who are non-employed and actively searching for a job, we end up with an unemployment rate in Uganda of only 3%. A better alternative would therefore consist in combining the job-seekers with the non-employed who are not actively searching for a job. Byrne and Strobl (2004) have indeed shown that, in a developing country like Trinidad and Tobago, job-seekers do not differ from the non-employed who are not actively searching for a job. Yet, in our dataset, this latter category is not distinguished among the broader category of inactive individuals. This limitation explains why we combine job-seekers with inactive when we create our “employment” variable. Yet, we verify in the robustness checks that our results hold if we focus on an alternative dependent variable that is equal to 1 if the respondent reports to be employed and 0 if she reports to be non-employed and actively searching for a job.

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 km² (approximately 65 km*65 km) each. We depict these in Figure A3 in the Appendix. 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) dataset⁹ 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 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

⁹This dataset is available at <http://www.cru.uea.ac.uk/cru/data/hrg/>.

(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 2002.¹⁰ 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 A4 in the Appendix. 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 Lake Victoria.

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

¹⁰This database is available at <http://www.agiweb.org/pubs/globalgis/description.html>.

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) to 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 A5 in the Appendix. 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.¹¹ 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,¹² 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

¹¹Note that the average region pair in our dataset exhibits bidirectional flows. This feature confirms that migration flows are not all induced by local push or pull factors. For instance, one expects migration for marriage to lead to bidirectional flows across regions, especially in the context of a developing country where marriage is notably used as “an implicit interhousehold contractual arrangement aimed at mitigating income risks and facilitating consumption smoothing” (see Rosenzweig and Stark (1989)). But this feature is also consistent with the fact that departures from (resp. arrivals to) a region may prompt some in-migration (resp. out-migration), in response for instance to an increase (resp. decrease) in available labor market opportunities accompanying out-migration (resp. in-migration). Such a pattern has been documented in the US during the Great Depression by Boustan, Fishback and Kantor (2010) and, as we show below, also seems to be at work in the Ugandan context.

¹²The “near normal” category concerns SPI values ranging from -0.99 to 0.99, i.e., rainfall shocks classified below “moderate”.

kilometer while the average nightlights intensity is equal to 15.208.

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 $n_{j,01-02}^{WP}$. This variable stands for the difference between the weather-predicted value of the in-migration rate in region j ($i_{j,01-02}^{WP}$ hereafter) and the weather-predicted value of the out-migration rate from region j ($o_{j,01-02}^{WP}$ hereafter) over this period. In the following, we describe how we compute $i_{j,01-02}^{WP}$ and $o_{j,01-02}^{WP}$, 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:

$$O_{k,01-02} = \alpha + \beta.SPI_{k,97-02} + u_k. \quad (1)$$

For each source region k , we then compute $P_{kj,01-02}$ which represents the share of people leaving source region k who settle in destination region j . We regress $P_{kj,01-02}$ on a function that is quadratic in $\frac{dist_{kj}}{road_{kj}}$, that is in the geographic distance between regions k and j , once this distance has been normalized by road density between these regions:

$$P_{kj,01-02} = \delta_k + \theta_k \left(\frac{dist_{kj}}{road_{kj}} \right) + \eta_k \left(\frac{dist_{kj}}{road_{kj}} \right)^2 + \mu_k. \quad (2)$$

More precisely, $\frac{dist_{kj}}{road_{kj}} = 0$ if the source and the destination regions are contiguous. For the case where regions k and j do not share a common border, we take the centroid of each source and destination region pair and consider an imaginary straight line between these two centroids. We compute the geographic distance related to the portion of this straight line that links the border of region k to the border of region j . We then divide this geographic distance by the average of the road density in regions that are crossed by this portion of the straight line. This average is weighted by distance, meaning that road densities in regions that stand for a higher (resp. lower) share of this geographic distance are assigned a higher (resp. lower) weight. In mathematical terms, $\frac{dist_{kj}}{road_{kj}} = [\frac{\sum_i dist_i}{\sum_i \sum_i dist_i road_i}]$ (where i indexes regions located between the source and the destination regions) when the source and the destination regions are not contiguous.

Gravity models originally designed to account for trade between economies have been used recently to also account for immigration flows across source and destination regions (see Lewer and van den Berg (2008), Grogger and Hanson (2011) and Alesina, Harnoss and Rapoport (2013) for an application of gravity models to immigration). These models suggest that immigration, like international trade, is driven by the attractive force between source and destination regions and hindered by the costs of moving from the source to the destination region. The geographic distance between the source and the destination region normalized by the road density between these regions clearly captures the cost of moving from one region to the other. We therefore expect the sign of θ_k in Equation (2) to be negative.

It is worthwhile emphasizing that $\frac{dist_{kj}}{road_{kj}}$ only depends on the size and road density of the regions that are located between the source and the destination regions. Put differently, it depends neither on the size nor on the road density in the destination region. This is an important requirement in order to ensure that our instrument will satisfy the exclusion restriction. The size and the road density in the destination region are indeed likely correlated with the employment probability of non-migrants in this region.

