Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Climate Variability and Internal Migration: A Test on Indian Inter-State Migration

Ingrid DALLMANN, Katrin MILLOCK

2013.45

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/bandeau-haut/documents-de-travail/
ISSN : 1955-611X

Climate Variability and Internal Migration: A Test on Indian Inter-State Migration *

Ingrid Dallmann†

Katrin Millock‡

May 15, 2013

Abstract

We match migration data from the Indian census with climate data to test the hypothesis of climate variability as a push factor for internal migration. The main contribution of the analysis is to introduce relevant meteorological indicators of climate variability, based on the standardized precipitation index. Gravity-type estimations derived from a utility maximization approach cannot reject the null hypothesis that the frequency of drought acts as a push factor on inter-state migration in India. The effect is significant for both male and female migration rates. Drought duration and magnitude as well as flood events are never statistically significant.

JEL codes: O15, Q54.

Keywords: climate change, India, internal migration, PPML, SPI.

---

*We thank Eric Strobl for providing climate data, including the SPI. We thank also Catherine Bros, Sudeshna Chattopadhyay, Miren Lafourcade and Céline Nauges for their help and advice, as well as participants in the 2nd International Conference on Environment and Natural Resources Management in Developing and Transition Economies (enrmdte), in particular Simone Bertoli. Any errors or omissions are only the authors' responsibility, naturally. Financial support from the French National Research Agency grant ANR-JCJC-0127-01 is gratefully acknowledged.

†University Paris-Sud 11 (ADIS); ingrid.dallmann-gamarra@u-psud.fr

‡Paris School of Economics, CNRS, Centre d'Economie de la Sorbonne; millock@univ-paris1.fr

1 Introduction

Negative effects linked to climate change are more and more apparent, not only through the increase in natural disasters that cause huge economic and human losses but also through its long-term consequences. But, does climate change affect migration? According to a report issued by the UK Government Office for Science (2011) [19], the response is affirmative: environmental change will affect migration in the present and in the future, but the influence will be principally through economic, social and political drivers. Climate variability may have direct effects, causing injury, death, crop damage and disruption of socio-economic activities, but also have indirect effects on the environment and the economy, hence inducing migration either directly or indirectly.

The purpose of this paper is to test the hypothesis that long-term climate variability acts as a push-factor on internal migration. Specifically, we investigate if the frequency, duration and magnitude of drought and flood events have induced inter-state migration flows in India. Since the environmental factor is not the only driver of migration, we control also for the most important social and economic drivers. In order to do so, we match data from the Indian census of 1991 and 2001 with climate data of the Intergovernmental Panel on Climate Change (IPCC). The econometric specification is based on a random utility model. The estimation results show that the frequency of drought events has a significant impact on inter-state migration flows. Each additional month of drought in the origin state during the five years preceding the year of migration increases the bilateral migration rate by 0.9%. The relative effect is rather small compared to the economic drivers of migration. In addition, barriers to inter-state migration have a much more important effect, as reflected in the low Indian inter-state migration rates.

The analysis contributes to a small but growing literature that analyzes the link between migration and climate change. Amongst these studies, the definition of migration (net migration, out-migration, immigration), the choice of the zone of study (rural-urban, local, internal, international, developing and developed countries), the aggregation level, the theoretical model, the indicators of climate change and the empirical methodology vary significantly and, as a result, they are inconclusive and hardly comparable. Macroeconomic studies on international migration flows such as Reuveny and Moore (2009) [36] and Coniglio and Pesce (2011) [13] show that both weather-related natural disasters and climate anomalies may (directly) induce increased migration into OECD countries, as suggested by the theoretical predictions of Marchiori and Schumacher (2011) [26]. Naudé (2008) [34] finds that the number of natural disasters during the 5 years preceding migration has an indirect but significant positive effect on net international migration in sub-Saharan Africa from 1965 to 2005. Beine and Parsons (2012) [6] who include both weather variables such as rainfall and temperature and natural disasters in their analysis find no evidence that climate change would induce an increase in international migration flows. This result is compatible with household level analyses, such as Gray (2009) [20] on data from the Andean Zone of Southern Ecuador. Gray (2009) finds that environmental factors influence local and internal migration, and that negative environmental conditions do not increase international migration necessarily, as predicted in the "environmental refugees" literature (Myers, 1997) [33]. International migration implies high costs that may be prohibitive for the poorest households that are the most vulnerable to environmental conditions.

Some recent studies on environmentally induced internal migration concern a well-known historical episode, the Dust Bowl[1] in the U.S. Great Plains in the 1930's. During this episode, weather and environmental conditions played a significant role in explaining internal migration (Gutmann et al., 2005) [?]. Hornbeck (2012) [23] concludes that the Dust Bowl had an immediate substantial and persistent negative impact on the value and income from farmlands (through soil erosion). The economy adjusted mainly through migration rather than through capital inflows or industrialization. Another study of urbanization, Barrios et al. (2006) [4] finds that rainfall variability had an important impact on urbanization in sub-Saharan Africa, but not in the other developing countries included in their sample.

The only existing studies on India have either analyzed cross-section data (Bhattacharya and Innes, 2008 [10]), focused on the indirect impact of climate on migration through its effect on agricultural yields (Viswanathan and Kumar, 2012 [43]) or used village data representing only some districts or states (Badiani and Safir, 2008 [3]). Bhattacharya and Innes (2008) [10] studied the relationship between population growth and environment in India, but the measures used to proxy environmental deterioration - net vegetation cover - are endogenous and dependent on agricultural production and behaviour of the households, contrary to exogenous measures such as the rainfall or temperature. We thus extend the existing literature on Indian internal migration by introducing new standardized exogenous measures of climate variability into a gravity-type model of internal migration. In doing so we also contribute to the migration literature using gravity-type models (Karemera et al., 2000 [24], Mayda, 2010 [28], Van Lottum and Marks, 2010 [42], and in particular Özden and Sewadeh, 2010 [35] on India).

[1]Dust storm series, considered an ecological catastrophe, that affected the U.S. and Canada great plains region during almost a decade in the 1930's.

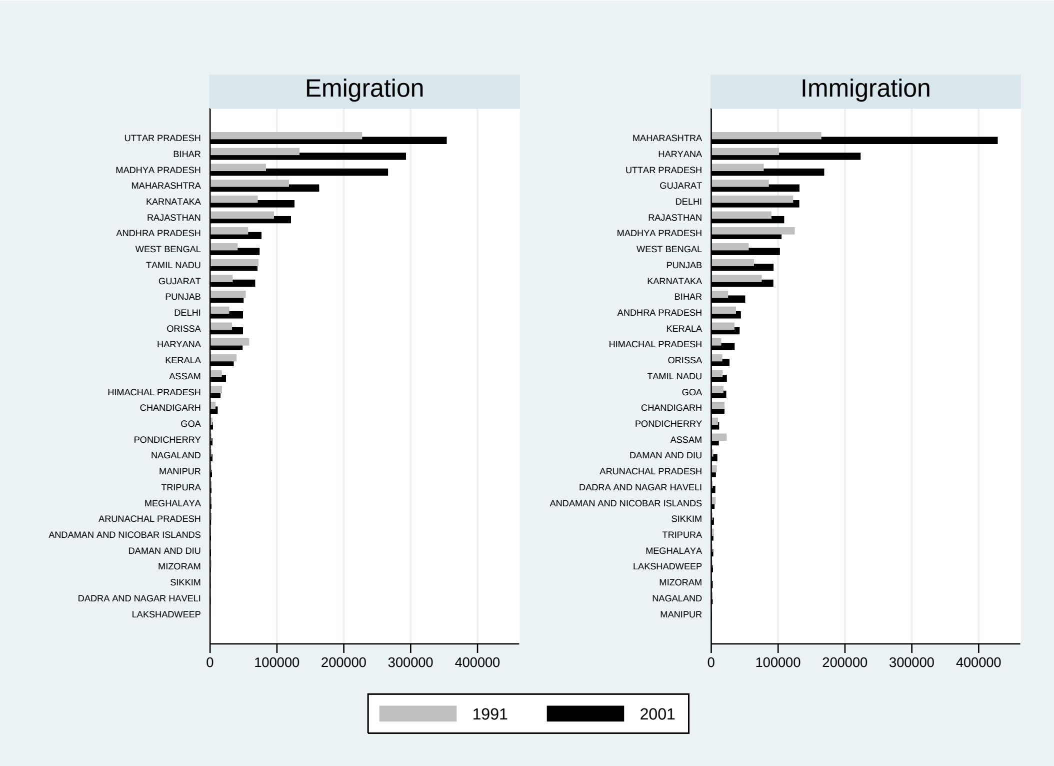
The relationship between climate change and migration is complex and many questions arise: who will be affected by climate change induced migration? Where and in which geographic space is this migration likely to occur? Will the migration be permanent or temporal? Which climatic conditions are the more influential? Through which channels will migration occur? And what will be its political implications? More detailed empirical and theoretical studies may clarify some of these questions. In this paper, we focus on internal (inter-state) migration in India, since climate change induced migration is more likely to occur within the internal borders of a country, because of migration costs, including legal barriers (Marchiori et al., 2012 [27], Beine and Parsons, 2012 [6]). In addition, low-income and lower-middle-income countries are also more vulnerable to climate change than high-income countries (Stern, 2007 [40]; Government Office for Science, 2011 [19]) due to their lower climate change adaptation capacity and their geographical location.

