

# Skilled migration and business networks\*

Frédéric Docquier<sup>a,b</sup> and Elisabetta Lodigiani<sup>b,c</sup>

<sup>a</sup> National Fund for Scientific Research

<sup>b</sup> IRES, Université Catholique de Louvain

<sup>c</sup> Department of Economics, Università degli Studi di Milano

October 11, 2006

## Abstract

The role of migrants' networks in promoting cross border investments has been stressed in the literature, possibly making migration and FDI complements rather than substitutes in the long run. In this paper, we estimate the magnitude of such business network externalities in dynamic empirical models of FDI-funded capital accumulation. We use original data on capital and migration stocks rather than flows. Regarding migrants, we distinguish the total and the skilled diasporas abroad. In both cross-sectional and panel frameworks, we find evidence of strong network externalities, mainly associated to the skilled diaspora. These network externalities are stronger for countries exhibiting intermediate corruption index.

**JEL Classification:** F2, O15, Z13

**Keywords:** FDI, Migration, Brain drain, Network, Diaspora

---

\*This paper is a part of the research projects "People and firms" and "Sustainable Development in a Diverse World" conducted by the Centro Studi Luca d'Agliano (Italy). We thank Giorgio Barba Navaretti, Alok Bhargava, Matteo Manera, Matteo Picchio and Hillel Rapoport for their helpful comments. The usual disclaimers apply. Correspondence address: Elisabetta Lodigiani, Université Catholique de Louvain, IRES, Department of Economics, 3 Place Montesquieu, B-1348 Louvain-La-Neuve, Belgium. Email: lodigiani@ires.ucl.ac.be.

# 1 Introduction

For the last decades, the pace of international migration has accelerated. According to the United Nations, the number of international migrants increased from 75 to about 200 million between 1960 and 2005. An increasing proportion of them is concentrated in high-income countries. The phenomenon is likely to further develop in the coming decades given the rising gap in wages and the differing demographic futures in developed and developing countries. The consequences of emigration for countries of origin have attracted the increased attention of policymakers, scientists and international agencies.

Many observers have emphasized the benefits from unskilled migration and the costs of skilled migration for developing countries. However, alongside the direct impact on the labor market, migrants generate multiple feedback effects on their origin countries. An important channel concerns remittances. The recent Global Economic Prospects (World Bank, 2006) stress the substantial welfare gains for migrants' families. Officially recorded remittances worldwide exceeded \$232 billion in 2005, twice the level of international aid. About 72 percent of this goes to developing countries. In addition, recent models in the brain drain literature emphasize the beneficial effects of skilled migration prospects on education enrollment and the benefits associated to return migration (after additional skills and knowledge have been acquired abroad)<sup>1</sup>. This literature shows that the global impact of skilled migration on human capital is ambiguous.

Network or diaspora externalities constitute an additional channel through which migration affects source countries. By creating trust, providing market information and reducing transaction costs, the diaspora abroad acts as promoting trade, investment and technology adoption in the origin country. The purpose of our paper is to evaluate the magnitude of these "business network" externalities on foreign direct investment (FDI). In a global context, FDI inflows constitute a major source of capital accumulation and technology diffusion in

---

<sup>1</sup>See Stark et al (1997, 1998), Mountford (1999), Beine et al (2001, 2003), Stark and Wang (2002).

developing countries. As suggested by various sectoral studies, the diaspora impact is likely to be linked to the presence of skilled migrants abroad. As a result, a beneficial brain drain can be obtained, even when depressing the average level of schooling in the emigration country.

Using an original data set on emigration stocks by educational attainment and FDI-funded capital stock, we empirically evaluate the relationship between FDI, the size and the educational structure of the diaspora. Our empirical study has three important characteristics:

- First, it relies on two original sources of data. Regarding FDI, we use a classical inventory method to evaluate the FDI-funded stock of capital per worker in a large number of countries. As FDI flows are very volatile and can hardly be interpreted in the long-run (long-run equilibria are usually defined in terms of stocks), working on stock data is much better than working on flows. Regarding migration, we distinguish skilled and unskilled migrants and, contrary to previous studies, we also use migration stocks instead of flows. We rely on a new comprehensive and consistent data set on international migration by educational attainment, that describes the loss of skilled workers to the OECD for all countries in 1990 and 2000 (see Docquier and Marfouk, 2006). Our analysis of business network is the first that distinguishes the role of migrants' education.
- Second, we compare cross-section and panel elasticities. Our analysis is guided by the availability of migration data. As migration stocks are only evaluated for 1990 and 2000, our core empirical model describes the change in the FDI-funded capital stock between 1990 and 2000. Such a cross-section regression raises multiple problems of endogeneity and omitted variables. In a second stage, we use an extended and simplified version of the migration data set<sup>2</sup> and test for the network effect in a panel regression with 4 observations by country. Although many controls are not available on a large period, the panel results confirm the existence of strong diaspora

---

<sup>2</sup>See Defoort (2006).

effects of similar intensity.

- Finally, instead of relying on bilateral data, our analysis is based on the aggregate stock of FDI-funded capital received by the world countries. Several reasons motivate our choice. Bilateral FDI data are only available for limited pairs of countries or for limited periods. Considering aggregate FDI inflows allows computing the total stock of capital and save us modeling the competition and/or coordination between foreign investors. In addition, since international migrants are strongly concentrated in one or two OECD countries, our analysis can be viewed as bilateral analysis between all the world countries and their main OECD partners.

We find evidence of important network externalities. Our analysis confirms that business networks are mostly driven by skilled migration. The elasticity of the FDI-funded capital growth rate to skilled migration is between 2 and 3 percent. Hence, the size of the diaspora matters. The recent literature on the brain drain reveals the human capital response to skilled migration is likely to be positive in large countries characterized by low rates of migration. This paper brings an additional channel through which large countries may benefit from skilled migration: having a large educated diaspora abroad stimulates physical capital accumulation. On the other hand, small countries are less likely to benefit from skilled migration. Finally, we show that diaspora externalities are stronger in countries where corruption is not too high and not too low.

The rest of the paper is organized as following. Section 2 gives a brief overview of the literature on network. Section 3 describes the simple framework we use to derive a dynamic equation of capital accumulation. Data are presented in Section 4. Cross-section results are discussed in Section 5. Section 6 gives the panel estimates. Finally, Section 7 concludes and discusses possible extensions.

## **2 Brief literature review**

Diaspora externalities have long been recognized in the sociological literature and, more recently, by economists in the field of international trade. In many

instances indeed, and contrarily to what one would expect in a standard trade theoretic framework, trade and migration appear to be complements rather than substitutes thanks to the participation of migrants to trade networks that reduce transaction and other types of information costs<sup>3</sup>. The same 'transaction cost' argument holds for the relationship between migration and FDI. To the extent that skilled migrants participate in business networks that contribute to reduce transaction costs between the host and home countries, skilled migration will encourage future FDI flows, which will foster activity and welfare in the emigration country.

Rauch (2003) explains the importance of networks/diasporas as conduits for trade, investment and technology transfer from North America and Europe to the less developed world. Also IOM (International Organization for Migration) stresses the importance of diaspora contributions within FDI and Trade. For example, it is estimated that 50 to 70 percent of FDI in China originated in the Chinese diaspora. The importance of Chinese networks is confirmed for instance by Gao (2003) and Tong (2003).

Why would diasporas be important in promoting international trade and investments? Rauch (2003) stresses two major channels through which the diaspora could promote international trade and investments . First, it creates (or substitutes for) trust in a weak international legal environment. Co-ethnic networks provide community enforcement of sanctions to deter opportunism and violations of contracts. If a party acts opportunistically, then its reputation would suffer within that network. Second, the diaspora provides market information or supplies matching and referral services. Co-ethnic networks can promote trade because they are familiar with the market needs in their country of origin. They can provide important information to foreign investors, which may otherwise be difficult or costly to obtain. In addition, they reduce communication barriers: migrants know the language, the culture, the values, the law and the practices of their home country. They know the way of thinking of

---

<sup>3</sup>See for example Gould (1994), Lopez and Schiff (1998), Rauch and Trindade (2002), Rauch and Casella (2003), Wagner, Head and Ries (2002).

their compatriots and they better understand who is well to trust or not to trust being more aware of potential business partners. The channels just described seem to apply mainly to skilled migrants, as it is confirmed by various sectoral case-studies, notably in the case of the software industry (Saxenian, 1999, 2001, Arora and Gambardella, 2004).

A few empirical studies aimed at measuring the magnitude of the diaspora externality. In his study on the role of ethnic Chinese networks in attracting FDI, Gao (2003) considered both the population share of ethnic Chinese and the log of the absolute population of ethnic Chinese in the source country.