The need to ensure that our instrument will satisfy the exclusion restriction prevents us from controlling in Equation (2) for the second critical variable that typically enters a gravity model of migration flow, namely population size in the source and destination regions (the larger the population in a source region, the more people are likely to migrate,

and the larger the population in the destination region, the larger is the labor market for immigrants). Moreover, we also cannot control for other standard predictors of migration flows. For instance, Friedberg and Hunt (1995), Card (2001) and Borjas (2003) have shown that the ratio of destination to source region per capita incomes should positively impact migration flows between these regions. Similarly, the share of migrants from the source region already established in the destination region should positively impact the migration flows between these regions. Munshi (2003) and McKenzie and Rapoport (2010) indeed demonstrate that the cost of adapting to a new country is mitigated by the presence of compatriots familiar with both the source and destination country cultures. Yet, the network size of migrants in the destination region is likely correlated with unobservables that also influence the employment probability of non-migrants in this region.

Once Equation (2) is estimated, one can compute the weather-predicted in-migration flow to destination region j denoted by $I_{j,01-02}^{WP}$. This flow is 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 :

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

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

$$i_{j,01-02}^{WP} = \frac{I_{j,01-02}^{WP}}{pop_{j,01}}.$$

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:

$$I_{k,01-02} = \alpha + \beta.SPI_{k,97-02} + u_k. \quad (3)$$

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 geographic distance between regions j and k , once this distance has been normalized by road density between these regions:

$$P_{jk,01-02} = \delta_k + \theta_k \left(\frac{dist_{jk}}{road_{jk}} \right) + \eta_k \left(\frac{dist_{jk}}{road_{jk}} \right)^2 + \mu_k. \quad (4)$$

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 :

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

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

$$o_{j,01-02}^{WP} = \frac{O_{j,01-02}^{WP}}{pop_{j,01}}.$$

Two sets of regressions allow us to construct $n_{j,01-02}^{WP}$, 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 $\frac{dist_{kj}}{road_{kj}}$ (resp. $\frac{dist_{jk}}{road_{jk}}$). 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. Perhaps somewhat surprisingly, they show that a high SPI generates both higher out- and in-migration flows. This pattern may be due to the fact that departures from (resp. arrivals to) a region due to weather shocks prompt some in-migration (resp. out-migration), in response for instance to an increase (resp. decrease) in available labor market opportunities accompanying out-migration (resp. in-migration). Such a pattern has already been documented in the US during the Great Depression by Boustan, Fishback and Kantor (2010). On net however, our results are fully intuitive: we find that a high SPI generates a higher net in-migration flow (see column 3).

As for the second set of regressions (Equation (2) and Equation (4)), due to their large numbers ($56 \times 2 = 112$) not reported here, we find that for all but 8 of these either the negative coefficient on the linear distance term is significantly different from zero at least 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 least 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 geographic distance between these regions, once this distance has been normalized by road density.

5 Results

Our objective below is to test Hypothesis 1 and Hypothesis 2.

5.1 Testing Hypothesis 1

To test Hypothesis 1, 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 in Uganda. We then turn to an IV probit approach that allows us to address this potential bias.

5.1.1 A naïve probit approach

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

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

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.

Column 1 of Table 4 reports the marginal effects of the probit estimation of Equation (5), where robust standard errors are bootstrapped at the regional level.¹³ We observe that an increase 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).

At this stage, it is important to ensure that these results are driven by in-migration, not by out-migration.¹⁴ It could indeed be the case that the decrease in the employment probability of the non-migrants is simply induced by the departure of individuals with the highest human and physical capital. In Column 2 of Table 4, following Docquier, Ozden and Peri (2011),

¹³An 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 99,025) – we consider bootstrapping at the regional level to be a more conservative approach than clustering.

¹⁴We thank an anonymous referee for raising this issue.

we analyze the correlation between the employment probability of the non-migrants on one hand and the in-migration and out-migration rates on the other hand. Consistent with our model, this check indicates that the negative correlation between the employment probability of the non-migrants and the net in-migration rate is due to in-migration, not out-migration: the coefficient of the in-migration rate is negative and significant while the coefficient of the out-migration rate is positive and significant.

Yet, 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.

5.1.2 An 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.

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

$$n_{j,01-02} = a + b.n_{j,01-02}^{WP} + c.D_{j,01} + d.SPI_{j,97-02} + \mathbf{X}_{ij,02}' \cdot \mathbf{e} + \epsilon_{ij,02}, \quad (6)$$

where $n_{j,01-02}$, $D_{j,01}$ and $\mathbf{X}_{ij,02}$ are defined as in Equation (5). The variable $n_{j,01-02}^{WP}$ is the instrument, that is the weather-predicted value of the net in-migration rate in destination region j that we defined in Section 4. The variable $SPI_{j,97-02}$ refers to the mean of the yearly values of the SPI in destination region j between 1997 and 2002. It is critical to control for this variable in the second stage (and therefore in the first stage) of the IV probit. As is apparent in Section 4., the instrument $n_{j,01-02}^{WP}$ indeed depends on the mean of the yearly values of the SPI between 1997 and 2002 in the source regions. Yet, it is likely that the SPI in the source regions positively depends on the SPI in the destination region when these source regions are geographically close to the destination region. This intuition is confirmed by

Figure A6 which clearly shows that the correlation between the SPI of two regions decreases with the geographic distance between the centroids of these regions. Obviously, the SPI in the destination region influences the employment probability of non-migrants in this region. Therefore, not controlling for the SPI in the destination region while performing the IV probit will surely violate the exclusion restriction and therefore undermine the validity of our instrument.