2 Inter-state migration and climate variability in India

Analyzing inter-state migration in India is particularly appropriate for a study of internal migration because of the size of the Indian states (the equivalent of European states), and the heterogeneity among them, especially as regards demography and climate. India has a large variety of climate regions, ranging from tropical in the South to temperate and alpine in the Himalayan North. This variation is maybe greater than any other area of similar size in the world. Nearly 75% of the annual rainfall is received during the monsoon season (June to September). The main natural disasters in India are drought, flood and tropical cyclones (Attri and Tyagi, 2010) [2]. India is also considered by the Environmental Vulnerability Index as extremely vulnerable, not only because of its climate vulnerability, but also because of its population density. In fact, India is after China the second most populated country in the world (1,210 million inhabitants in 2011 that represents 17.5% of the world population with only 2.4% of the world surface area), with a population growth between 2001 and 2011 of 17.6%, which exceeds the world population growth (12.9%) [12]. Its population is mainly rural, of 72.2 % in 2001 (this represents 742.5 million people).[2] Even if its rural population has dropped in percentage since 1901 and with an accelerating rate from the 1970's and onwards, India remains a country with a low urbanisation level (Datta, 2006) [14]. Besides, the population densities contrast very much be-

[2]Indian Census: www.censusindia.gov.in

Figure 1: India interstate out-migration and in-migration by state, 1991 and 2001



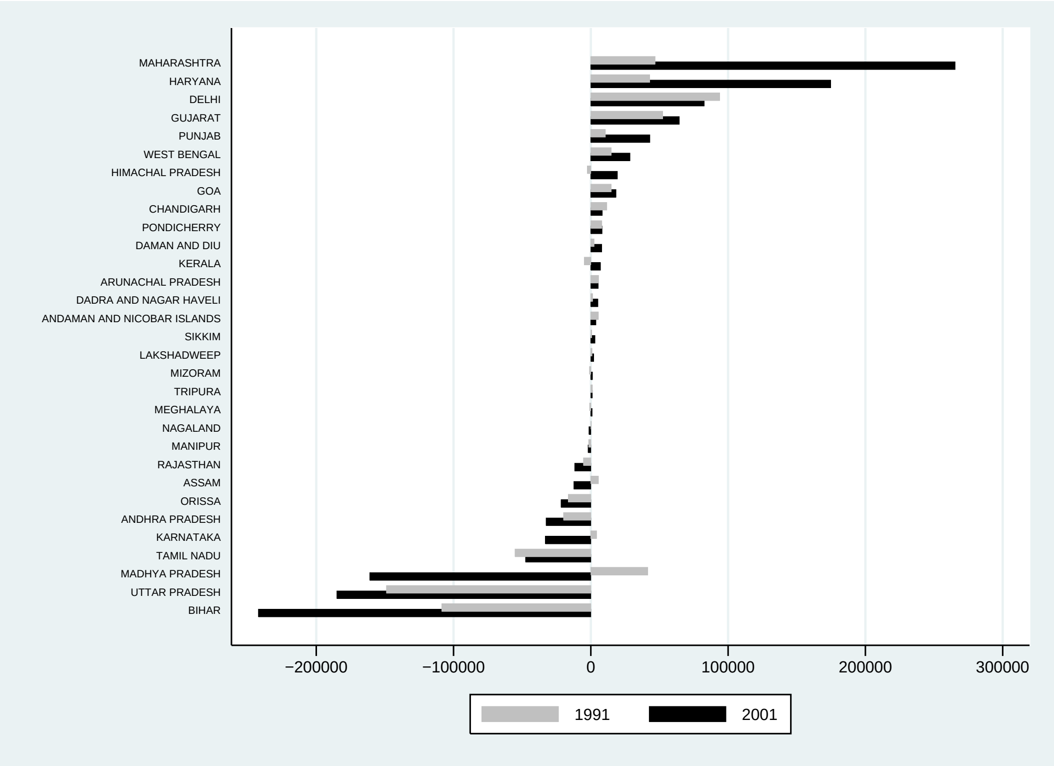
The definition of migrants is that of individuals declaring the last place of residence in $t-1$ to be different from the place of enumeration in the Census.

tween states, for instance it ranges from 17 to 11,297 people per square km in 2011 (Arunachal Pradesh and Delhi respectively). In 1991 26.7 % of the total population was an internal migrant, in 2001 this proportion increased to 30.1% (310 million persons) with 11.8% and 13.4% (41.6 million persons) of the migrants being inter-state migrants. These statistics motivate the interest in better understanding the migration pattern and the potential influence of climate change as a determinant of migration.

We use the definition of migrants as individuals declaring the last place of residence in $t-1$ to be different from the place of enumeration in the years 1991 and 2001. Figure 1 thus shows the number of emigrants and immigrants by their origin and destination states according to this definition.[3]

[3] These figures are thus much lower than the total number of migrants, which includes also durations of stay of 1-4 years or even 5-9 years. We focus on the duration of one year or below in order to match the data more precisely in time with the available socio-

FIGURE 2: India net interstate migration by state, 1991 and 2001



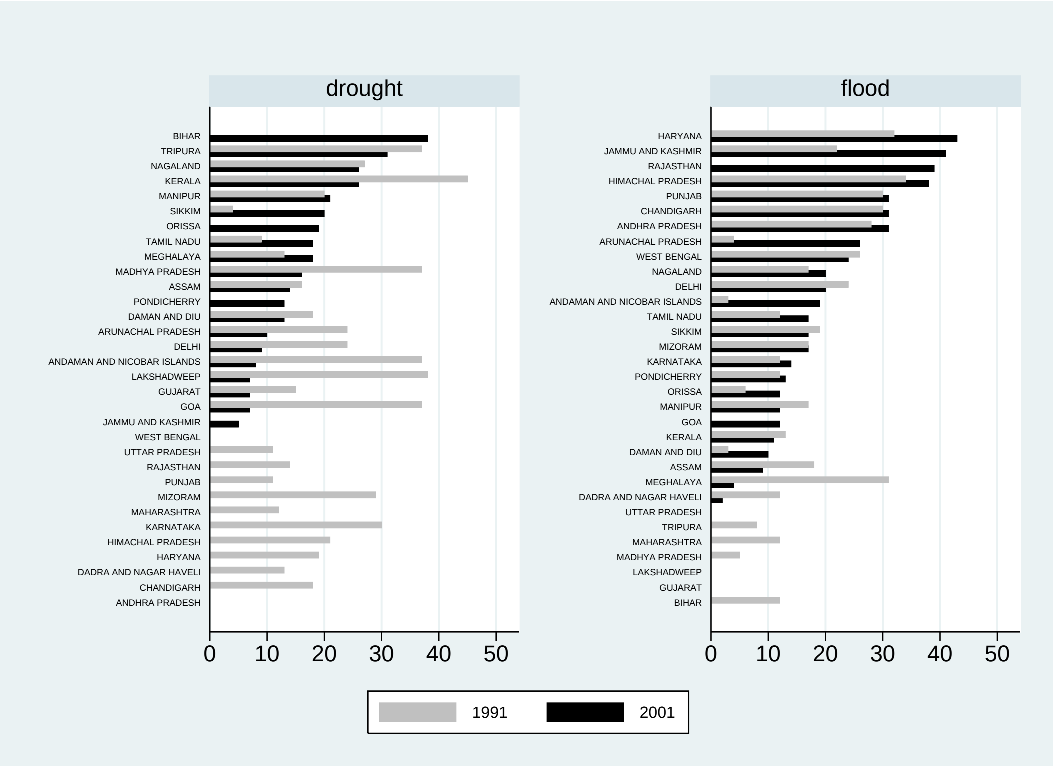
The definition of migrants is that of individuals declaring the last place of residence in $t-1$ to be different from the place of enumeration in the Census.

Figure 1 confirms the description in Özden and Sewadeh (2010) [35] of the major migration corridors based on the National Sample Survey data from 1999-2000. The states with the highest numbers of out-migrants are Uttar Pradesh, Bihar, Maharashtra and Madhya Pradesh, with Madhya Pradesh overtaking Maharashtra in 2001 in absolute number of migrants with duration of residence of one year or less. Incidentally, Maharashtra is also the state with the largest inter-state in-migration in absolute numbers, resulting in a positive net migration, compared to the other states with large gross out-migration flows (Figure 2).