In a gravity model framework, Tong (2003) studied the role of ethnic Chinese in promoting bilateral investments by using the product of the numbers of ethnic Chinese in pairs of countries in 1990.

Focusing on the mechanisms through which NAFTA-related variables might work to reduce migration to the US, Aroca and Maloney (2004) used data on migration flows from and within Mexico without distinguishing between skilled and unskilled migrants. They found that both FDI and trade variables are substitutes for labor flows (FDI and trade reduce migration).

In another study on Germany, Buch, Kleiner and Toubal (2003) used data on inward and outward migration of Germans and foreigners. As they could not have information on the stocks of Germans living abroad or on foreigners living in Germany, they computed gross and net stocks of migrants in order to obtain proxies for the community of Germans living abroad and of foreigners living in Germany, respectively. They find that FDI are complements to migration: there is a relatively strong link between the stocks of German migrants and the stocks of German FDI abroad. For the immigration of the foreigners and FDI inflows, the evidence is weaker. They tested also for causality: they found that with regard to outward FDI and emigration of Germans, the causality seems to run from migration to FDI; with regard to FDI inflows and immigration of foreigners, the causality seems to run from FDI to migration.

Only in a more recent study Kugler and Rapoport (2005) combined US Census data on immigration stocks by country of origin and education level for

1990 and 2000 with data from the US Bureau of Economic Analysis on FDI outflows by destination country and sector. They model the relationships of substitutability or complementarity between migration (by skill level) and the sectoral composition of FDI. They find that skilled migration and FDI inflows are negatively correlated contemporaneously but past skilled migration is associated with an increase in current FDI inflows. Moreover, they find evidence of substitutability between current migration and FDI for migrants with secondary education and of complementarity between past migration and FDI for unskilled migrants.

### 3 What should be tested?

In a couple of papers, Razin et al. (2005a,b) develop a stylized neo classical model to explain FDI bilateral flows between countries. Introducing setup fixed costs, they explain why zero reported flows can be observed between pairs of countries. Here, focusing on aggregate FDI (rather than bilateral flows), we consider a similar stylized model abstracting from setup fixed costs but accounting for risk premium and informational costs. These costs are assumed to be decreasing in the size of the emigration network, especially the skilled diaspora abroad. Testable implications can be derived from this simple framework.

Let us consider a small open economy characterized by a production function à la Barro, Mankiw and Sala-I-Martin (1995):

$$Y_{i,t} = A_{i,t}F(K_{i,t}, H_{i,t}, N_{i,t}) \quad (1)$$

where  $A_{i,t}$  determines the total factor productivity,  $K_{i,t}$  is the stock of capital,  $H_{i,t}$  is the stock of human capital,  $N_{i,t}$  is the size of the labor force and  $F(\cdot)$  is the production function, homogenous of degree one (constant returns to scale).

As capital is mobile across countries, the optimal marginal productivity of capital is equal to the international interest rates  $r_t^*$  augmented of a country-specific risk premium  $\pi_{i,t}$ . This gives

$$\frac{\partial Y_{i,t}}{\partial K_{i,t}} = A_{i,t}F'_K(K_{i,t}^*, H_{i,t}, N_{i,t}) = r_t^* + \pi_{i,t} = \rho_{i,t}$$

Since  $F'_K$  is homogenous of degree zero, the first order condition determines the optimal quantity of capital per worker,  $k_{i,t}^*$ , and the optimal capital stock,  $K_{i,t}^*$ :

$$K_{i,t}^* \equiv k_{i,t}^* N_{i,t} = \kappa(A_{i,t}, h_{i,t}, \rho_{i,t}) N_{i,t}$$

where  $h_{i,t}$  is the country-wide average level of human capital and  $\kappa$  is the inverse function of  $F'_K$ . We have  $\kappa'_A > 0$ ;  $\kappa'_h > 0$  and  $\kappa'_\rho < 0$ .

For many reasons (imperfect information about the determinants of the optimal capital stock, transaction and adjustment costs, time to build private infrastructure, etc.), adjusting the capital stock to its optimal value is not an instantaneous process. For simplicity, we consider that capital adjustments follow a simple Cobb-Douglas process:

$$K_{i,t} = K_{i,t-1}^{1-\beta} K_{i,t}^{*\beta} \quad (2)$$

Taking logs and working in units of labor, we have:

$$\ln k_{i,t} = (1 - \beta) \ln k_{i,t-1} - (1 - \beta) \ln m_{i,t} + \beta \ln k_{i,t}^*$$

where  $m_{i,t}$  is one plus the growth rate of the labor force.

The log-ratio of physical capital is then given by the following equation

$$\ln \frac{k_{i,t}}{k_{i,t-1}} = \beta \ln \kappa(A_{i,t}, h_{i,t}, \rho_{i,t}) - \beta \ln k_{i,t-1} - (1 - \beta) \ln m_{i,t} \quad (3)$$

Several elements determining the optimal level of capital are endogenous:

- Network effects influence the country-risk premium. Denoting by  $M_{i,t}$  the total number of migrants and by  $h_{i,t}^M$  the share of high-skill among migrants, we can write

$$\rho_{i,t} = \rho(M_{i,t}, h_{i,t}^M, Risk_{i,t}),$$

where  $Risk_{i,t}$  is an indicator of economic and political climate that can be captured by the democracy index, the corruption perception index, political instability, etc.



- The total factor productivity is likely to depend on human capital (according to the Schumpeterian model of technology diffusion), lagged physical capital per worker (under learning by doing), the population size and growth rate (due to increasing returns) and variables capturing the economic and political climate:

$$A_{i,t} = A(h_{i,t-1}, k_{i,t-1}, N_{i,t}, m_{i,t}, Risk_{i,t}),$$

- The average level of schooling could also be treated as endogenous.

Plugging all these effects into (3), we obtain the following  $\beta$ -convergence empirical model:

$$\ln \frac{k_{i,t}}{k_{i,t-1}} = a_0 + a_1 \cdot \ln k_{i,t-1} + a_2 \cdot h_{i,t} + a_3 \cdot \ln m_{i,t} + a_4 \cdot \ln N_{i,t} \quad (4)$$

$$+ a_5 \cdot \ln M_{i,t} + a_6 \cdot h_{i,t}^M + a_7 \cdot Risk_{i,t} + a_8 \cdot X_{i,t} + \epsilon_{i,t}$$

where

- $a_0$  is a constant;
- $a_1$  captures the convergence speed towards the optimal amount of capital per worker: we expect a negative effect ( $a_1 = -\beta$ );
- $a_2$  captures the effect of human capital on capital accumulation: the effect is ambiguous. On the one hand, the proportion of high-skill workers has a positive effect on labor productivity. On the other hand, it is more and more usual to relocate part of the production process in countries endowed in unskilled labor when the cost of labor is low;
- $a_3$  and  $a_4$  capture the dilution effect of population growth and size as well as the market size. We expect dilution effect to play negatively on capital per worker: since capital adjustments take time, a rise in the labor force has a negative impact on capital per worker in the short-run. However, a large labor force or a high population growth rate induces an increasing market size which is more attractive for investments;

- $a_5$  measures the intensity of the migration business network. We expect a positive sign if network effects are significant;
- $a_6$  determines the importance of high-skill workers in determining the business network externality. We expect a positive estimate;
- $a_7$  controls for the political environment;
- $a_8$  is a vector of parameters associated to a set of addition controls  $X_{i,t}$ ,
- $\epsilon_{i,t}$  is the residual.

In Section 5, this empirical model will be applied to data on the FDI-funded stock of capital. In Section 6, we will use a panel regression model.

Are migration and FDI substitutes or complements? From equation (4), the general impact of migration on capital accumulation is quite difficult to evaluate. Using  $K_{i,t} = k_{i,t} \cdot N_{i,t}$ , we have:

$$\begin{aligned} \ln K_{i,t} = & a_0 + (1 + a_1) \cdot (\ln K_{i,t-1} - \ln N_{i,t-1}) + a_2 \cdot h_{i,t} + a_3 \cdot \ln m_{i,t} \\ & + (1 + a_4) \cdot \ln N_{i,t} + a_5 \cdot \ln M_{i,t} + a_6 \cdot h_{i,t}^M + a_7 \cdot Risk_{i,t} + a_8 \cdot X_{i,t} \end{aligned}$$

Obviously, a new migrant (leaving her country between  $t - 1$  and  $t$ ) induces a one-for-one decrease in  $N_{i,t}$  and a one-for-one increase in  $M_{i,t}$ . Older migrants (who left before  $t - 1$ ) impacted on  $N_{i,t-1}$  and in turn on  $K_{i,t-1}$ . Migration also affects the structure of the labor force ( $h_{i,t}$ ), the education structure of the network ( $h_{i,t}^M$ ) and the growth rate of the labor force ( $m_{i,t}$ ). Therefore, deriving the global impact of contemporaneous migration on capital is a complex task. Hence, we will not address the issue of global substitutability or complementarity between migration and FDI. Our purpose will only be to check for the existence of network effects in capital accumulation and to examine whether such diaspora relationships are skill biased.