The estimates of the first stage of the IV probit are reported in Table 5. Since $n_{j,01-02}^{WP}$ 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 t-statistic related to the coefficient of $n_{j,01-02}^{WP}$ is equal to 174.67).

The second stage of the IV probit approach entails computing the probit estimates of Equation (7):

$$P(E_{ij,02} = 1) = G(a + b.\widehat{n_{j,01-02}} + c.D_{j,01} + d.SPI_{j,97-02} + \mathbf{X}_{ij,02}' \cdot \mathbf{e} + \epsilon_{ij,02}), \quad (7)$$

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

It is important at this stage to discuss our empirical strategy in view of our theoretical model. More precisely, our empirical strategy aims at estimating $\epsilon_{w,L}$ (the elasticity of the wage w with respect to the labor supply L), although, because of lack of data on wages, we proxy the impact of a change in L on w by the impact of a change in L on the employment probability (an adequate approach given that the labor supply increases with w). Our model thus suggests regressing the change in the employment status of the non-migrants between 2001 and 2002 on the change in the labor supply between 2001 and 2002, captured by the net in-migration rate during this period (i.e., $n_{j,01-02}$). Since $n_{j,01-02}$ is likely endogenous, we implement an IV strategy that basically amounts to instrumenting $n_{j,01-02}$ with rainfall shocks in the source regions. However, cross-sectional data only provide information on the

employment probability of the non-migrants in 2002, meaning that the employment probability of the non-migrants in 2001 is left as an omitted variable. Yet, since our instrument is built such that it affects the employment probability of the non-migrants in the destination region only through its impact on migration, the employment probability of the non-migrants in 2001 is correlated with our instrument only through $n_{j,01-02}$. This feature rules out the possibility that our estimates suffer from an omitted variable bias. Nevertheless, we include in our regression an indicator of the economic development in the destination region in 2001 ($D_{j,01}$) to serve as a proxy for the employment probability of the non-migrants in this region in 2001.

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 7.8 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. More importantly, it validates Hypothesis 1, according to which the impact of an influx of migrants on the employment probability of the non-migrants is more negative in a developing country like Uganda characterized by low road density and therefore low capital mobility, than in developed countries.¹⁵

5.2 Testing Hypothesis 2

Our results so far show a much larger negative impact of the net in-migration rate on the employment probability of non-migrants in Uganda than the one documented for developed countries. We now want to test Hypothesis 2, by examining whether this negative impact is significantly stronger in Ugandan regions characterized by lower road density (and therefore lower prospects of capital mobility).

To this end, we estimate Equations (6) and (7) on two sub-samples: the sub-sample of regions characterized by a below-median road density, and the sub-sample of regions

¹⁵It is worth noting that the employment probability of both migrants and non-migrants decreases following an influx of migrants (results available upon request). Therefore, our results are not driven by a pure crowding out effect, whereby the labor of migrants substitutes for the labor of non-migrants. Put differently, an influx of migrants does have a negative impact on wages.

characterized by an above-median road density. Results for the second-stage are reported in Table 7. They confirm Hypothesis 2. 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 5.3 percentage points in regions showing above-median road density, the employment probability decreases by more than 10 percentage points in regions showing below-median road density, i.e., arguably a very large impact.¹⁶

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. In other words, a negative impact of an inflow of migrants that is lower in regions with higher road density (Hypothesis 2) could be accounted for by the fact that, in these regions, the inflow of migrants is accompanied by a larger outflow of residents. Recall that we guard against this possibility by focusing on the impact of the net in-migration rate (i.e., on the impact of the in-migration rate once the impact of the out-migration rate has been netted out). An alternative way to rule out this mechanism would be to analyze whether individuals living in regions with higher road density are indeed more likely to migrate out of these regions than are individuals living in regions with lower road density. We do so by computing the probit estimates of Equation (8):

$$P(O_{ij,01-02} = 1) = G(a + b.R_j + c.D_{j,01} + \mathbf{X}_{i,02}' \cdot \mathbf{d} + \epsilon_{ij,02}), \quad (8)$$

where G is the cumulative distribution function for a standard normal density. The dummy $O_{ij,01-02}$ stands for the decision of individual i to migrate out of 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 nightlights intensity of region j in 2001. Finally, $\mathbf{X}_{i,02}$ is a vector of socio-economic characteristics of individual i in 2002 (gender, age, and education).

Table A1 in the Appendix 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

¹⁶A test of the equality of the coefficients related to $n_{j,01-02}$ in both equations shows that these coefficients significantly differ at the 1% confidence level.

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.

6 Robustness checks

In the following, we perform two robustness checks. One investigates whether our results hold when we rely on a more standard (though less suitable for developing countries) definition of unemployment. The other examines the robustness of our results when we control for an additional variable in the IV probit approach which, when it is omitted, could lead to a violation of the exclusion restriction.

6.1 An alternative dependent variable

Our dependent variable is equal to 1 if the respondent reports to be employed and 0 if she reports to be unemployed (i.e., non-employed and actively searching for a job) or inactive. Yet, although suitable for developing countries (see ILO (1982)), our definition of unemployment (being non-employed and actively searching for a job or being inactive) is not standard. Unemployment in developed countries indeed does not include the inactive population. We therefore investigate whether our results hold when we rely on an alternative dependent variable that is equal to 1 if the respondent reports to be employed and 0 if she reports to be non-employed and actively searching for a job.