Our objective is to test whether drought or flood events measured on a normalized scale over the long run have influenced the gross out-migration flows. Figure 3 illustrates the data that we use in the analysis. The measure is the number of months with one standard deviation or more of either low

economic data, such as net state product per capita, and climate data. If we include other durations of stay, we could only analyze average figures over a longer time period.

Figure 3: Drought and flood frequency by state, 1991 and 2001



The definition of drought/flood frequency is the number of months with the standardized precipitation index (SPI) at least one standard deviation below/above its long run mean.

rainfall ("drought") or excess rainfall ("flood") in the five years preceding the census in either 1991 and in 2001 (see Section 4.3 for a detailed description of the data and its calculation). The first thing to note is that the months with drought events by state varied much between 1991 and 2001, whereas there is less variation over time for the number of months with flood events by state. Overall, several of the states record no drought or flood events at all in the five years preceding 2001 when using the rainfall measures standardized with respect to the long term mean (1901-2001). The states with a high number of drought events in the five years preceding 1991 were Kerala and Madhya Pradesh, in addition to several small states and island states, and Bihar, Tripura and Nagaland in 2001.[4] Madhya Pradesh and Bihar are also important out-migration states. The states with the highest number of

[4]The analysis will account for the differences in population by using the migration rates defined as bilateral migrants over the number of individuals who stayed in the state over the same time period.

months with flood events were Himachal Pradesh, Haryana, Meghalaya, Punjab, Chandigarh and Andhra Pradesh in the five years preceding 1991, and Haryana, Jammu and Kashmir, Rajasthan, Himachal Pradesh and Punjab in the years preceding 2001.

3 Empirical specification and method

3.1 Theoretical framework and econometric specification

We base the econometric specification on the random utility model used recently by Beine et al. (2011) [7], amongst others, and in particular by Beine and Parsons (2012) [6] for analyzing climate change and international migration. People choose to stay in their residence place or to migrate to one state among all possible destinations by maximizing their utility. The utility of staying in the residence place is assumed linear in the log of income and the residence state characteristics. The utility of moving depends on the log of the income in the potential destination state, the potential destination state characteristics and the cost of migration. Assuming that the error term follows an iid extreme value distribution, and taking logs of the utility differential between migrating to state j or staying in state i results in the following gravity-type specification:

$$\ln \frac{m_{ij,t}}{pop_{ii,t}} = \ln \frac{w_{j,t}}{w_{i,t}} + S_{j,t} - S_{i,t} - C_{ij,t} \tag{1}$$

where $m_{ij,t}$ is the bilateral migration flow from state i to state j and $pop_{ii,t}$ is the population initially located in state i and staying in state i. The income differential between states is represented by the relation between $w_{j,t}$ and $w_{i,t}$, the per capita income of the destination and the origin states. $S_{j,t}$ represents time-varying destination state characteristics, like employment and education possibilities. The origin state characteristics $S_{i,t}$ include origin state characteristics that vary little over time, such as amenities, geographic vulnerability and irrigation infrastructure, as well as time varying characteristics like climate, education or safety net programs. $C_{ij,t}$ is the migration cost, that includes monetary costs (that may vary with the distance between origin and destination states) and psychological costs (from moving to a state that does not share the same culture and traditions).

In our specification the income ratio is proxied by the ratio of the Net State Domestic Product per capita in the destination state compared to the origin state. Recent work has established evidence that temperature and rainfall affect income growth, although not always absolute levels of income (Dell et al., 2009 [16]; Barrios et al., 2010 [5]). Here we use the income ratio, which is less correlated with climate variability in the origin state. Rather than studying the indirect effect of climate variability working through income we aim at testing if there is a direct effect on migration from the direct utility-decreasing effects of climate variability. As shown in the correlation matrix (Table 9 in Appendix B) our main climate variable - frequency of droughts - has less than a 10 % correlation with the income ratio. If there was concern that the correlation was larger, it would indeed be difficult to identify a separate direct effect of climate variability on bilateral migration flows.

The cost of migration is represented by distance, and a common border or language between states. We also control for caste (or ethnic) similarity between states by controlling for the ratios of scheduled castes and scheduled tribes in the destination state compared to the origin state. The principal variables of interest are the ones representing climate variability. Our hypothesis is that adverse weather events act as a push factor on migration. In particular, this is the case in developing countries where poor people do not move by comparing origin and destination climate conditions but rather escape from adverse climate events that affect their well-being. Accordingly, all our variables representing climate variability act only in the origin state.[5] We include origin state fixed effects (D_i) that are invariable in time to capture the vulnerability of the geographic zone, especially mountains, low elevation coasts and arid lands, but also to catch the effect of long-term climate change adaptation strategies adopted by the state, such as irrigation infrastructure. This dummy controls also for the states affected by the Armed Forces (Special Powers) Act of 1958. The Act gives special power to armed forces (military and air forces) in the so called "disturbed" areas. The states and Union Territories affected are: Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland and Tripura. These states have experienced violence that may have induced migration.

Destination state and time fixed effects (D_{jt}) capture characteristics varying in time like employment and education potentials in the destination state.

The resulting econometric specification is thus the following:

[5] Lewer and Van den Berg (2008) [25] and Bodvarsson and Van den Berg (2009) [11] discuss the potential sources of bias in the gravity model. One of them is that the presence of unilateral variables, like the climate variables in this specification, can result in standard error clustering. They cite Feenstra (2004) [18], who argues that adding fixed effects dummies eliminates this bias.

$$
\begin{aligned}
\ln \frac{m_{ij,t}}{pop_{ii,t}} =& a_0 + a_1 \ln \frac{w_{j,t-1}}{w_{i,t-1}} \\
&+ a_2 \ln \frac{SC_{jt}+1}{SC_{it}+1} + a_3 \ln \frac{ST_{jt}+1}{ST_{it}+1} \\
&+ a_4 \ln dist_{ij} + a_5 border_{ij} + a_6 language_{ij} \\
&+ a_7 \sum_{t-5}^{t} clim_{i,t} + D_i + D_{jt} + u_{ijt}
\end{aligned} \tag{2}
$$

where

State i:	Origin state.
State j:	Destination state.
m_{ijt}:	Migration flow from state i to state j during year $t-1$ to t.
$pop_{ii,t}$	Population of state i staying in the state during year $t-1$ to t.
$\frac{w_{j,t-1}}{w_{i,t-1}}$:	Ratio of the Net State Domestic Product per capita in state j and in state i at time $t-1$.
SC_{it}, SC_{jt}:	Scheduled caste rate in state i/j at time t.
ST_{it}, ST_{jt}:	Scheduled tribe rate in state i/j at time t.
$dist_{ij}$:	Distance from state i to state j.
$border_{ij}$:	Dummy variable for common border between state i and j.
$language_{ij}$:	Dummy variable for common language between state i and j.
$clim_{i,t}$:	Frequency, duration or magnitude of drought/flood in state i, during the five years preceding t.
D_i:	Time-invariant fixed effect for state i.
D_{jt}:	Destination-time fixed effect for state j.

The expected signs are: $a_1>0$, $a_4<0$, $a_5>0$ and $a_6>0$. All else equal, the larger the differential of the per capita income between states, the larger the incentive to migrate.The relation of migration with distance is negative, since it proxies migration travel costs. Common border and language are viewed as facilitators for migration (or factors that reduce the cost of migrating).

We include an additional cost factor for migration in the form of differences in scheduled caste and tribe ratios in the destination state compared to the origin state, to account for similarity between states. Scheduled castes and tribes may be the most vulnerable parts of the population to climate variability given that they often are day labourers and hence likely to be the first affected by climate events. These variables capture network effects in the sense that, for an individual belonging to the scheduled castes (or tribes) population, moving to a state with a higher ratio of scheduled castes (or tribes) compared to the origin state would imply lower costs of migration because of the network in the destination state, whereas moving to a state with a lower ratio of scheduled castes (or tribes) would imply higher costs of migration because of the smaller network. Ex ante, the coefficients a_2 and a_3 could thus be either positive or negative.

For the variables representing climate variability, we expect a positive sign (a_7>0) for the different measures of drought and flood. More drought or flood events (in quantity, duration and magnitude) are likely to increase migration.

3.2 Estimation method

The specification (2) is based on a semi log form. This represents a problem for those state pairs where the migration flows equal zero, since dropping such observations from the data set may generate selection bias. On the Indian sample such state pairs represent 10% of the total number of observations. One method to avoid sample selection problems by excluding the observations with migration equal to zero, is to add one to each bilateral migration rate observation. Nevertheless, the problem remains that the log-linear specification will cause OLS estimation of the elasticities to be inconsistent in the presence of heteroskedasticity[6] (Santos Silva and Tenreyro, 2006) [38]. Instead Santos Silva and Tenreyro (2006) [38] demonstrate that a Poisson Pseudo Maximum Likelihood estimator (PPML) with robust standard errors produces consistent estimates in a non-linear model. The assumption of equality between the standard deviation and the mean of the dependent variable that is characteristic of the standard Poisson maximum likelihood estimator (Poisson MLE) is no longer necessary in the PPML method. We thus follow these authors and recent applications on migration (Beine and Parsons, 2012) [6]) and report our results with the PPML estimator.