## 4 Data issues

In this section, we describe the data used to test the relationship between the diaspora size and the foreign capital stock in activity. In particular, we describe

how we have built measures for the capital stock per worker (FDI-funded or total) and for the network size (by educational attainment).

**Capital stock data.** Data on foreign direct investments and the gross formation of physical capital are taken from the World Development Indicators. Regarding FDI, this data set gives the total inflows by country, abstracting from the origin of the inflows and the type of FDI. Hence, our analysis will focus on the diaspora impact on aggregate FDI inflows rather than on bilateral exchanges. In addition, we will not distinguish between vertical FDI that aim at relocating a part of the activity, and horizontal FDI that aim at launching a good on a new market.<sup>4</sup>

Data on FDI and total investments are mainly available from the late 1970's (about 100 observations in 1975 for both variables) and are available for about 150 countries in the recent years. Tables 1 and 2 give a broad pictures of the data. We compute the average growth rate of FDI and the share of FDI in total investment by income group, by region of particular interest and by level of corruption. Regarding corruption, we distinguish four groups of equal size according to the quartiles of the distribution. The first group is made of countries whose corruption index is less than or equal to 2.4; the second group by countries with a level comprised between 2.4 and 3.3; the third one with a level comprised between 3.3 and 5.8, the fourth, the less corrupted one, with a level higher than 5.8.

Table 1 compares the average annual growth rate of FDI and total investments evaluated in constant 1995 \$US between 1980 and 2000. With globalization, the growth rate of FDI have been much stronger than the growth rate of total investments in the last twenty years, except in the early nineties. Strong disparities can be observed across groups of countries. High growth rates are observed in high-income, OECD countries, comforting the fact that economic activity tends to concentrate where initial productivity is high, i.e. where human capital is high or where the number of firms is large. However, as many stages of the production process can be relocated in countries where the cost of labor

---

<sup>4</sup>For a more precise definition see Barba Navaretti and Venables, 2004

is low and as new markets are emerging all around the world, other less developed countries have also benefited from large investments. FDI grew rapidly in low-income countries in the early eighties and nineties, especially in Asian countries. Sub-Saharan and Latin American countries also exhibit high growth rates in the early nineties. It is also worth noticing that FDI growth rates have been stronger in less corrupted, especially in the early stages of globalization.

**Tab 1. Average annual growth rate of FDI and total investment**

	Growth rate of total investment (1995 US\$, average per year)				Growth rate of gross FDI (1995 US\$, average per year)			
	1980-1984	1985-1989	1990-1994	1995-2000	1980-1984	1985-1989	1990-1994	1995-2000
<i>Total</i>	0.8%	4.8%	0.5%	3.2%	4.3%	19.1%	-1.1%	27.1%
<i>By income group</i>								
High Income Countries	1.2%	4.7%	0.6%	3.7%	4.6%	19.9%	-3.9%	29.5%
Upper-Middle Income Countries	-4.8%	7.3%	5.5%	2.4%	0.8%	6.8%	16.4%	16.8%
Lower-Middle Income Countries	0.2%	1.9%	-10.9%	-3.1%	1.5%	8.4%	16.6%	9.2%
Low Income Countries	4.4%	4.9%	8.7%	3.1%	10.7%	15.1%	44.6%	2.8%
<i>By Region</i>								
Mena	3.0%	-3.8%	1.1%	1.2%	7.0%	8.3%	8.2%	9.9%
Sub-Saharan Countries	-2.8%	1.1%	1.9%	0.6%	-5.3%	3.3%	20.6%	-2.0%
East Asia	1.6%	7.0%	0.6%	1.0%	18.7%	27.3%	-6.2%	23.9%
South East Asia	4.0%	8.8%	7.7%	-5.1%	-0.7%	24.3%	6.0%	6.0%
Other Asia	3.8%	7.2%	0.5%	3.1%	10.1%	15.9%	7.9%	8.4%
Latin America	-7.5%	4.0%	7.3%	3.0%	-4.4%	3.9%	22.8%	17.7%
OECD Countries	1.0%	4.8%	0.5%	3.9%	4.7%	19.8%	-3.8%	28.3%
<i>By Corruption Index Level</i>								
1 level	-1.2%	1.5%	-4.5%	-3.3%	-6.7%	19.1%	15.8%	10.6%
2 level	-0.6%	5.0%	-8.7%	1.5%	-9.6%	11.2%	17.1%	11.3%
3 level	-2.2%	4.1%	4.3%	3.0%	6.5%	3.8%	20.1%	10.2%
4 level	1.5%	5.0%	0.9%	3.5%	4.3%	20.3%	-3.6%	29.6%
<i>By size</i>								
Large ad Upper Middle Countries	0.9%	4.9%	0.5%	3.2%	2.7%	20.7%	-1.9%	25.2%
Small and Lower Middle Countries	-0.5%	4.4%	-0.1%	3.4%	22.3%	11.5%	3.4%	35.4%

Source: World Bank Development Indicators (2003). Own calculations.

Table 2 gives a broad picture of the share of FDI in the gross formation of physical capital. This proportion indicates whether or not globalization affected the ownership of capital all around the world. On the whole sample, the share of FDI increased from 5.4 percent in 1980 to 39.3 percent 2000. Remarkable increases were observed between 1995 and 2000 and to a lesser extent, between 1985 and 1990. The largest changes are obtained for high-income OECD countries. Nevertheless, important relative changes are also observed in South-East Asia, East Asia, Latin America. A remarkable fact is that changes in the FDI proportion vary with country size (usually capturing the degree of openness): small countries have experienced a drastic increase in FDI over the last years.

**Tab 2. Share of FDI in total investment (selected years)**

	1980	1984	1985	1989	1990	1994	1995	2000
<i>Total</i>	5.4%	6.4%	6.0%	11.3%	10.8%	10.0%	11.3%	39.3%
<i>By income group</i>								
High Income Countries	5.8%	6.9%	6.4%	12.7%	13.1%	10.5%	12.0%	45.7%
Upper-Middle Income Countries	4.5%	5.6%	4.5%	4.4%	4.2%	6.9%	8.6%	19.0%
Lower-Middle Income Countries	2.1%	2.3%	4.0%	5.5%	1.5%	5.6%	6.3%	13.0%
Low Income Countries	0.9%	1.2%	1.9%	3.0%	2.8%	11.7%	10.8%	10.7%
<i>By Region</i>								
Mena	7.3%	8.8%	4.0%	7.3%	6.8%	9.5%	9.5%	15.6%
Sub-Saharan Countries	1.9%	1.7%	6.3%	7.0%	5.1%	11.9%	13.9%	11.9%
East Asia	0.8%	1.8%	2.0%	4.7%	5.0%	3.5%	3.6%	12.2%
South East Asia	7.0%	5.5%	5.7%	11.1%	14.5%	13.4%	11.7%	22.7%
Other Asia	5.0%	6.7%	2.0%	2.9%	2.6%	3.7%	5.3%	7.2%
Latin America	3.8%	4.5%	4.8%	4.7%	4.4%	8.7%	9.9%	22.0%
OECD Countries	5.7%	6.8%	6.4%	12.5%	12.7%	10.3%	12.0%	42.6%
<i>By Corruption Index Level</i>								
1 level	1.5%	1.2%	2.4%	5.4%	2.5%	6.5%	7.6%	16.9%
2 level	2.9%	1.8%	2.9%	3.9%	1.2%	4.1%	5.4%	9.4%
3 level	3.0%	4.6%	3.8%	3.7%	4.7%	9.6%	10.3%	15.5%
4 level	6.0%	6.9%	6.6%	13.0%	13.2%	10.6%	12.0%	46.2%
<i>By size</i>								
Large ad Upper Middle Countries	5.6%	6.1%	5.2%	10.5%	10.1%	8.9%	10.4%	33.1%
Small and Lower Middle Countries	3.3%	9.1%	14.4%	19.9%	19.3%	22.9%	21.8%	99.7%

*Source: World Bank Development Indicators (2003). Own calculations.*

As investment flows are extremely volatile and cannot capture the long-run trends of nations (in the long-run, flows just compensate for depreciation and demographic growth), our analysis is based on stock data. Obviously, there is no data set providing series of capital stock (*a fortiori*, FDI-funded capital stock) by country. We thus use investment data to construct capital stock series for 114 countries in 1990 and 2000. We distinguish the FDI-funded capital stock and the total capital stock. We use a classical inventory method based on the standard equation of capital accumulation:

$$K_{i,t} = K_{i,t-1}(1 - d) + I_{i,t-1}$$

where  $d$  is the depreciation rate (fixed at 4 percent a year) and  $I_{i,t-1}$  is the amount of FDI or total investment alternatively.

