Diasporas

AFD Workshop, 30 June 2008

Michel BEINE[1], Frédéric Docquier[2] and Caglar Ozden[3]

[1] University of Luxembourg and University of Brussels (Belgium) and CES-Ifo;

michel.beine@uni.lu

[2] FNRS and IRES, Université Catholique de Louvain

[3] World Bank, DECRG

Quotation

"On the day I left Nigeria, I felt sad because I was leaving my family behind. I believed I would return eight years later, probably marry an Igbo girl, and then spend the rest of my life in Nigeria. But 25 years ago, I fell in love with an American girl, married her three years later, and became eligible to sponsor a Green Card visa for my 35 closest relatives, including my parents and all my siblings, nieces and nephews. The story of how I brought 35 people to the United States exemplifies how 10 million skilled people have emigrated out of Africa during the past 30 years. We came to the United States on student visas and then changed our status to become permanent residents and then naturalized citizens. Our new citizenship status helped us sponsor relatives, and also inspired our friends to immigrate here." (Philip Emeagwali)

Main goal of the paper

- This paper provides the first comprehensive analysis of diasporas
- *Diaspora*: stock of migrants who gather in number in a given destination country.
- This paper analyses the driving forces of size , skill composition and relative concentration between skills
- We show that existing diasporas are strong factors influencing future migrations flows and their composition in terms of skills.

Related literature on factors of migration

- Self selection mechanisms (wage differentials, probability of finding a job, migration costs, ...) : Borjas (1987) $\rightarrow$ Possible negative selection of migrants.
- Out-selection factors (immigration policies : Hanson, 2007)
- Income Maximization model (Grogger and Hanson, 2007) : importance of skill premium and base wage differentials
- Endogeneity of migration wrt to education (Chiquiar and Hanson, 2005) : migration costs decrease with education $\rightarrow$ possible positive, negative of intermediate selection of migrants.

Major point

- Migration costs are endogeneous and depend on the size and the structure of the diaspora
- Early evidence of Carrington, Detragiache and Vishwanath, 1996) from US internal black migration migrants'network tends to decrease costs for new migrants.
- Generalize the statistical evidence to all countries and study the dynamics of disaporas.
- Able to address important questions regarding diasporas.

Questions of this paper

- What are the determinants of migration flows ?
- What are the determinants of the educational quality of a diaspora ?
- What are the determinants of the concentration levels of diasporas by educational level.
- While there are a large set of determinants, we argue and find in this paper that diaspora effects (network effects, family reunification) are important determinants $\longrightarrow$ important implications for policies such as selective immigration policies .

Examples from country pairs

- The migration flows, the educational quality and the relative concentration (skilled vs unskilled) are obviously determined by a set of important (bilateral and unilateral) factors (theoretical expected signs from theory)
- Distance (-,+,+) : 336 Turks in NZ , 41 prct skilled, 34 prct unskilled; 5470 Turks in Norway, 9 pct skilled, 44 pct unskilled
- Colonial links (+,-,-) : Pakistan with US (64 pct skilled; 19 pct unskilled; Pakistan-UK : 18 pct skilled-60 pct unskilled; Pakistan-Germany : 43 pct skilled-43 pct unskilled)
- Selective immigration policies (-,+,+) : Pakistan-Australia (74 pct skilled; 13 pct unskilled)

Examples from country pairs

- Wage differential (+,+,+) : see recent works of Rozensweig (2008) and Grogger and Hanson (2008)
- Diaspora effects (+,-,-) : Turkey-Spain (33 pct skilled; 29 pct unskilled; Turkey-Germany : 6 pct skilled-86 pct unskilled ; Turkey-Luxembourg : 44 pct skilled-26 pct unskilled)
- $\longrightarrow$ Diasporas matter ; but what is the share of the total impact associated to diasporas effect ? $\longrightarrow$ quantification of the impact of diaspora effects accounting for the other factors.