Another potential econometric problem has been labelled multilateral resistance in the application of gravity models (Anderson, 2011 [1]). This

[6]The Breusch-Pagan/Cook-Weisberg test on heteroskedasticity in an OLS regression on the data leads to a test statistic of 133.45 and a p-value of 0. So we can conclude that the null hypothesis of homoskedasticity is rejected.

means that the migration decision takes into account not only the comparison between the origin and the destination state characteristics, but also the opportunities in all the alternative destinations. By assuming an extreme value distribution of the error term, we have assumed away this possible problem, but this assumption needs to be tested, and if necessary, corrected. Mayda (2010) [28], for example, includes opportunities of other countries in her migration gravity model by adding a "multilateral pull" variable, which is the average of the log ratio between per worker GDP and distance of all other destination possibilities. Bertoli and Fernández-Huertas Moraga (2013) [8] address multilateral resistance with a more general method, a common correlated effects estimator, but this method is not possible in our case because of the short period of the data. If the specification presented here is correct, the choice of one state as destination should not be affected by the presence or not of other states (according to the assumption on the Independence of Irrelevant Alternatives). We thus re-estimate the econometric model, removing one state at a time, and compare the main parameter estimates in these estimations with the parameter in the estimations including all the states (following Grogger and Hanson, 2008 [21]).

4 Data and variable definitions

4.1 Area and period studied

We use bilateral inter-state migration data from the Indian census of 1991 and 2001. Between 1991 and 2001, India changed the territorial administrative division of its states. In 1991, India counted 27 states and 5 Union Territories. In 2001, 3 states were divided in two[7], resulting in a total of 30 states and 5 Union Territories. To unify the database, we use the territorial administrative division of 1991. Hence, for 2001, we aggregate the data of the divided states as they were defined in 1991. We analyse the Union Territories as states. Since we do not have data from 1991 on the state of Jammu and Kashmir[8], we removed this state from the sample. Jammu and Kashmir represent only 1% of the Indian population. The final sample thus counts 31 states for 1991 and 2001. As the analysis of migration is made in a bilateral manner, we have 930 observations (31x30, migration between the same states being 0) for each year.

[7]Uttar Pradesh, Bihar and Madhya Pradesh, that have given rise to the states Uttaranchal, Jharkhand and Chhattisgarh respectively.

[8]The census was not conducted in the state of Jammu and Kashmir in 1991.

4.2 Dependent variable: Bilateral migration rate

Several studies use net migration when data are not available on in- or out-migration. This is the case especially in studies of international migration, but at the level of countries, the census is a rich source of information for the analysis of local or internal migration. We thus use the bilateral gross migration rates between states, rather than net migration, to not lose information unnecessarily.

According to the Indian Census, inter-state migration occurs "if the place of enumeration of an individual differs from the place of birth or last residence and these lie in two different States, the person is treated accordingly as an inter-State migrant with regard to birth place or last residence concept" and a migrant is defined as “a person who has moved from one politically defined area to another similar area. ... Thus a person who moves out from one village or town to another village or town is termed as a migrant provided his/her movement is not of purely temporary nature on account of casual leave, visits, tours, etc.” It is thus a definition based on intent of staying rather than on a minimum duration of stay. We use data on migration flows from the census of India of 1991 and 2001.[9] Migration flows are identified by the current place of residence (destination state), by the place of residence of provenance (origin state) and with different duration of stay (1 year, 1-4 years, 5-9 years).

Our dependent variable is the gross migration flow m_{ijt} from state i to state j between time $t-1$ and time t, divided by the population that did not move in the same period, and multiplied by 100,000 for scaling purposes.

4.3 Climate variables: The Standardized Precipitation Index (SPI)

To test the hypothesis of climate variability acting as a push factor for internal migration, we compute normalized measures of scarcity of water ("droughts") and excess water ("floods").

Rainfall is the main factor of vulnerability to water availability. The scarcity of water had negative consequences on food availability and human health historically, and caused diseases and displacement of populations (Barrios et al., 2006) [4]. The consequences in urban areas can be the difficulty to cover the requirements in drinking water in quantity as well as in quality. In rural areas, the principal problem is that the output and quality of the

[9]The population census in India is taken every ten years, but we only had access to computerized data from 1991 onwards. Data from 2011 on inter-state migration flows are not yet available.

crops are affected. The fact that these data are accessible and reliable over a long period further motivates their use as a measure of climate variability.

We compute climate variability measures based on the IPCC rainfall data. The IPCC data was constructed by assimilating the observations from meteorological stations across the world in 0.5 degrees latitude by 0.5 degrees longitude grids covering the land surface of the earth. Each grid was then allocated to a single country (for more details see Mitchell et al., 2002) [30]. For India, we have data by district and by month from 1901 to 2006.

From the rainfall data, we calculate the Standardized Precipitation Index (SPI), developed by McKee et al. (1993) [29] with the objective to define and to capture the length of a drought episode. By using the SPI we can determine a drought or a flood (excess of wetness) event for a period in a given place. Conceptually, the SPI represents a z-score or the number of standard deviations above or below that an event is from the mean, for which the mean and the standard deviation are calculated over past periods (here 1901 to 2001). It is used as a standardized measure of drought and is constructed as a deviation from a precipitation gamma distribution within a defined scale (here 12 months). Its values are between -3 and 3 and a (moderate) drought begins when the SPI has a value of -1 (rain falls one standard deviation below its historical mean) and goes on in time until the SPI becomes positive again. In that way, we know the beginning and end date and can calculate the length of a given drought episode. We also know the intensity of the drought according to the value of the SPI. An excess of wetness can be measured following the same logic. It begins with a value of +1 (rainfall increases by one standard deviation above its historical mean) and continues until the SPI becomes negative. Table 1 illustrates the definition of intensity of a drought or a flood with this method.[10]

The main advantages of this measure is that it takes into account the space and temporal deviation and that it gives us a measure of the start, length and intensity of drought, rather than only the absolute value of the temperature or rainfall. Additionally, it allows us to have a measure with a fixed mean and variance, which makes the SPI of different meteorological stations comparable.[11]

The raw data are on a district level and to aggregate the data on a state level, we calculate the average of the SPI in every state (a principal

[10]For more details on the SPI, see McKee et al. (1993) [29]

[11]Indeed, the Lincoln Declaration on Drought Indices (11 December 2009, Lincoln, USA) recommended that *The National Meteorological and Hydrological Services (NMHSs) around the world use the SPI to characterize meteorological droughts and provide this information on their websites, in addition to the indices currently in use. WMO was requested to take the necessary steps to implement this recommendation.*

Table 1: *Definitions of drought and flood according to the SPI*

SPI values	**Category**
0 to -0.99	Mild drought
-1 to -1.49	Moderate drought
-1.5 to -1.99	Severe drought
$<=$ -2	Extreme drought
0 to 0.99	Mild flood
1 to 1.49	Moderate flood
1.5 to 1.99	Severe flood
$>=2$	Extreme flood

Source: McKee et al. (1993) [29] for drought and Guerreiro et al. (2008) [22] for flood.

component analysis is presented in Appendix A as a test of this procedure). We create five variables based on the SPI to measure the frequency, the duration and the magnitude of a drought or a flood:

1. *Frequency*: First, we define a binary variable by state which takes the value of 1 if there was a drought/flood event in a month in that state, and 0 otherwise. The final measure is the number of months with drought/flood in the origin state during the five years preceding migration.[12] The measures count total months of either severe or moderate drought/flood. Extreme events are not common on the state level data. Aggregation at a state level takes out any extreme events at a finer district-level and may lead to less precise results.[13]

2. *Maximal duration*: In the aim to catch the impact of a long drought or flood duration, we compute the maximal number of months that a drought/flood lasted in the five years preceding migration.

3. *Magnitude*: This variable is defined as the sum of the absolute values of the SPI for a drought or a flood five years preceding migration.

4. *Average monthly magnitude*: The *magnitude* divided by the *frequency.*

[12]Barrios et al. (2006) [4], Naudé (2008) [34] and Strobl and Valfort (2012) [41] also use a lag of five years for the impact of natural disasters and climate variables.

[13]We would like to control for exogenous natural disasters other than climate-driven ones (such as cyclones and other natural disasters), but the data we studied from EM-DAT, collected by the Centre for Research on the Epidemiology of Disasters (CRED), did not seem reliable at the state level (as compared to the country level).

5. *Longest drought/flood magnitude*: The sum of the absolute values of the SPI of the longest drought or flood in the five years preceding migration.

Our measures are strictly exogenous and not influenced by economic activity, in contrast to other environmental variables like soil degradation or air pollution.