The estimates of the second stage of the IV probit analysis are reported in Table A2 in the Appendix and confirm that our results are robust to relying on this alternative dependent variable. The orders of magnitude are much lower however. This finding is consistent with the idea that unemployment needs to be defined more broadly in developing countries. Removing the inactive from our analysis indeed provides a very partial picture of the consequences of internal migration on local labor markets in Uganda.

6.2 Avoiding the violation of the exclusion restriction

It may be that weather shocks in the source regions have an impact on the economic conditions in the destination region (and therefore on the employment probability of the non-migrants in this region) that does not only transmit through internal migration, but through other channels, as for instance, when the destination region and the source regions are trade partners (or trade competitors). To control for these other channels, we include a proxy for economic growth in the destination region between 1997 and 2002 (the period over which the instrument could have an impact on the dependent variable that does not transmit through internal migration). This proxy is the growth rate in the nightlights intensity between 1997 and 2002. The estimates of the second stage of the IV probit analysis are reported in Table A3 in the Appendix and show that our results are roughly unchanged by this control.

7 Conclusion

This paper investigates the impact of weather-induced internal net in-migration rates on the employment probability of non-migrants in destination regions in Uganda. Consistent with the prediction of a simple theoretical model, our results reveal a larger negative impact 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 7.8 percentage points. Our results further show that this negative impact is significantly stronger in Ugandan 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 10 percentage points.

Our findings suggest that the development of road infrastructure which ranks high on the World Bank's agenda¹⁷ 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

¹⁷See World Bank (2009). See also Hulten, Bennathan and Srinivasan (2006) and Freund and Rocha (2011) for evidence on the positive impact of road infrastructure on development.

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.

8 Funding

This work was supported by the Agence Nationale de Recherche [Grant ANR-JCJC-0127-01].

9 Acknowledgements

We thank two anonymous referees, Milo Bianchi, Pierre Cahuc, Marie Gachassin, Mathilde Maurel, Katrin Millock, François Rycx, Ilan Tojerow, Philip Verwimp, and participants at the DIAL Development Conference “Shocks in developing countries” (Paris, 2011) and at the Research Seminar in Applied Economics and Management (Université Libre de Bruxelles) for their generous guidance and helpful comments.

References

- [1] Alesina, Alberto, Johann Harnoss and Hillel Rapoport. 2013. “Birthplace diversity and economic growth.” Working Paper.
- [2] 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.
- [3] 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.

- [4] Banerjee, Abhijit, Esther Duflo and Nancy Qian. 2012. "On the road: access to transportation infrastructure and economic growth in China." Working Paper.
- [5] Beegle, Kathleen, Joachim De Weerd 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] Bloom, David E. and Jeffrey D. Sachs. 1998. "Geography, demography, and economic growth in Africa." *Brookings Papers on Economic Activity* 1998(2): 207-295.
- [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 demand curve 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] 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.
- [13] US Bureau of Transportation Statistics. 2006. *Freight in America*. United States Department of Transportation.

- [14] Buys, Piet, Uwe Deichmann and David Wheeler. 2010. "Road network upgrading and overland trade expansion in Sub-Saharan Africa." *Journal of African Economies* 19(3): 399-432.
- [15] Byrne, David and Eric Strobl. 2004. "Defining unemployment in developing countries: evidence from Trinidad and Tobago." *Journal of Development Economics* 73(1): 465-476.
- [16] 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.
- [17] Card, David. 1990. "The impact of the Mariel Boatlift on the Miami labor market." *Industrial and Labor Relations Review* 43(2): 245-257.
- [18] Card, David. 2001. "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration." *Journal of Labor Economics* 19(1): 22-64.
- [19] Card, David. 2005. "Is the new immigration really so bad?" *Economic Journal* 115(507): F300-F323.
- [20] Carrington, William and Pedro de Lima. 1996. "The impact of 1970s repatriates from Africa on the Portuguese labor market." *Industrial and Labor Relations Review* 49(1): 330-347.
- [21] Chen, Chien-Hsun. 1996. "Regional determinants of foreign direct investment in Mainland China." *Journal of Economic Studies* 23(2): 18-30.
- [22] Cieslik, Andrzej. 2005. "Regional characteristics and the location of foreign firms within Poland." *Applied Economics* 37: 863-874.
- [23] Datta, Saugato. 2012. "The impact of improved highways on Indian firms." *Journal of Development Economics* 99(1): 46-57.
- [24] Dercon, Stefan. 2004. "Growth and shocks: evidence from rural Ethiopia." *Journal of Development Economics* 74(2): 304-329.

