

IZA DP No. 1668

## Brain Drain in Developing Regions (1990-2000)

Frédéric Docquier  
Olivier Lohest  
Abdeslam Marfouk

July 2005

# **Brain Drain in Developing Regions (1990-2000)**

**Frédéric Docquier**

*University of Lille 2, World Bank,  
IWEPS and IZA Bonn*

**Olivier Lohest**

*IWEPS (Regional Govt. of Wallonia, Belgium)*

**Abdeslam Marfouk**

*Free University of Brussels and IWEPS*

Discussion Paper No. 1668

July 2005

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

Email: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### Brain Drain in Developing Regions (1990-2000)\*

In this paper, we analyze the distribution of the brain drain in the LAC region (Latin America and the Caribbean), Asia and Africa. We rely on an original data set on international migration by educational attainment for 1990 and 2000. Our analysis reveals that the brain drain is strong in Eastern, Middle and Western Africa, Central America and the Caribbean. However, the Kernel approach suggests that the dispersion and the intradistribution dynamics of skilled migration rates strongly differ across regions. We then tautologically disentangle the brain drain into two multiplicative components, the global migration rate and the selection bias. Among the most affected countries, LAC countries suffer from high migration rates whilst most African countries suffer from high selection biases. Finally, exploratory Moran's tests reveal strong spatial, political and cultural autocorrelations in migration rates and selection biases. The latter result suggests that skilled workers react differently than unskilled workers to a large set of variables.

JEL Classification: F22, O15, J11, J24

Keywords: international migration, brain drain, human capital, spatial autocorrelation

Corresponding author:

Frédéric Docquier  
University of Lille 2  
1 Place Déliot  
F-59024 Lille  
France  
Email: [f.docquier@skynet.be](mailto:f.docquier@skynet.be)

---

\* This article benefited from helpful comments by Michel Beine, David Smith, Sébastien Laurent, Caglar Ozden and Maurice Schiff. The first authors thank the World Bank for financial support (Contract PO. 7620076 - UPI 269656). Special thanks to Anthon Eff for providing his data set on spatial and cultural proximities between countries. The usual disclaimers apply.

# Contents

1	Introduction . . . . .	3
2	Measurement, notations, overview . . . . .	5
2.1	Emigration stocks by region . . . . .	6
2.2	Emigration rates by region . . . . .	7
3	Regional disparities . . . . .	9
3.1	The distribution of the brain drain . . . . .	9
3.2	The 1990-2000 dynamics . . . . .	16
4	A tautological decomposition . . . . .	20
5	A step toward empirical analysis . . . . .	25
6	Conclusion . . . . .	30
7	Appendix . . . . .	31
7.1	The DM05 dataset - global overview . . . . .	31
7.2	Applying the Kernel method to the decomposition . . . . .	37

# 1 Introduction

The international migration of skilled workers (the so-called brain drain) has attracted considerable attention in the recent years. Given recent developments of immigration policies conducted in receiving countries and the booming demand for highly skilled workers, many case studies support the view that the migration of the educated has intensified over the 1990s. By the next decade, the size of brain drain is unlikely to fall given the expansion of the IT sector and the dark demographic prospects faced by most industrialized nations (see United Nations, 2001). Today, industrial countries such as Canada, the UK and Germany are worrying about the magnitude of the emigration flow of skills. In these countries, recent immigration policies are designed to replace the personnel losses with highly skilled foreigners primarily attracted from poorer countries. Unsurprisingly, it is mainly for less developed countries that the detrimental consequences of the brain drain have been stressed in literature. Between 1990 and 2000, the stock of skilled immigrants in the OECD increased by 64 percent. The rise was stronger for immigrants coming from less developed countries (+93 percent), especially those coming from Africa (+113 percent) and the LAC<sup>1</sup> region (+97 percent).

Is the brain drain a curse or a boon for economic development? What are the best policy responses, both from an internationalist and a nationalist point of view? There is no clear and straightforward answer to these questions. The reason is that, until recently and despite many anecdotal evidence, there were no reliable databases documenting the brain drain for a large set of countries and for different years. The debate thus remained theoretical, with two contrasting strands of literature:

- Dating back to the 1960s, the "earlier literature" focuses on the ex-post effect of the brain drain (the human capital stock is considered as given) and investigates all of its consequences for remaining residents. By reducing the number of educated remaining in the country, the brain drain unambiguously reduces the average level of education and generates a loss for those left behind.
- On the contrary, the "new literature" emphasizes the positive impact of skilled migration flows and migration prospects on the ex-ante stock of human capital (before migration is netted out)<sup>2</sup>. In the long-run, the global impact of the brain drain balances its ex-ante beneficial effects and the ex-post detrimental effects.

---

<sup>1</sup>Henceforth, LAC stands for Latin America (South America and Central America) and the Caribbean

<sup>2</sup>Taking account of some indirect economic effects, one can reasonably consider that past migration flows or migration prospects have positive effects on human capital accumulation. The potential channels potentially at work are return migration, remittances and/or the impact of migration prospects on the expected return on education. This literature is surveyed in Commander et al. (2004) or Docquier and Rapoport (2004). See Schiff (2005) for a critical appraisal.

Given the fast evolving process of international migration and the policy issues at stake, the international community must be prepared to address the major challenges raised by the brain drain. In particular, assessing the size of these ex-post and ex-ante effects requires a better knowledge of the international mobility of skilled workers. The purpose of our paper is to offer the first comprehensive and accurate picture to date of the brain drain distribution across developing regions. Our study can be used to evaluate the pertinency of theoretical models and their policy implications.

We focus on three regions which mostly contain the least developed countries: the LAC region, Asia and Africa<sup>3</sup>. Our analysis relies on a new harmonized and exhaustive data set on migration stocks and rates by educational attainment. The methodology used to compute these data is detailed in Docquier and Marfouk (2005). It builds on the collection of Census information on the structure of immigration in all OECD countries. Aggregating these data allows to evaluate the stock of emigrants from developing countries to the OECD area by level of diploma. In a second step, the educational structure of emigrants is compared to that of residents in each possible sending country. Such a comparison allows to compute emigration rates by educational attainment for 1990 and 2000.

In this article, we (i) characterize the distribution of brain drain (as measured by the migration rate of tertiary educated workers) across developing countries in 2000, (ii) study the intradistribution dynamics of skilled migration between 1990 and 2000, (iii) disentangle the driving forces leading to skilled migration rates and (iv) explore the structural links between these forces and the characteristics of sending countries.

The first step builds on descriptive statistics as well as on kernel density estimates. Descriptive statistics show that the average brain drain is especially strong in Central America and the Caribbean, and Eastern, Middle and Western Africa. Dispersion indicators reveal a strong heterogeneity within regions. Applying non parametric techniques to skilled migration rates provides many insights compared to usual descriptive statistics. Despite common multimodality properties, the range and the dispersion of skilled migration rates strongly differ across regions. At the aggregate, a Gaussian mixture model allows us to identify three major clusters in 2000.

In the second step, we rely on stochastic conditional kernel densities to derive the evolution of the brain drain between 1990 and 2000. Our analysis provides striking insights on the intradistribution dynamics. Several clubs are emerging in each region. Some exhibit an increase in skilled migration; others exhibit a significant decrease. Whilst the dynamic is relatively simple in the LAC region and Asia, it is more complex in Africa where at least six major clubs are distinguished.

In the next step, we provide a tautological decomposition of the brain drain. The skilled migration rate is expressed as the product of two important factors: the total migration rate and the selection bias (the latter is defined as the ratio of the share of tertiary educated among emigrants to the same proportion calculated on

---

<sup>3</sup>Our full sample is not far from the group of less developed countries identified by the United Nations. We include Japan in Asia and we exclude less developed islands from Oceania.

all natives). These two components are likely to depend on specific determinants. Abstracting from the endogeneity and multicollinearity issues, such a decomposition enables the characterization of the position of all regions and clusters in regard of these components. It comes out that countries suffering from the brain drain exhibit either high global migration rates or high selection biases, not both. The average selection bias is an important cause of skilled migration in Asia and in most African countries, but the bias is very heterogeneous within region. On the contrary, the dispersion of LAC countries is essentially due to their degree of openness. The degree of openness is highly sensitive to the population size; the selection bias is strongly linked to the supply of human capital of natives. Of course, other determinants should be used to characterize the driving force of migration and selection.

Finally, we conduct spatial, cultural, economic and political autocorrelation tests on skilled migration rates and selection biases. Our analysis should be considered as exploratory. Our main objective is to assess whether migration rates and selection biases can be structurally associated to the traditional repulsive factors of migration. We use several weight matrices capturing the proximity between countries in terms of culture, political regime, level of development, inequality or geography. We leave intact the basic issues of causality and analytical form of these relationships (such issues should clearly be addressed with panel econometric techniques). However we offer important stylized facts and important insights for further empirical works. Our analysis suggests that selection biases are structurally linked to a large set of variables such as fractionalization, political features and distances. Although our hypothesis needs to be validated by econometric studies, our autocorrelation tests indicate that skilled workers react differently than unskilled workers to most traditional "push" factors.

The paper is organized as follows. Section 2 describes the data sources and the methodology used to measure international migration by educational attainment. The regional characteristics and a clustering analysis are presented in Section 3. Section 4 provides the tautological decomposition and analyzes the distribution of the components. Autocorrelation tests are conducted in section 5. Finally, section 6 concludes.

## **2 Measurement, notations, overview**

Our analysis builds on a new comprehensive and consistent database on international migration by educational attainment. A more detailed exposition of the methodology and the assumptions can be found in Docquier and Marfouk (2005). The major hypotheses are summarized in Appendix 7.1. The data set (labelled as DM05) describes the loss of skilled workers to the OECD for all countries in 1990 and 2000. The data set relies on two steps. Emigration stocks are first computed by educational attainment from each country. In a second step, these numbers are expressed in percentage of the total labor force born in the sending country (including migrants

themselves). A similar objective is pursued in Carrington and Detragiache (1998) who relied on a set of tentative assumptions to estimate the rate of emigration of tertiary educated workers from 61 developing countries in 1990<sup>4</sup>. Whilst Carrington and Detragiache’s study clearly initiated new debates on skilled migration, their estimates are quite poor to allow empirical studies of the causes and the consequences for developing countries. Generalizing this study, The DM05 data set provides the first comprehensive and accurate picture to date of skilled emigration rates in all the world countries.

Here we focus on three major sending areas, the LAC region (including Central America, South America and the Caribbean), Asia and Africa. Since the world configuration changed between 1990 and 2000, the number of observations varies: we have 136 observations in 2000 (53 in Africa, 50 in Asia and 33 in the LAC region) and 125 in 1990 (52 in Africa, 41 in Asia and 33 in the LAC region).

## 2.1 Emigration stocks by region

Stock data on emigration can only be captured by aggregating consistent immigration data collected in receiving countries. Regarding stocks, we count as migrants all working-aged (25 and over) foreign born individuals living in an OECD country; considering the working-aged population (aged 25 and over) maximizes the comparability of the immigration population with data on educational attainment in the source countries. It also excludes a large number of students who temporarily emigrate to complete their education. By restricting the set of receiving countries to the OECD area, we focus on the South-North and North-North brain drain. We are aware that a brain drain can be observed outside the OECD area (to the Gulf countries, South Africa, Malaysia, Hong-Kong, Singapore, Taiwan, etc.). Given Census data collected from various non OECD countries, we estimate that about 90 percent of high-skill international emigrants are living OECD countries. Three levels of schooling are distinguished. Low-skill workers are those with primary education (0 to 8 years of schooling completed), medium skilled workers are those with secondary education (9 to 12 years of schooling) and high-skilled workers are those with tertiary education (13 years and above).

Information about the origin and skill of natives and immigrants is available from national censuses and population registers. More specifically, country  $i$ ’s census usually identifies individuals on the basis of age, country of birth  $j$ , and skill level  $s$ . Our method consists in collecting Census or register data from a large set of receiving countries, with the highest level of detail on birth countries and (at least) three levels

---

<sup>4</sup>The strongest assumption is that they transpose the skill structure of US immigrants on the total OECD immigration stock. For example, Surinamese migrants to the Netherlands are assumed to be distributed across educational categories in the same way as Surinamese migrants to the US. Since the US immigration policy differs from that of many countries, this assumption is highly tentative for countries with a low migration rate to the USA (Africa, most Asian countries, Oceania or Europe).

of educational attainment:  $s = h$  for high skilled,  $s = m$  for medium skilled,  $s = l$  for low skilled. Let  $M_{t,s}^{i,j}$  denote the stock of working-aged individuals born in  $j$ , of skill  $s$ , living in country  $i$  at time  $t$ . The emigration stock of educational level  $s$  from country  $j$  at time  $t$  amounts to  $M_{t,s}^j = \sum_i M_{t,s}^{i,j}$ .<sup>5</sup>

Table 1 provides the total stock of emigrants by origin area and by region of destination. It comes out that developing countries are more and more concerned by the migration of skilled workers. Between 1990 and 2000, the total stock of emigrants has been multiplied by 2.13 in Africa, 1.84 in Asia and 1.97 in the LAC region. The major regions experiencing a booming emigration are Middle, Southern and Western Africa, South-eastern Asia and Central America. About 50 percent of emigrants chose the US as destination. The percentage varies across continents: it amounts to 80.2 percent for the LAC region, 59.6 percent in Asia and only 32.4 percent in Africa. For historical reasons, the European Union EU-15 is the major destination for African migrants (46.2 percent in 2000), especially for Middle Africa and Northern Africa. Canada also plays an important role for Asian and African migrants, as well as for people from the Caribbean. Australia attracts a large fraction of Southern African emigrants. Regarding the other OECD countries, a large number of Eastern-Asian and South-American (mainly Brazilian) native born opt for Japan.

## 2.2 Emigration rates by region

Migration rates by educational attainment are obtained by comparing the emigration stocks to the total number of people born in the source country and belonging to the same educational category (as in Carrington and Detragiache, 1998, and Adams, 2003). Calculating the brain drain as a proportion of the total educated labor force is a better strategy to evaluate the pressure imposed on the local labor market. It is indeed obvious that the pressure exerted by 1,070,000 Indian skilled emigrants (4.4% of the Indian educated total labor force) is less important than the pressure exerted by 30,000 of skilled emigrants from Grenada (65% of their educated labor force). Denoting  $N_{t,s}^j$  as the stock of individuals aged 25+, of skill  $s$ , living in country  $j$ , at time  $t$ , we define the emigration rates by

$$m_{t,s}^j = \frac{M_{t,s}^j}{N_{t,s}^j + M_{t,s}^j}$$

---

<sup>5</sup>A global overview of immigration statistics and data sources is provided in table A1 in appendix 7.1.

**Table 1. Main destination of skilled emigrants by region of origin**

	Stock of skilled emigrants	Percentage living in the USA	Percentage living in Canada	Percentage living in Australia	Percentage living in the EU-15	Percentage living in other OECD	Percentage living in the UK	Percentage living in Germany	Percentage living in France
<i>Situation in 2000</i>									
AFRICA	1387966	32.4%	12.1%	6.4%	46.2%	2.9%	18.5%	3.0%	15.1%
Eastern Africa	347379	29.6%	17.3%	6.9%	43.8%	2.4%	31.1%	1.4%	4.6%
Middle Africa	96994	16.8%	7.6%	0.5%	72.6%	2.4%	6.6%	3.9%	28.0%
Northern Africa	445718	26.6%	13.5%	4.6%	52.7%	2.6%	5.4%	4.7%	30.7%
Southern Africa	171397	23.7%	13.2%	24.1%	31.1%	8.0%	27.2%	0.4%	0.5%
Western Africa	326478	52.5%	5.3%	0.7%	40.2%	1.2%	21.9%	3.3%	8.8%
ASIA	7041367	59.6%	13.6%	7.3%	13.7%	5.9%	5.8%	3.5%	1.5%
Eastern Asia	2314310	61.3%	15.0%	6.3%	6.0%	11.4%	2.6%	1.0%	0.9%
South-central Asia	1871266	56.5%	15.3%	5.9%	20.0%	2.4%	11.9%	3.9%	1.0%
South-eastern Asia	2134311	66.2%	11.4%	9.6%	8.9%	3.9%	3.0%	1.6%	1.6%
Western Asia	721481	42.3%	10.8%	6.9%	36.8%	3.2%	8.1%	16.5%	4.4%
LAC REGION	3655136	80.2%	8.1%	0.9%	7.6%	3.2%	1.9%	0.9%	0.6%
Caribbean	1161806	79.8%	13.5%	0.1%	5.7%	0.8%	3.6%	0.5%	0.2%
Central America	1344209	94.3%	2.9%	0.3%	1.4%	1.0%	0.2%	0.3%	0.2%
South America	1149121	64.0%	8.8%	2.4%	0.0%	24.8%	2.2%	2.1%	1.6%
<i>Situation in 1990</i>									
AFRICA	651916	31.4%	11.8%	9.0%	45.0%	2.8%	17.3%	2.8%	12.3%
Eastern Africa	178901	23.8%	20.1%	9.9%	44.1%	2.1%	33.3%	1.5%	3.9%
Middle Africa	41839	14.3%	4.3%	0.8%	77.1%	3.4%	5.1%	1.4%	27.6%
Northern Africa	242710	27.3%	13.5%	7.7%	48.5%	3.0%	4.6%	4.5%	21.2%
Southern Africa	64118	32.6%	0.4%	32.0%	29.1%	6.0%	25.8%	0.5%	0.6%
Western Africa	124348	55.4%	5.1%	1.1%	36.8%	1.6%	18.7%	2.9%	8.0%
ASIA	3836581	57.4%	11.6%	9.0%	12.8%	9.2%	4.7%	3.7%	1.4%
Eastern Asia	1421042	58.1%	11.1%	6.0%	5.1%	19.7%	1.9%	1.6%	0.6%
South-central Asia	834064	54.1%	14.6%	7.7%	21.3%	2.3%	11.2%	5.7%	1.2%
South-eastern Asia	1183789	64.4%	10.7%	13.4%	8.2%	3.4%	2.7%	1.3%	1.3%
Western Asia	397685	40.6%	10.5%	9.5%	36.6%	2.8%	7.4%	14.1%	4.6%
LAC REGION	1856287	77.3%	10.0%	1.4%	7.4%	3.9%	1.9%	0.7%	0.7%
Caribbean	663205	78.8%	15.7%	0.2%	4.7%	0.7%	3.5%	0.1%	0.1%
Central America	596014	92.3%	3.2%	0.6%	1.5%	2.4%	0.1%	0.3%	0.2%
South America	597067	60.7%	10.5%	3.6%	16.3%	8.9%	2.0%	1.8%	1.9%

Source: Own calculations based on Docquier and Marfouk (2005).