4.4 Net State Domestic Product (NSDP)

The NSDP per capita is used as a measure of the income per capita of the state. We use the database of the Reserve Bank of India and calculate the deflated NSDP at constant price for the two years of interest (1990 and 2000).

The variable used is the ratio of the NSDP per capita of the destination state divided by that of the origin state, in the year preceding migration $(t-1)$, in order to reduce any endogeneity with the migration flows.

4.5 Distance between states

The distance between states (or countries in international studies) is commonly used as a measure of migration costs, notably in those based upon the gravity model. We calculate the distance between different states, by taking the most populated city as reference city, most often the capital of the state, but in some cases the economic center of the state, according to the great circle formula.[14]

$$d_{ij} = R * cos^{-1}(sin(a)sin(b) + cos(a)cos(b)cos(c-d)) \tag{3}$$

where

d_{ij}:	distance between state i and state j
R:	equatorial radius, equal to 6,378 km
a:	latitude degree of state i
b:	latitude degree of state j
c:	longitude degree of state i
d:	longitude degree of state j

As explanatory variable we use the distance between two states, measured in km.

[14]The latitudes and longitudes of the largest cities in every state can be found on the website "Maps of India". See www.mapsofindia.com

4.6 Common border and common language

We introduce a dummy variable to control for neighboring states. It takes the value of one for bilateral migration where the origin and destination states have a common border, and zero otherwise.

One of the specificities of India is that there are 18 different native languages (English excluded) inside the country. As another proxy of the cost of migration, we introduce a language dummy variable. It takes the value of one for bilateral migration where the origin and destination states share a common language, and zero otherwise. To assign a language to a state, we took the major language spoken in the state. The source of this variable is "Maps of India".

These two variables are proxies for cultural and traditional similarities between states.

4.7 Scheduled castes (SC) and scheduled tribes (ST)

In India, 16.2 % of the population belong to a scheduled caste (also called "the untouchables") and 8.2% to scheduled tribes in 2001. In the literature of Indian migration, these two factors are almost always taken into account to examine the role of social factors in the migration decision.[15] Indeed, they play an important role in Indian social structure. The "Hindu Varna" System, who establishes the classification of the society in India, excludes, categorizes and isolates groups of population, on the basis of the caste, the ethnicity and the religion. This discrimination persists in the labour force participation (Dubey et al., 2006) [17]. In this social stratification, the SC and ST are the most discriminated against and were the object of policies of positive discrimination. The ST are isolated, partly because of their geographic locations, often in hills and woods with weak density of population. But unlike the SC, they had limitless access to the natural resources of land, water and forests where they live (Dubey et al., 2006) [17]. If these groups of individuals experience discrimination from upper castes and dominant groups, we may hypothesize that they would like to stay within their communities and be more likely to migrate (if they do so) where they can find their pairs. Indeed, Bhattacharya (2002) [9] find that scheduled castes are less likely to migrate (from rural to urban areas) but if they do so, they go where they can find other scheduled caste population. This suggests a social network effect.

We include the ratios of the scheduled caste and tribe rates in the destination state compared to the origin state as another control for social similarities between states.

[15] See for example Bhattacharya (2002) [9] and Mitra and Murayama (2008) [31].

4.8 Descriptive statistics

Table 2 presents the mean, the standard deviation and the minimum and maximum of each variable. The total number of observations is 1860, representing bilateral migration flows across 31 Indian states in two years (1991 and 2001).

The average of 8 migrants per 100,000 individuals may seem very small, but the variable measures the bilateral rate for a unique origin-destination pair in one year. For example, 8 per 100,000 individuals migrate from Assam to West Bengal between 1990 and 1991, which represents a total of almost 1800 individuals[16]. We have 930 possible combinations like this and we can analyze an accumulated migration in a longer period than one year. It is also important to note that the dispersion is very large (the standard deviation is almost 4 times the mean) and that the bilateral migration rate can take values from 0 and up to 455 migrants per 100,000 individuals.

The average number of months (at any time) with a drought or flood event is almost 15 months (out of a total of 5*12 months), but the descriptive statistics show large variation in the variable, as indeed for all climate variability measures tested here. The longest duration of a drought over the period studied was on average 12 months, just as for a flood episode. Over the time period studied the average drought and flood were of moderate size - an absolute value of the SPI of 0.81 for droughts and 0.83 for floods - but higher for droughts than for floods in the sum of the absolute values of the SPI (16.42 compared to 15.15).

5 Results

In Table 3 we present six regressions with the PPML estimator. The six regressions include origin state fixed effects and destination-time fixed effects. Regression (1) is without the climate variability measures and in the regressions (2)-(6) the variables corresponding to drought events are included one at a time. We introduce the five types of variables (drought frequency, longest drought duration, drought magnitude, average drought magnitude per month, magnitude of the longest drought) separately because of the high correlation between them (see Table 9).

In all six regressions, the results show that the economic motivations, proxied by the ratio of the net state domestic product per capita between the destination and the origin state are important, together with the variables

[16]There are 22,408,756 individuals that did not move in 1990 from West Bengal.

Table 2: *Summary statistics*

Variable	Unit	Mean	Std. Dev.	Min.	Max.
bilateral mig rate	x100,000	7.97	28.23	0.00	455.30
NSDP ratio	-	1.27	0.95	0.13	7.45
distance	km	1,368	672	33	2,846
border	1/0	0.12	0.32	0	1
language	1/0	0.10	0.30	0	1
SC rate	#/capita	0.11	0.08	0.00	0.29
ST rate	#/capita	0.23	0.31	0.00	0.95
drought frequency	# of months	14.52	12.47	0.00	45.00
flood frequency	# of months	14.66	11.70	0.00	43.00
drought duration	# of months	11.85	9.94	0.00	37.00
flood duration	# of months	12.05	10.27	0.00	43.00
drought magnitude	SPI	16.42	14.91	0.00	55.76
flood magnitude	SPI	15.15	12.94	0.00	55.83
drought avg. magnitude	SPI	0.81	0.57	0.00	1.88
flood avg. magnitude	SPI	0.83	0.47	0.00	1.80
longest drought magnitude	SPI	13.90	12.43	0.00	42.52
longest flood magnitude	SPI	12.41	11.70	0.00	55.83

representing the cost of migration. An increase of 1% in the per capita income ratio between the destination state and the origin state increases the bilateral migration rate by about 0.6 to 0.9%.

Bilateral migration rates between contiguous states are 2.4 times larger than for states that do not share a common border. States that share a common language have 50% larger bilateral migration rates.[17] Geographical distance is also statistically significant with a 1% larger distance decreasing the bilateral migration rate by 0.7%. The differences in scheduled caste and scheduled tribe rates between the destination and the origin state are not significant. Maybe the origin state fixed effects catch part of their significance, because these factors vary little over time.

Among the five drought measures tested, the role of push factor for migration is rejected but for the frequency of drought events (regression (2)). An additional month of drought during the five years preceding migration would increase the bilateral migration rate by 0.9 % at a 10% level of significance.

None of the flood variables are statistically significant (results presented in Table 4). All the other variables in the estimations with the flood variables are robust with respect to the size and significance of the coefficients.

[17]The marginal effects for dummy variables are calculated as $(e^{b_i} - 1)$ where b_i is the estimated coefficient of the variable.

It thus seems that drought episodes are more relevant as push variables related to climate variability for inter-state migration in India, compared to flood episodes. The four states with the highest out-migration in the years studied are Uttar Pradesh, Bihar, Madhya Pradesh and Maharashtra. These states all had less than 12 months of moderate flood episodes in the five years preceding the 1991 census and none in the five years preceding the 2001 census. By comparison, they all had experienced drought episodes, in particular the major out-migration states Bihar and Madhya Pradesh. Maharashtra and Uttar Pradesh had relatively low numbers of months with a moderate drought, but these states are also characterized by high interstate in-migration flows that in the case of Maharashtra compensate for the out-migration and results in net in-migration. The results regarding episodes with excess water are thus not surprising given the climate variability in the period studied.

To further test the relationship between climate variability and internal migration in India, we did separate estimations on male and female migration rates (presented in Tables 5 and 6). The Indian census incorporates a question on the reason for migration, with the possible answers being work/employment, business, education, marriage, moved after birth, moved with household and other. Marriage was cited as the predominant reason for migration among women (64% of women) and work for men (38% of men). The estimations show that economic considerations, as proxied by the relative wage ratio between the origin and the destination state, are significant only for male migration. All the other significant explanatory variables are of about the same size as in the estimations on the total migration rates. Drought frequency positively affects the bilateral migration rates for both men and women and the magnitude is slightly larger than that estimated on the total sample, implying that one month of additional drought increases the bilateral male or female migration rate by 1 %. These results may be interpreted as evidence that migration of women, even if the primary *stated* reason is marriage, forms part of a larger risk-coping strategy of the household in line with the early evidence in Rosenzweig and Stark, 1989 [37], who found that rural households used marriage of daughters as an insurance strategy to handle spatially covariant risk.