- [25] Deschenes, Olivier and Enrico Moretti. 2009. "Extreme weather events, mortality and migration." *Review of Economics and Statistics* 91(4): 659-681.
- [26] Docquier, Frederic, Caglar Ozden and Giovanni Peri. 2011. "The Labor Market effects of immigration and emigration in OECD Countries." IZA Discussion Paper No. 6258.
- [27] Donaldson, Dave. 2013. "Railroads of the Raj: estimating the impact of transportation infrastructure." *American Economic Review*, forthcoming.
- [28] Escribano, Alvaro, J. Luis Guash and Jorge Pena. 2010. "Assessing the impact of infrastructure quality on firm productivity in Africa: cross-country comparisons based on investment climate surveys from 1999 to 2005." World Bank Policy Research Working Paper 5191. Washington D.C: The World Bank.
- [29] Fafchamps, Marcel and Mans Söderbom. 2013. "Network proximity and business practices in African manufacturing." *World Bank Economic Review*, forthcoming.
- [30] 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.
- [31] Freund, Caroline and Nadia Rocha. 2011. "What Constrains Africa's Exports?" *World Bank Economic Review* 25(3): 361-386.
- [32] Friedberg, Rachel M. 2001. "The impact of mass migration on the Israeli labor market." *Quarterly Journal of Economics* 116(4): 1373-1408.
- [33] 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.
- [34] Gray, Clark and Valerie Mueller. 2012. "Drought and population mobility in rural Ethiopia." *World Development* 40(1): 13-145.
- [35] Grogger, Jeffrey and Gordon H. Hanson. 2011. "Income maximization and the selection and sorting of international migrants." *Journal of Development Economics* 95(1): 42-57.

- [36] 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.
- [37] Hanson, Gordon H. 2009. "The economic consequences of the international migration of labor." *Annual Review of Economics* 1(1): 179-208.
- [38] Henderson, J. Vernon, Adam Storeygard and David N. Weil. 2012. "Measuring economic growth from outer space." *American Economic Review* 102(2): 994-1028.
- [39] 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.
- [40] Holl, Adelheid. 2004a. "Manufacturing location and impacts of road transport infrastructure: empirical evidence from Spain." *Regional Science and Urban Economics* 34(3): 341-363.
- [41] Holl, Adelheid. 2004b. "Transport infrastructure, agglomeration economies, and firm birth. Empirical evidence from Portugal." *Journal of Regional Science* 44(4): 693-712.
- [42] Hulten, Charles R., Esra Bennathan and Sylaja Srinivasan. 2006. "Infrastructure, externalities, and economic development: a study of the Indian manufacturing industry." *World Bank Economic Review* 20(2): 291-308.
- [43] 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.
- [44] ILO. 1982. "Resolution concerning statistics of the economically active population, employment, unemployment and underemployment." Thirteenth International Conference of Labour Statisticians.
- [45] Jedwab, Rémi and Alexander Moradi. 2012. "Transportation infrastructure and development in Ghana." Working Paper.

- [46] Lewer, Joshua J. and Hendrik Van den Berg. 2008. "A Gravity Model of Immigration." *Economics Letters* 99(1): 164-167.
- [47] Limão, Nuno and Anthony J. Venables. 2001. "Infrastructure, geographical disadvantage, transport costs, and trade." *World Bank Economic Review* 15(3): 451-479.
- [48] 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.
- [49] Lucas, Robert E.B. 1997. "Internal migration in developing countries." *Handbook of Population and Family Economics* 13: 721-798.
- [50] 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.
- [51] McIntosh, Molly Fifer. 2008. "Measuring the labor market impacts of Hurricane Katrina migration: evidence from Houston, Texas." *American Economic Review* 98(2): 54-57.
- [52] McKenzie, David and Hillel Rapoport. 2010. "Self-selection patterns in Mexico-U.S. migration: the role of migration networks." *Review of Economics and Statistics* 92(4): 811-821.
- [53] McPherson, Michael A. 1995. "The hazards of small firms in Southern Africa." *Journal of Development Studies* 32(1): 31-54.
- [54] Michaels, Guy. 2008. "The effect of trade on the demand for skill: evidence from the interstate highway system." *Review of Economics and Statistics* 90(4): 683-701.
- [55] 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.
- [56] Miguel, Edward and Shanker Satyanath. 2011. "Re-examining economics shocks and civil conflict." *American Economic Journal: Applied Economics* 3(4), 228-232.

- [57] 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.
- [58] Munschi, Kaivan. 2003. "Networks in the modern economy: Mexican migrants in the U.S. labor market." *Quarterly Journal of Economics* 118(2): 549-599.
- [59] Pischke, Jörn-Steffen and Johannes Velling. 1997. "Employment effects of immigration to Germany: an analysis based on local labor markets." *Review of Economics and Statistics* 79(4): 594-604.
- [60] Renkow, Mitch, Daniel G. Hallstrom and Daniel D. Karanja. 2004. "Rural infrastructure, transactions costs and market participation in Kenya." *Journal of Development Economics* 73(1): 349-367.
- [61] Rijkers, Bob, Mans Söderbom and Josef L. Loening. 2010. "A rural-urban comparison of manufacturing enterprise performance in Ethiopia." *World Development* 38(9): 1278-1296.
- [62] Rosenzweig, Mark R. and Oded Stark. 1989. "Consumption smoothing, migration, and marriage: evidence from rural India." *Journal of Political Economy* 97(4): 905-926.
- [63] Rothenberg, Alex. 2012. "Transport infrastructure and firm location choice in equilibrium: evidence from Indonesia's highways." Working Paper.
- [64] Ruist, Joakim and Arne Bigsten. 2013. "Wage effects of labour migration with international capital mobility." *World Economy* 36(1): 31-47.
- [65] Shiferaw, Admasu, Mans Söderbom, Eyerusalem Siba and Getnet Alemu. 2013. "Road infrastructure and enterprise dynamics in Ethiopia." Working paper.
- [66] Suen, Wing. 2000. "Estimating the effects of immigration in one city." *Journal of Population Economics* 13(1): 99-112.
- [67] Tybout, James R. 2000. "Manufacturing firms in developing countries: how well do they do, and why?" *Journal of Economic Literature* 38(1): 11-44.