We start from an hypothetical long-run value given by

$$K_{i,1980} = \frac{I_{i,75-80}}{d}$$

where  $I_{i,75-80}$  is the growth-corrected average amount of investment between 1975 and 1980.

We then apply the capital accumulation function sequentially to compute annual stocks from 1980 to 2000 . Series of capital per worker  $k_{i,t}$  are obtained by dividing the capital stock by the labor force, measured as the population aged 25 and more in the country.

**Migration and human capital data.** Data on the population aged 25 and more (proxy of the labor force) are provided by the United Nations. The labor force is splitted across educational group using international human capital indicators. 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),
- high-skilled workers are those with tertiary education (13 years and above).

Several sources are combined. Following Docquier and Marfouk (2006), we use De la Fuente and Domenech for OECD countries and Barro and Lee (2000) data for other countries. For countries where Barro and Lee measures are missing, we use Cohen and Soto's available indicators (2001) or we transpose the skill sharing of the neighboring country with the closest rate of enrollment in education. Hence, data on the labor force by educational attainment are available for all the world countries.

Regarding migration, our analysis builds on a new comprehensive and consistent data set on international migration by educational attainment (see Docquier and Marfouk, 2006). This data set describes the loss of skilled workers to the OECD for all countries in 1990 and 2000. They distinguish the same educational groups as in the human capital data above. Emigration stocks by educational attainment are computed for every country of the world. These

stocks are obtained by aggregating consistent immigration data collected in receiving countries. Docquier and Marfouk 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, they focus on the South-North and North-North brain drain. Although a brain drain can be observed outside the OECD area (to the Gulf countries, South Africa, Malaysia, Hong-Kong, Singapore, Taiwan, etc.), they estimate that about 90 percent of high-skill international migrants are living OECD countries. Data are available for all the world countries. They measure the size of the diaspora residing in the OECD, by educational attainment.

These data can be used for the cross-section analysis of the determinants of FDI. For extended panel regressions, we use the estimates provided in Defoort (2006). Focusing on the six major destination countries (USA, Canada, Australia, Germany, UK and France), they have computed skilled emigration stocks and rates from 1975 to 2000 (one observation every 5 years). On the whole, the six destination countries represent about 75 percent of the OECD total immigration stock. However, for some origin countries, the coverage is quite low. For example, Surinamese emigrants mainly live in the Netherlands. About 3 percent of Surinamese emigrants live in the six major receiving countries. The panel analysis is then based on much reliable econometric techniques, but less reliable data.

**Other data.** As for FDI and total investments, the world development indicators provide many information about other country characteristics such as population size and growth, level of income.

Data on corruption are taken from the Transparency International data set. The Corruption Perception Index (CPI) is a composite index reflecting the perceptions of business people, country analysts, both resident and non-resident. It draws on 16 different polls from 10 independent institutions. The CPI score



relates to perceptions of the degree of corruption as and ranges between 10 (non corrupted country) and 0 (highly corrupted country). A number of countries - including some which could be among the most corrupt - are missing because not enough survey data is available. To maximize the number of observation and reduce to risk of selection bias, we use the 2004 CPI scores which are available for 146 countries. The highest scores (around 9.5) are obtained in Scandinavian countries, New Zealand, Iceland, Singapore. The worst scores are obtained in Haiti and Bangladesh (1.5), Nigeria (1.6), Burma and Chad (1.7), Paraguay (1.9).

Data on political regime are taken from the POLITY IV data set. The indicator of democracy ranges from 0 in dictatorial regimes to 1 in democratic regimes. It measures the general openness of political institutions and combines variables such as the regulation of Executive Recruitment (institutionalized procedures regarding the transfer of executive power), the competitiveness of executive recruitment (extent to which executives are chosen through competitive elections), the openness of executive recruitment (opportunity for non-elites to attain executive office), executive constraints (operational independence of chief executives), the regulation of participation: development of institutional structures for political expression) and the competitiveness of participation (extent to which non-elites are able to access institutional structures for political expression). The worst scores (0) are obtained in Afghanistan, Burma, Cuba, Equatorial Guinea, Iraq, Libya, Saudi Arabia, Sudan, Syria, Turkmenistan.

## 5 Empirical analysis

Our general  $\beta$ -convergence model is given by (4) in which  $k_{i,t}$  measures the FDI-funded capital stock per worker in country  $i$  at time  $t$ . The dependent variable is the average annual real growth rate of  $k_{i,t}$  between 1990 and 2000. Building on (4), we introduce a set of controls  $X_{i,t}$  which were shown to influence investment decision in existing empirical studies. As argued by Barba Navaretti and Venables (2004), "explanatory variables can be a vector of firm

and/or industry characteristics, of home country characteristics, of host country characteristics and of bilateral relationships between home and host countries, such as the distance between them. Choice of variable to use depends partly on the hypothesis being investigated and partly on data availability”.

To avoid serious multicollinearity problems, we do not incorporate all potential controls simultaneously. We compare several regressions and try to end up with the most reasonable model in which only significant variables are kept. In these regressions, we will consider the corruption index, the democracy index, the distance with two important industrialized regions (the USA and the EU15) as well as international trade.

Geographical distance can be used as a proxy for trade costs (Gao, 2003). Trade costs can have opposite implications for the pattern of FDI. Vertical FDI are negatively affected by distances as they involve trade. Horizontal FDI are likely to increase with distances (one of the main reason of horizontal FDI is to serve foreign markets minimizing trade costs). Usually geographical distance is considered as one of the most important obstacles to FDI, meaning that (i) there could be a dominance of vertical FDI, but also that (ii) setup fixed costs involved by horizontal FDI can be positively correlated with distance (Markusen and Venables, 2000). The recent literature also assimilates greater geographic distance to greater 'cultural' distance and thus larger communication and information costs (Butch et al., 2005). In this sense, greater distance could have a direct (negative) effect on both vertical and horizontal FDI.

Similarly, the degree of trade openness has an ambiguous impact on FDI, depending if the type of investment. In the case of horizontal FDI, more openness induces less investments. In the case of vertical FDI, the opposite correlation is expected. Considering the importance of vertical investment towards developing countries that occurs from the 1990's onward, we include as a measure of trade openness the log of the trade (imports + exports) with OECD countries in percentage of the 1990 GDP. However more trade could not only be an indicator of vertical FDI, but also an indicator of openness (related to the country size), competitiveness and therefore attractiveness of the country. Open economies

are likely to be more attractive for FDI since transnational corporations can reap economies of scale and scope, even in countries where the market size is small. That could be one of the reasons why in the latest years developing countries increased their participation in regional integration scheme.

We will also introduce a dummy variable to underline that high income countries are more attractive for capital investments.

**General model.** Table 3 gives the results for our general specification. Five alternative models are distinguished. Since heteroskedasticity can be important across countries, the standard errors for the coefficients are based on White's heteroskedasticity-consistent covariance matrix.

**Tab 3. Cross-section - general specification**

Dependent variable = Growth rate of FDI-funded capital stock per worker

	Model 1	Model 2	Model 3	Model 4	Model 5
logfdi90pc	-0.040 (3.58)***	-0.040 (4.23)***	-0.042 (4.15)***	-0.042 (4.21)***	-0.041 (4.06)***
dlogpop	-0.064 (0.37)	-0.016 (0.09)	-0.005 (0.02)	-0.043 (0.24)	-0.022 (0.12)
logMSTOT90	0.021 (1.90)*	0.013 (2.04)**	0.015 (2.11)**	0.017 (2.51)**	-
SHM90	0.207 (2.27)**	0.179 (2.10)**	0.291 (2.58)**	0.174 (2.04)**	-
logMSHIGH90	-	-	-	-	0.019 (2.52)**
HIGHincome	0.097 (2.77)***	0.083 (2.16)**	0.031 (-0.68)	0.090 (2.33)**	0.088 (2.29)**
logPAC90	-0.008 (0.66)	-	-	-	-
DemocracyIndex	-	0.077 (1.72)*	0.115 (2.41)**	0.082 (1.95)*	0.077 (1.74)*
logtrade	-	-	0.036 (2.07)**	0.034 (2.13)**	0.036 (2.27)**
logdistUSA	-	-	-0.011 (0.48)	-	-
logdistEU15	-	-	-0.040 (1.55)	-	-
Constant	0.162 (0.88)	0.076 (0.63)	0.344 (1.06)	-0.040 (0.34)	0.013 (0.13)
Observations	114	113	96	109	109
R-squared	0.50	0.50	0.57	0.53	0.52

Robust t statistics in parentheses  
\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

The main results can be summarized as follows:

*Convergence speed.* In every specification, the estimated coefficient of the lagged dependent is highly significant and very stable. We find a convergence speed of about 4 percent per year.