We also tested a number of additional potential explanatory variables. One of the most important variables for Indian migration may be poverty rates or inequality. The difficulty with such data is to obtain a perfect match between those variables and the years of migration (1990-91 and 2000-01). Several measures of head count ratios were tested, for example, but they are never significant in the migration rate estimations. In Table 7 in Appendix B we present the results controlling for inequality, as measured by the Gini coefficient for rural areas of the origin state.[18] The sign of the estimated coefficient is positive but never significant. The effect of a 1% increase in the relative income ratio still varies between a 0.6 to 0.9% increase in the bilateral migration rate (although it is no longer significant in regression (2)). The impact of drought frequency is also robust.

As a final robustness test, we re-estimate the base specification (in Table 3) removing one state at a time, to check whether the implicit assumption of the econometric specification of independence of irrelevant alternatives is acceptable. In Table 8 we present the coefficients of the income ratio for each of these 31 estimations. The income ratio is in all cases positive and significant, although somewhat lower in magnitude and at a lower level of significance in the estimation where Uttar Pradesh was removed from the sample. The size of the impact of the income ratio remains around 0.6-0.9 otherwise, the major change occurring when the model is re-estimated without Tamil Nadu - in this case the impact of the income ratio is higher (1.13). The test thus confirms the validity of the chosen specification.

[18]The majority of the internal migration flows in India are rural-rural (46%) or rural-urban (25%), compared to migration originating from urban areas.

Table 3: *Internal migration and drought*

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \frac{w_{j,t-1}}{w_{i,t-1}}$	0.885**	0.652*	0.802**	0.741*	0.905**	0.852**
	(0.399)	(0.396)	(0.391)	(0.402)	(0.393)	(0.398)
ln distance$_{ij}$	-0.676***	-0.676***	-0.676***	-0.676***	-0.676***	-0.676***
	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)
border	1.221***	1.223***	1.222***	1.222***	1.222***	1.222***
	(0.147)	(0.148)	(0.148)	(0.147)	(0.147)	(0.147)
language	0.404**	0.402**	0.403**	0.403**	0.404**	0.404**
	(0.160)	(0.159)	(0.159)	(0.159)	(0.159)	(0.159)
$\ln \frac{SC_{jt}+1}{SC_{it}+1}$	0.484	-1.115	-2.209	0.558	-0.823	-0.202
	(18.567)	(18.727)	(18.845)	(18.625)	(18.747)	(18.708)
$\ln \frac{ST_{jt}+1}{ST_{it}+1}$	-7.118	-6.745	-7.113	-6.479	-6.742	-6.895
	(6.691)	(6.738)	(6.734)	(6.758)	(6.808)	(6.744)
drought frequency$_{it}$		0.009*				
		(0.005)				
longest drought dur$_{it}$			0.008			
			(0.007)			
drought magnitude$_{it}$				0.005		
				(0.004)		
drought avg magn$_{it}$					0.050	
					(0.096)	
longest drought magn$_{it}$						0.003
						(0.005)
Origin state dummies D_i						
Destination state/year dummies D_{jt}						
N	1860	1860	1860	1860	1860	1860
R^2	0.692	0.696	0.695	0.694	0.692	0.693

The dependent variable is the bilateral migration rate from state i to state j between $t-1$ and t.
Robust standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.001

TABLE 4: *Internal migration and flood*

	(2)	(3)	(4)	(5)	(6)
$\ln \frac{w_{j,t-1}}{w_{i,t-1}}$	0.899**	0.880**	0.891**	0.813*	0.882**
	(0.396)	(0.407)	(0.397)	(0.418)	(0.409)
ln distance$_{ij}$	-0.677***	-0.676***	-0.677***	-0.676***	-0.676***
	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)
border	1.220***	1.221***	1.221***	1.222***	1.221***
	(0.147)	(0.147)	(0.147)	(0.147)	(0.147)
language	0.405**	0.404**	0.405**	0.406**	0.404**
	(0.159)	(0.159)	(0.160)	(0.160)	(0.159)
$\ln \frac{SC_{jt}+1}{SC_{it}+1}$	-0.024	0.350	-1.901	-0.229	0.094
	(18.569)	(18.639)	(18.810)	(18.500)	(19.036)
$\ln \frac{ST_{jt}+1}{ST_{it}+1}$	-7.700	-7.264	-7.557	-6.601	-7.172
	(6.650)	(6.655)	(6.653)	(6.769)	(6.667)
flood frequency$_{it}$	-0.005				
	(0.006)				
longest flood duration$_{it}$		-0.002			
		(0.006)			
flood magnitude$_{it}$			-0.004		
			(0.006)		
flood avg magn$_{it}$				-0.117	
				(0.145)	
longest flood magn$_{it}$					-0.001
					(0.005)
Origin state dummies D_i					
Destination state/year dummies D_{jt}					
N	1860	1860	1860	1860	1860
R^2	0.692	0.692	0.692	0.692	0.692

The dependent variable is the bilateral migration rate from state i to state j between $t-1$ and t.
Robust standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.001

Table 5: *Internal male migration and drought*

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \frac{w_{j,t-1}}{w_{i,t-1}}$	0.884**	0.614	0.770**	0.729*	0.907**	0.846**
	(0.376)	(0.377)	(0.370)	(0.379)	(0.373)	(0.375)
ln distance$_{ij}$	-0.694***	-0.694***	-0.693***	-0.694***	-0.694***	-0.694***
	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)
border	1.100***	1.102***	1.102***	1.102***	1.101***	1.101***
	(0.150)	(0.150)	(0.150)	(0.150)	(0.150)	(0.150)
language	0.495**	0.493**	0.494**	0.493**	0.494**	0.494**
	(0.163)	(0.162)	(0.162)	(0.162)	(0.163)	(0.162)
$\ln \frac{SC_{jt}+1}{SC_{it}+1}$	7.751	9.045	7.646	9.421	6.877	7.794
	(21.228)	(21.370)	(21.252)	(21.352)	(21.036)	(21.269)
$\ln \frac{ST_{jt}+1}{ST_{it}+1}$	-4.287	-4.172	-4.408	-4.108	-4.102	-4.266
	(3.483)	(3.377)	(3.448)	(3.389)	(3.434)	(3.448)
drought frequency$_{it}$		0.010*				
		(0.005)				
longest drought dur$_{it}$			0.008			
			(0.006)			
drought magnitude$_{it}$				0.006		
				(0.004)		
drought avg magn$_{it}$					0.056	
					(0.095)	
longest drought magn$_{it}$						0.003
						(0.005)
Origin state dummies D_i						
Destination state/year dummies D_{jt}						
N	1860	1860	1860	1860	1860	1860
R^2	0.668	0.672	0.671	0.671	0.668	0.667

The dependent variable is the male bilateral migration rate from state i to state j between $t-1$ and t.
Robust standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.001

Table 6: *Internal female migration and drought*

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \frac{w_{j,t-1}}{w_{i,t-1}}$	0.508	0.295	0.447	0.372	0.582	0.484
	(0.415)	(0.408)	(0.402)	(0.409)	(0.404)	(0.410)
ln distance$_{ij}$	-0.644***	-0.643***	-0.643***	-0.644***	-0.644***	-0.644***
	(0.078)	(0.079)	(0.078)	(0.078)	(0.078)	(0.078)
border	1.381***	1.384***	1.383***	1.382***	1.381***	1.382***
	(0.148)	(0.149)	(0.149)	(0.148)	(0.148)	(0.148)
language	0.285*	0.284*	0.285*	0.284*	0.284*	0.285*
	(0.160)	(0.160)	(0.160)	(0.160)	(0.160)	(0.160)
$\ln \frac{SC_{jt}+1}{SC_{it}+1}$	-6.085	-5.172	-6.282	-4.571	-8.499	-6.072
	(15.554)	(15.614)	(15.568)	(15.597)	(15.474)	(15.574)
$\ln \frac{ST_{jt}+1}{ST_{it}+1}$	-4.091	-4.748	-4.789	-4.379	-4.013	-4.385
	(3.310)	(3.310)	(3.358)	(3.256)	(3.263)	(3.304)
drought frequency$_{it}$		0.011*				
		(0.006)				
longest drought dur$_{it}$			0.009			
			(0.007)			
drought magnitude$_{it}$				0.007		
				(0.004)		
drought avg magn$_{it}$					0.096	
					(0.097)	
longest drought magn$_{it}$						0.005
						(0.005)
Origin state dummies D_i						
Destination state/year dummies D_{jt}						
N	1860	1860	1860	1860	1860	1860
R^2	0.718	0.721	0.720	0.720	0.718	0.719

The dependent variable is the female bilateral migration rate from state i to state j between $t-1$ and t.
Robust standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.001