This step requires data on the size and the skill structure of the working-aged population in the countries of origin ( $N_{t,s}^j$ ). Population data by age are provided by the United Nations<sup>6</sup>. We focus on the population aged 25 and more. Data are missing for a couple of countries but can be estimated using the CIA world factbook<sup>7</sup>. Population data are split across educational group using international human capital indicators. We use Barro and Lee (2000) data for most countries. For countries where Barro and Lee measures are missing, we transpose the skill sharing of the neighboring country with the closest human development index regarding education. This method gives good approximations of the brain drain rate, broadly consistent with anecdotal evidence. Detailed emigration rates of skilled workers are given in table A.2 in Appendix. Figure A.1.a maps skilled migration rates at the world level. The highest rates are clearly observed in Africa and Central America.

### 3 Regional disparities

This section describes the distribution of the brain drain across developing countries, compares the mobility of skilled and unskilled workers and analyzes the intradistribution dynamics of skilled migration over the last decade.

#### 3.1 The distribution of the brain drain

Regional descriptive statistics on skilled migration rates are provided in Table 2. We distinguish weighted and unweighted average migration rates. Weighted average rates are obtained by aggregating immigration and population data on the whole region. It gives the emigration rate of the region treated as a single entity. Hence, the weighted average is strongly affected by the largest countries. For example, China, India and Indonesia altogether represent 70 percent of the Asian population; Brazil represent 33 percent of the LAC region. On the opposite, the unweighted average allocates an identical weight to each regional observation whatever their relative demographic size. It gives the mean of country-specific emigration rates. We also report the (unweighted) quartiles of the regional distribution. The variables Q0 and Q100 are the minimal and maximal regional rates; Q50 is the regional median rate; Q25, and Q75 respectively correspond to the quartiles at 25 and 75 percent. Finally, four measures of dispersion are provided: the max/min ratio (Q100/Q0), the interquartile ratio (Q75/Q25), the standard error and the coefficient of variation (standard error in percent of the mean).

At the world level (i.e. including wealthy countries), the weighted average rate of migration of skilled workers amounts to 5.3 percent in 2000 (4.6 percent in 1990), about five times as large as the rate of migration of unskilled workers (1.1 percent

---

<sup>6</sup>See <http://esa.un.org/unpp>.

<sup>7</sup>See <http://www.cia.gov/cia/publications/factbook>.

in 2000). In our full sample (the LAC region, Asia and Africa), the average skilled emigration rate (7 percent) is 26 times as large as the average unskilled migration rate (0.3 percent). The unweighted average skilled migration rate in Asia (5.5 percent) is very similar to the world benchmark value. On the contrary, the average rate is twice as large in the LAC region and Africa (11.0 and 10.4 percent). Between 1990 and 2000, the brain drain has increased in all regions, especially in the LAC region. It is worth noticing that the ratio of "skilled/unskilled" rates is high in Africa and Asia, suggesting that these two regions are characterized by a stronger selection of their emigrants (see below). Data by detailed area exhibit stronger disparities<sup>8</sup>. The most affected region is the Caribbean (mostly grouping small opened islands) where the average brain drain reaches 42.8 percent. Other remarkable regions are Eastern Africa (18.6 percent), Middle Africa (16.1 percent), Western Africa (14.8 percent) and Central America (16.9 percent). It is mainly in these regions that the brain drain is severe. Note that South-Eastern Asia (9.8 percent) is also affected by the emigration of their educated workers. On the contrary, the brain drain is very low in South America, Northern and Southern Africa, and Eastern and South-central Asia.

The unweighted average rate (giving the same weight to small and large countries) is much higher than the weighted average. It amounts to 19.7 percent in 2000 (22.2 percent in 1990). As we shall see below, this is due to the fact that small countries tend to be more opened to migration. Compared to our conclusions based on weighted averages, the most affected regions are identical: the brain drain is severe in the Caribbean (67.2 percent), Eastern Africa (26.0 percent), Middle Africa (16.1 percent), Western Africa (25.6 percent) and Central America (26.6 percent).

These regional average rates are hiding a very strong heterogeneity within regions. In each region (even in the less affected), some countries exhibit drastic values. For example, whilst the South American average brain drain is extremely small, Guyana peaks at 89 percent. The Max/Min ratio ( $Q_{100}/Q_0$ ) is extremely large in all regions. It is difficult to derive a clear and unambiguous message from standard dispersion indicators. Focusing on the unweighted distribution, the Max/Min ratio is very strong in Asia (especially in South Central Asia) and in Africa. Regarding the interquartile range ( $Q_{75}/Q_{25}$ ), the highest dispersion indicators are obtained for Asia and the LAC region. The standard error statistic offers another ranking with the LAC region and Africa above Asia. Finally, the coefficient of variation completely transforms the ranking compared to the standard error given the effect of the mean in the LAC region and Africa. Such results suggests that the distribution of skilled migration rates is complex.

---

<sup>8</sup>See Tables A.2 and A.3 for the description of detailed areas.

**Table 2. Migration rates in developing countries - Descriptive statistics**

	Nb. obs.	Weighted Average	Unweighted Average	Q0 (Minimum)	Q25 (Quartile 25%)	Q50 (median)	Q75 (Quartile 75%)	Q100 (Maximum)	Ratio Q100/Q0	Ratio Q75/Q25	Standar Error	Coef. of Variation
<i>Skilled migration rate in 2000</i>												
WORLD	(195)	5.3%	19.7%	0.2%	5.8%	11.8%	25.2%	89.0%	536.6	4.4	0.211	106.9%
AFRICA	(53)	10.4%	20.2%	0.5%	7.6%	13.7%	26.0%	67.5%	140.6	3.4	0.170	84.4%
<i>Eastern Africa</i>	(17)	18.6%	26.0%	7.6%	12.4%	21.2%	35.6%	56.2%	7.4	2.9	0.157	60.2%
<i>Middle Africa</i>	(9)	16.1%	16.1%	2.4%	13.0%	14.6%	22.0%	33.0%	13.5	1.7	0.084	52.4%
<i>Northern Africa</i>	(6)	7.3%	8.8%	2.4%	5.1%	8.2%	11.8%	17.0%	7.1	2.3	0.049	55.6%
<i>Southern Africa</i>	(5)	6.8%	3.9%	0.5%	3.5%	3.6%	4.3%	7.5%	15.6	1.2	0.022	57.9%
<i>Western Africa</i>	(16)	14.8%	25.6%	2.6%	11.1%	16.4%	45.5%	67.5%	26.3	4.1	0.210	82.0%
ASIA	(50)	5.5%	9.6%	0.2%	1.7%	5.9%	14.3%	38.6%	233.1	8.3	0.101	104.9%
<i>Eastern Asia</i>	(7)	3.9%	9.7%	1.1%	2.5%	5.6%	13.6%	28.8%	27.3	5.4	0.092	95.4%
<i>South-central Asia</i>	(14)	5.3%	7.1%	0.2%	0.7%	2.8%	10.8%	29.7%	178.9	15.9	0.091	129.0%
<i>South-eastern Asia</i>	(11)	9.8%	14.8%	2.1%	7.6%	15.2%	17.0%	37.4%	18.1	2.2	0.101	68.4%
<i>Western Asia</i>	(18)	6.9%	8.4%	0.6%	2.1%	6.0%	7.7%	38.6%	68.0	3.7	0.099	118.5%
LAC REGION	(33)	11.0%	38.8%	2.2%	8.1%	28.7%	66.8%	89.0%	39.5	8.3	0.314	80.8%
<i>Caribbean</i>	(13)	42.8%	67.2%	21.6%	63.5%	71.1%	83.6%	85.1%	3.9	1.3	0.198	29.5%
<i>Central America</i>	(8)	16.9%	26.6%	7.2%	15.8%	24.3%	30.0%	65.5%	9.1	1.9	0.165	61.8%
<i>South America</i>	(12)	5.1%	16.2%	2.2%	3.8%	6.0%	9.8%	89.0%	39.5	2.6	0.249	153.6%
<i>Unskilled migration rate in 2000</i>												
WORLD	(195)	1.1%	4.3%	0.0%	0.3%	1.6%	5.2%	43.8%	2959.1	18.9	0.070	161.7%
AFRICA	(53)	0.9%	1.6%	0.0%	0.1%	0.3%	1.6%	18.7%	1261.5	11.7	0.032	198.6%
ASIA	(50)	0.4%	1.4%	0.0%	0.1%	0.5%	1.5%	15.6%	593.4	12.9	0.027	186.7%
LAC REGION	(33)	2.8%	8.4%	0.1%	0.8%	5.1%	12.1%	39.0%	298.6	14.5	0.097	114.5%
<i>Skilled migration rate in 1990 (*)</i>												
WORLD	(175)	4.6%	22.2%	0.2%	5.0%	10.4%	30.5%	96.7%	556.8	6.1	0.251	112.8%
AFRICA	(52)	10.0%	16.2%	0.2%	5.4%	9.4%	17.5%	80.4%	463.0	3.2	0.181	111.3%
ASIA	(42)	5.1%	11.5%	0.2%	2.4%	7.4%	20.6%	43.9%	196.9	8.5	0.112	97.9%
LAC REGION	(33)	9.8%	38.3%	1.5%	6.9%	29.7%	67.7%	91.1%	58.8	9.8	0.311	81.1%

Source: Own calculations based on Docquier and Marfouk (2005).

(\*) In 1990, the Ex-USSR is totally included in Europe

Another picture of the empirical distribution emerges from the graphical representation of a kernel density estimator (Silverman 1986, Härdle 1990). This non parametric approach consists in letting the data speak for themselves in determining the shape of the density. It has been used by many authors to compute the distribution of income, unemployment, GDP or Gini index. The Kernel method allows to estimate the distribution of any variable without imposing restrictions on its functional form (as with a parametric approach). We use the method to depict the salient features of the distribution of skilled migration rates, to compare the distribution of skilled and unskilled migration rates and to distinguish the main clusters of countries. We use the Gaussian kernel functional form and compute the optimal bandwidth using the plug-in rule of Sheater and Jones (1991). As shown by Park and Turloch (1992), this plug-in rule is particularly efficient in cases of complex multimodal densities.<sup>9</sup>

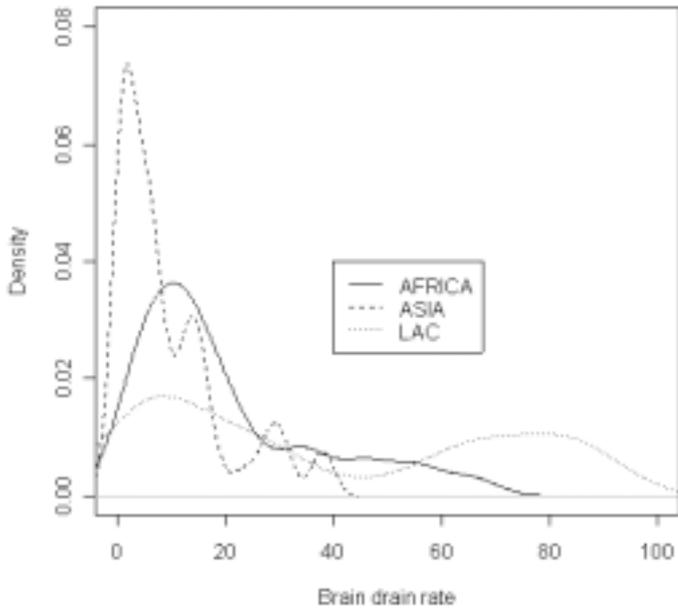
Figure 1a compares the density estimates of skilled migration rates for Africa, Asia and the LAC region in 2000. It shows that the distributions are globally multimodal in all developing regions. the LAC region and Africa exhibit a bimodal density whilst a richer multimodal structure is detected in Asia. All densities are right skewed, especially in Asia and Africa. There is clearly a dominant mode at low rates of migration and one or several less important modes at high rates. However, the range, the concentration and the position of the modes are very different. Asia exhibits the smallest range and the highest degree of concentration. We clearly distinguish one dominant mode around 4 and 6 percent (grouping Nepal, Bangladesh, Syria, Bahrain, Burma, Turkey, Korea, Yemen, India, China and Korea), a less important mode around 15 percent (Singapore, Iran, Philippines, East Timor, Brunei and Singapore), and two minor modes around both 30 (Vietnam, Sri Lanka, Hong Kong and Cyprus) and 38 percent (Laos and Lebanon). The Asian density is close to zero around 40 percent (the maximum, 38.6 percent, is reached in Lebanon). Compared to Asia, the African pattern exhibits a larger distribution range and less concentration around the modes. We observe one dominant mode at 15 percent (Gabon, Mali, Zambia, Malawi and Togo), and a highly significant proportion of countries between 25 and 65 percent. This reveals that a significant number of African countries are severely affected by the brain drain. The maximum is reached in Cape Verde (67.5 percent), but countries such as Gambia, Seychelles, Mauritius, Sierra Leone, Ghana, Mozambique and Liberia exhibit high rates. Finally, the LAC region distribution is more uniform and less concentrated. Two modes of similar magnitude are detected. The first mode is slightly more important. Basically, the left mode depicts South America and the right one depicts the Caribbean. Most Central American countries are located between these two modes. The non parametric approach clearly suggests that the dispersion of the brain drain is high in the LAC region, intermediate in Africa and lower in Asia.

---

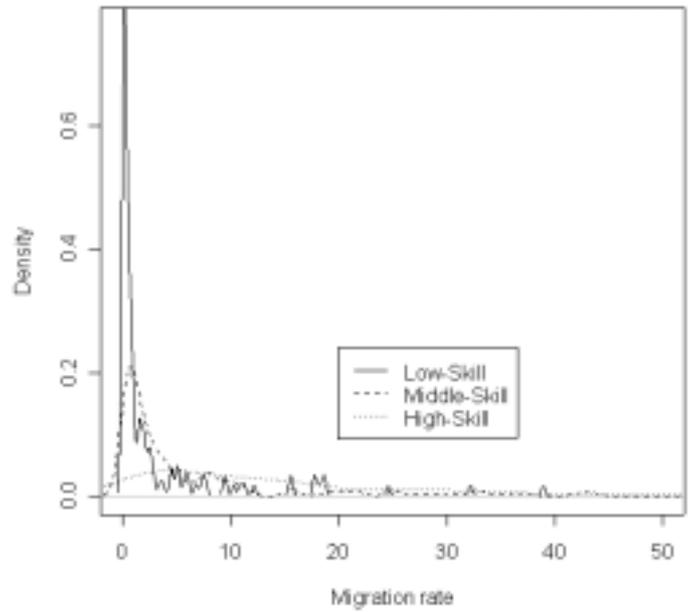
<sup>9</sup>Our kernel estimates are robust to the choice of the bandwidth rule.

**Figure 1. Kernel density estimators by region, education level and cluster**

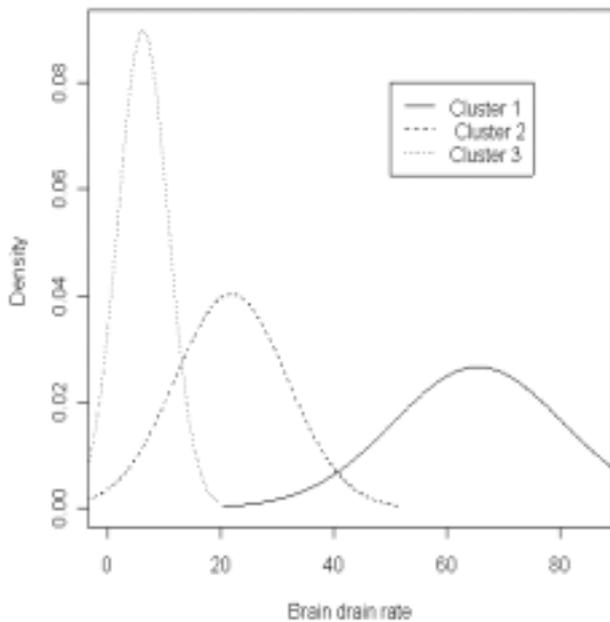
*1.a. Skilled migration rates by region*



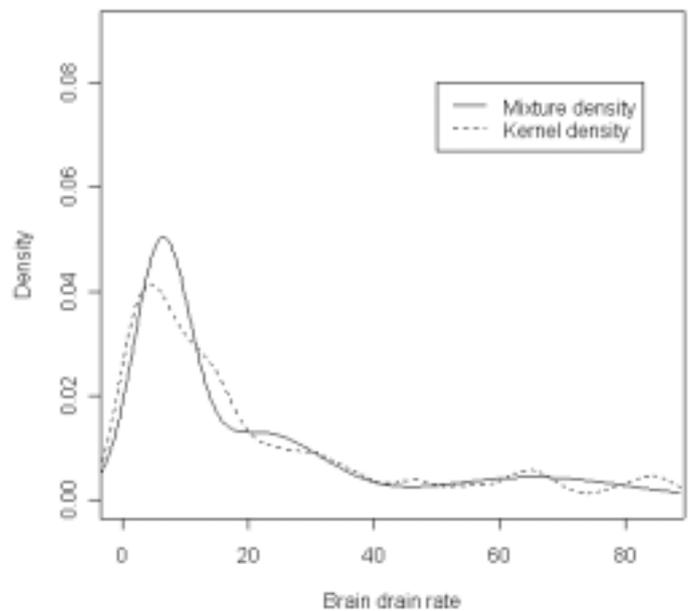
*1.b. Migration rates by education level (full sample)*



*1.c. Mixture components of skilled migration rates – clusters (full sample)*



*1.d. Kernel and mixture densities of skilled migration rates (full sample)*



These three regions are aggregated on Figure 1b which compares skilled and unskilled migration rates. It shows that the multimodality property of skilled migration rates is less clear at the aggregated level. This is due to the fact the regional modes correspond to different regional rates. However, skilled migration is clearly more diffused than low-skill and medium-skill migration. The range of the distribution obviously increases with educational attainment, whilst the concentration (captured by the density at the highest mode) is decreasing with education. Note that the distribution of low skilled migration is heavy tailed and strongly right skewed.