*Market size.* A potentially important determinant of FDI is the market size. Since our dependent variable is capital per worker, the size of the market is neutralized on the left hand-side. Anyway, under increasing returns (which can be related to fixed setup costs), the market size may positively affect the capital stock per worker. In model 1, we control for the log of the working-aged population in 1990 as a proxy to the market size. This variable is not significant. Similar results were obtained with log of the total population (regression not reported). We did not consider the log of the GDP because of endogeneity problems. There is no evidence of additional market size effect on the right hand-side. The estimated coefficient of the growth rate of the labor force is negative (as expected) but statistically not significant. In separate regressions, to avoid multicollinearity problems, we ran regressions considering the rate of growth of the labor force by skill level: in all cases, this variable is never significant.

*Structure of the labor force.* The structure of the labor force is potentially important in predicting FDI inflows. As argued in section 2, the proportion of high-skill workers has a positive effect on labor productivity. Nevertheless, part of the production process in countries endowed in unskilled labor when the cost of labor is low. We obtain evidence that the average level of schooling has a positive effect on FDI inflows. However, including the share of high-skill workers causes serious problems of stability given the strong multicollinearity with many variables such as the lagged capital stock per worker. By adding a dummy for high-income countries, we capture the strong attractiveness of human capital. The coefficients are very stable across samples and specifications. The coefficient for this dummy is positive and generally significant.

*Country openness.* As trade costs and various types of trade barriers are crucial in explaining the pattern of FDI, we introduce the distance with the most important countries and trade in model 3. The estimated coefficients for

distance are negative, but statistically not significant. This can be due to the fact that our study focuses on total FDI inflows rather than on bilateral exchanges. On the contrary, the estimated coefficient of trade openness is positive and statistically significant.

*Political climate.* In models 2 to 5, we control for democracy as a potential determinant. The estimated coefficient is positive (between 0.077 and 0.115) and statistically significant at 5 or 10 percent. This measure serves as our proxy for the domestic investment environment, assuming that a stable macroeconomic environment generates more investment. In separate regressions, we considered also a variable measuring the size of the informal market (component of the corruption index). The estimated coefficient was negative sign but it was never significant. Similarly, political instability coefficients were not statistically significant. Another important candidate is the corruption index. Nevertheless, as with human capital indicators, perception corruption index induce serious problems of multicollinearity and stability of the model. The corruption index is highly correlated with the high income dummy (0.8225) and with the lagged dependent capital stock per worker (0.5745). For this reason, we will come back to corruption below.

*Network effects.* The estimated coefficients of the log of the stock of total expatriates and of the share of high skilled migrants are always positive and highly significant. The migration stock is the only significant variable capturing the size of the country. However, given the discussion about the market size, we have strong reasons to believe that such an effect is related to diaspora rather than to market size. For example, in model 1, we consider both the size of the labor force and the stock of expatriates. We obtain a positive diaspora effect despite a strong correlation (0.68 ) between the network size and the labor force (which turns out to be insignificant). In further regressions, by excluding the labor force and reducing the risk of multicollinearity, the diaspora effect becomes very significant. Model 4 is our finest specification with all our statistically significant variables. The elasticity of capital per worker to migration amounts to 1.7 percent. A 10 percentage point rise in the number of migrants increases

the growth rate of capital per worker by 1.74 percent. The proportion of skilled migrants is also an important factor of business externality. Our results reveal that business networks are mostly driven by skilled migration. In model 5, we impose diaspora effect to transit through skilled workers and obtain a 2 percent elasticity of capital growth to migration.

In the rest of this section, we start from our best specifications (models 4 and 5) and allow for non linear effects. We examine whether the intensity of diaspora externalities varies with the level of corruption and compares the impacts on FDI-funded capital and on the total capital stock per worker.

**Business networks and corruption.** In the general specification, we exclude corruption for multicollinearity and stability reasons. Let us now examine whether corruption affects the size of diaspora effects. The literature stresses the importance of networks both in weak and strong international legal environment (i.e. corrupted and not corrupted countries). While a good investment climate is likely to be more attractive for FDI, ethnic networks are more likely to be crucial in weak environment, where information costs are higher and where community enforcement of sanctions are important. In countries with clean institutions, the role of ethnic networks consists in limiting information costs.

To capture the relationship between networking and the level of corruption, we classify countries according to their perception corruption index. As in Tables 2 and 3, the minimal value is 1.5 for the most corrupted countries and 9.7 for the less one. We distinguish four country groups of equal size according to the quartiles of the distribution. The first group is made up by countries whose corruption index is less than 2.4; the second group by countries with a level comprised between 2.4 and 3.3; the third one with a level comprised between 3.3 and 5.8, the fourth, the less corrupted one, with a level higher than 5.8. We create four dummies, one for each group, and then generate four interaction terms by multiplying the log of the total number of the expatriates (or the log of high skilled migrants) by each dummy. These variables are used as predictors in the regressions.

Table 4 gives the results. Models 1 and 2 are based on the total number

of expatriates and abstract from the skill structure of migrants (to avoid multicollinearity problems); models 3 and 4 are based on the number of skilled expatriates.

**Tab 4. Cross-section - diaspora effect by level of corruption**

Dependent variable = Growth rate of FDI-funded capital stock per worker

	Model 1	Model 2	Model 3	Model 4
logfdi90pc	-0.040 (3.61)***	-0.039 (3.55)***	-0.041 (3.81)***	-0.04 (3.81)***
dlogpop	-0.070 (0.37)	-0.108 (0.54)	-0.042 (0.35)	-0.073 (0.35)
logMSTOT90IP1	0.011 (1.69)*	0.012 (1.79)*	-	-
logMSTOT90IP2	0.013 (2.01)**	0.015 (2.29)**	-	-
logMSTOT90IP3	0.014 (2.21)**	0.017 (2.60)**	-	-
logMSTOT90IP4	0.011 (1.55)	0.014 (2.07)**	-	-
logMSHIGH90IP1	-	-	0.016 (2.00)**	0.016 (2.11)**
logMSHIGH90IP2	-	-	0.019 (2.27)**	0.021 (2.58)**
logMSHIGH90IP3	-	-	0.020 (2.59)**	0.022 (3.00)***
logMSHIGH90IP4	-	-	0.018 (2.08)**	0.021 (2.57)**
Democracy Index	0.080 (1.88)*	-	0.066 (1.56)*	-
logtrade	0.032 (2.25)**	0.029 (1.94)*	0.032 (2.33)**	0.030 (2.07)**
HIGHIncome	0.099 (2.54)**	0.104 (2.77)***	0.088 (2.30)**	0.092 (2.49)**
Constant	0.077 (0.78)	0.109 (1.04)	0.042 (0.39)	0.071 (0.62)
Observations	109	109	109	109
R-squared	0.51	0.50	0.53	0.52
Robust t statistics in parentheses				
* significant at 10%; ** significant at 5%; *** significant at 1%				

All other things being equal, does corruption affect the diaspora externality? Considering the total number of expatriates, the estimate for the most corrupted countries is 0.011 with a significance level of 10 percent; for the second and third groups, we have respectively 0.013 and 0.014 with a significance level of 5 percent; in the less corrupted countries, the estimated coefficient is positive (0.011), but statistically insignificant. The latter result deserves an econometric

extension. In model 1, the VIF (variance inflation factor) for the four group is 14.61. Usually 10 is the tolerated value. A value greater than 10 can reflect multicollinearity problems and inflated standard errors. As democracy index is correlated to corruption and comes out to be the less significant variable in the regression (significant at 10%), we estimate model 2 by excluding this variable. Then, the estimated coefficient for the less corrupted countries becomes positive and significant at 5% and the VIF has a tolerated value. Anyway, the coefficients indicate that network effects are stronger in countries exhibiting an intermediate corruption index.

Considering the number of skilled expatriates, all the coefficients are positive and statistically significant at 5 percent. This result holds with or without controlling for democracy. We obtain the same profile than in models 1 and 2. Networks effects play a minor role into the most corrupted countries. The major effect is obtained for the third and the second groups characterized by intermediate corruption levels. We conclude that corruption induces a inverted U-shape impact on networking. Diaspora externalities seems to play a major role in community enforcement of sanction, at least when the country is not too corrupted. On the contrary the role played by ethnic networks is in providing information is weaker (although highly significant) if the level of corruption is very low.