6 Conclusions

The objective of the paper is to test the hypothesis that long term climate variability acts as a push-factor on internal migration. We apply an econometric specification based on utility maximization on Indian census data from 1991 and 2001 matched with climate data. To the best of our knowledge this is one of few attempts to investigate the impact of climate variability on internal migration in a gravity-type estimation on the level of such a large and diverse country as India. The main contribution of the analysis is to introduce relevant objective meteorological indicators of climate variability, based on the standardized precipitation index. The base model estimation results provide a rather good fit of bilateral migration flows between states and confirm the impact of income differences between states and the cost to migrate. We then augment the base model to include climate variability in the form of drought episodes, measured either through their frequency, duration or magnitude. The estimation results do not reject the hypothesis of drought frequency acting as a push factor for inter-state migration in India. Even if the statistical significance is only at a 10% level, we show that the significance and size of the effect are robust. The effect is verified while controlling for origin and destination state fixed effects. It holds for both male and female bilateral migration rates, confirming the importance of climate variability as a push factor for internal migration. By comparison, economic motivations for migration were only significant in the estimations on male bilateral migration rates. Drought duration and magnitude were never statistically significant in the estimations. We suggest that the findings may be interpreted as evidence of the expectations of future drought inducing migration. Observed frequency of droughts tends to reinforce future expectations of drought and may hence induce migration. This does not mean that the other proposed measures are of less interest in an application on other geographical zones and time periods. Indeed, the insignificance of the drought magnitude is probably explained by the actual events on a state level in India over the years that we study here. In fact, the drought and flood episodes on a state level were of moderate nature over the time studied. As regards the duration of a drought episode, it may encourage adaptation through other measures than migration, such as participation in the non-climate dependent economic sectors. This is a topic to be studied in future research.

We also control for the econometric problems that arise when applying a gravity-type model on bilateral migration flows. In particular, we apply the pseudo poisson maximum likelihood estimator to correct for the presence of zero migration flows between certain states and control for heteroskedasticity. As to the size of the induced increase in migration rates, the estimation results indicate that an additional month of drought in the origin state during the five years preceding migration increases the bilateral migration rate by 0.9%. Such an effect may seem small, especially when compared to the important role of barriers to migration in the Indian context that explain the low Indian inter-state migration rates. Sharing a common language, for instance, would increase the bilateral migration rate by 50 %. The impact of drought frequency is thus moderated by the barriers to migration. Nevertheless, the results show that an increase in the frequency of drought events can induce additional large numbers of inter-state migrants in absolute values.

Detailed analysis of the rainfall data shows that aggregation on the state level masks important variability between districts. A more detailed modelling of (rural-urban) migration flows at the district level thus seems appropriate in order to further test the hypothesis that drought or flood episodes may induce migration flows in excess of those normally observed. This is the subject of ongoing research.

References

[1] Anderson, J.E. (2011). The gravity model. *Annual Review of Economics*, 3, p. 133-160.

[2] Attri, S.D. & Tyagi, A. (2010). Climate profile of India. Government of India Ministry of Earth Sciences India Meteorological Department. New Delhi.

[3] Badiani, R. & Safir, A. (2010). Coping with aggregate shocks: Temporary migration and other labor responses to climatic shocks in rural India. Mimeo World Bank.

[4] Barrios, S., Bertinelli L. & Strobl, E. A. (2006). Climatic change and rural-urban migration: The case of sub-Saharan Africa. *Journal of Urban Economics*, 60(3), 357-371.

[5] Barrios, S., Bertinelli L. & Strobl, E. A. (2010). Trends in rainfall and economic growth in Africa: A neglected cause of the African growth tragedy. *The Review of Economics and Statistics*, 92(2), 350-366.

[6] Beine, M. & Parsons, C. (2012). Climate factors as determinants of international migration. CESifo Working Papers No. 3747.

[7] Beine, M., Docquier, F. & Özden, C. (2011). Diasporas. *Journal of Development Economics*, 95(1), 30-41.

[8] Bertoli, S. & Fernández-Huertas Moraga, J. (2013). Multilateral resistance to migration. *Journal of Development Economics*, forthcoming.

[9] Bhattacharya, P.C. (2002). Rural to urban migration in LDCS: a test of two rival models. *Journal of International Development*, 14(7), 951-972.

[10] Bhattacharya, H. & Innes, R. (2008). An empirical exploration of the population-environment nexus in India. *American Journal of Agricultural Economics*, 90(4), 883-901.

[11] Bodvarsson, Ö. & Van den Berg, H. (2009). *The economics of immigration.* Berlin Heidelberg: Springer-Verlag.

[12] Census of India (2011). Provisional Population Totals. Office of the Registrar General and Census Commissioner, India. Paper 1 of 2011, India Series 1.

[13] Coniglio, N. & Pesce, G. (2011). Climate variability and international migration: what are the links? Paper presented at the annual conference of the European Association of Environmental and Resource Economists, July 1, Rome.

[14] Datta, P. (2006). Urbanisation of India. Paper presented at The Regional and Sub-Regional Population Dynamic - Population Process in Urban Areas - European Population Conference, June 21-24, Liverpool.

[15] DeGaetano, A. T. (2001). Spatial grouping of United States climate stations using a hybrid clustering approach. *International Journal of Climatology*, 21(7), 791-807.

[16] Dell, M., Jones, B.F. & Olken, B.A. (2009). Temperature and income: Reconciling new cross-sectional and panel estimates. *American Economic Review, Papers and Proceedings* 99(2), 198-204.

[17] Dubey, A., Palmer-Jones, R. & Sen, K. (2006). Surplus labour, social structure and rural to urban migration: Evidence from Indian data. *The European Journal of Development Research*, 18(1), 86-104.

[18] Feenstra, R.C. (2004). *Advanced international trade: Theory and evidence*. Princeton, NJ: Princeton University Press.

[19] The Government Office for Science, London (2011). *Foresight: Migration and Global Environmental Change*, Final Project Report.

[20] Gray, C. (2009). Environment, land, and rural out-migration in the Southern Ecuadorian Andes. *World Development*, 37(2), 457-468.

[21] Grogger, J. & Hanson, G.H. (2008). Income maximization and the selection and sorting of international migrants. NBER working paper No. 13821.

[22] Guerreiro, M., Lajihna, T. & Abreu I. (2008). Flood analysis with the standardized precipitation index. *Revista da Faculdade de Ciencia e Tecnologia de la Universidade Fernando Pessoa*, 4, 8-14.

[23] Hornbeck, R. (2012). The enduring impact of the American Dust Bowl: Short and long-run adjustments to environmental catastrophe. *American Economic Review*, 102(4), 1477-1507.

[24] Karemera, D., Oguledo, V. & Davis, B. (2000). A gravity model analysis of international migration to North America. *Applied Economics*, 32(13), 1745-1755.

[25] Lewer, J. & Van den Berg, H. (2008). A gravity model of immigration. *Economics Letters*, 99(1), 164-167.

[26] Marchiori, L. & Schumacher I. (2011). When nature rebels: international migration, climate change, and inequality. *Journal of Population Economics*, 24(2), 569-600.

[27] Marchiori, L., Maystadt, J.-F. & Schumacher I. (2012). The impact of weather anomalies on migration in sub-Saharan Africa. *Journal of Environmental Economics and Management*, 63, 355-374.

[28] Mayda, A. M. (2010). International migration : A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics*, 23(4), 1249-1274.

[29] McKee, T., Doesken, N. & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. Paper presented at the 8th Conference on Applied Climatology, January 17-22, Anaheim, California, USA.

[30] Mitchell, T.D., Hulme, M. & New, M. (2002). Climate data for political areas. *Area*, 34(1), 109-112.

[31] Mitra, A. & Murayama, M. (2008). Rural to urban migration: A district level analysis for India. IDE Discussion Paper. No. 137.

[32] Munoz-Diaz, D. & Rodrigo, F. (2004). Spatio-temporal patterns of seasonal rainfall in Spain (1912-2000) using cluster and principal component analysis: comparison. *Annales Geophysicae*, 22, 1435-1448.

[33] Myers, N. (1997). Environmental refugees. *Population and Environment*, 19(2), 167-182.

[34] Naudé, W. (2008). Conflict, disasters, and no jobs : Reasons for international migration from Sub-Saharan Africa. United Nations University, Research Paper No. 2008/85.

[35] Özden, Ç. & Sewadeh, M. (2010). How important is migration?, In E. Ghani (Ed.) *The poor half million in South Asia: What is holding back lagging regions?* (pp. 294-322). New Delhi, India: Oxford University Press.

[36] Reuveny, R. & Moore, W. (2009). Does environmental degradation influence migration? Emigration to developed countries in the late 1980s and 1990s. *Social Science Quarterly*, 90(3), 461-479.

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

[38] Santos Silva, J.M.C. & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4), 641-658.

[39] Singh, K., & Singh, H. (1996). Space-time variation and regionalization of seasonal and monthly summer monsoon rainfall of the sub-Himalayan region and Gangetic plains of India. *Climate Research*, 6, 251-262.

[40] Stern, N. (2007). *The economics of climate change: the Stern review.* Cambridge: Cambridge University Press.