At the full sample level, formal tests such as Hartigan and Hartigan’s (1985) Dip test or Silverman (1981) test confirm the multimodality. It is worth investigating whether multimodality can be related to the existence of several groups of countries sharing similar characteristics (in terms of geography, demography, economy, political regime, etc.). We explore that possibility by applying a hierarchical Gaussian mixture model (Fraley and Raftery, 2002) that disentangles the global distribution into unimodal distributions of a given Gaussian form<sup>10</sup>. Our rationale is that multiple modes are not purely random, but result from the aggregation of various standard parametric distributions of different mean and variance. The mixture model reveals the existence of three components (i.e. three clusters of countries). It also determines the parameters of these three distributions and the probability that each country belongs to a particular group. We consider that a country belongs to one cluster when the corresponding probability exceeds 50 percent. Figure 1c depicts the densities within each group. Figure 1d compares the true world distribution (estimated by the Kernel method above) and the mixture density obtained by aggregating the three densities in Figure 1c. This comparison shows that our mixture decomposition matches the true distribution in a very efficient way. The groups are defined as follows:

- Group 1 (22 countries) is depicted by a Gaussian distribution with a mean and a standard error of 65.6 percent and 226.0 respectively. It is mainly made of small countries (such as the Caribbean countries, Cape Verde, etc.) and two medium-sized African countries (Ghana and Mozambique). Group 1 contains Antigua and Barbuda, the Bahamas, Barbados, Belize, Cape Verde, Dominica, Gambia, Ghana, Grenada, Guyana, Haiti, Jamaica, Liberia, Mauritius, Mozambique, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Seychelles, Sierra Leone, Suriname, Trinidad and Tobago.
- Groupe 2 (39 countries) is depicted by a Gaussian distribution with a mean and a standard error of 21.9 percent and 98.1 respectively. It is mainly made of small and relatively wealthy countries (Brunei, Macao and Cyprus), the majority of medium-sized countries and several large and relatively poor countries (Kenya, Uganda and Vietnam). Group 2 contains Afghanistan, Angola, Brunei,

---

<sup>10</sup>Formally, Fraley and Raftery’s method combines hierarchical agglomeration clustering, maximization algorithm for mixture models and Bayes factor for selection model. For the mixture model, we used the MCLUST package in R statistical computing environment.

Cambodia, Cameroon, Hong Kong, Macao, Comoros, Rep. of the Congo, Cuba, Cyprus, Dominican Republic, East Timor, El Salvador, Eritrea, Gabon, Guatemala, Guinea-Bissau, Honduras, Iran, Kenya, Laos, Lebanon, Malawi, Mali, Mexico, Morocco, Nicaragua, Panama, Rwanda, Sao Tome and Principe, Senegal, Singapore, Somalia, Sri Lanka, Togo, Uganda, Vietnam and Zambia.

- Group 3 (75 countries) is depicted by a Gaussian distribution with a mean and a standard error of 6.2 percent and 19.8 respectively. It is mainly made of oil producers (such as the Gulf countries), ex-USSR countries, very large countries and medium-sized countries with relatively high human capital level. Group 3 contains Algeria, Argentina, Armenia, Azerbaijan, Bahrain, Bangladesh, Benin, Bhutan, Bolivia, Botswana, Brazil, Burkina Faso, Burma (Myanmar), Burundi, Central African Republic, Chad, Chile, China, Colombia, Dem. Rep. of the Congo, Costa Rica, Cote d'Ivoire, Djibouti, Ecuador, Egypt, Equatorial Guinea, Ethiopia, Georgia, Guinea, India, Indonesia, Iraq, Israel, Japan, Jordan, Kazakhstan, Korea, Kuwait, Kyrgyzstan, Lesotho, Libya, Madagascar, Malaysia, Maldives, Mauritania, Mongolia, Namibia, Nepal, Niger, Nigeria, Occupied Palestinian Territory, Oman, Pakistan, Paraguay, Peru, Philippines, Qatar, Saudi Arabia, South Africa, Sudan, Swaziland, Syria, Taiwan, Tajikistan, Tanzania, Thailand, Tunisia, Turkey, Turkmenistan, United Arab Emirates, Uruguay, Uzbekistan, Venezuela, Yemen and Zimbabwe.

The unweighted average skilled migration rate is extremely high within the first group (66.2 percent, i.e. 3.3 times as large as the sample mean). The variance is quite large as well. Group 2 exhibits intermediate average migration rate (23.7 percent) and variance. The variance and the average migration rate (6.0 percent) are lower within group 3. The weighted average rates respectively amount to 72.4, 21.0 and 4.7 percent in Group 1, 2 and 3. The main difference between these groups concerns their average demographic size. Whilst the average population size in the full sample amounts to 36.9 million of inhabitants, the average size is 2.8 million in Group 1, 14.2 million in group 2 and 58.7 million in Group 3. Not surprisingly, large countries tend to be less opened to migration. Nevertheless, the population size is not the only important determinant of skilled emigration rates. Small countries such as Comoros, Cyprus, Sao Tome and Principe belong to group 2 – despite a population size lower than one million; Maldives, Qatar and Djibouti belong to group 3 despite small population sizes as well. On the contrary, Ghana belongs to group 1 despite a 20 million population size. Large countries such as Vietnam (78 million), Kenya (30 million), Morocco (29 million), Uganda (23 million) and Afghanistan (21 million) belong to group 2, with a relatively high emigration rate given their size. We will argue below that skilled emigration rates are likely to result from various potential factors.

### 3.2 The 1990-2000 dynamics

Many anecdotal evidence reveal that skilled migration is a fast-evolving process which is likely to further accelerate in the coming decades as a part of the global transformation affecting the world economy. Table 2 confirms these trends by showing that, despite a general rise in the educational attainment of natives, the weighted average rate of skilled migration has increased by 0.7 point of percentage between 1990 and 2000 at the world level (note that the stock of skilled immigrants in the OECD increased by 64 percent). Such a rise is observed in all developing regions. Nevertheless, looking at the unweighted averages, things are less clear. Whilst the unweighted average has increased in the LAC region, it has clearly decreased in Asia and Africa. These features suggest that the dynamics of skilled migration is a complex process affecting countries in a very heterogeneous way.

Building on Quah (1997), the intradistribution dynamic pattern can be characterized on the basis of the stochastic kernel method. This approach estimates the probability to move from one point to another in the distribution between two periods. We now introduce time and compute the distribution change between 1990 and 2000. This approach has been applied to various economic problems. To the best of our knowledge, there is no study applying the technique to migration rates. We apply this method to the countries for which observations are available in 1990 and 2000 (this eliminates countries from the former Soviet block, former Yugoslavia, former Czechoslovakia, East Timor and Eritrea). Figure 2a, 2c and 2e depicts the kernel transition densities estimated by continent. At each point, the 3D-function measures the probability (vertical axis) to exhibit a given migration rate in 2000 (left horizontal axis) conditional to its migration rate in 1990 (right horizontal axis). Figures 2b, 2d and 2f give the contour plots associated to the kernel transition density. Each center of the contour plot corresponds to a peak in the transition density. Contour plots usually offer a complementary picture compared to Table 2. We conclude to small dynamic movements when the graph is concentrated along the diagonal. Conversely, a stronger gap with the diagonal reveals important changes in country rankings.

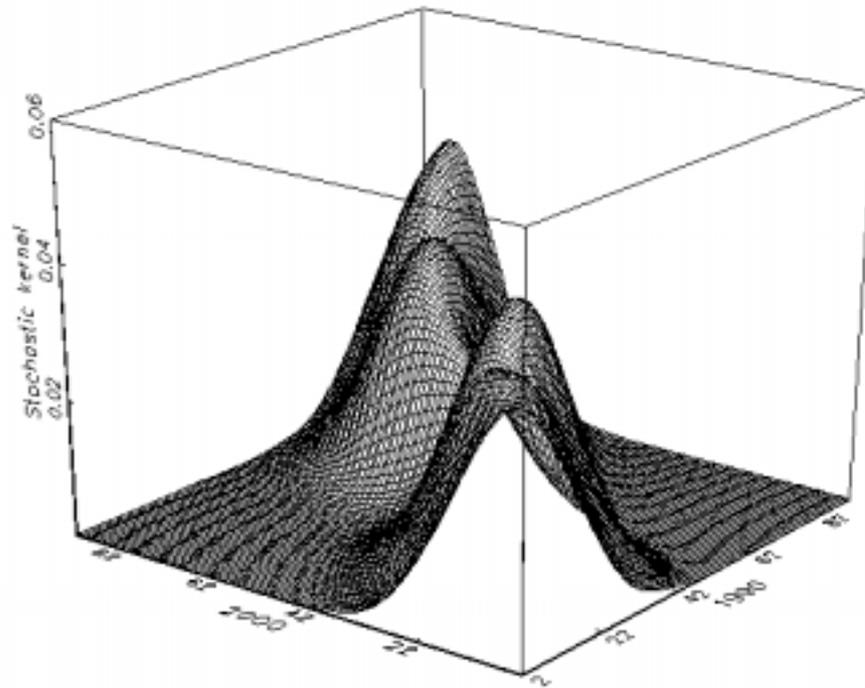
The LAC pattern reveals four clubs of comparable size. Changes are rather small in the LAC region. The conditional density is very close from the diagonal, suggesting that the brain drain is very persistent<sup>11</sup>. The conditional distribution exhibits a peak at (88;85), one at (62;64), one at (34;36) and the other at (4;6). As in other regions, the unweighted average rate has slightly decreased at the top of the distribution (Guyana and Jamaica) and has increased among the least affected countries (Brazil and Paraguay). The Q100/Q0 and Q75/Q25 ratios in table 2 confirm a minor convergence process between the extremities of the distribution.

---

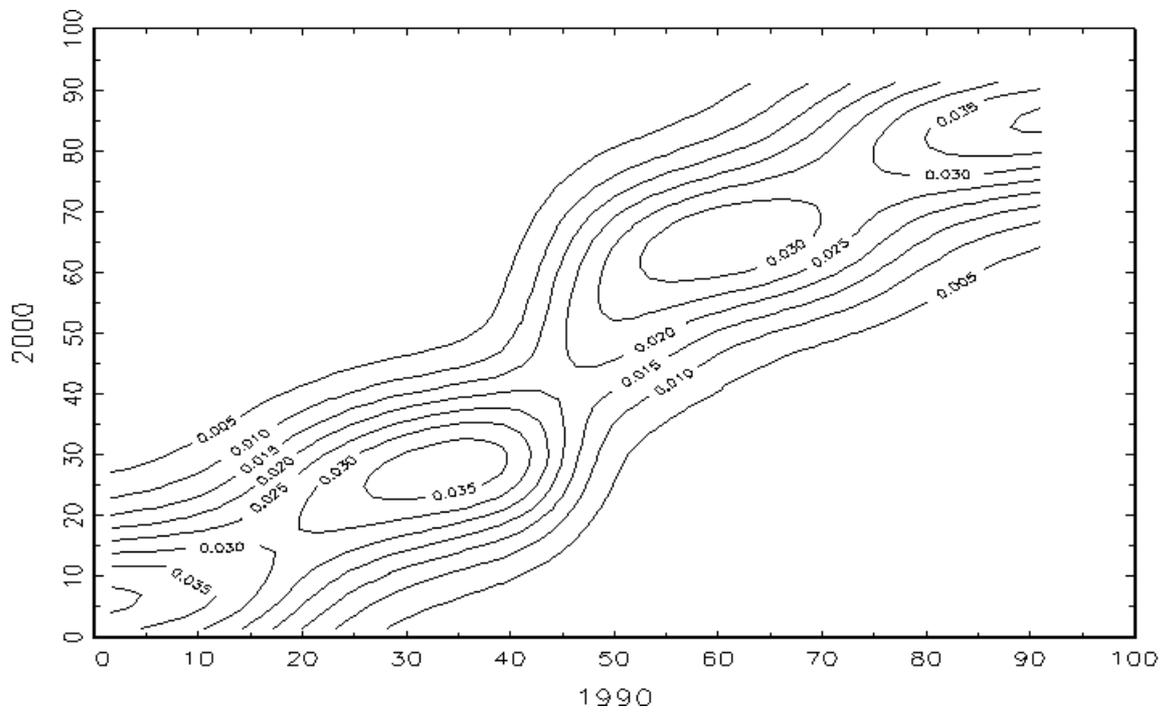
<sup>11</sup>Each pair gives the skilled migration rate in 1990 and 2000.

## Figure 2. Dynamic pattern of skilled migration

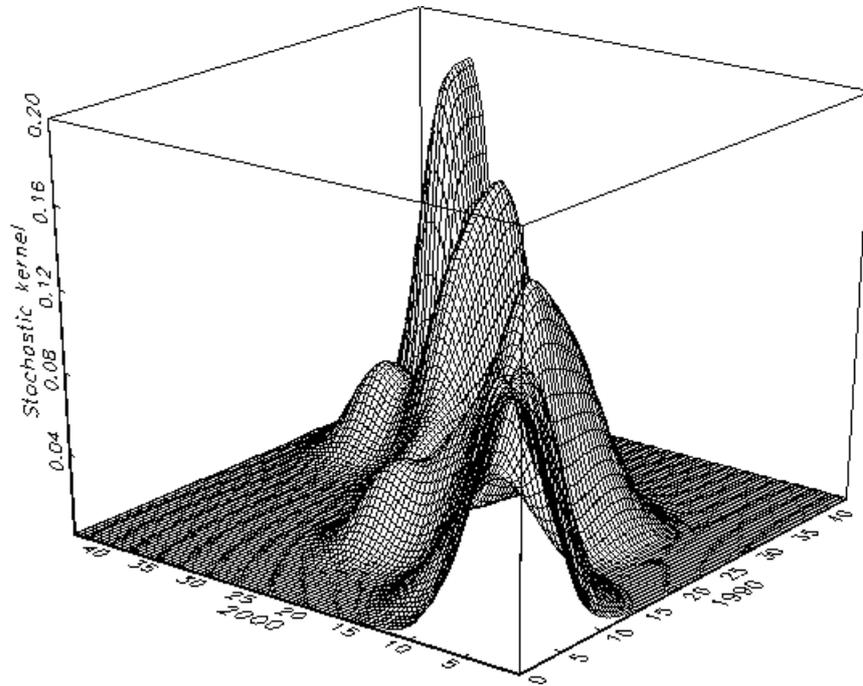
### 2.a. Stochastic kernel in the LAC region



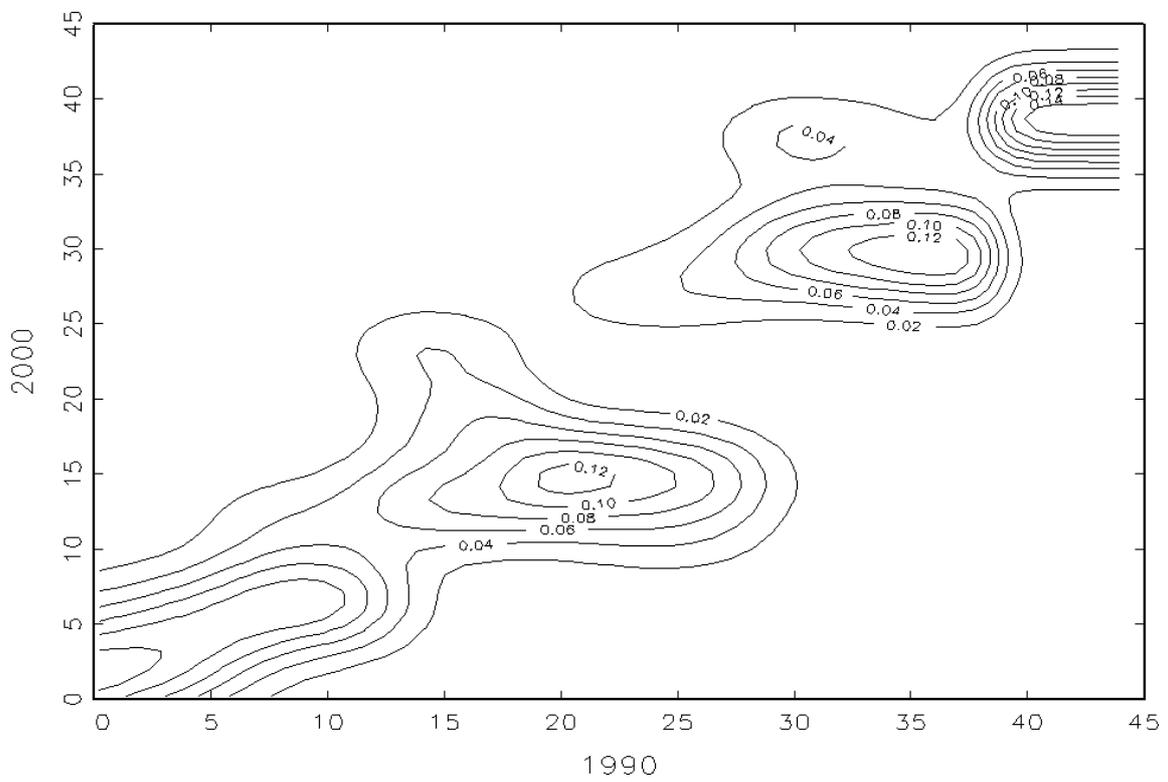
### 2.b. Countour plot in the LAC region