**Business network and total investment.** For the matter of comparison, we apply our general  $\beta$ -convergence model (4) to the total stock of physical capital per worker (an alternative measure of  $k_{i,t}$ ) rather than to the FDI-funded capital stock per worker. Such an analysis allows us to confirm the existence of diaspora effect at the global level (the FDI-funded capital stock is a component of the total capital stock) or to highlight some compensating effect due, for example, to an joint increase in FDI inflows and outflows.

Basically, we use the same specifications as in table 3. Table 5 describes the results.



**Tab 5. Cross-section - general specification**

Dependent variable = Growth rate of the total capital stock per worker

	Model 1	Model 2	Model 3	Model 4
logk190	-0.015 (2.38)**	-0.017 (2.58)**	-0.017 (2.41)**	-0.017 (2.30)**
dlogpop	-0.088 (3.41)***	-0.072 (2.52)**	-0.084 (2.80)***	-0.081 (2.70)***
logMSTOT90	0.008 (2.60)**	0.006 (1.87)*	0.008 (2.64)***	-
SHM90	0.072 (3.04)***	0.063 (2.38)**	0.055 (1.91)*	-
logMSHIGH90	-	-	-	0.009 (2.62)**
HIGHincome	0.029 (2.12)**	0.024 (1.52)	0.022 (1.4)	0.021 (1.35)
logPAC90	-0.007 (2.68)***	-0.005 (1.76)*	-0.005 (1.45)	-0.005 (1.34)
Democracy Index	-	0.034 (2.49)**	0.032 (2.87)***	0.033 (3.01)***
logtrade	-	-	0.010 (2.21)**	0.011 (2.40)**
Constant	0.196 (2.97)***	0.178 (2.65)***	0.134 (1.60)	0.151 (1.64)
Observations	103	101	95	95
R-squared	0.36	0.41	0.50	0.48
Robust t statistics in parentheses				
* significant at 10%; ** significant at 5%; *** significant at 1%				

Four interesting results are emerging.

*Lower convergence speed.* A first remarkable result is that we obtain a much slower convergence speed (1.7 percent a year instead of 4 percent) for the total capital stock. Over the period 1990-2000, it seems that FDI movements have been much more rapid than local investments. This can be explained either by the general trend of increasing exchanges between countries (globalization) or by stronger imperfections in capital adjustment.

*Imperfections matter.* Although globalization is an undeniable phenomenon affecting the openness of the world countries, imperfections on the local market for capital seem to be stronger. Indeed, the growth rate of the population has a negative effect the growth rate of capital per worker, indicating that the total stock of capital adjusts more slowly to demographic changes. We did not find evidence of such an effect with FDI-funded capital.

*Less sensitivity to trade and political regime.* We also note that the impact of the democracy index and of trade are also divided by 3 compared to table 3. The total stock of capital per workers is less sensitive to the economic and political environment.

*Smaller diaspora externality.* Finally, the network effects are smaller although significant. Compared to table 3, the estimates are divided by 3. This can reflect (i) the fact that the FDI-funded capital stock remains a small fraction of the total capital stock, or (ii) a general tendency towards increased specialization and exchanges between countries. Inflows can be partly compensated by outflows.

## 6 Panel data analysis

In a cross-section setting, the standard ordinary least square estimator with heteroskedasticity consistent standard errors reveals a 2 percent elasticity of the growth rate of the FDI-funded capital stock per worker to the stock of skilled emigrants. These cross-section results can be biased and inconsistent given the dynamic nature of the growth equation. In order to obtain more accurate results, we extend our analysis in a panel setting using a more sophisticated econometric method which take into account the possible endogeneity of explanatory variables and omitted variables. As mentioned above, the quality of the panel data on migration is lower on a large time period. Our objective is to confirm the existence of business network externalities when a robust econometric technique is applied.

There is a large debate about the most accurate methodology to estimate growth equations (see Islam, 1993, 2003, Caselli et al, 1996, Bhargava and Sargan, 1983, Bhargava et al, 2003). Here, we use a GMM system estimator for dynamic panel data model. This technique exploits both the cross-sectional and the time dimension of the data. It accounts for unobserved fixed effects. It controls for the potential endogeneity of all the explanatory variables and allows for the inclusion of the lagged dependent variable.

**Econometric methodology.** Let us briefly present the technique used. Consider the simplified version of the regression equation (4) in which all explanatory variables (except the lagged dependent) are grouped:

$$\ln\left(\frac{k_{i,t}}{k_{i,t-1}}\right) = a_0 + a_1 \cdot \ln(k_{i,t-1}) + \beta' X_{i,t} + \eta_i + \varepsilon_{i,t}$$

where  $X_{it}$  represents the set of the explanatory variables other than the lagged dependent,  $\eta_i$  represents the unobserved country-specific fixed effect,  $\varepsilon_{it}$  is the error term.

This equation can be re-written in the standard dynamic panel form

$$\ln(k_{i,t}) = a_0 + (1 + a_1) \cdot \ln(k_{i,t-1}) + \beta' X_{i,t} + \eta_i + \varepsilon_{i,t} \quad (5)$$

A general approach to estimate such an equation is to use a transformation that removes unobserved effects and that uses for instrumental variables. Anderson and Hsiao (1992) propose to work with first differences and then to search for instruments. They proposed for the lagged dependent either the two period lagged difference or the two period lagged level of the dependent variable. A generalization of that method was proposed by Arellano-Bond (1991). They suggest using the entire set of instruments in a GMM procedure to reach significant efficiency gains.

Differentiating (5) yields

$$\begin{aligned} \ln(k_{i,t}) - \ln(k_{i,t-1}) &= (1 + a_1) \cdot [\ln(k_{i,t-1}) - \ln(k_{i,t-2})] \\ &\quad + \beta' (X_{it} - X_{it-1}) + (\varepsilon_{it} - \varepsilon_{it-1}) \end{aligned}$$

in which the unobserved country fixed effect is eliminated.

By construction the error term  $(\varepsilon_{it} - \varepsilon_{it-1})$  is correlated with the lagged dependent in first differences  $[\ln(k_{i,t-1}) - \ln(k_{i,t-2})]$ . Hence, instrumental variables are required to deal with both the potential endogeneity of all the explanatory variables and the bias due to the presence of the lagged dependent among the regressors. In the Arellano-Bond method, the first-difference of the explanatory variables are instrumented by the lagged values of the explanatory variables in levels. Under the assumptions that the error term is not serially

correlated and that the explanatory variables are weakly exogenous or predetermined (i.e. the explanatory variables are not correlated with future realizations of the error term), the following moment conditions are applied for the first difference equations:

$$\begin{aligned} E[\ln(k_{i,t-s}) \cdot (\varepsilon_{it} - \varepsilon_{it-1})] &= 0 \text{ for } s \geq 2; t = 3, \dots, T \\ E[X_{it-s} \cdot (\varepsilon_{it} - \varepsilon_{it-1})] &= 0 \text{ for } s \geq 2; t = 3, \dots, T \end{aligned}$$

The problem with this method is that taking first differences of the level equation, explanatory variables which are constant over time cannot be taken into account. Moreover, as Bond, Hoeffler and Temple (2001) point out, when time series are persistent, the first-difference GMM estimator can behave poorly: estimates can be seriously biased. To overcome these problems Bond et al. (2001) suggest to use a more informative set of instruments within the framework developed by Arellano and Bover (1995) and Blundell and Bond (1998). We use this new estimator that combines the regression in differences with the regression in levels in a single system. The instruments used in the first differentiated equation are the same as above, but the instruments for the equation in level are the lagged differences of the corresponding variables.

For the level equation the following moments condition are to be satisfied:

$$\begin{aligned} E[(\ln(k_{i,t-s}) - \ln(k_{i,t-s-1}))(\eta_i + \epsilon_{it})] &= 0 \text{ for } s = 1 \\ E[(X_{it-s} - X_{it-s-1})(\eta_i + \epsilon_{it})] &= 0 \text{ for } s = 1 \end{aligned}$$

The validity of the instruments can be tested using a Sargan-Hansen overidentification test (that is a specification test) and a test on the serial correlation of the error term (see Arellano and Bond (1991), Arellano and Bover (1995), Blundell and Bond (1998), Bond, Hoeffler and Temple (2001) )<sup>5</sup>.