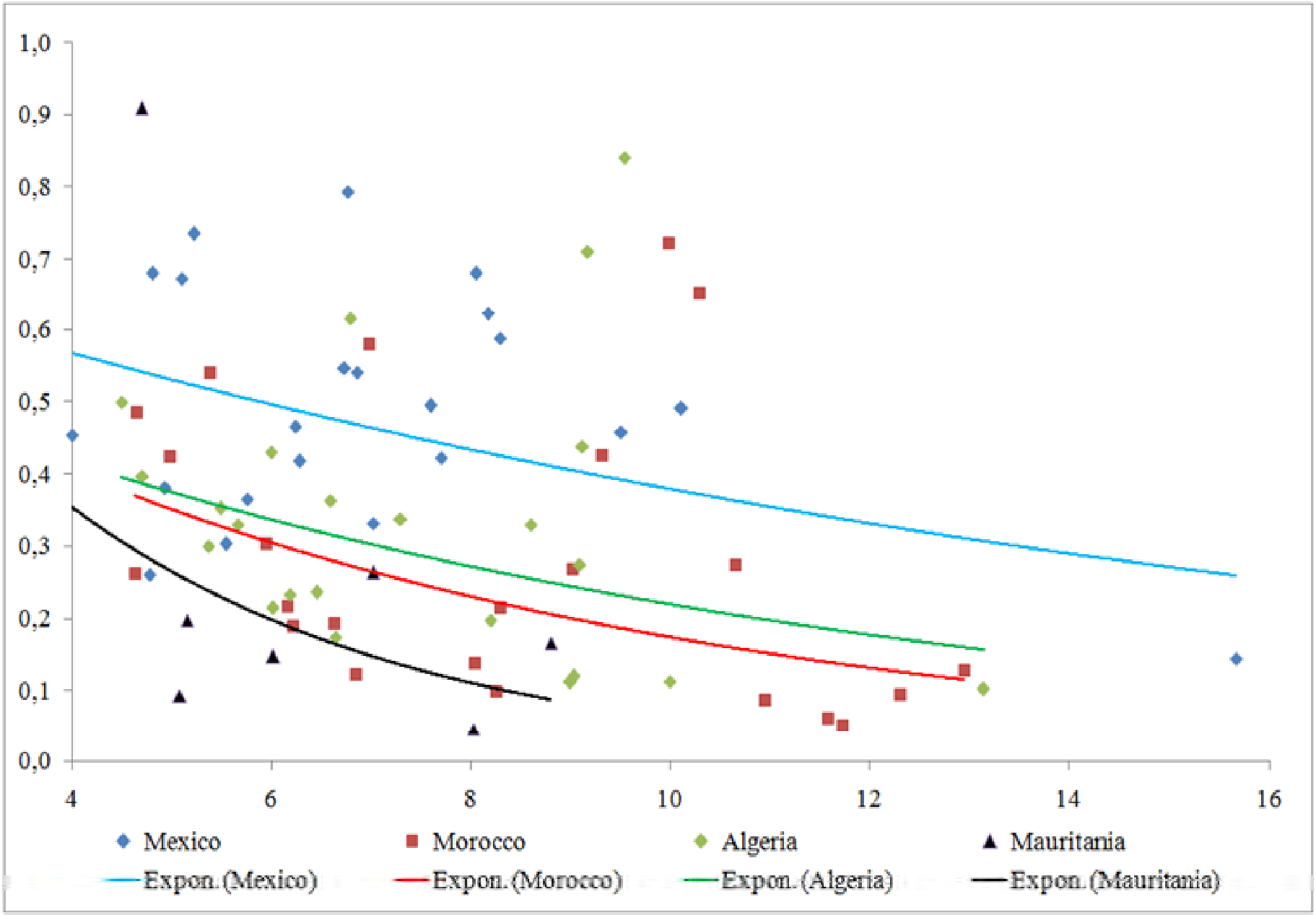
Stylized facts from aggregate data

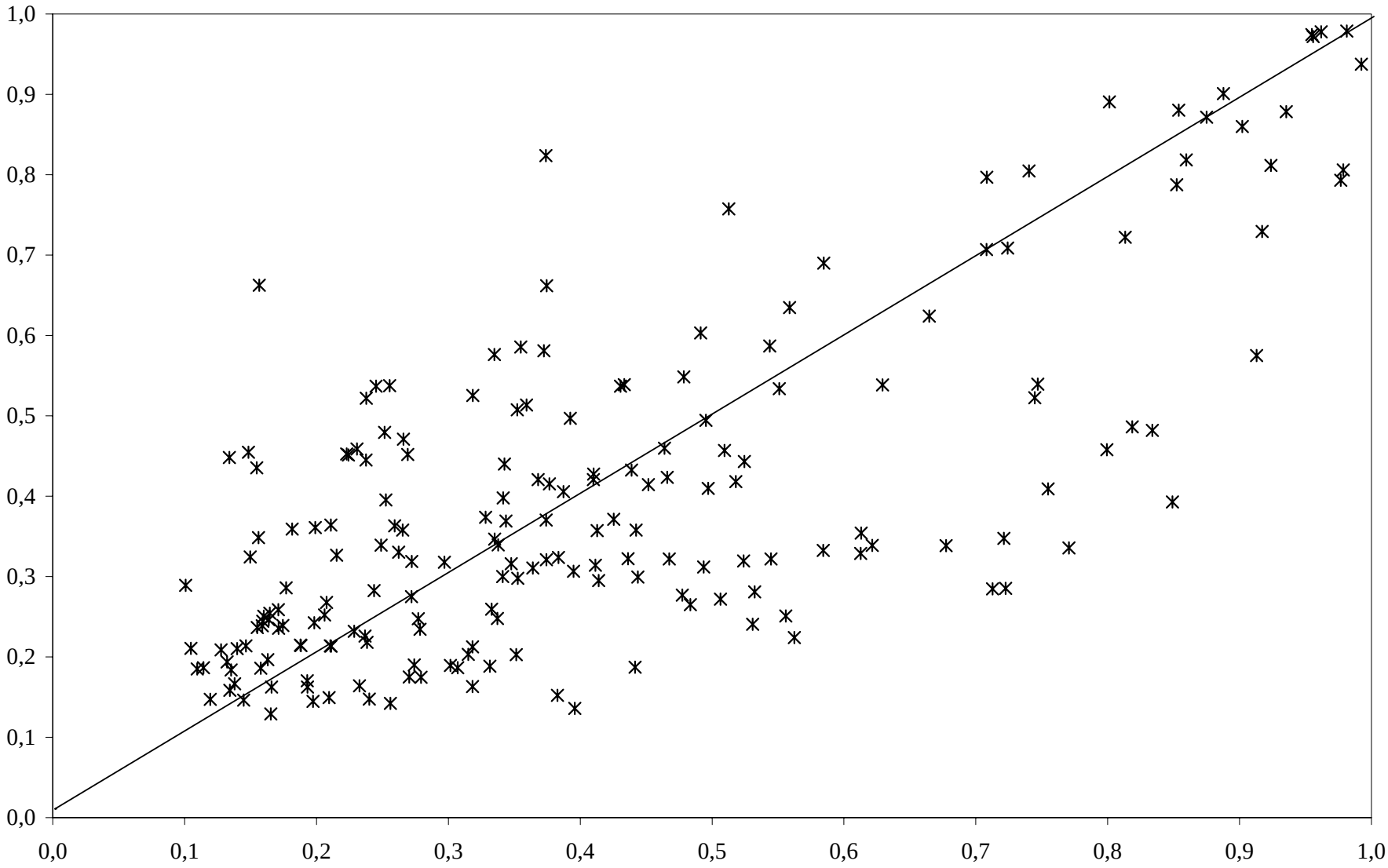
- No clear relationship between diaspora size and its educational structure (either after dropping zeros or not). Log Skill Ratio-Log Size.
- Nevertheless, at the bilateral level , there seems to be a negative relationship.Percentage of Skilled-Log Size
- No clear pattern for relative concentration between skilled and unskilled :
  Concentration of high vs low-skilled diasporas

The data

- Most important data : diaspora size and composition in terms of education (3 levels) at the bilateral level for almost all sending countries and 30 major receiving countries; 2 years (1990 and 2000)
- Distribution of size: highly skewed : about 33 percent of zeros; big diasporas : more than 6 millions legal Mexican migrants in the US.
- Big variation of the skill composition of diasporas
- Big variation of the concentration of skilled and unskilled diasporas.


12
10
8
6
4
2
0
-2
-4
-2
0
2
4
6
8
10
12
14
16


1,0
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0,0
4
6
8
10
12
14
16
Mexico
Morocco
Algeria
Mauritania
Expon. (Mexico)
Expon. (Morocco)
Expon. (Algeria)
Expon. (Mauritania)


1,0
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0,0
0,0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1,0

Econometric analysis

- We study 3 dimensions
- Dynamics of size : population age 25 born in country k and living in country i
- Education level : log ratio of proportion of high-skilled to low-skilled
- (Relative) Concentration of diaspora by education level : Herfindahl index applied to bilateral distribution diasporas.