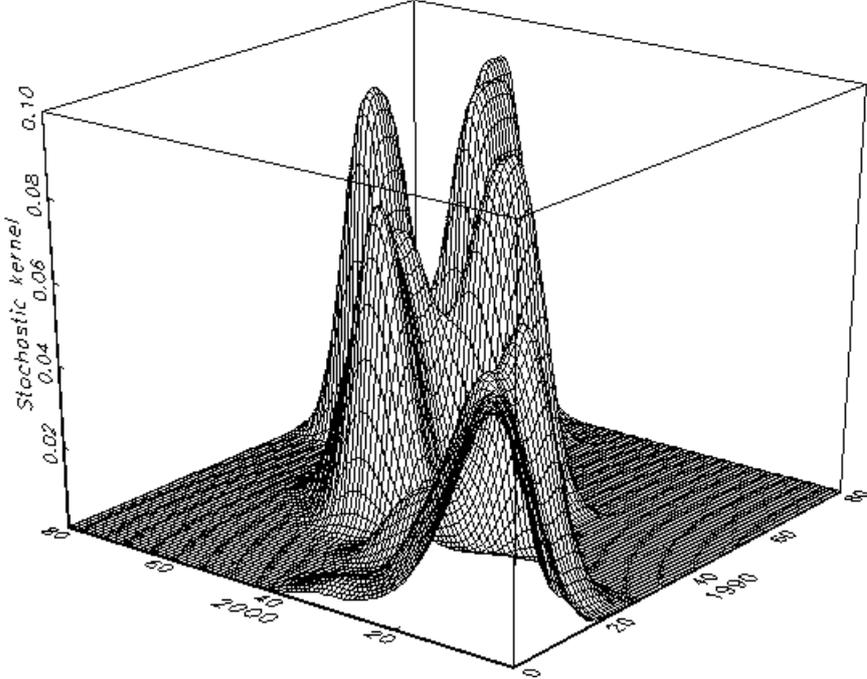
2.c. Stochastic kernel in Asia



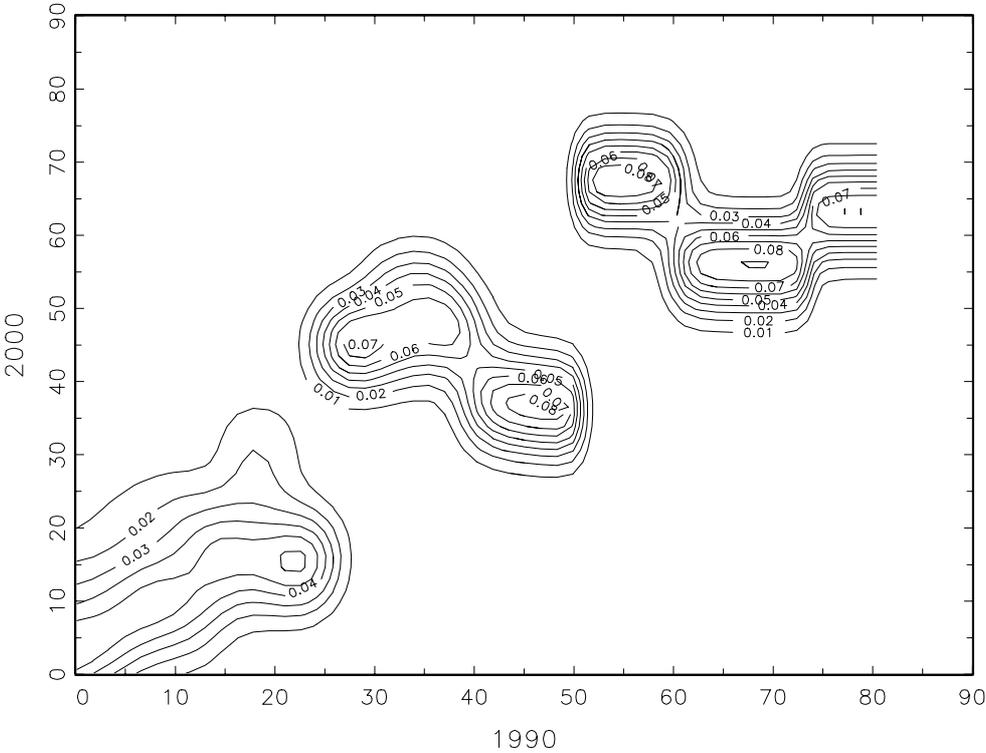
2.c. Countour plot in Asia



2.e. Stochastic kernel in Africa



2.f. Countour plot in Africa



The Asian patterns reveal similar changes. Four main clubs can be distinguished. The brain drain has increased (from 1.25 to 2.5 percent) in the least affected countries. This group contains two important countries (India and China) as well as Nepal and Gulf countries such as Kuwait and Qatar. The fact that the brain drain has increased in India and China explains why the weighted average rate increased in Asia as a whole. Brain drain has decreased in the three other clubs. Peaks are observed at (21;14) with countries such as Brunei and Macao, and (35;30) or (40;37) with Lebanon, Cyprus and Hong Kong.

Finally, the richest pattern is observed in Africa where about six clubs can be distinguished. A sharp rise in skilled migration was experienced in two clubs at peaks (35;45) and (55;65) representing countries such as Ghana, Sierra Leone, Mozambique, Mauritius, Somalia and Cape Verde. On the contrary, the brain drain has decreased in the most affected clubs at (75;60) and (67;55), as well as in a couple of intermediate countries at (45;35) and (22;15). These groups contain the Gambia, Seychelles, Rep. of Congo, Morocco, Tunisia, Uganda and Kenya. In addition to these clubs, and as in the other areas, the brain drain has increased in the least affected countries (such as Libya, Chad, Burkina Faso, Equatorial Guinea and Swaziland). We note that the average skilled migration rates of African countries which experienced a recent genocide (10.3 percent) is higher than the African average and has increased by 2.3 percentage points between 1990 and 2000<sup>12</sup>.

## 4 A tautological decomposition

A high brain drain can be the result of several (endogenous) factors. It can be due to a high degree of openness, to a strong selection among migrants (i.e. to self-selection or out-selection policies) or simply to the fact that skilled migrants are drawn from a small number of educated native born. What are the salient features of the brain drain in the developing regions? A simple tautological decomposition allows to emphasize the basic factors at play in the LAC region, Asia and Africa. Tautologically, the emigration rate of skilled workers can be disentangled into three components:

$$\begin{aligned} m_{t,h}^j &= \frac{M_{t,h}^j}{N_{t,h}^j + M_{t,h}^j} = \frac{\sum_s M_{t,s}^j}{\sum_s (N_{t,s}^j + M_{t,s}^j)} \times \frac{M_{t,h}^j}{\sum_s M_{t,s}^j} \times \frac{\sum_s (N_{t,s}^j + M_{t,s}^j)}{N_{t,h}^j + M_{t,h}^j} \\ &= \bar{m}_t^j \times c_t^j \times d_t^j \end{aligned}$$

The first component ( $\bar{m}_t^j$ ) is the mean or global emigration rate calculated over all educational levels; the second component ( $c_t^j$ ) is a composition rate measuring the share of tertiary educated among emigrants; the third component ( $d_t^j$ ) is an index of human capital deficiency given by the inverted share of skilled workers among native born (including the natives living abroad). Intuitively, the second and third

---

<sup>12</sup>These 10 countries are identified in Easterly (2000).

components are obviously related. If a country exhibits a high share of educated native born, it is more likely that its proportion of educated among emigrants will also be high. To neutralize this supply effect, we define the selection bias,  $\beta_t^j$ , as the product of the composition rate by the human capital deficiency index. This product measures both the ratio of skilled/total migration rates and the ratio of the shares of educated among emigrants/native-born:

$$\beta_t^j = c_t^j \times d_t^j = \frac{m_{t,h}^j}{\bar{m}_t^j}$$

Table 3 provides descriptive statistics on these various components by region<sup>13</sup>. In terms of global migration ( $\bar{m}_t^j$ ), the strongest rates are observed in the LAC region (in the Caribbean, Central America and a pair of South American countries: Guyana and Suriname), far above the Asian and African rates (despite high average migration rates in Western Asia and Northern Africa). Regarding selection biases ( $\beta_t^j$ ), Asia (7.1) and Africa (7.0) exhibit very important biases compared to the LAC region (2.1). On average, the most important values are observed in Eastern (18.9), Middle (16.4) and Western Africa (15.1) and, to a lower extent, in South-Central Asia (10.0). Looking at the unweighted average, Africa suffers from severe biases. The maxima are observed in Niger (93.0), Rwanda (71.8), Malawi (58.7) or Mozambique (49.9), Lesotho (47.5). In Asia, the highest biases are in Nepal (17.8), China (17.1) and Burma (16.4). Haiti (7.3), Paraguay (5.3) and the Bahamas (5.1) are the most affected LAC countries. All dispersion indicators show that heterogeneity is high in Africa, intermediate in Asia and low in the LAC region.

Differences in selection biases are related to the composition of migration and to the human capital structure of residents. As far as composition is concerned, the highest proportions of educated migrants are found among Asian emigrants, whilst Africa and the LAC region exhibit lower composition rates. Unsurprisingly, composition rates tend to be stronger in the less poor regions as Southern Africa or Eastern Asia. Regarding human capital deficiency, we note the dramatic African position (unweighted average index of 52.2, with a maximal index of 276.9 observed in Mozambique) compared to Asia and the LAC region (unweighted averages of 13.1 and 7.2 respectively)<sup>14</sup>. At the world level, migration rates and selection biases exhibit very different patterns within and across regions. What about the heterogeneity across and within the clusters distinguished in the previous section? Are selection biases and global migration rates homogeneous within clusters? Looking at the interquartile range and the coefficient of variation, Table 3 reveals a relative homogeneity of global migration rates compared to the regional distributions. The migration component is strong in group 1, intermediate in group 2 and low in group 3. On the contrary, the selection bias is more heterogeneous within clusters than within regions.

<sup>13</sup>Table A.2 in appendix gives the detailed numbers by country and by region.

<sup>14</sup>In appendix 7.2, we apply the Kernel method to each component by continent in 2000. Using the mixture model, we distinguish three clusters in terms of migration rates and selection biases.

**Table 3. Tautological decomposition - Descriptive statistics**

	Weighted Average	Unweighted Average	Q0	Q25	Q50	Q75	Q100	Ratio Q100/Q0	Ratio Q75/Q25	Standar Error	Coef. of Variation
<b>I. Migration effect (MIG)</b>											
AFRICA	1.5%	2.6%	0.0%	0.5%	0.9%	2.8%	25.1%	695.8	6.1	0.044	173.3%
Eastern Africa	1.0%	2.9%	0.3%	0.5%	0.9%	2.5%	19.5%	61.2	5.2	0.048	167.7%
Middle Africa	1.0%	2.1%	0.1%	0.6%	0.8%	3.0%	5.6%	55.4	5.4	0.019	91.7%
Northern Africa	2.9%	3.2%	0.3%	0.9%	2.7%	5.1%	7.6%	26.5	5.9	0.027	83.9%
Southern Africa	1.0%	0.4%	0.0%	0.1%	0.3%	0.4%	1.3%	36.9	4.7	0.005	108.3%
Western Africa	1.0%	2.9%	0.1%	0.5%	1.2%	2.4%	25.1%	392.3	4.3	0.058	198.7%
ASIA	0.8%	2.5%	0.0%	0.3%	1.4%	3.4%	17.2%	346.3	10.8	0.035	139.0%
Eastern Asia	0.5%	3.0%	0.2%	0.4%	2.8%	4.0%	9.1%	40.6	9.4	0.029	98.7%
South-central Asia	0.5%	0.7%	0.0%	0.2%	0.3%	0.9%	2.8%	56.4	4.3	0.008	114.3%
South-eastern Asia	1.6%	3.2%	0.2%	1.1%	3.4%	3.7%	10.0%	44.2	3.4	0.027	83.4%
Western Asia	3.5%	3.4%	0.1%	0.5%	1.8%	3.2%	17.2%	151.5	6.8	0.048	140.1%
LAC REGION	5.3%	17.8%	0.4%	2.9%	11.5%	32.3%	53.7%	121.2	11.3	0.165	92.6%
Caribbean	15.3%	29.3%	10.5%	13.1%	32.3%	37.6%	53.7%	5.1	2.9	0.141	48.1%
Central America	11.9%	12.1%	3.1%	7.6%	8.9%	14.3%	29.0%	9.3	1.9	0.079	65.6%
South America	1.6%	9.3%	0.4%	1.0%	2.4%	3.8%	47.4%	106.9	3.6	0.160	172.1%
CLUSTERS											
Group 1	9.4%	24.7%	0.9%	10.9%	25.2%	37.3%	53.7%	58.4	3.4	0.166	67.3%
Group 2	6.0%	5.4%	0.3%	1.8%	3.4%	8.5%	20.0%	62.6	4.8	0.049	91.3%
Group 3	0.7%	1.3%	0.0%	0.3%	0.6%	1.8%	6.6%	182.7	6.2	0.015	115.4%
<b>II. Selection bias (BIAS)</b>											
AFRICA	7.0	18.7	2.1	7.4	15.6	23.3	93.0	43.9	3.2	17.378	93.1%
Eastern Africa	18.9	23.8	2.9	9.8	20.3	24.0	71.8	25.1	2.5	18.813	79.1%
Middle Africa	16.4	15.0	2.9	7.4	17.4	21.8	24.5	8.4	2.9	8.156	54.5%
Northern Africa	2.5	6.5	2.1	2.3	2.4	4.6	24.3	11.5	2.0	8.037	124.3%
Southern Africa	6.8	17.4	5.6	8.4	12.3	13.3	47.5	8.4	1.6	15.294	87.7%
Western Africa	15.1	20.2	2.7	10.8	15.6	22.2	93.0	34.6	2.1	19.880	98.2%
ASIA	7.1	5.9	1.0	2.8	4.0	7.5	17.9	17.3	2.7	4.510	76.7%
Eastern Asia	8.5	4.9	2.0	2.7	3.2	3.5	17.1	8.4	1.3	4.999	101.3%
South-central Asia	10.0	8.3	2.3	3.1	9.1	12.0	17.9	7.8	3.9	5.062	60.8%
South-eastern Asia	6.0	6.3	2.7	4.2	4.6	7.6	16.4	6.0	1.8	3.679	58.3%
Western Asia	1.9	4.1	1.0	2.5	3.1	4.8	16.2	15.7	1.9	3.205	78.5%
LAC REGION	2.1	2.8	1.0	1.8	2.4	3.1	7.2	7.2	1.8	1.323	47.9%
Caribbean	2.8	2.8	1.6	1.6	2.3	3.1	7.2	4.6	1.9	1.599	57.5%
Central America	1.4	2.4	1.2	2.1	2.5	2.8	3.3	2.7	1.3	0.663	27.9%
South America	3.2	3.0	1.0	2.1	3.0	3.5	5.3	5.3	1.7	1.263	42.2%
CLUSTERS											
Group 1	7.7	8.0	1.0	2.0	2.8	6.7	49.1	48.6	3.4	11.275	141.4%
Group 2	3.5	11.4	1.2	3.1	5.5	14.4	71.8	58.4	4.7	14.539	128.0%
Group 3	6.6	10.1	1.0	2.9	5.1	13.9	93.0	89.7	4.8	12.875	127.8%
<b>IIa. Composition effect (SEL)</b>											
AFRICA	30.9%	37.9%	10.9%	21.9%	40.8%	50.4%	65.0%	6.0	2.3	0.156	41.3%
Eastern Africa	40.8%	40.2%	13.4%	37.1%	43.4%	48.5%	55.0%	4.1	1.3	0.115	28.6%
Middle Africa	30.9%	35.1%	12.4%	18.4%	40.5%	48.0%	52.9%	4.3	2.6	0.144	41.1%
Northern Africa	19.6%	34.5%	12.9%	14.3%	33.5%	53.6%	58.9%	4.6	3.8	0.206	59.8%
Southern Africa	62.1%	50.7%	33.9%	49.6%	51.1%	56.1%	62.6%	1.8	1.1	0.095	18.8%
Western Africa	42.0%	34.1%	10.9%	19.5%	30.4%	49.7%	65.0%	6.0	2.6	0.169	49.4%
ASIA	46.8%	49.8%	8.8%	41.5%	50.3%	59.3%	78.0%	8.8	1.4	0.133	26.7%
Eastern Asia	55.5%	60.3%	48.0%	54.6%	61.1%	62.9%	78.0%	1.6	1.2	0.088	14.6%
South-central Asia	52.5%	46.6%	29.8%	39.9%	46.9%	51.9%	60.5%	2.0	1.3	0.089	19.0%
South-eastern Asia	51.4%	46.0%	25.3%	32.8%	46.3%	58.2%	67.1%	2.6	1.8	0.145	31.5%
Western Asia	22.9%	50.4%	8.8%	44.3%	51.3%	61.5%	69.6%	7.9	1.4	0.145	28.8%
LAC REGION	26.3%	40.0%	14.4%	37.8%	42.7%	46.9%	60.1%	4.2	1.2	0.111	27.9%
Caribbean	38.6%	41.9%	26.5%	38.6%	42.7%	44.4%	52.4%	2.0	1.2	0.065	15.5%
Central America	16.6%	32.3%	14.4%	19.1%	31.1%	42.3%	57.7%	4.0	2.2	0.145	44.7%
South America	41.2%	43.0%	18.4%	41.6%	44.9%	47.6%	60.1%	3.3	1.1	0.103	23.9%
CLUSTERS											
Group 1	40.0%	38.9%	15.2%	36.9%	42.0%	44.4%	58.5%	3.9	1.2	0.116	29.9%
Group 2	25.5%	35.8%	10.9%	21.4%	39.5%	47.3%	61.9%	5.7	2.2	0.153	42.6%
Group 3	45.5%	47.5%	8.8%	41.1%	49.2%	55.8%	78.0%	8.8	1.4	0.136	28.7%
<b>IIb. Human Capital Deficiency Index (HCD)</b>											
AFRICA	22.8	52.2	4.7	23.4	42.0	51.8	276.9	58.7	2.2	48.618	93.1%
Eastern Africa	46.4	64.6	7.7	34.6	44.8	52.2	276.9	35.9	1.5	64.314	99.6%
Middle Africa	53.1	41.9	18.2	23.4	43.9	51.8	69.0	3.8	2.2	18.059	43.1%
Northern Africa	12.7	18.1	4.7	10.5	15.4	16.9	46.7	9.9	1.6	13.502	74.7%
Southern Africa	10.9	35.5	9.0	23.7	24.1	24.8	95.8	10.6	1.0	30.718	86.6%
Western Africa	35.9	63.0	17.7	36.1	48.5	80.3	188.2	10.6	2.2	45.012	71.5%
ASIA	15.2	13.1	3.4	5.6	7.9	18.1	47.2	14.0	3.2	11.279	85.8%
Eastern Asia	15.4	9.1	3.8	4.3	4.7	5.5	35.7	9.5	1.3	10.879	119.6%
South-central Asia	19.1	19.3	5.5	5.9	16.5	30.9	39.8	7.2	5.3	13.117	67.8%
South-eastern Asia	11.6	14.6	4.1	8.5	14.5	18.9	29.2	7.1	2.2	6.934	47.6%
Western Asia	8.5	9.0	3.4	5.6	6.6	7.9	47.2	14.0	1.4	9.471	104.9%
LAC REGION	7.9	7.2	3.6	5.0	6.1	8.6	18.6	5.2	1.7	3.364	46.7%
Caribbean	7.3	6.7	3.6	4.2	5.6	7.1	18.6	5.2	1.7	3.870	57.8%
Central America	8.5	8.4	4.5	5.4	8.3	9.7	14.3	3.2	1.8	3.314	39.3%
South America	7.8	6.9	4.3	5.1	5.8	8.0	11.7	2.7	1.6	2.489	35.9%
CLUSTERS											
Group 1	19.2	28.4	3.6	4.9	6.9	18.5	276.9	77.5	3.7	58.3	205.2%
Group 2	13.6	33.7	4.5	8.4	18.2	40.1	148.6	32.9	4.8	36.7	109.0%
Group 3	14.5	23.0	3.4	5.8	10.6	33.6	188.2	56.1	5.8	27.8	120.9%