---

<sup>5</sup>In our analysis we use the command `xtabond2` implemented in STATA. We used the robust two-step variant. We know that, though asymptotically more efficient, the two-step estimates can be downward biased. But `xtabond2` makes available a finite-sample correction to the two-step covariance matrix derived by Windmeijer. STATA guide suggests this variant for system GMM estimator, because more efficient. However, we tried all the regressions using only the robust one-step variant. The main results of interest did not change very much

**Empirical results.** The period of analysis is divided into 4 sub periods of 5 year each (1980-85, 1985-90, 1990-95, 1995-00). We have 83 countries for a total of 332 observations in a balanced panel data set. One of the most difficult issue to apply the above dynamic panel technique is to identify the nature of the explanatory variables (they can be endogenous, exogenous, weakly exogenous or predetermined<sup>6</sup>). We tried several specification considering the value of the Hansen test and the serial correlation test. At the end, we consider the time dummies and the high income dummy as exogenous variables; all the other time-varying explanatory variables are considered as predetermined (we used their one lagged and earlier values as instruments).

Starting from the best cross-section specification (model 4 and 5 in Table 3), Table 6 gives the results of the panel regressions. We have added interaction terms between high income dummy, democracy and trade. Doing this, we can better understand the different effects that these two variables can have according to the different type of FDI (vertical or horizontal).

---

<sup>6</sup>For the exogenous variables they enter as their own instruments in the regressions, two periods and earlier lagged values of endogenous variables, one period and earlier lagged values of predetermined or weakly exogenous variables can be used as instruments

**Tab 6. Panel - general specification**

Dependent variable = Growth rate of the FDI-funded capital stock per worker

	Model 1	Model 2	Model 3	Model 4
Initial FDI-funded capital in logs	-0,026 (4.04)***	-0,032 (3.34)***	-0,027 (3.96)***	-0,033 (3.08)***
High income dummy	-0,309 (1.22)	-	-0,15 (0.56)	-
Initial GDP per capita in logs	-	0,024 (1.07)	-	0,021 (0.75)
Labor force growth rate	0,007 (0.03)	0,064 (0.29)	-0,035 (0.15)	0,005 (0.02)
Total lagged migration stock in logs	0,031 (2.19)**	0,025 (1.72)*	-	-
Lagged Share of skilled migrants	(0.26) (2.32)**	(0.22) (1.97)*	-	-
Skilled lagged migration stock in logs	-	-	0,029 (2.21)**	0,025 (1.80)*
Democracy score (lagged)	-0,151 (1.93)*	-0,156 (2.04)**	-0,181 (2.18)**	-0,179 (2.27)**
Democracy score (lagged) x High income dummy	0,457 (2.24)**	0,181 (2.10)**	0,382 (1.78)*	0,232 (2.43)**
Lagged trade in logs	0,059 (2.43)**	0,05 (2.44)**	0,07 (2.74)***	0,068 (2.75)***
Lagged trade in logs x High income dummy	0,004 (0.1)	-0,03 (1.27)	-0,028 (0.68)	-0,046 (1.84)*
d90	0,024 (1.26)	0,027 (2.12)**	0,034 (1.98)*	0,036 (2.65)***
d95	0,06 (2.01)**	0,061 (2.48)**	0,073 (2.32)**	0,076 (2.72)***
d00	0,055 (2.84)***	0,063 (3.45)***	0,073 (4.02)***	0,082 (5.23)***
Constant	-0,355 (1.86)*	-0,399 (2.69)***	-0,216 (1.47)	-0,291 (2.34)**
Observations	332	332	332	332
Number of countries	83	83	83	83
Hansen test- Prob > chi2	0,305	0,512	0,105	0,225
Arellano-Bond test for AR(1) in first differences- Pr > z	0,081	0,083	0,081	0,082
Arellano-Bond test for AR(2) in first differences- Pr > z	0,385	0,413	0,392	0,411

Note: System GMM, Robust two-step; t- statistic in parenthesis  
\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

The main results can be summarized as follows:

*Convergence speed.* In every specification, the estimated coefficient of the lagged dependent variables is highly significant. We find a slower convergence speed than in the cross section analysis (about 3 percent per year instead of 4 percent). A possible explanation is that OLS estimates are biased upwards since the lagged dependent variable captures country-specific effects. Here it is instrumented.

*Market size, structure of the labour force and attractiveness of the country.* Working on the capital per worker, the size of the market is implicitly neutralized

on the left-hand side. As in the cross-section, the effect of the growth rate of the labour force is not statistically significant. As we can control for endogeneity, we use an alternative measure of the market size, the log of GDP per capita: the coefficient associated to that variable is positive but not statistically significant. This variable is very correlated with the lagged dependent and with interaction terms (for high income dummy and its interactions with democracy and trade openness, the correlations respectively amount to 0.7166, 0.7115 and 0.7769)<sup>7</sup>. Also the high income dummy alone is not statistically significant and it has not the right sign; it is very correlated with its interaction terms.

*Country openness.* The estimated coefficient of trade openness is positive and statistically significant (indicator of vertical FDI and attractiveness of the country). The interaction with the high income dummy is almost always negative (indicator of horizontal FDI), but in general not statistically significant. One of the possible explanations can be that our model better refers to vertical FDI and captures the exchanges between developed and developing countries.

*Political climate.* We have two variables that indicates the political climate. Democracy and its interaction with the high income dummy. The estimated coefficients are in general statistically significant at 5 or 10 percent. The sign is positive for the interaction term (in general, high income countries are more democratic) and a negative sign for the overall effect. On the aggregate, stronger FDI growth rates have been observed in the developing (less democratic) countries.

*Network effects.* The estimated coefficient of the log of the stock of the total number of expatriates and of the share of high skilled workers are positive and statistically significant. The magnitude of the coefficient, at least for the log of the total stock of migrants, is in general higher than in the cross-section analysis (2.5 to 3 percent, instead of 2 percent). Also the log of the high skilled is

---

<sup>7</sup>We prefer the model with the High Income dummy to the model with GDP per capita since GDP per capita exhibit a strong correlation with other regressors. Hence, standard errors can be inflated. Moreover, the two-step estimated covariance matrix of moment conditions was singular in model 2. The number of instruments may be large relative to the number of groups. A generalized inverse must be used to calculate optimal weighting matrix in the two-step estimation.

positive and statistically significant. This confornts the results of the cross-section analysis.

*Time dummies.* We include in each regression the year time dummies to allow for aggregate time effects that have the same influence on the dependent variable for all countries. Time dummies are positive and in general statistically significant.

**Time stability of diaspora externalities.** We added time dummies in each regression to capture aggregate time effects (a change in the constant over time) that seem to play an important role in explaining the data. It is worth investigating whether the magnitude of network externalities has increased over time. In Table 7, we allow also for interactions between migration stocks and the time dummies<sup>8</sup>. We started the analysis considering the general case with time dummies and interaction terms. Since the time dummies were not significant (and the interaction terms too), we then ran the regressions without considering separated time dummies<sup>9</sup>. In that case, interaction terms become statistically significant (see both the individual t-statistic and the F-test for the joint significance). It suggests that the diaspora externality increased in the nineties.

---

<sup>8</sup>Instead of using the lagged value of the interaction terms, we separately use as instruments the time dummies and the lagged migration stocks (in logs).

<sup>9</sup>The significance of the interaction variables may simply reflect omitted variables and specification errors, so in this case it can only mean that we don't have omitted variables and specification problems and not that the impact on migration on FDI growth has been stable over time



**Tab 7. Panel - Specification with time interaction terms**

Dependent variable = Growth rate of the FDI-funded capital stock per worker

	Model 1	Model 2	Model 3	Model 4
Initial FDI-funded capital in logs	-0,025 (3.37)***	-0,026 (3.94)***	-0,027 (4.03)***	-0,027 (4.10)***
High income dummy	-0,406 (1.56)	-0,343 (1.33)	-0,174 (0.61)	-0,184 (0.71)
Labor force growth rate	0,040 (0.17)	-0,007 (0.03)	0,004 (0.02)	-0,042 (0.18)
Total lagged migration stock in logs	0,026 (1.94)*	0,027 (2.13)**	-	-
Lagged Share of skilled migrants	0,221 (1.86)*	0,248 (2.25)**	-	-
Total lagged migration stock in logs x d90	-0,003 (0.24)	0,002 (1.72)*	-	-
Total lagged migration stock in logs x d95	0,015 (1.5)	0,006 (2.30)**	-	-
Total lagged migration stock in logs x d00	0,013 (1.92)*	0,005 (3.32)***	-	-
logmighig_1	-	-	0,021 (1.65)	0,023 (1.85)*
logmighig_190	-	-	-0,001 (0.1)	0,004 (2.47)**
logmighig_195	-	-	0,017 (1.38)	0,008 (2.39)**
logmighig_100	-	-	0,012 (1.28)	0,008 (4.38)***
Democracy score (lagged)	-0,131 (1.97)*	-0,151 (2.00)**	-0,151 (2.15)**	-0,182 (2.25)**
DEMHI	0,487 (2.37)**	0,476 (2.28)**	0,368 (1.69)*	0,409 (1.95)*
logtrade_1	0,058 (2.29)**	0,059 (2.39)**	0,068 (2.37)**	0,071 (2.77)***
logtradeHI	0,026 (0.61)	0,01 (0.26)	-0,015 (0.34)	-0,022 (0.55)
d90	0,059 (0.36)	-	0,047 (0.32)	-
d95	-0,110 (0.97)	-	-0,096 (0.83)	-
d00	-0,098 (1.16)	-	-0,048 (0.48)	-
Constant	-0,296 (1.68)*	-0,316 (1.74)*	-0,15 (1.04)	-0,162 (1.16)
Observations	332	332	332	332
Number of countries	83	83	83	83
Hansen test- Prob > chi2	0,379	0,28	0,193	0,075
Arel-Bond test for AR(1) in first differences- Pr > z	0,081	0,079	0,08	0,078
Arel-Bond test for AR(2) in first differences- Pr > z	0,372	0,381	0,375	0,394
Interaction terms (P-value)		0,0103		0,005