- [68] Wei, Yingqi, Xiaming Liu, David Parker and Kirit Vaidya. 1998. "The regional distribution of foreign direct investment in China." *Regional Studies* 33(9): 857-867.
- [69] Wooldridge, Jeffrey. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.
- [70] World Bank. 2009. *World Development Report 2009: Reshaping economic geography*. The World Bank: Washington D.C.
- [71] World Bank. 2011. *Africa's transport infrastructure. Mainstreaming maintenance and management*. The World Bank: Washington D.C.
- [72] Yang, Dean and HwaJung Choi. 2007. "Are remittances insurance? Evidence from rainfall shocks in the Philippines." *World Bank Economic Review* 21(2): 219-248.
- [73] 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

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 (1)	In-migration flow (2)	Net in-migration flow (3)
SPI (yearly average between 1997 and 2002)	2737.011** (1069.263)	4558.948** (1719.766)	1821.937** (782.179)
R ²	0.205	0.295	0.091
Observations	56	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	
	(1)	(2)
Net in-migration rate	-0.295*** (0.017)	
In-migration rate		-0.253*** (0.016)
Out-migration rate		0.665*** (0.039)
Economic development in 2001	-0.100*** (0.002)	-0.095*** (0.002)
Male	0.231*** (0.001)	0.232*** (0.001)
Age	0.204*** (0.001)	0.204*** (0.001)
Age ²	-0.009*** (0.000)	-0.009*** (0.000)
Education	-0.005*** (0.001)	-0.005*** (0.001)
Urban area	-0.076*** (0.002)	-0.088*** (0.002)
Pseudo-R ²	0.237	0.237
Observations	530,827	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 the instrument. OLS analysis.

	Net in-migration rate
Instrument (weather-predicted value of the net in-migration rate)	1.350*** (0.008)
Economic development in 2001	0.005*** (0.000)
SPI (yearly average between 1997 and 2002)	0.029*** (0.000)
Male	0.002*** (0.000)
Age	0.000*** (0.000)
Age ²	-0.000*** (0.000)
Education	0.007*** (0.000)
Urban area	0.009*** (0.000)
R ²	0.254
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
Instrumented value of the net in-migration rate	-0.782*** (0.033)
Economic development in 2001	-0.108*** (0.002)
SPI (yearly average between 1997 and 2002)	0.002 (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
Instrumented value of the net in-migration rate	-1.031*** (0.067)	-0.527*** (0.037)
Economic development in 2001	-0.155*** (0.003)	-0.057*** (0.003)
SPI (yearly average between 1997 and 2002)	-0.038*** (0.003)	0.018*** (0.003)
Male	0.225*** (0.002)	0.236*** (0.002)
Age	0.199*** (0.001)	0.208*** (0.001)
Age ²	-0.009*** (0.000)	-0.009*** (0.000)
Education	-0.003 (0.002)	-0.001 (0.002)
Urban area	-0.067 (0.004)	-0.074*** (0.002)
Observations	225,490	305,337

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

Proof of the proposition of the theoretical model

The purpose of this section is to show that

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

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

$$dw = [F_{KL}(K, L) \frac{dK}{dL} + F_{LL}(K, L)] dL,$$

which can be rewritten as follows:

$$\epsilon_{w,L} \equiv \frac{dw}{dL} \frac{L}{w} = \left[\frac{F_{KL}(K, L)}{F_L(K, L)} \frac{dK}{dL} L + \frac{F_{LL}(K, L)L}{F_L(K, L)} \right].$$

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:

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

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: $dw = F_{LL}(K, L)dL$, which can be rewritten as follows: $\epsilon_{w,L}|_{K=cst} \equiv \frac{dw}{dL} \frac{L}{w}|_{K=cst} = \frac{F_{LL}(K, L)L}{F_L(K, L)}$.

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:

$$dK[c' - F_{KK}(K, L)] = dL F_{KL}(K, L),$$

which can be rewritten as follows:

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

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:

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

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

$$Kc' \equiv c\epsilon_{c,K}. \quad (2)$$

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:

$$dc = dkF_{KK}(k, 1),$$

and therefore:

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

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:

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

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

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

Figures

Figure A1: The impact of an influx of migrants on the employment probability of the non-migrants: the case of perfect capital mobility.