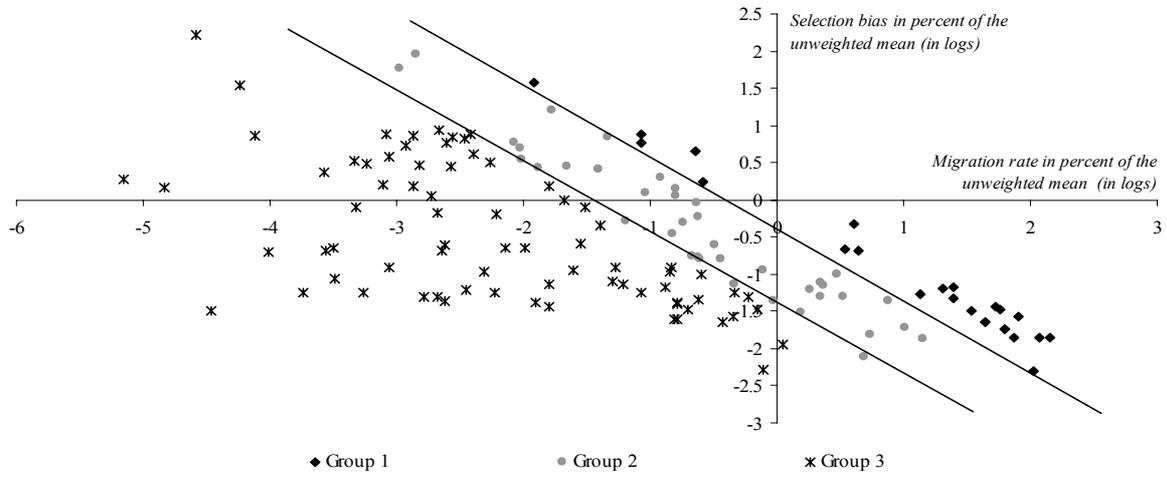
Figure 3a shows that each cluster contains both countries with high migration rates and countries with high selection biases. We evaluate the selection bias and the total migration rate in percentage of the full sample (unweighted) mean, and then take the log of these percentages. Figure 3a characterizes the position all the countries and depicts the dispersion within and across clusters. It is worth noting that no country simultaneously suffers from a high selection bias and a high migration rate. Group 1 is essentially made of two types of countries; Caribbean countries suffering from high migration rates and relatively low selection biases, and a couple of African countries characterized by intermediate migration rates and strong selection biases (Gambia, Ghana, Liberia and Mozambique).

A similar distinction is observed within cluster 2. We distinguish African countries with low migration rates and high selection biases (Malawi, Rwanda, Cameroon, Zambia, Gabon and Mali), from countries with medium migration rates and low or medium selection biases (Mexico, Dominican Republic, Lebanon, Cyprus and El Salvador). Finally, group 3 mainly contains countries with relatively low migration rates such as countries from former USSR, the largest countries and the richest medium-sized countries in which the incentive to leave is small.

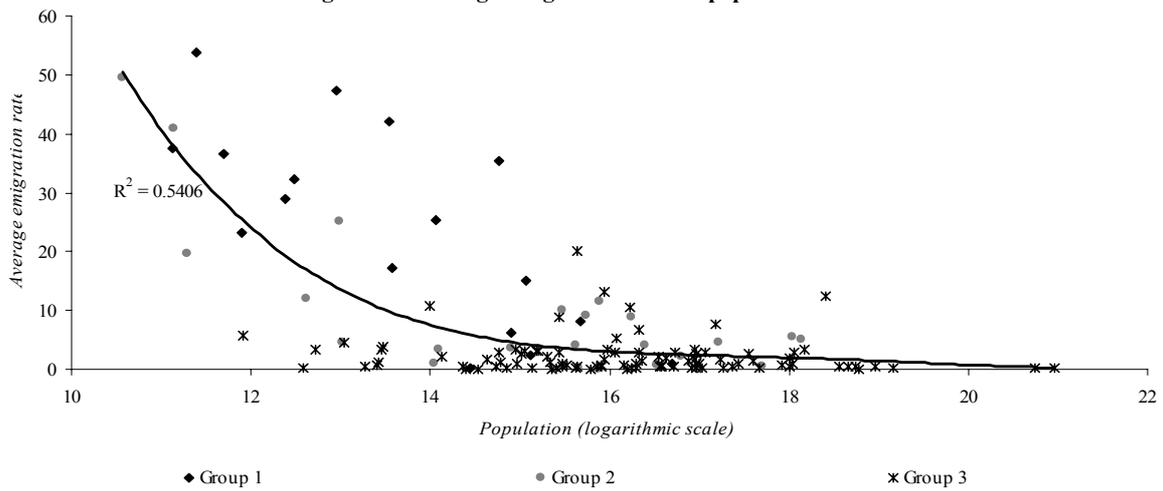
Understanding the driving forces which explain the openness to migration and the selection bias is a major challenge for economists and policymakers. Figure 3b shows that the global migration rate, reflecting the degree of openness of the country, is strongly related to the country size. An adjusted polynomial of degree 3 in the log of population size captures 54 percent of the variance of the global migration rate. Obviously, the economic literature on international migration distinguishes many potential determinants that could explain the remaining share of the variance. The neoclassical approach, further reflected in human capital theories or in the Harris-Todaro model, predicts that migration responds to economic variables such as differences in net wages, in employment rates and in other amenities (see Sjaastad, 1962, Harris and Todaro, 1970). The political conditions (political regime, degree of freedom, civil liberties and political rights) in the source country may also act as push factors. Non economic factors such as distances, climate and random events are also key ingredients in determining the costs of moving to a foreign country. Since migration costs can be of psychological or monetary natures, the concept of distance should include geographic, linguistic or cultural disparities: the more distant the new culture is, the higher the migration barriers are. These migration costs are clearly counteracted by network externalities, i.e. by a set of interpersonal ties connecting migrants and former migrants at destination (see Massey et al, 1993, Carrington et al., 1996). If the ethnic group is largely represented in the destination country, psychological and informational costs tend to be smaller.

**Figure 3. Clusters, global migration rates and selection biases**

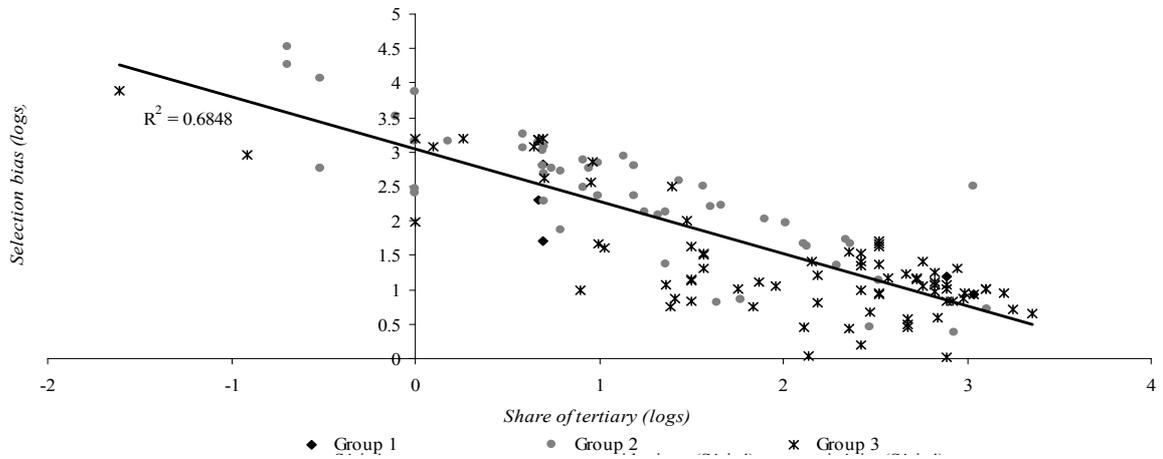
**Figure 3.a. Migration rate and selection bias by cluster**



**Figure 3.b. Average emigration rate and population size**



**Figure 3.c. Selection bias and Human capital stock (share of tertiary)**



Whilst the factors of migration have been largely investigated in empirical studies (see Pedersen et al, 2004, for a recent panel data analysis), understanding the macroeconomic factors which lead to selection biases is a critical issue. As argued by Borjas (1999), the skill composition of migration flows is endogenous. Wage dispersion and welfare policies at destination are seen as potential determinants. Analyzing the selection bias allows to explain whether skilled migrants react more or less strongly to the traditional push and pull factors. Figure 3c shows that the selection bias is strongly related to the supply of human capital of natives (measured by the share of educated). A log-linear relationship between these two variables captures 68 percent of the selection bias variance. The higher the level of education of natives, the lower the relative pressure of skilled migration. Of course, this picture leaves intact the issue of endogeneity of human capital (in particular its relationship with migration rates), and the characterization of the remaining share of the variance.

In the next section, we make one step in that direction by characterizing the spatial, cultural, political and economic autocorrelations in the components of the brain drain.

## 5 A step toward empirical analysis

Obviously, an empirical assessment of the skilled migration factors requires regressing our estimated rates on their potential determinants. Since current emigration stocks depend on present and past decisions about migration, a panel data analysis based on the historical path of push and pull factors is particularly suitable. Collecting panel data on explanatory variables for a large set of developing countries is a complex task. The purpose of this section is less ambitious whilst revealing a number of important insights for further econometric analysis. Disregarding time lags, causality and possible interactions between the push and pull factors of migration, we build on Eff (2004) to generalize spatial autocorrelation tests and assess whether proximate countries in terms of geography, culture, language, diversity, political regime, demography and economic performances exhibit similar migration rates and selection biases in 2000. Whilst global migration rates are expected to be correlated with the traditional migration factors, our main purpose is to detect the potential factors of selection bias. Detecting a positive autocorrelation in selection biases means that the behavior of skilled migrants is structurally linked to its own set of variables, at least in the reduced form. Lucas (2001) recently recommended applying spatial autocorrelation techniques to migration rates, as is was done in the trade literature.

We apply autocorrelation tests to the 2000 skilled migration rates and to the two major components distinguished in the previous section. Regarding proximity criteria, we consider a set of variables which are likely to affect migration rates and selection biases.

**Fractionalization proximity.** To characterize the proximity regarding fractionalization, we use Alesina et al. (2003) indicators of ethnic, linguistic and religious

fractionalization. These variables give the probability that two random individuals from country  $j$  belong to the same ethnic group, speak the same language or share the same religion. As shown in Alesina, the diversity is particularly strong in Africa. As argued by Collier and Gunning (1996) or Easterly and Levine (1997), diversity may be detrimental in non democratic countries where minorities are more likely to be discriminated. It is then relevant to assess the link between fractionalization and migration in general (and skilled migration, in particular). We note that selection biases are strong in ethnically divided countries such as Niger (Hausa 56%, Djerma 22%, Fula 8.5%, Tuareg 8%, Beri Beri 4.3%, Arab, Toubou, and Gourmantche 1.2%, etc.), Nigeria (Africa's most populous country composed of more than 250 ethnic groups: Hausa and Fulani 29%, Yoruba 21%, Igbo 18%, Ijaw 10%, Kanuri 4%, Ibibio 3.5%, Tiv 2.5%, etc.) or in Afghanistan (Pashtun 42%, Tajik 27%, Uzbek 9%, Aimak 4%, Turkmen 3%, Baloch 2%, etc.). Religious fractionalization is particularly high in Rwanda (Roman Catholic 57%, Protestant 26%, Adventist 11%, Muslim 5%, Indigenous beliefs .1%, etc.), one of the most affected countries. It is thus worth evaluating the links between fractionalization, migration rates and selection biases.

**Political proximity.** We consider political indicators such as the number of armed conflicts provided in the PRIO data set (as an indicator of political instability)<sup>15</sup>; the POLITY IV indicator of democracy (ranging from -10 in dictatorial regimes to +10 in democratic regimes), the Freedom House index of civil liberties and political rights ranging from 1 to 7 (low values for democratic regimes) and the Freedom House democracy index ranging from 0 to 1 (high values for democratic regimes) combining civil liberties and political rights scores<sup>16</sup>. Whilst several studies attempted to endogenize political factors, diversity and political indicators can reasonably be seen as potential causes of emigration, bad economic performances and selection bias. Note that the strongest selection biases are observed in countries with instable regimes and low political rights. Let us give a couple of stylized facts. After holding its first free and open elections in 1993, Niger experienced serious coups d'état in 1996 and 1999. Since the 1960s, many fights between the Hutus and Tutsi have been observed in Rwanda. Between the 1960s and the late 1990s, Lesotho relied on military regimes. Corruption poses major problems in Malawi. In the LAC region, Haiti (the poorest country in the Western Hemisphere) has been plagued by political violence for most of its history. Until 1989, Paraguay was under the 35-year dictatorship of Alfredo Stroessner. Jamaica experienced recurrent violence, political violence and civil unrest, including gang violence fueled by the drug trade. In Asia, Afghanistan's recent history is a story of war and civil unrest; after World War II,

---

<sup>15</sup>Formally, we combine several indicators. The weighted number of conflicts at periode  $t$  is given by  $C_t = C_{2t} + C_{3t}(T_{end} - T_{beg} + 1)$  where  $C_{it}$  is the number of conflicts of intensity  $i$  at period  $t$ ,  $T_{end}, T_{beg}$  are ending year and starting year. Then we sum up time observation in  $CC = C_{1975+} + 0.7 \times C_{1960-74} + 0.3 \times C_{1945-59}$

<sup>16</sup>The democracy index is calculated as 14 minus the score of Political rights, minus the score of civil liberties. Dividing the sum by 12 gives an index ranging from 0 to 1.

Mao Zedong's dictatorship imposed strict controls over everyday life and cost the lives of tens of millions of people in China. The political climate has always been tense in Burma where, in 1990, the ruling junta refused to hand over power after multiparty elections in 1990.

**Demography.** We use two demographic criteria. The log of population size is seen as a good indicator of openness: small countries tend to be more opened to migration than large countries. Its impact on selection is less clear since strong selection biases affect both large and small countries (e.g. Nepal and China). Second, we consider the life expectancy at birth, which captures life conditions in the origin countries (a push factor) and potentially affects the incentive to educate. These demographic factors can reasonably act as determinants of emigration.

**Economic proximity.** Three socio-economic variables are distinguished: the level of GDP per capita in PPP values (World Bank database), the UN human index indicator (UN) and the Gini index (UN world inequality database combined with Domenech and Castello index of human capital inequality<sup>17</sup>). The potential relationship between these variables and the brain drain involves multiple channels of influence, multiple causality links and a complex time structure. Anyhow, it is interesting to test whether (and by how much) skilled migration rates are linked to economic variables. In particular, one could expect the degree of inequality to affect the selection bias through various channels such as liquidity constraints (in countries where inequality is strong, the poor are unlikely to afford paying the moving costs) or incentives (the poorer the low class, the higher is the incentive to emigrate). Borjas (1987) and Chiswick (1999) have initiated an interesting debate on inequality and self-selection. These channels can be counteracted by an ex-post effect in high emigration countries: as the richest leave the country, inequality decreases.

**Distances.** Finally, we use classical indicators of distance between the countries. Our weight matrices are all taken from the recent study by Eff (2004). Physical proximity is based on classical distance between national centroids. Linguistic distance relies on Mathematical anthropologists' works on language phylogeny. Phylogenetic language graphs provide a measure of proximity between each pair of languages. Since various languages can be spoken in one country, every language in country  $j$  must be compared to every language in country  $k$ : as for the fractionalization index, this measure gives the expected similarity of the languages spoken by two persons; one drawn at random from each country. Religious distance is based on CIA estimates of the percentage of each nation's population that adheres to one of the seven major religions. The index gives the probability that two persons from  $i$  and  $k$  drawn at random share the same religion. Finally, since colonial links are likely to influence migration choices, we use an index of colonial proximity between countries. The colonial link is simply modeled as a binary variable equal to 1 if both countries were dominated by the same power, and to 0 otherwise.

---

<sup>17</sup>When the income inequality index is not available, we estimate it using the human capital inequality index provided by Domenech and Castello.