Note: System GMM, Robust two-step; t- statistic in parenthesis  
\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

## 7 Conclusion

In this paper, we estimate a dynamic empirical model of FDI-funded capital accumulation. In a cross-section model focusing on the period 1990-2000, our sample of 114 countries reveals the existence of strong network effects, mainly associated to the skilled diaspora. These network effects are stronger in democratic countries as well as in countries exhibiting intermediate corruption index. Very corrupted regimes face strong difficulties to attract foreign investments. In a panel extension with 83 countries and 4 periods of 5 years, we confirm the existence of business network externalities. The elasticity of the capital growth rate to the stock of skilled emigrants is between 2 and 3 percent.

Our empirical analysis confirms that diaspora externalities constitute an important channel through which the brain drain positively affect sending countries. Even when the brain drain depresses the average level of schooling, it is likely to increase FDI inflows. The size of the diaspora matters. Business externalities are likely to be stronger in large countries. The positive effect of their diaspora on FDI reinforces the potentially beneficial effect that migration prospects induce on human capital formation. On the contrary, small countries are less likely to benefit from skilled migration both in terms of human capital (as demonstrated in the new brain drain literature) and physical capital. Finally, we show that diaspora externalities are stronger in countries where corruption is not too high and not too low.

## 8 References

- Anderson, T.W., and C. Hsiao (1982), Formulation and Estimation of Dynamic Model Using Panel Data, *Journal of Econometrics* 18, 67-82.
- Arellano, M., and S. R. Bond (1991), Some Specification Tests for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies* 58, 277-298.
- Arellano, M. and O. Bover (1995), Another Look at the Instrumental Variable

Estimation of error-Components Models, *Journal of Econometrics* 68, 29-51.

- Aroca, P. and W.F. Maloney (2004), Migration, Trade and FDI in Mexico, World Bank WP.
- Arora, A. and A. Gambardella (2004), The Globalization of the Software Industry: Perspectives and Opportunities for Developed and Developing Countries; in A.B. Jaffe, J. Lerner and S. Stern, eds.: *Innovation Policy and the Economy*, Vol. 5, Cambridge, Ma.: The MIT Press.
- Barba Navaretti, G. and J.A. Venables (2004), *Multinational firms in the world economy*, Princeton University Press, Princeton.
- Barghava, A., Jamison D.T., Lau J.L. and C.J.L. Murray, Modeling the effects of health on economic growth, *Journal of Health Economics*, 20 (2001), 423-440.
- Barghava, A. and J.D. Sargan (1983), Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods, *Econometrica*, Vol. 51, No. 6 (Nov., 1983), 1635-1660.
- Barro, R.J., G. Mankiw and X. Sala-I-Martin (1995), Capital Mobility in Neoclassical Models of Growth, *American Economic Review* 85(1), 103-115.
- Beine, M., Docquier, F. and H. Rapoport (2001), Brain Drain and Economic Growth: Theory and Evidence, *Journal of Development Economics*, 64, 1: 275-289.
- Beine, M., Docquier, F. and H. Rapoport (2003), Brain Drain and LDCs' Growth: Winners and Losers, Discussion Paper No 819, IZA Bonn.
- Bhagwati, J. N. and K. Hamada (1974), 'The brain drain, international integration of markets for professionals and unemployment', *Journal of Development Economics*, 1, 1: 19-42.
- Blundell, R. and S. Bond (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics* 87, 115-43.

- Bond S., Hoeffler A. and J. Temple (2001), GMM Estimation of Empirical Growth Models, CEPR Discussion Paper No. 3048, London: CEPR.
- Buch, C.M., Kleinert, J., Lipponer, A. and F. Toubal (2005), Determinants and effects of foreign direct investment: evidence from German firm-level data, *Economic Policy*, January 2005: 51-110.
- Buch, C.M., Kleinert, J. and F. Toubal (2003), Where Enterprises Lead, People Follow? Links between Migration and German FDI, Kiel WP 1190.
- Caselli F., Esquivel G. and F. Lefort (1996), Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics, *Journal of Economic Growth*, 1, 363-389.
- Cohen, D. and M. Soto (2001), Growth and human capital: good data, good results, CEPR Discussion Paper, n. 3025.
- Defoort, C. (2006), Tendances de long terme en migrations internationales. Analyse à partir des 6 principaux pays receveurs, Manuscript, Catholic University of Louvain.
- Docquier, F. and A. Marfouk (2006), 'International migration by education attainment, 1990 2000', In: Caglar Ozden and Maurice Schiff (eds). *International Migration, Brain Drain and Remittances*, McMillan and Palgrave: N.Y.
- Gao, T.(2003), Ethnic Chinese Networks and International Investment: Evidence from Inward FDI in China, *Journal of Asian Economics*, 14: 611-629.
- Gould, D. (1994), Immigrants Links to the Home Countries: Empirical Implication for U.S. bilateral Trade Flows, *The Review of Economics and Statistics* 76(2): 302-316.
- Islam, N. (1995), Growth Empirics: A Panel data Approach, *Quarterly Journal of Economics*, CX, 1127-70.
- Islam, N. (2003), What have We Learnt from the Convergence Debate? , *Journal of Economic Surveys*, 17 (3), 309-362.

- Kugler M. and H. Rapoport (2005), Skilled Emigration, Business Networks and Foreign Direct Investment, Cesifo Working Paper No. 1455.
- López, R. and M. Schiff (1998), Migration and the Skill composition of the Labor Force: The Impact of Trade Liberalization in LDCs, *Canadian Journal of Economics*, 31(2), 318-336.
- Markusen, J.R. and J.A. Venables. (2000), The Theory of Endowment, Intra-industry and Multinational Trade, *Journal of International Economics* 52: 209-234.
- Mountford, A. (1997), 'Can a brain drain be good for growth in the source economy?', *Journal of Development Economics*, 53, 2: 287-303.
- Rauch, J. (2003), *Diasporas and development: Theory, Evidence and Programmatic Implications*, August 2003.
- Rauch, J. and V. Trindade (2002), Ethnic Chinese Networks In International Trade, *The Review of Economics and Statistics*, MIT Press, vol. 84(1):116-130.
- Rauch, J. and A. Casella (1998), Anonymous Market and Group ties in International Trade, *Journal of International Economics*, 58(1):19-47.
- Razin A., Rubinstein A. and E. Sadka (2005a), Fixed Costs and Bilateral FDI Flows: Conflicting Effects of Country Specific Productivity Shocks, Mimeo, Tel Aviv University.
- Razin A., Y Rubinstein and E. Sadka (2005b), Corporate Taxation and Bilateral FDI with Threshold Barriers, Mimeo, Tel Aviv University.
- Saxenian, A. (1999), *Silicon Valley's New Immigrant Entrepreneurs*, Public Policy Institute of California, San Francisco, CA.
- Saxenian, A. (2001) *Bangalore, the Silicon Valley of India?*, CREDPR.
- Stark, O., Helmenstein C. and A. Prskawetz (1997), 'A brain gain with a brain drain', *Economics Letters*, 55: 227-34.

- Stark, O., Helmenstein C. and A. Prskawetz (1998), 'Human capital depletion, human capital formation, and migration: a blessing or a 'curse'?', *Economics Letters*, 60, 3: 363-7.
- Stark, O. and Y. Wang (2002), 'Inducing human capital formation: migration as a substitute for subsidies', *Journal Public Economics*, 86(1): 29-46.
- Tong, S. (2003), *Ethnic Networks in FDI and the Impact of Institutional Development*, Working Paper, University of Hong Kong.
- Wagner D., Head K. and J. Ries (2002), *Immigration and the Trade of Provinces*, *Scottish journal of Political Economy*, Vol. 49, no.5, pp.507-525
- World Bank (2006), *Global economic prospects 2006*, World Bank - Washington.