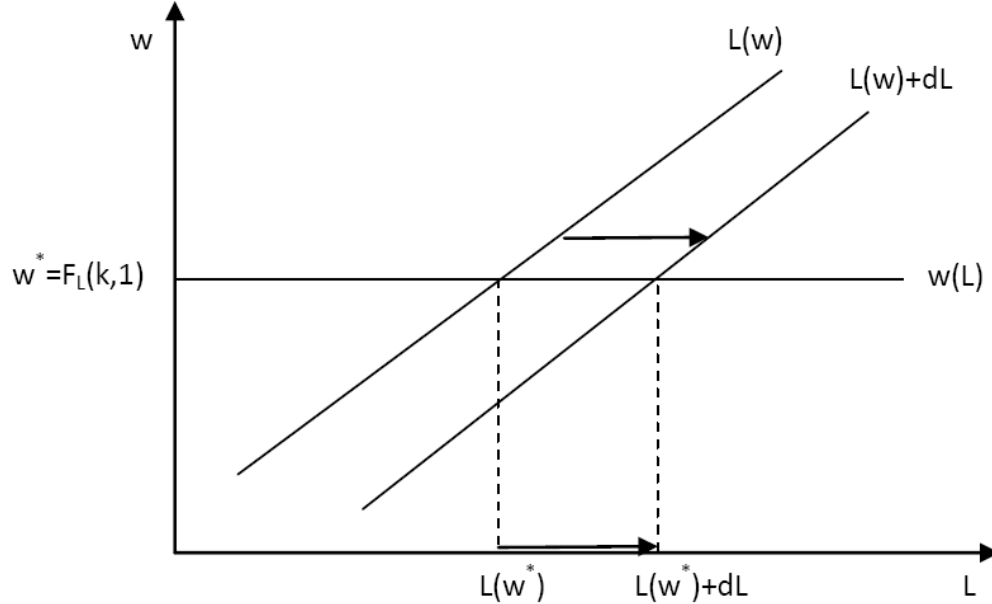


Figure A2: The impact of an influx of migrants on the employment probability of the non-migrants: the case of imperfect capital mobility.

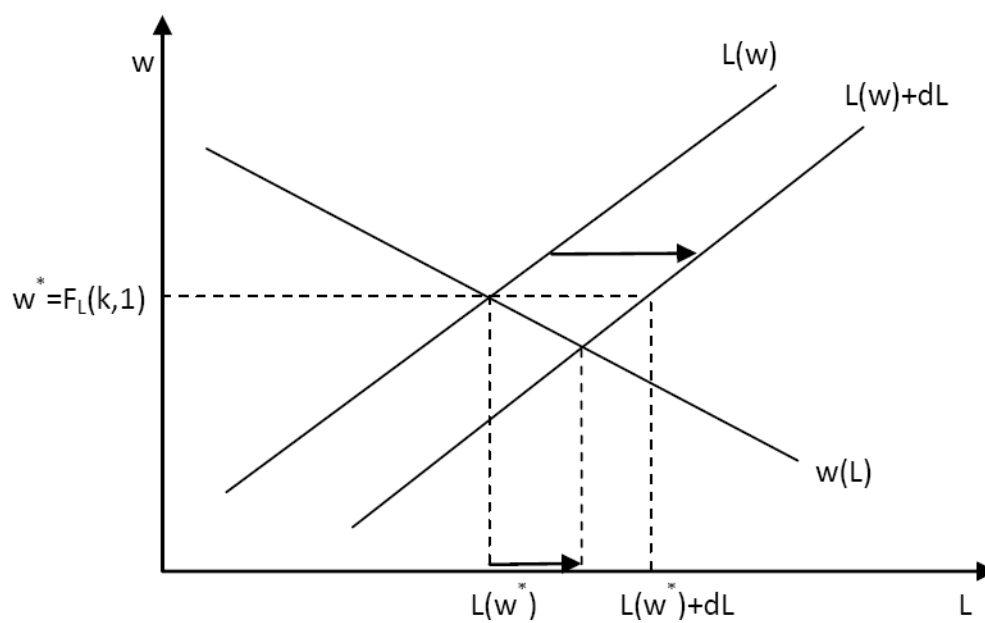


Figure A3: The 56 Ugandan districts (2002).



Figure A4: Road density in Uganda (2002).



Figure A5: Nightlights in Uganda (2001).