Our analysis builds on Anselin (1988) who recommend using Moran’s  $I$  statistic to test for spatial autocorrelation. Moran’s  $I$  differs from the usual correlation coefficient in that it compares the value of a single variable for each pair of locations arrayed according to their degree of (spatial, cultural, political, etc.) relatedness. The null hypothesis of the Moran’s statistics is the absence of autocorrelation. In that case, the Moran’s  $I$  is close from its mean value. A positive weighted autocorrelation is obtained when the Moran’s  $I$  significantly exceeds the mean. This implies that proximate countries in terms of a criterion  $x$  exhibit similar migration rates. Whilst it says nothing about causality, analytical form or time structure of the relationships, positive autocorrelations mean that a reduced-form-structural relationship is likely to hold between the criterion  $x$  and the brain drain. The Moran’s test can thus be seen as a precious guide for empirical research. Except for the distance indicators which are directly available in the matrix format (see Eff, 2004, for details), the other weights are calculated as  $\omega_{j,k} \equiv (1 + d_{j,k})^{-1}$ , where  $d_{j,k}$  is the absolute difference  $|x_j - x_k|$  between country  $j$  and country  $k$  in regards of the criterion  $x$ .

Results are provided in Table 4. Autocorrelation occurs when the  $p$ -value of the Moran test is lower than a critical threshold fixed at 1 percent<sup>18</sup>. In case of positive autocorrelation, we provide the standardized Moran statistics  $z$ . Our autocorrelation tests suggest the existence of reduced-form relationships between skilled migration rates and religious diversity, the political regime (democracy index, civil liberties and political rights), the population size (reflecting the degree of openness) and inequality. Autocorrelation exists at the .99 level of significance when using classical distance indicators. Autocorrelation is particularly strong with the demographic size, the political regime at origin and spatial and linguistic distances (religious and colonial history are also significant).

This confirms Eff’s study which focuses on distances and demonstrates that autocorrelation is more likely than not in international data<sup>19</sup>. Regarding the components of the brain drain, it comes out that a link exists between global migration rate and most variables, except colonial ties. It does not mean that colonial ties cannot be used to predict bilateral migration between countries. It signifies that two proximate countries in terms of colony usually do not exhibit similar rates of emigration.

---

<sup>18</sup>We use a simulation method to infer the distribution of the statistic. The empirical pseudo-significance is based on 10000 random permutations.

<sup>19</sup>Given the non linearity of our decomposition and possible compensation effects, it is noteworthy that a positive autocorrelation in the components (even in all components) do not necessarily induce a positive autocorrelation of the brain drain. Alternatively, the absence of autocorrelation in the brain drain does not necessarily induce the absence of autocorrelation in the components.

**Table 4. Results of autocorrelation tests (2000 - all countries)**

Weighting matrix		Number of Observ.	Brain drain			Global migration rate			Selection Bias		
			Moran I	Prob	z	Moran I	Prob	z	Moran I	Prob	z
<b>Diversity</b>	<i>Ethnic Fractionalization</i>	129	-0.0072	0.2223	.	0.0010	0.0016 **	5.9	0.0044	0.0004 **	8.6
	<i>Language Fractionalization</i>	129	-0.0079	0.3178	.	0.0078	0.0003 **	8.3	0.0230	0.0001 **	16.8
	<i>Religion Fractionalization</i>	131	0.0024	0.0013 **	6.9	-0.0032	0.0199 *	3.2	-0.0044	0.0369 *	2.5
<b>Stability</b>	<i>Political Instability</i>	133	-0.0034	0.2183	0.7	0.0116	0.0167 *	2.9	-0.0100	0.4048	.
	<i>Political Regime</i>	128	0.0318	0.0070 **	3.7	0.0949	0.0001 **	9.7	0.0342	0.0027 **	4.2
	<i>Democracy Index</i>	130	0.0114	0.0001 **	10.1	0.0307	0.0001 **	20.8	-0.0010	0.0079 **	4.0
	<i>Civil Liberties</i>	130	0.0599	0.0001 **	10.9	0.1170	0.0001 **	20.7	0.0181	0.0029 **	4.6
<b>Demography</b>	<i>Political rights</i>	130	0.0389	0.0003 **	6.4	0.0942	0.0001 **	14.6	0.0138	0.0119 *	3.2
	<i>Population Size (in log)</i>	132	0.0717	0.0001 **	16.1	0.0972	0.0001 **	21.8	0.0072	0.0103 *	3.2
	<i>Life expectancy</i>	132	-0.0042	0.3161	.	0.0717	0.0001 **	7.3	0.2345	0.0001 **	23.2
<b>Economy</b>	<i>GDP per capita (in log)</i>	132	-0.0073	0.3481	.	0.0231	0.0001 **	9.1	0.0837	0.0001 **	28.6
	<i>UN human dev. Indicator</i>	132	-0.0085	0.2853	.	0.0052	0.0001 **	9.9	0.0360	0.0001 **	33.9
	<i>Gini Index</i>	133	0.0281	0.0049 **	3.7	0.0292	0.0038 **	3.8	0.0118	0.0386 *	2.0
<b>Distance</b>	<i>Spatial distance</i>	112	0.3510	0.0001 **	5.3	0.5736	0.0001 **	8.9	0.3470	0.0013 **	5.5
	<i>Religious distance</i>	110	0.0988	0.0146 *	2.9	0.1270	0.0072 **	3.7	0.2057	0.0016 **	5.9
	<i>Colonial history</i>	110	0.1041	0.0057 **	3.5	0.0088	0.2258	.	0.1319	0.0030 **	4.6
	<i>Language distance</i>	112	0.1817	0.0008 **	5.1	0.2679	0.0007 **	7.8	0.3124	0.0006 **	8.9

Note: (\*\*) = significant at 0.01; (\*) = significant at 0.05; (.) = not significant  
 $z = \text{standardized Moran } I = (\text{Moran } I - \text{Mean}) / \text{Standard error}$

Focusing on selection biases, it is structurally related to a large set of variables. The standardized Moran statistic is especially high when the proximity is defined in terms of linguistic and ethnic fractionalizations, life expectancy (an indicator of the quality of life) and linguistic distances. The relationship with the political regime is also significant. Although further econometric studies are obviously required to confirm our exploratory tests, they suggest that skilled workers react differently to distances and political and diversity features in the origin country.<sup>20</sup>

## 6 Conclusion

In this paper, we present new estimates of the brain drain experienced by the countries from three regions containing the large majority of developing countries (the LAC region, Asia and Africa). The new data set relies on Census data collected in all OECD countries. It provides consistent and reliable information about the loss of human capital in these regions. We present the data and analyze the distribution of the brain drain on the basis of descriptive statistics, non parametric kernel methods and Moran's tests. Several important messages and stylized facts emerge:

- On average, the most affected regions are the Caribbean, Central America and Eastern, Middle and Western Africa. In Asia, the South-Eastern Asian region also exhibit important rates.
- The distribution of skilled migration rates is globally multimodal in all developing regions. On the whole, the non parametric approach indicates that the dispersion of the brain drain is high in the LAC region, intermediate in Africa and rather low in Asia. At the aggregate, the global density can be decomposed into three clusters of various average intensity and variance.
- Stochastic kernels reveal a complex dynamic process between 1990 and 2000. The global rise in skilled migration has diffused heterogeneously across countries. We emphasize different dynamic patterns across regions. Four major clubs are distinguished in the LAC region and Asia. The richest pattern is observed in Africa where six clubs can be distinguished.
- Using a simple tautological decomposition of skilled migration rates, we demonstrate that countries suffering from the brain drain exhibit either high global migration rates or high selection biases, not both. The average selection bias

---

<sup>20</sup>Comparing the clusters' average characteristics and their dispersion to the world benchmark value, we observe that groups 1, 2 and 3 significantly differ in terms of population size, political instability, political regime, democracy index, and GDP per capita. Differences in terms of Gini index, civil and political rights and fractionalization are smaller. It does not mean that these variables are or are not good predictors of the brain drain: they could explain the variance within cluster but not across clusters. Results are available upon request to the authors.

is an important cause of skilled migration in Asia and Africa, but the range of the distribution of the bias is very large in Africa. On the contrary, the LAC brain drain is essentially due to the degree of openness. The degree of openness is highly sensitive to the the population size; the selection bias is closely linked to the supply of human capital of natives.

- Finally, using Moran’s tests reveals intuitive spatial and cultural autocorrelations in skilled migration rates. These tests provide evidence of important reduced-form relationships between selection biases and a large set of variables such as political indicators, fractionalization, the quality of life and distances. Indeed, stylized facts reveal that selection biases are strong in countries characterized by instable regimes, high degree of fractionalization. This suggests that skilled workers exhibit specific reactions to the traditional determinants of migration.

Naturally, assessing the determinants of the brain drain as well as consequences for the source country requires incorporating additional panel data from developing countries. Before such a complex task is performed, we believe that our study provides important stylized facts about the size and the source of the brain drain. It delivers many insights that can be useful for further econometric regressions

## 7 Appendix

### 7.1 The DM05 dataset - global overview

Table A1 describes the data sources and provides a global overview of immigration statistics in OECD countries. The DM05 data set devotes a special attention to the homogeneity and the comparability of the data. This induces several methodological choices.

- To allow comparisons, it considers the same number of receiving countries in 1990 and 2000. Consequently, Czechoslovakia, Hungary, Korea, Poland, Mexico and Turkey are considered as receiving countries in 1990 despite the fact that they were not members of the OECD.
- Migration is defined on the basis of the foreign-born concept rather than citizenship. Whilst the definition of foreign born is not fully comparable across countries, the homogeneity of our data sources is maximized. In a limited number of cases, immigrants are only classified by citizenship.
- The set of origin country contains all the members of the United Nations (191 in 2000 and 170 in 1990), The Vatican, the Occupied Palestinian territory, Taiwan, Macao and Hong Kong. Since most of Korean migrants to the USA did not

accurately report their origin, North and South Korea cannot be distinguished (estimates are provided for Korea as a whole). The world configuration has changed between 1990 and 2000. The former USSR seceded into 15 countries (7 on the European continent and 8 on the Asian continent), former Yugoslavia seceded into 5 countries, former Czechoslovakia seceded into 2 countries, Eritrea emerged as a new country in 1993, and we have no information about East Timor in 1990 (East Timor became independent of Indonesia in 2002). On the contrary, the former Democratic Republic and the former Republic of Yemen were unified, as well as East Germany and West Germany. For homogeneity reasons, we consider Yemen and Germany as single countries in 1990. Although they became independent after January 1990, data are available for Micronesia, Palau and Namibia in 1990.

Information about the origin country of migrants is available in all OECD countries. Information on the country of birth is available for the large majority of countries, representing 88 percent of the OECD immigration stock. Information on citizenship is used for the other countries (Italy, Germany, Greece, Japan, Korea)<sup>21</sup>. On the contrary, data on educational attainment are missing in a couple of cases. The educational structure can be obtained in 24 countries and can be estimated in 3 additional countries on the basis of the European Labor Force Survey (Belgium<sup>22</sup>, Greece and Portugal). The Labor Force survey provides less detailed information about immigrants' origins, but it is noteworthy that these data are only used to characterize 2 percent of the OECD migration stock in 2000 (and 0.7 percent in 1990). In 2000, the number of migrants whose educational attainment is not described amounts to 1.287 million, i.e. 2.2 percent of the total stock. In that case, we transpose the skill distribution observed in the rest of the OECD area or in the neighboring region. For example, if we have no information about the skill structure of immigrants to Japan, Chinese emigrants to Japan are assumed to be distributed in the same way as Chinese emigrants to the rest of the OECD. More precisely, the educational structure in 2000 is extrapolated on the basis of the Scandinavian countries (for Iceland) or the rest of the OECD (for Japan and Korea). Compared to 2000, educational attainment in 1990 is not available in Poland, the former Czechoslovakia and Hungary. For these three countries, we use proportions observed in the rest of Europe.

---

<sup>21</sup>In 1990, European council data must be used for the former Czechoslovakia, Hungary and Italy. These data are based on the concept of citizenship. Hence, the concept of citizenship is used for 88.3 percent of OECD immigrants in 1990.

<sup>22</sup>However the Belgian 1991 Census is available. It provides information by country of birth and educational attainment.

**Table A1. Data sources and global overview**

Country & Age group	Year 1990		Year 2000	
	Number and origin of immigrants	Structure by educational attainment	Number and origin of immigrants	Structure by educational attainment
Australia (25+)	Census 3 284 279	Census (35.1% ; 29.2% ; 35.7%)	Census 4 075 721	Census (24.3% ; 33.3% ; 42.4%)
Austria (25+)	Census 324 201	Census (58.4% ; 33.3% ; 8.3%)	Census 816 001	Census (47.5% ; 39.9% ; 12.7%)
Belgium (25+)	Census 748 543	Census (69.4% ; 18.0% ; 12.6%)	Improved E.C. (**) 867 620	LFS (Eurostat) (55.8% ; 22.6% ; 21.5%)
Canada (25+)	Census 3 709 285	Census (37.7% ; 11.8% ; 50.5%)	Census 4 661 330	Census (29.6% ; 11.6% ; 58.8%)
Czech Rep (25+)	See Slovakia	See Slovakia	Census 410 249	Census (38.7% ; 46.7% ; 14.6%)
Denmark (25+)	Register 93 934	Register (65.8% ; 22.1% ; 12.2%)	Register 169 664	Register (49.5% ; 31.7% ; 18.8%)
Finland (25+)	Register 34 305	Register (63.0% ; 24.8% ; 12.3%)	Register 90 511	Register (48.8% ; 27.4% ; 23.8%)
France (25+)	Census 3 480 664	Census (85.4% ; 6.0% ; 8.6%)	Census 3 755 514	Census (74.6% ; 9.0% ; 16.4%)
Germany (25-65)	Microcensuz* (Cit) 3 262 057	Microcensuz* (Cit) (68.2% ; 15.0% ; 16.8%)	Microcensuz* (Cit) 4 746 000	Microcensuz* (Cit) (66.4% ; 12.3% ; 21.2%)
Greece (25+)	EC Register (Cit) 112 805	LFS (Eurostat) (Cit.) (34.4% ; 40.3% ; 25.2%)	EC Register (Cit) 106 041	LFS (Eurostat) (Cit.) (36.0% ; 41.2% ; 22.7%)
Hungary	EC Register (Cit) 211 715	As the rest of Europe (70.0% ; 14.7% ; 15.3%)	Census 251 715	Census (43.4% ; 35.0% ; 21.6%)
Iceland (All)	Register 10 565	As Scandinavian countries (41.3% ; 37.5% ; 21.2%)	Register 16 927	As Scandinavian countries (31.9% ; 41.5% ; 26.7%)
Ireland (25+)	Census 130 940	Census (22.4% ; 52.7% ; 24.9%)	Census 281 232	Census (13.5% ; 45.8% ; 40.7%)
Italy (25+)	EC Register (Cit) 533 312	As the rest of EU15 (70.0% ; 14.7% ; 15.3%)	Census (Cit.) 923 788	Census (Cit.) (53.1% ; 31.6% ; 15.3%)
Japan (All/25+)	Register (Cit) 1 075 317	As the rest of OECD (43.3% ; 26.0% ; 30.7%)	Census (Cit) 951 302	As the rest of OECD (36.4% ; 29.0% ; 34.6%)
Korea (All)	Register (Cit) 49 500	As the rest of OECD (43.3% ; 26.0% ; 30.7%)	Register (Cit) 150 812	As the rest of OECD (36.4% ; 29.0% ; 34.6%)
Luxemburg (25+)	Census 83 398	Census (81.3% ; 6.6% ; 12.0%)	Census 114 625	Census (43.4% ; 31.9% ; 24.8%)
Mexico (25+)	IPUMS 10% 363 626	IPUMS 10% (49.6% ; 16.5% ; 33.9%)	IPUMS 10.6% 417 371	IPUMS 10.6% (37.4% ; 28.6% ; 34.0%)
Netherland (15-74)	Census* 961 662	Census* (70.0% ; 14.7% ; 15.3%)	Census* 1 320 320	Census* (54.2% ; 26.6% ; 19.2%)
New Zealand (15+)	Census 456 792	Census (32.6% ; 24.8% ; 42.7%)	Census 603 606	Census (31.3% ; 30.2% ; 38.5%)
Norway (25+)	Register 136 241	Register (30.2% ; 45.1% ; 24.6%)	Register 204 182	Register (13.6% ; 55.0% ; 31.5%)
Poland (13+)	Census 661 517	As the rest of Europe (70.0% ; 14.7% ; 15.3%)	Census 741 517	Census (60.7% ; 25.4% ; 13.8%)
Portugal (25+)	Census 170 390	LFS (Eurostat) (91.9% ; 3.3% ; 4.9%)	Census 207 476	LFS (Eurostat) (64.4% ; 20.9% ; 14.7%)
Slovak Rep (25+)	EC Register (Cit) 196 205	As the rest of Europe (70.0% ; 14.7% ; 15.3%)	Census 426 072	Census (78.0% ; 5.4% ; 16.6%)
Spain (25+)	Census 845 977	Census (56.6% ; 26.0% ; 17.4%)	Census 1 370 657	Census (33.1% ; 50.6% ; 16.4%)
Sweden (25+)	Census 617 449	Census (38.7% ; 39.1% ; 22.2%)	Census 805 143	Census (30.9% ; 41.7% ; 27.4%)
Switzerland (18+)	Census 1 463 670	Census (14.0% ; 72.1% ; 13.9%)	Census 1 704 948	Census (17.0% ; 63.1% ; 19.8%)
Turkey (15+)	Census 596 045	Census (75.4% ; 16.8% ; 7.8%)	Census 826 110	Census (53.1% ; 30.6% ; 16.3%)
United Kingdom (15+)	Census* 2 778 527	Census* (68.1% ; 11.4% ; 20.5%)	Census* 3 639 907	Census* (36.6% ; 28.8% ; 34.6%)
United States (25+)	IPUMS 5% 15 472 972	IMPUS 5% (25.5% ; 34.1% ; 40.3%)	Census* 24 366 085	Census* (23.1% ; 34.5% ; 42.5%)
OECD Total	41 844 960	(44.9% ; 25.3% ; 29.8%)	59 022 443	(36.4% ; 29.0% ; 34.6%)

Notes: EC = European Council (register data); LFS (Eurostat) = Labor Force Survey; (\*) = limited level of detail.

(\*\*) European council data corrected by the country specific "foreign born/foreign citizen" ratio in 1990.