[41] Strobl, E.A. & Valfort, M. (2012). The effect of weather-induced internal migration on local labor markets: Evidence from Uganda. IZA Discussion Paper No. 6923.

[42] Van Lottum, G. & Marks, D. (2010). The determinants of internal migration in a developing country: Quantitative evidence for Indonesia, 1930-2000. *Applied Economics*, 44(34), 4485-4494.

[43] Viswanathan, B. & Kumar, K.S.K. (2012). Weather variability, agriculture and rural migration: Evidence from state and district level migration in India. Paper presented at the 2nd International Conference on Environment and Natural Resources Management in Developing and Transition Economies (enrmdte), Clermont-Ferrand, October 17.

A Principal Component Analysis (PCA)

In order to match the climate and the census data, we have to aggregate the climate data to state level. The spatial grouping of observations is standard practice in the climatological literature (Munoz-Diaz and Rodrigo, 2004) [32]. These groupings serve to summarize climate data in a concise way (DeGaetano, 2001) [15]. PCA can be used to identify the most important correlations between different variables, so as to obtain a description of the major part of the overall variance, with a reduced number of linear combinations based on the original variables (Munoz-Diaz and Rodrigo, 2004) [32]. We apply PCA to test if aggregating rainfall across states imply losing important information or not.

We did a PCA between states and then between districts for the rainfall data, after having normalized the variables on the available period from 1901 to 2006. We applied an oblique rotation to the unrotated eigenvectors, according to the methodology of Barrios et al. (2010) [5].[19] In the PCA applied to the states, we find 3 big rain zones with a loading of 0.1 (by having one single state which belongs to no zone and no state which belongs to more than one zone). By comparison, with a loading of 0.4, the states rainfall patterns are completely independent, implying that there is no correlation between them (no regrouping of states were possible). The choice of the threshold for the loading is very subjective: Singh et Singh (1996) [39] take values included between 0.2 and 0.5, Barrios et al. (2010) [5] take a value of 0.2 for their inter country analysis on sub-Saharan Africa and 0.05 for their intra country analysis; Munoz-Diaz and Rodrigo (2004) [32] between 0.2 and 0.9.

We also check whether rainfall patterns are homogenous within states. When applying PCA to districts, we have 13 main rain zones with a loading of 0.1. The states contain between 1 to 3 different zones maximum, except for the states of Madhya Pradesh and Uttar Pradesh (regrouped in 5 and 6 zones respectively), but those are very large states. We checked the distribution of these zones on a map of India and confirmed that the states which belong to the same groups are indeed bordering, except in one case. We conclude that the climate analysis at the state level seems relevant.

[19] Given that the PCA is for us only a preliminary analysis, we will not develop the technical details further, but deeper applications can be seen in Munoz-Diaz and Rodrigo (2004) [32], Barrios et al. (2010) [5] and Singh and Singh (1996) [39].

B Additional estimation tables and correlation matrix

TABLE 7: *Internal migration, rural inequality and drought*

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \frac{w_{j,t-1}}{w_{i,t-1}}$	0.876**	0.653	0.800**	0.739*	0.895**	0.846**
	(0.398)	(0.397)	(0.392)	(0.402)	(0.391)	(0.398)
Rural Gini$_{it}$	1.685	0.914	0.989	1.276	1.653	1.476
	(2.814)	(2.746)	(2.752)	(2.765)	(2.821)	(2.781)
ln distance$_{ij}$	-0.676***	-0.676***	-0.676***	-0.676***	-0.676***	-0.676***
	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)
border	1.221***	1.223***	1.222***	1.222***	1.221***	1.222***
	(0.147)	(0.148)	(0.148)	(0.147)	(0.147)	(0.147)
language	0.405**	0.402**	0.403**	0.403**	0.404**	0.404**
	(0.159)	(0.158)	(0.159)	(0.159)	(0.159)	(0.159)
$\ln \frac{SC_{jt}+1}{SC_{it}+1}$	5.588	1.694	0.905	4.422	4.220	4.330
	(17.695)	(17.859)	(17.900)	(17.800)	(18.096)	(17.816)
$\ln \frac{ST_{jt}+1}{ST_{it}+1}$	-8.234	-7.362	-7.770	-7.350	-7.847	-7.893
	(7.283)	(7.318)	(7.328)	(7.355)	(7.436)	(7.357)
drought frequency$_{it}$		0.009*				
		(0.005)				
longest drought dur$_{it}$			0.008			
			(0.007)			
drought magnitude$_{it}$				0.005		
				(0.004)		
drought avg magn$_{it}$					0.049	
					(0.097)	
longest drought magn$_{it}$						0.003
						(0.005)
Origin state dummies D_i						
Destination state/year dummies D_{jt}						
N	1860	1860	1860	1860	1860	1860
R^2	0.693	0.696	0.695	0.695	0.693	0.694

The dependent variable is the bilateral migration rate from state i to state j between $t-1$ and t.
The Gini coefficients are for years 1993-1994 and 1999-2000. Source: The National Human Development Report 2001 (Estimated from NSS 50th & 55th Rounds on Household Consumer Expenditure).
Robust standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.001

Table 8: *Income ratio coefficient omitting one destination state*

State ID	Omitted destination	$\ln \frac{w_j}{w_i}$	SE	R^2
2	ANDHRA PRADESH	0.897**	(0.408)	0.698
3	ARUNACHAL PRADESH	0.887**	(0.400)	0.692
4	ASSAM	0.944**	(0.408)	0.694
5	BIHAR	0.846**	(0.401)	0.699
6	GOA	0.904**	(0.401)	0.693
7	GUJARAT	0.866**	(0.409)	0.687
8	HARYANA	0.847**	(0.410)	0.716
9	HIMACHAL PRADESH	0.886**	(0.403)	0.692
11	KARNATAKA	0.885**	(0.414)	0.691
12	KERALA	0.840**	(0.422)	0.717
13	MADHYA PRADESH	0.879**	(0.415)	0.697
14	MAHARASHTRA	0.798**	(0.394)	0.754
15	MANIPUR	0.885**	(0.399)	0.691
16	MEGHALAYA	0.888**	(0.399)	0.692
17	MIZORAM	0.875**	(0.399)	0.692
18	NAGALAND	0.884**	(0.399)	0.692
19	ORISSA	0.884**	(0.403)	0.694
20	PUNJAB	0.938**	(0.401)	0.657
21	RAJASTHAN	0.858**	(0.411)	0.705
22	SIKKIM	0.879**	(0.399)	0.691
23	TAMIL NADU	1.13**	(0.413)	0.695
24	TRIPURA	0.894**	(0.399)	0.692
25	UTTAR PRADESH	0.671*	(0.365)	0.724
26	WEST BENGAL	0.891**	(0.405)	0.696
27	ANDAMAN & NICOBAR ISLANDS	0.899**	(0.401)	0.691
28	CHANDIGARH	0.905**	(0.401)	0.699
29	DADRA & NAGAR HAVELI	0.875**	(0.401)	0.693
30	DAMAN & DIU	0.894**	(0.398)	0.713
31	DELHI	0.967**	(0.421)	0.694
32	LAKSHADWEEP	0.874**	(0.401)	0.692
33	PONDICHERRY	0.885**	(0.400)	0.692

The number of observations for each specification is 1800.
The dependent and independent variables are the same as in estimation (1) in Table 3.
Robust standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.001

TABLE 9: *Correlation matrix*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) migration rate												
(2) NSDP ratio	-0.0915*											
(3) Rural Gini ratio	0.0093	-0.0807*										
(4) distance	-0.2778*	-0.0060	-0.0124									
(5) border	0.3589*	-0.0257	0.0255	-0.5082*								
(6) language	0.2145*	0.0022	-0.0447*	-0.3970*	0.2741*							
(7) SC rate	0.0062	0.0918*	0.3357*	-0.1348*	0.0884*	0.0230						
(8) ST rate	-0.0169	-0.0299	-0.3876*	0.0696*	-0.0658*	0.0678*	-0.6739*					
Drought variables												
(9) frequency	0.0047	0.0936*	-0.0500*	0.1018*	-0.0426*	0.0317	-0.2884*	0.1819*				
(10) longest duration	-0.0120	0.1252*	-0.0568*	0.0808*	-0.0252	0.0154	-0.2885*	0.1875*	0.9526*			
(11) magnitude	0.0161	0.0218	-0.0921*	0.0980*	-0.0463*	0.0403*	-0.3107*	0.1872*	0.9316*	0.8833*		
(12) avg. magn.	0.0153	-0.0615*	-0.1113*	0.0715*	-0.0441*	0.0169	-0.3462*	0.1897*	0.6537*	0.6677*	0.7819*	
(13) longest magn.	0.0011	0.0335	-0.1184*	0.0841*	-0.0324	0.0269	-0.3190*	0.1947*	0.8888*	0.9049*	0.9713*	0.8131*

Number of observations: 1860
* $p < 0.1$