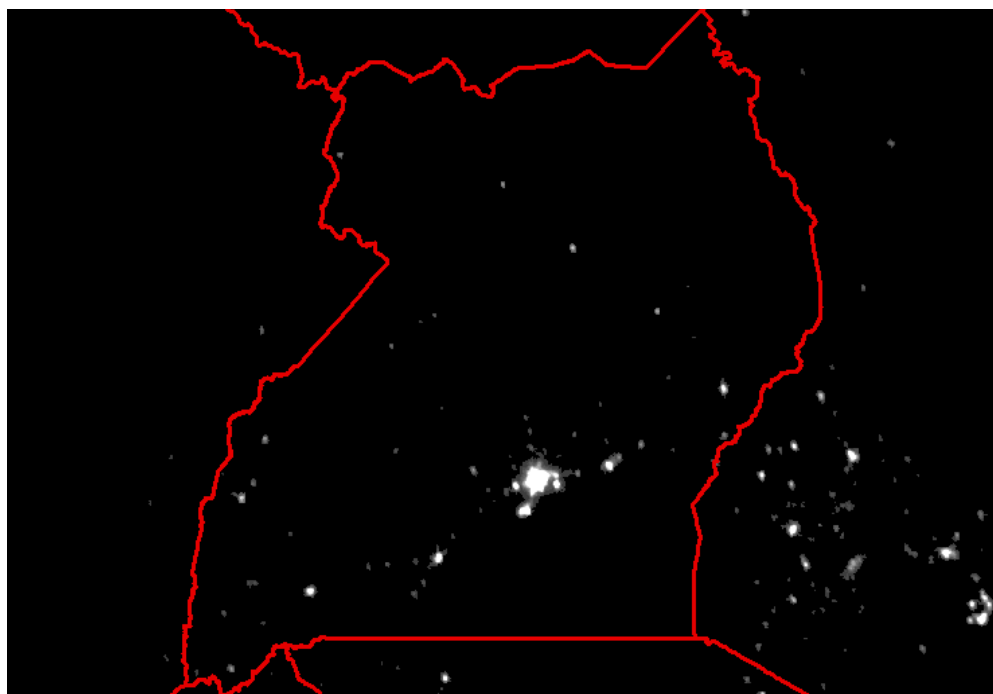
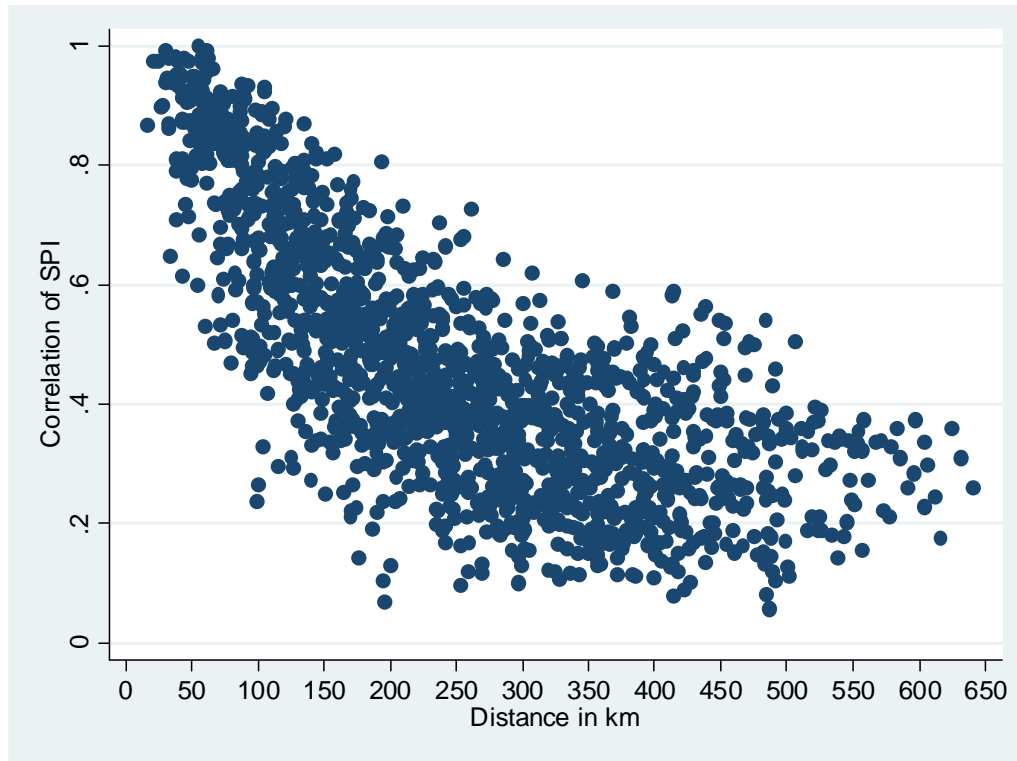


Figure A6: Relationship between the correlation of SPI and geographic distance, for each region pair.



Tables

Table A1: The relationship between the probability to migrate out of a region and the road density in this region. 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 A2: The impact of the net in-migration rate on the employment probability of the non-migrants when one relies on an alternative dependent variable. IV probit analysis.

	Employment probability (alternative definition)		
	All	Below-median road density	Above-median road density
Instrumented value of the net in-migration rate	-0.103*** (0.008)	-0.218*** (0.019)	-0.077*** (0.011)
Economic development in 2001	-0.005*** (0.001)	-0.004*** (0.001)	-0.003** (0.001)
SPI (yearly average between 1997 and 2002)	-0.015*** (0.001)	-0.009*** (0.001)	-0.023*** (0.001)
Male	-0.002*** (0.001)	-0.008*** (0.001)	0.003*** (0.001)
Age	0.007*** (0.000)	0.005*** (0.000)	0.009*** (0.000)
Age ²	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Education	-0.004*** (0.000)	-0.004*** (0.001)	-0.005*** (0.001)
Urban area	-0.018*** (0.001)	-0.012*** (0.001)	-0.020*** (0.001)
Observations	315,991	136,228	179,763

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 A3: The impact of the net in-migration rate on the employment probability of the non-migrants when one controls for economic growth in the destination region between 1997 and 2002. IV probit analysis.

	Employment probability		
	All	Below-median road density	Above-median road density
Instrumented value of the net in-migration rate	-0.745*** (0.033)	-1.003*** (0.062)	-0.423*** (0.041)
Economic development in 2001	-0.143*** (0.002)	-0.154*** (0.003)	-0.113*** (0.007)
SPI (yearly average between 1997 and 2002)	0.022*** (0.002)	-0.039*** (0.003)	0.038*** (0.005)
Economic growth between 1997 and 2002	0.237*** (0.010)	-0.011 (0.015)	0.271*** (0.026)
Male	0.231*** (0.001)	0.225*** (0.002)	0.235*** (0.002)
Age	0.204*** (0.001)	0.199*** (0.001)	0.208*** (0.001)
Age ²	-0.009*** (0.000)	-0.009*** (0.000)	-0.009*** (0.000)
Education	0.000 (0.001)	-0.003 (0.002)	-0.002 (0.002)
Urban area	-0.068*** (0.002)	-0.068*** (0.004)	-0.076*** (0.003)
Observations	530,827	225,490	305,337

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.