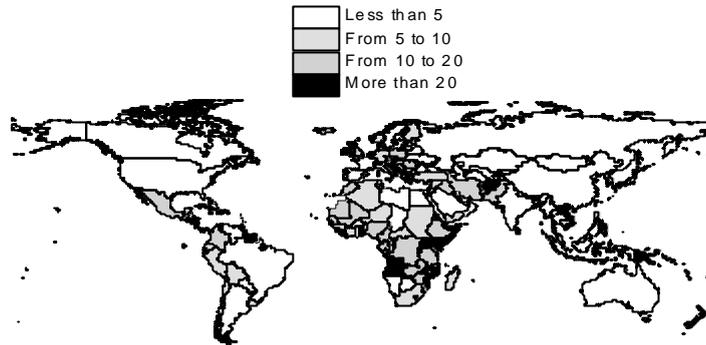
**Table A2. Skilled migration rate in 1990 and 2000 (ranked decreasingly with respect to 1990)**

LAC REGION		BD 1990	BD 2000	Change	AFRICA	BD 1990	BD 2000	Change	ASIA	BD 1990	BD 2000
Antigua and Barbuda	Caribbean	64.85%	66.77%	1.92%	Burundi	9.48%	8.51%	-0.97%	China	3.08%	3.83%
Bahamas, The	Caribbean	57.30%	61.28%	3.97%	Comoros	7.04%	21.19%	14.14%	Hong Kong	32.49%	28.79%
Barbados	Caribbean	67.25%	63.49%	-3.77%	Djibouti	7.62%	10.96%	3.34%	Japan	1.31%	1.17%
Cuba	Caribbean	31.04%	28.67%	-2.37%	Eritrea	-	33.98%	-	Korea	9.42%	5.65%
Dominica	Caribbean	68.57%	64.21%	-4.36%	Ethiopia	8.02%	10.06%	2.03%	Macao	20.58%	14.44%
Dominican Republic	Caribbean	18.21%	21.58%	3.37%	Kenya	42.68%	38.44%	-4.24%	Mongolia	0.22%	1.05%
Grenada	Caribbean	77.74%	85.11%	7.37%	Madagascar	5.69%	7.55%	1.87%	Taiwan	15.20%	12.69%
Haiti	Caribbean	78.57%	83.57%	4.99%	Malawi	16.80%	18.72%	1.92%	Afghanistan	13.47%	23.30%
Jamaica	Caribbean	85.09%	85.05%	-0.04%	Mauritius	64.68%	56.17%	-8.51%	Bangladesh	2.07%	4.33%
Saint Kitts and Nevis	Caribbean	77.81%	78.49%	0.68%	Mozambique	26.61%	45.13%	18.52%	Bhutan	0.72%	0.59%
Saint Lucia	Caribbean	67.67%	71.08%	3.41%	Rwanda	17.34%	25.99%	8.65%	India	2.85%	4.31%
St Vincent & Gren.	Caribbean	80.67%	84.51%	3.85%	Seychelles	66.21%	55.87%	-10.34%	Iran	25.26%	14.47%
Trinidad and Tobago	Caribbean	77.80%	79.28%	1.48%	Somalia	17.38%	32.65%	15.28%	Kazakhstan	-	1.17%
Belize	Central Am.	67.32%	65.47%	-1.85%	Tanzania	11.64%	12.36%	0.72%	Kyrgyzstan	-	0.70%
Costa Rica	Central Am.	8.25%	7.21%	-1.04%	Uganda	44.17%	35.56%	-8.61%	Maldives	1.16%	1.20%
El Salvador	Central Am.	32.75%	30.97%	-1.79%	Zambia	16.72%	16.77%	0.05%	Nepal	1.79%	5.28%
Guatemala	Central Am.	20.44%	24.15%	3.71%	Zimbabwe	7.51%	12.74%	5.23%	Pakistan	6.86%	12.58%
Honduras	Central Am.	22.21%	24.44%	2.23%	Angola	4.58%	32.95%	28.37%	Sri Lanka	28.66%	29.66%
Mexico	Central Am.	10.66%	15.30%	4.64%	Cameroon	13.20%	17.17%	3.97%	Tajikistan	-	0.43%
Nicaragua	Central Am.	29.74%	29.62%	-0.12%	Central African Rep.	4.05%	7.08%	3.04%	Turkmenistan	-	0.17%
Panama	Central Am.	22.45%	15.98%	-6.47%	Chad	2.14%	2.44%	0.30%	Uzbekistan	-	0.67%
Argentina	South Am.	3.64%	2.51%	-1.13%	Congo, Dem. Rep.	20.97%	13.69%	-7.29%	Brunei	22.95%	15.59%
Bolivia	South Am.	6.37%	5.81%	-0.56%	Congo, Rep.	9.38%	22.16%	12.79%	Burma (Myanmar)	4.25%	4.02%
Brazil	South Am.	1.55%	2.25%	0.70%	Equatorial Guinea	1.10%	12.97%	11.87%	Cambodia	15.64%	18.32%
Chile	South Am.	6.74%	6.14%	-0.59%	Gabon	5.45%	14.65%	9.20%	East Timor	-	15.48%
Colombia	South Am.	9.13%	10.41%	1.27%	Sao Tome & Principe	3.62%	21.96%	18.34%	Indonesia	3.95%	2.06%
Ecuador	South Am.	5.51%	9.55%	4.04%	Algeria	7.24%	9.42%	2.18%	Laos	29.86%	37.41%
Guyana	South Am.	91.08%	88.96%	-2.12%	Egypt	5.87%	4.55%	-1.32%	Malaysia	24.68%	11.11%
Paraguay	South Am.	3.55%	3.92%	0.36%	Libya	2.29%	2.37%	0.09%	Philippines	12.95%	13.73%
Peru	South Am.	5.56%	5.81%	0.25%	Morocco	21.65%	16.95%	-4.70%	Singapore	24.84%	15.16%
Suriname	South Am.	52.67%	47.87%	-4.80%	Sudan	5.18%	6.93%	1.75%	Thailand	2.41%	2.41%
Uruguay	South Am.	6.87%	8.08%	1.21%	Tunisia	17.79%	12.55%	-5.24%	Vietnam	24.51%	27.07%
Venezuela	South Am.	3.67%	3.40%	-0.27%	Botswana	2.21%	3.61%	1.41%	Armenia	-	8.83%
					Lesotho	10.39%	4.32%	-6.08%	Azerbaijan	-	1.97%
					Namibia	3.18%	3.46%	0.29%	Bahrain	4.31%	4.91%
					South Africa	9.88%	7.51%	-2.37%	Cyprus	31.53%	31.18%
					Swaziland	0.17%	0.48%	0.31%	Georgia	-	1.64%
					Benin	7.34%	11.34%	4.00%	Iraq	8.12%	11.14%
					Burkina Faso	1.46%	2.56%	1.10%	Israel	8.56%	7.90%
					Cape Verde	56.77%	67.46%	10.70%	Jordan	8.59%	7.22%
					Cote d'Ivoire	2.89%	5.74%	2.85%	Kuwait	2.80%	7.11%
					Gambia, The	80.42%	63.29%	-17.14%	Lebanon	43.91%	38.64%
					Ghana	37.65%	46.86%	9.21%	Oman	0.39%	0.57%
					Guinea	13.45%	11.25%	-2.19%	Palestine	7.40%	7.18%
					Guinea-Bissau	9.34%	24.42%	15.08%	Qatar	1.82%	2.50%
					Liberia	32.36%	44.99%	12.63%	Saudi Arabia	0.92%	0.91%
					Mali	8.23%	15.00%	6.77%	Syria	7.03%	6.05%
					Mauritania	2.79%	11.78%	8.99%	Turkey	8.32%	5.82%
					Niger	6.45%	5.96%	-0.49%	United Arab Emirates	0.39%	1.01%
					Nigeria	8.00%	10.72%	2.72%	Yemen	5.49%	6.04%
					Senegal	12.25%	17.70%	5.45%			
					Sierra Leone	34.18%	52.52%	18.35%			
					Togo	11.06%	18.74%	7.68%			

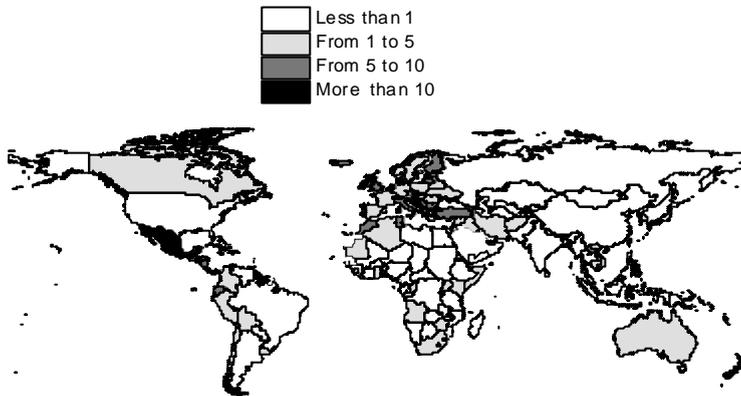
**Table A3. Country decomposition in 2000**

LAC REGION		MIG	BIAS	COMP	HCD	AFRICA				ASIA							
Antigua and Barbuda	Caribbean	37.6%	1.78	48.3%	3.67	Algeria	Northern Af.	4.5%	2.12	14.1%	15.05	Afghanistan	South-cent. As.	1.5%	15.16	42.8%	35.41
Argentina	South Am.	1.0%	2.41	48.2%	5.00	Angola	Middle Af.	2.8%	11.68	16.9%	69.00	Armenia	Western As.	3.4%	2.62	47.0%	5.58
Bahamas, The	Caribbean	12.0%	5.12	52.4%	9.77	Benin	Western Af.	0.4%	26.01	52.6%	49.47	Azerbaijan	Western As.	0.7%	2.91	49.9%	5.83
Barbados	Caribbean	32.3%	1.96	43.3%	4.54	Botswana	Southern Af.	0.4%	8.42	33.9%	24.82	Bahrain	Western As.	1.3%	3.91	50.7%	7.70
Belize	Central Am.	29.0%	2.26	41.3%	5.46	Burkina Faso	Western Af.	0.2%	14.58	30.1%	48.48	Bangladesh	South-cent. As.	0.4%	10.58	36.3%	29.11
Bolivia	South Am.	1.7%	3.39	50.9%	6.65	Burundi	Eastern Af.	0.4%	23.99	51.1%	46.96	Bhutan	South-cent. As.	0.0%	11.87	29.8%	39.78
Brazil	South Am.	0.4%	5.08	43.4%	11.69	Cameroon	Middle Af.	0.8%	21.76	49.5%	43.94	Brunei	South-east As.	3.4%	4.59	59.3%	7.73
Chile	South Am.	2.1%	2.88	47.4%	6.07	Cape Verde	Western Af.	25.1%	2.69	15.2%	17.73	Burma (Myanmar)	South-east As.	0.2%	16.35	56.1%	29.16
Colombia	South Am.	2.7%	3.88	41.7%	9.30	Central African Rep	Middle Af.	0.3%	21.13	40.8%	51.79	Cambodia	South-east As.	4.0%	4.55	25.7%	17.73
Costa Rica	Central Am.	3.1%	2.32	45.0%	5.15	Chad	Middle Af.	0.1%	24.03	48.0%	50.08	China	Eastern As.	0.2%	17.12	48.0%	35.70
Cuba	Caribbean	10.5%	2.72	38.6%	7.06	Comoros	Eastern Af.	3.8%	5.51	13.4%	40.98	China, Hong Kong SAR	Eastern As.	9.1%	3.17	61.9%	5.12
Dominica	Caribbean	40.8%	1.57	37.8%	4.17	Congo, Dem. Rep.	Middle Af.	0.6%	24.47	36.6%	66.77	China, Macao SAR	Eastern As.	4.5%	3.23	55.2%	5.85
Dominican Republic	Caribbean	13.1%	1.65	26.5%	6.22	Congo, Rep.	Middle Af.	3.0%	7.39	40.5%	18.24	Cyprus	Western As.	17.2%	1.81	37.2%	4.86
Ecuador	South Am.	6.6%	1.45	28.0%	5.18	Cote d'Ivoire	Western Af.	0.7%	8.37	30.7%	27.25	East Timor	South-east As.	3.4%	4.61	25.3%	18.22
El Salvador	Central Am.	20.0%	1.55	19.1%	8.14	Djibouti	Eastern Af.	0.7%	16.83	37.6%	44.81	Georgia	Western As.	0.5%	3.03	51.9%	5.85
Grenada	Caribbean	53.7%	1.58	40.9%	3.88	Egypt	Northern Af.	0.9%	5.30	58.9%	9.00	India	South-cent. As.	0.4%	12.10	60.5%	20.01
Guatemala	Central Am.	8.8%	2.73	19.0%	14.35	Equatorial Guinea	Middle Af.	4.5%	2.90	12.4%	23.36	Indonesia	South-east As.	0.2%	9.10	46.3%	19.63
Guyana	South Am.	42.1%	2.11	44.2%	4.77	Eritrea	Eastern Af.	2.5%	13.63	40.5%	33.64	Iran	South-cent. As.	1.9%	7.61	58.5%	13.01
Haiti	Caribbean	11.5%	7.25	39.0%	18.58	Ethiopia	Eastern Af.	0.5%	21.84	48.7%	44.89	Iraq	Western As.	2.7%	4.10	38.6%	10.62
Honduras	Central Am.	8.1%	3.01	23.8%	12.65	Gabon	Middle Af.	0.8%	17.40	52.9%	32.88	Israel	Western As.	4.1%	1.94	57.6%	3.36
Jamaica	Caribbean	35.3%	2.41	42.8%	5.63	Gambia, The	Western Af.	3.3%	19.33	20.4%	94.89	Japan	Eastern As.	0.5%	2.57	63.8%	4.03
Mexico	Central Am.	12.4%	1.23	14.4%	8.56	Ghana	Western Af.	2.2%	21.76	44.1%	49.37	Jordan	Western As.	2.8%	2.54	55.6%	4.57
Nicaragua	Central Am.	8.9%	3.34	38.5%	8.68	Guinea	Western Af.	0.5%	23.35	26.2%	89.18	Kazakhstan	South-cent. As.	0.4%	2.72	49.2%	5.54
Panama	Central Am.	6.1%	2.61	57.7%	4.52	Guinea-Bissau	Western Af.	2.2%	11.01	14.2%	77.30	Korea	Eastern As.	2.8%	2.03	54.1%	3.76
Paraguay	South Am.	0.7%	5.32	45.6%	11.66	Kenya	Eastern Af.	1.6%	23.38	44.8%	52.16	Kuwait	Western As.	1.8%	4.06	67.8%	5.98
Peru	South Am.	2.9%	2.03	46.9%	4.33	Lesotho	Southern Af.	0.1%	47.53	49.6%	95.77	Kyrgyzstan	South-cent. As.	0.2%	2.92	52.7%	5.55
Saint Kitts and Nevis	Caribbean	49.4%	1.59	44.4%	3.57	Liberia	Western Af.	3.5%	12.82	58.5%	21.93	Laos	South-east As.	10.0%	3.73	25.7%	14.50
Saint Lucia	Caribbean	23.1%	3.08	36.8%	8.36	Libya	Northern Af.	0.9%	2.55	54.1%	4.72	Lebanon	Western As.	15.0%	2.57	44.5%	5.78
St Vincent & Gren.	Caribbean	36.5%	2.32	42.7%	5.42	Madagascar	Eastern Af.	0.5%	15.69	43.4%	36.17	Malaysia	South-east As.	1.6%	7.12	59.2%	12.04
Suriname	South Am.	47.4%	1.01	18.4%	5.50	Malawi	Eastern Af.	0.3%	58.68	43.2%	135.90	Maldives	South-cent. As.	0.3%	4.08	46.5%	8.77
Trinidad and Tobago	Caribbean	25.2%	3.14	51.0%	6.16	Mali	Western Af.	1.0%	15.62	10.9%	143.04	Mongolia	Eastern As.	0.4%	2.73	61.1%	4.47
Uruguay	South Am.	2.6%	3.11	41.2%	7.55	Mauritania	Western Af.	1.2%	10.04	21.9%	45.78	Nepal	South-cent. As.	0.3%	17.92	47.2%	38.00
Venezuela	South Am.	1.0%	3.26	60.1%	5.42	Mauritius	Eastern Af.	10.7%	5.25	28.9%	18.18	Oman	Western As.	0.1%	5.00	62.7%	7.96
						Morocco	Northern Af.	7.6%	2.24	12.9%	17.36	Pakistan	South-cent. As.	1.0%	12.19	38.6%	31.55
						Mozambique	Eastern Af.	0.9%	49.09	17.7%	276.91	Palestine	Western As.	2.9%	2.52	55.0%	4.57
						Namibia	Southern Af.	0.3%	12.30	51.1%	24.06	Philippines	South-east As.	5.0%	2.74	67.1%	4.09
						Niger	Western Af.	0.1%	93.04	49.4%	188.21	Qatar	Western As.	0.5%	5.45	69.6%	7.84
						Nigeria	Western Af.	0.6%	18.84	65.0%	28.97	Saudi Arabia	Western As.	0.2%	5.13	64.6%	7.94
						Rwanda	Eastern Af.	0.4%	71.82	48.3%	148.56	Singapore	South-east As.	3.2%	4.73	57.1%	8.27
						Sao Tome and Principe	Middle Af.	5.6%	3.91	18.4%	21.20	Sri Lanka	South-cent. As.	2.8%	10.58	39.5%	26.80
						Senegal	Western Af.	2.8%	6.44	16.7%	38.47	Syria	Western As.	1.9%	3.24	44.3%	7.31
						Seychelles	Eastern Af.	19.5%	2.86	37.1%	7.72	Taiwan	Eastern As.	3.4%	3.69	78.0%	4.73
						Sierra Leone	Western Af.	2.1%	24.43	50.4%	48.52	Tajikistan	South-cent. As.	0.1%	2.92	49.6%	5.89
						Somalia	Eastern Af.	3.3%	9.76	28.2%	34.62	Thailand	South-east As.	0.6%	3.88	44.6%	8.69
						South Africa	Southern Af.	1.3%	5.64	62.6%	9.01	Turkey	Western As.	5.6%	1.04	8.8%	11.74
						Sudan	Northern Af.	0.3%	24.27	52.0%	46.67	Turkmenistan	South-cent. As.	0.1%	2.29	41.1%	5.57
						Swaziland	Southern Af.	0.0%	13.29	56.1%	23.70	United Arab Emirates	Western As.	0.2%	5.34	67.3%	7.93
						Tanzania	Eastern Af.	0.5%	23.09	51.3%	45.02	Uzbekistan	South-cent. As.	0.2%	3.50	59.5%	5.88
						Togo	Western Af.	1.2%	15.66	40.0%	39.16	Vietnam	South-east As.	3.4%	8.01	40.0%	20.04
						Tunisia	Northern Af.	5.4%	2.33	14.9%	15.66	Yemen	Western As.	0.4%	16.25	34.5%	47.16
						Uganda	Eastern Af.	1.1%	33.41	46.2%	72.37						
						Zambia	Eastern Af.	0.8%	20.33	48.5%	41.96						
						Zimbabwe	Eastern Af.	1.4%	9.19	55.0%	16.70						

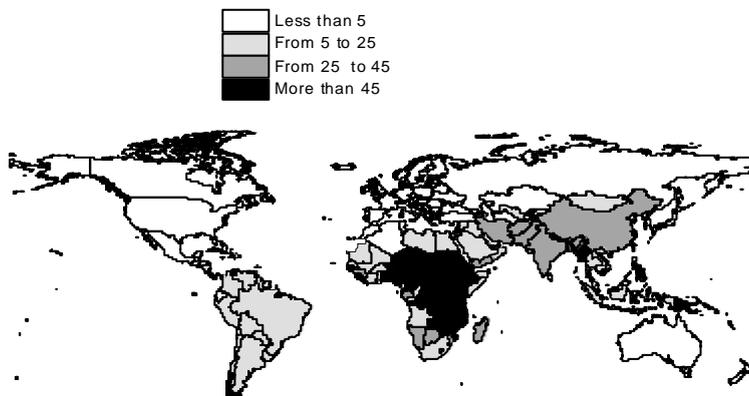
**Figure A.1. World maps**  
*A.1.a Skilled migration rate in percent*



*A.1.b Global migration rate in percent*



*A.1.c Selection bias*



As shown in Table A.1, the total stock of immigrants in the OECD area has increased from 41.8 million to 59.0 million between 1990 and 2000. The average share of tertiary educated immigrants has risen from 29.8 to 34.6 percent. The share of tertiary educated is particularly high in traditional immigration countries such as Canada, the US and Australia. It sharply increased in Ireland, the UK, Norway, Finland and Luxemburg. Note that for Poland, Hungary, and former Czechoslovakia in 1990, we use the structure observed in the rest of Europe; which probably slightly overestimates the reality.

The next two tables A.2 and A.3 provide country-specific migration rates of skilled workers (1990-2000) and their decomposition in 2000. Figures A.1.a to A.1.c map the skilled migration rates, the global migration rates and selection biases in the world countries.

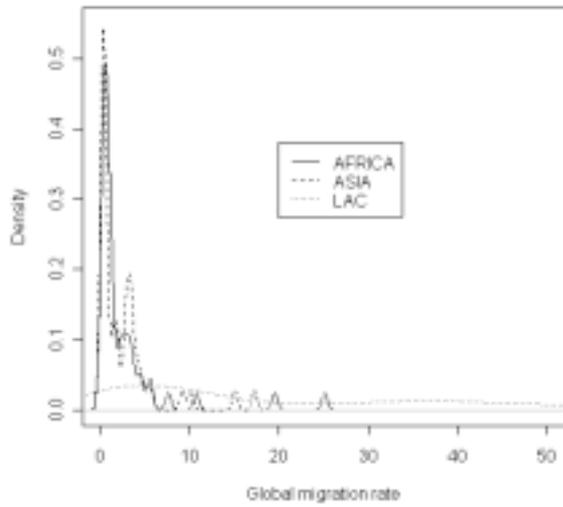
## 7.2 Applying the Kernel method to the decomposition

The non parametric approach allows us to draw a better picture of regional disparities. For each continent in 2000, we apply the Kernel method to the four components distinguished in section 4. Figure A.2.a clearly shows that the skilled migration dispersion in LAC countries is related to the dispersion in global migration rates, i.e. in the degree of openness. The LAC distribution is the most uniform, ranging from 0.5 to about 47.5 percent with two minor peaks associated with South American countries (at 5 percent) and to a group including the Caribbean countries, Guyana and Suriname (at 36 percent). The Asian density is more concentrated between 0 and 6 percent (with a few exceptions such as Hong Kong, Laos, Lebanon and Cyprus at higher rates). The migration rate is lower than 1 percent in 24 Asian countries (out of 50). The African density is even more concentrated between 0 and 8 percent. The rate is below 1 percent in 35 countries (out of 53). Mauritius, Seychelles and Cape Verde are the most affected countries.

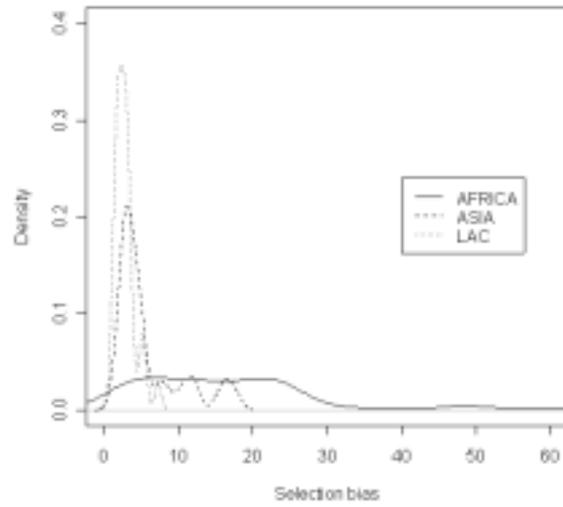
Figure A.2.b depicts the selection bias density. The LAC density is very concentrated on the interval [1;5]. The range of the Asian density is larger and a significant number of countries exhibit a selection bias higher than 10. The African distribution is defined on a very important interval [1;50], revealing that a large number of countries suffer from a strong selection bias. On the whole, the selection bias is an important factor of brain drain in Africa and Asia, whilst the position of LAC countries is essentially due to their degree of openness.

**Figure A.2. Distribution of the components**

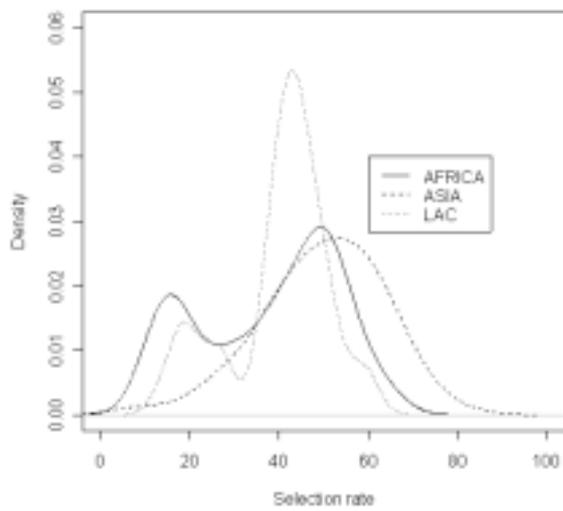
*A.2.a. Average migration rate in percent*



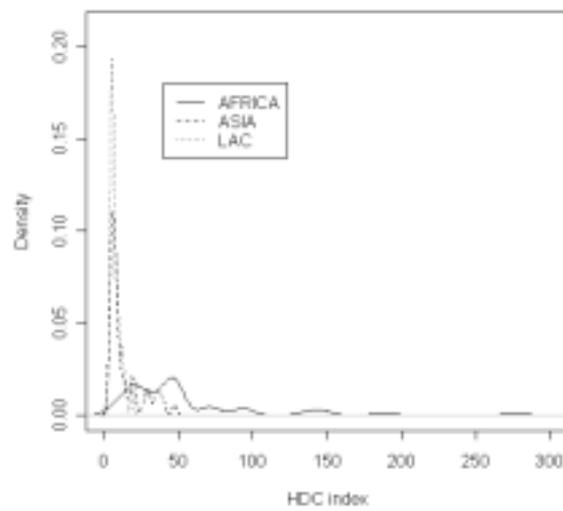
*A.2.b. Selection bias*



*A.2.c. Composition rate in percent*



*A.2.d. Human capital deficiency index*



**Table A.4. Explaining clusters by applying mixture to the decomposition**

Selection / Migration	High Selection Bias	Medium Selection Bias	Low Selection Bias
High Migration Rate			ANTIGUA AND BARBUDA, BARBADOS, BELIZE, CAPE VERDE, <i>Cyprus</i> , DOMINICA, <i>El Salvador</i> , Grenada, GUYANA, JAMAICA, <i>Lebanon</i> , SAINT KITTS AND NEVIS, SAINT LUCIA, SAINT VINCENT AND THE GRENADINES, SEYCHELLES, SURINAME, TRINIDAD AND TOBAGO
Medium Migration Rate		HAITI	Algeria, BAHAMAS, <i>Cambodia</i> , <i>Hong Kong</i> , <i>Macao</i> , <i>Cuba</i> , <i>Dominican Republic</i> , Ecuador, Equatorial Guinea, <i>Guatemala</i> , <i>Honduras</i> , Israel, <i>Laos</i> , MAURITIUS, <i>Mexico</i> , <i>Morocco</i> , <i>Nicaragua</i> , <i>Panama</i> , Philippines, <i>Sao Tome and Principe</i> , Tunisia, Turkey
Low Migration Rate	Lesotho, <i>Malawi</i> , MOZAMBIQUE, Niger, <i>Rwanda</i>	<i>Afghanistan</i> , <i>Angola</i> , Bangladesh, Benin, Bhutan, Botswana, Burkina Faso, Burma (Myanmar), Burundi, Cameroon, Central African Republic, Chad, China, Dem. Rep. of the Congo, <i>Rep. of the Congo</i> , Cote d'Ivoire, Djibouti, <i>Eritrea</i> , Ethiopia, <i>Gabon</i> , GAMBIA, GHANA, Guinea, <i>Guinea-Bissau</i> , India, Indonesia, <i>Iran</i> , <i>Kenya</i> , LIBERIA, Madagascar, Malaysia, <i>Mali</i> , Mauritania, Namibia, Nepal, Nigeria, Pakistan, <i>Senegal</i> , Sierra Leone, Somalia, Sri Lanka, Sudan, Swaziland, Tanzania, <i>Togo</i> , <i>Vietnam</i> , Yemen, <i>Zambia</i> Zimbabwe	Argentina, Armenia, Azerbaijan, Bahrain, Bolivia, Brazil, <i>Brunei</i> , Chile, Colombia, <i>Comoros</i> , Costa Rica, <i>East Timor</i> , Egypt, Georgia, Iraq, Japan, Jordan, Kazakhstan, Korea, Kuwait, Kyrgyzstan, Libya, Maldives, Mongolia, Occupied Palestinian Territory, Oman, Paraguay, Peru, Qatar, Saudi Arabia, <i>Singapore</i> , South Africa, Syria, Taiwan, Tajikistan, Thailand, Turkmenistan, <i>Uganda</i> , United Arab Emirates, Uruguay, Uzbekistan, Venezuela

Legend: GROUP 1 (High brain drain; Bold capital characters); GROUP 2 (Medium brain drain: Bold italic characters): GROUP 3 (Low brain drain: Normal characters)

Applying formal tests of multimodality to the distribution of global migration rates and selection biases reveals that both densities are multimodal. Developing a hierarchical Gaussian mixture model allows to disentangle each of these distributions into three Gaussian distributions. We thus define nine possible clusters of different migration rates (low, medium, high) and different selection biases (low, medium, high). These clusters are presented in table A.4. Confirming Figure 4a, it is worth noticing that there is no country suffering from both high migration rate and selection bias. Haiti is the only case combining both medium migration rate and selection bias. The most affected countries (Group 1) usually exhibit high migration rates, except several African countries such as Gambia, Ghana, Liberia and Mozambique.

The distribution of composition rates is provided on Figure A.2.c. The density is unimodal in Asia. It peaks at about 60 percent and a significant number of countries (several high-income countries such as Qatar, Oman and Eastern Asian developed countries) exhibit very large selection rates. The African density is bimodal with a dominant mode at 47 percent and a less important one at 15 percent. The bimodality is more pronounced in the LAC region, where the less important mode concerns Central American countries, and the dominant one characterizes South African countries and some cases in the Caribbean.

The human capital deficiency index is distributed as in Figure A.2.d. The density is strongly tailed in the LAC region where most South and Central American countries lie around 5. The highest index is observed in Haiti (18). The concentration is important in Asia where the index is lower than 10 in 31 countries. Note that a significant set of poor Asian countries is characterized by an index between 29 and 25 (this set contains China, Pakistan, Bangladesh, etc.). The highest index is observed in Yemen (47). On the contrary, the African distribution is extremely large, ranging from 5 to 277. The large majority of countries (37 out of 53) lie between 10 and 55. However, five poor countries (Mozambique, Niger, Rwanda, Mali and Malawi) exhibit an index higher than 100 (the share of tertiary educated is lower than 1 percent).

## References

- [1] Adams, R. (2003): International migration, remittances and the brain drain: a study of 24 labor-exporting countries, World Bank Policy Research Working Paper, n. 2972.
- [2] Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and, R. Wacziarg (2003): Fractionalization, *Journal of Economic Growth*, 8: 155-194.
- [3] Anselin L. (1988b): Spatial dependence and spatial structural instability in applied regression analysis, *Journal of Regional Science*, 30, 185-207.
- [4] Barro, R.J. and J.W. Lee (2000): International Data on Educational Attainment: Updates and Implications, CID Working Papers 42, Center for International Development (Harvard University).

- [5] Borjas, G.J. (1987): Self-selection and the earnings of immigrants, *American Economic Review*, 77(4).
- [6] Carrington W.J., E. Detragiache, T. Vishwanath (1996): Migration with endogenous moving costs, *American Economic Review*, 86(4), 909-930.
- [7] Carrington, W.J. and E. Detragiache (1998): How big is the brain drain?, IMF Working paper WP/98/102.
- [8] Carrington, W.J. and E. Detragiache (1999): How extensive is the brain drain, *Finance and Development*, June: 46-49.
- [9] Chiswick, B. (1999): Are immigrants favourably self-selected?, *American Economic Review*, 89(2).
- [10] Commander, S., M. Kangasniemi and L.A. Winters (2004): The brain drain: a review of theory and facts, *Brussels Economic Review*, 47(1).
- [11] Collier, P. and J.W. Gunning (1999): Explaining African economic performance, *Journal of Economic Literature*, 37(1), 64-111.
- [12] De la Fuente, A. and R. Domenech (2002): Human capital in growth regressions: how much difference does data quality make? Un update and further results, CEPR Discussion Paper, n. 3587.
- [13] Docquier, F. and A. Marfouk (2005): Measuring international migration by educational attainment in 1990-2000 - Release 1.1, The World Bank, Mimeo.
- [14] Docquier, F. and H. Rapoport (2004): Skilled migration: the perspective of developing countries, Policy Research Working paper, 3382, The World Bank: Washington.
- [15] Domenech, R. and A. Castello (2002): Human Capital Inequality and Economic Growth: Some New Evidence, *Economic Journal*, 112, C187-200.
- [16] Easterly, W. and R. Levine (1997): Africa's growth tragedy: policies and ethnic divisions, *Quarterly Journal of Economics*, 112(4), 1203-1250.
- [17] Easterly, W. (2000): Can institutions resolve ethnic conflict, Policy Research Working Paper, 2482, The World Bank.
- [18] Eff, A.E. (2004), Spatial and cultural autocorrelation in international datasets, Department of Economics and Finance Working Paper Series, Feb.
- [19] Fraley, C. and A.E. Raftery (2002): Model-based clustering: discriminant analysis and density estimation, *JASA*, 97(458): 611-631.
- [20] Härdle, W. (1990): *Applied Nonparametric Regression*, Econometric Study Monographs, Cambridge University Press.

- [21] Hartigan, J.A. and P.M. Hartigan (1985): The DIP test of unimodality, *Annals of Statistics*, 13(1): 70-84.
- [22] Harris, J.R. and M.P. Todaro (1970): Migration, unemployment and development: a two-sector analysis, *American Economic Review*, 60(5), 126-142.
- [23] Lucas, R.E.B. (2001): The effects of proximity and transportation on developing country population migrations, *Journal of Economic Geography*, 1: 323-339.
- [24] Massey, D.S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino and J.E. Taylor (1993): Theories of international migration: review and appraisal, *Population and Development Review*, 19(3), 431-466.
- [25] OECD (2002): Trends in international migration, Paris: OECD Editions.
- [26] Park and Turloch (1992): Practical performance of several data driven bandwidth selectors, *Computational Statistics*, 7, 251-270.
- [27] Pedersen, P.J., M. Pytlikova and N. Smith (2004), Selection or network effects? Migration flows into 27 OECD countries, 1990-2000, IZA DP, n. 1104.
- [28] Quah D. (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", *Journal-of-Economic-Growth*, 2(1), March 1997, pages 27-59.
- [29] Schiff, M. (2005): Brain Gain: Claims about Its Size and Impact on Welfare and Growth Are Greatly Exaggerated, IZA discussion paper, n. 1599.
- [30] Sheater, S. and M. Jones (1991): A reliable data-based bandwidth selection method for kernel density estimation, *Journal of the Royal Statistical Society*, 53: 683-690.
- [31] Silverman, B. (1981): Using kernel density estimates to investigate multimodality, *Journal of the Royal Statistical Society*, 43, 97-99.
- [32] Silverman, B. (1986): Density estimation for statistics and data analysis, Chapman & Hall, London.
- [33] Sjaastad, L.A. (1962): The costs and returns of human capital, *Journal of Political Economy*, 70(5), 80-93.
- [34] United Nations (2001): World population ageing: 1950-2050, Department of Economic and Social Affairs - Population division.