Econometric methodology

- Gravity type models : include push factors and migration costs (distance, colonial links, language, ...)
- Two types of dependent variable : levels and change between 1990 and 2000 $\rightarrow$ impact of size of diasporas on flows
- For regressions on the levels : long run values of the variables
- Include origin country dummies with and without destination country dummies (inclusion of destination specific observable variables): factors such as immigration policy

Econometric Issues

- Zero values for dependent variables $\rightarrow$ issue of selection bias (dropping observations with zero values creates a sub-sample different from the population) $\rightarrow$ Heckman 2-step estimation
- Instrument in Heckman procedure : diplomatic representation in level regressions
- Intuition : if no diplomatic representation in source countries, migration cost prohibitive for every potential migrant.
- Log-linearization induces biases in estimation $\rightarrow$ Poisson regression models for size as a robustness check
- In general : findings very robust to estimation techniques and specifications

Results

- Diasporas are strong attractors of future migrants
  impact on size
- Diasporas tend to be more profitable to unskilled workers : unskilled workers more sensitive to migration costs
- Diasporas tend to affect the skill composition of future diasporas : impact on selection
- Diasporas tend to favour the concentration of unskilled vs skilled migrants : impact on concentration

Table 2: Determinants of stock and change : total and low skilled

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Tot Stock	Tot stock	Tot stock	Tot stock	Tot change	Tot change	Total Low	Low
Col links	1.989	1.117	2.138	0.910	0.127	-0.051	2.336	1.188
	(14.79)***	(8.01)***	(17.86)***	(5.77)***	(1.10)	(0.29)	(17.34)***	(7.79)***
linguage	1.100	1.984	1.259	1.467	0.496	1.056	1.049	1.362
	(13.13)***	(20.77)***	(14.95)***	(10.98)***	(6.48)***	(8.34)***	(10.94)***	(10.60)***
Log(dist)	-1.016	-0.545	-1.035	-0.580	-0.448	-0.095	-1.133	-0.624
	(20.49)***	(10.34)***	(24.66)***	(11.83)***	(10.69)***	(1.63)	(23.70)***	(12.76)***
Shengen	0.020	0.098	-0.157	-0.022	0.277	0.599	-0.464	-0.121
	(0.15)	(0.59)	(1.09)	(0.11)	(2.02)**	(2.56)**	(2.82)***	(0.62)
Immig. pol		0.025		0.012		0.035		0.026
		(5.49)***		(2.29)**		(6.71)***		(4.97)***
Pop. at dest		1.007		0.982		0.109		0.942
		(40.41)***		(36.42)***		(2.30)**		(32.40)***
Social exp		0.790		0.232		0.175		0.706
		(7.28)***		(1.79)*		(1.28)		(4.81)***
Wage at dest.		0.181		0.112		0.040		0.116
		(25.96)***		(9.19)***		(4.51)***		(9.49)***
Lagged diasp					0.699	0.831		
					(43.91)***	(23.44)***		
Constant	9.986	-15.903	8.847	-10.239	2.365	-6.119	8.422	-10.995
	(12.21)***	(17.66)***	(13.36)***	(7.88)***	(4.02)***	(5.07)***	(9.51)***	(7.09)***
Dest dum	yes	no	yes	no	yes	no	yes	no
Orig dum	yes	yes	yes	yes	yes	yes	yes	Yes
Method	OLS	OLS	Heckman	Heckman	Heckman	Heckman	Heckman	Heckman
Mills ratio			1.075	-1.692	-0.848	-0.823	1.109	-0.898
			(9.77)***	(7.50)***	(5.57)***	(3.01)***	(8.75)***	(4.12)***
Observations	3961	3405	5760	4992	5760	4992	5760	4992
R-squared	0.82	0.72						

Absolute values of *Robust t* statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Determinants of educational structure

	(1)	(2)	(3)	(4)	(5)	(6)
	Skill ratio	Skill ratio	Skill ratio	Skill ratio	Change SR	Change SR
Lagged diasp	-0.171	-0.088	-0.194	-0.132	-0.143	-0.108
	(16.19)***	(8.47)***	(20.62)***	(11.83)***	(17.62)***	(11.47)***
Col. links	-0.042	-0.439	-0.022	-0.410	0.101	0.096
	(0.62)	(6.08)***	(0.32)	(5.21)***	(1.67)*	(1.46)
linguage	0.466	0.703	0.460	0.721	0.176	0.257
	(9.38)***	(11.03)***	(9.37)***	(11.68)***	(4.17)***	(4.95)***
Log(dist)	0.096	0.273	0.090	0.263	0.086	0.116
	(3.35)***	(10.17)***	(3.40)***	(9.96)***	(3.78)***	(5.25)***
Shengen	0.502	0.305	0.519	0.303	0.390	0.117
	(5.65)***	(3.14)***	(6.26)***	(2.97)***	(5.48)***	(1.37)
Immig pol		-0.014		-0.015		0.001
		(4.98)***		(5.52)***		(0.30)
Soc exp		-1.206		-1.253		-0.756
		(16.11)***		(20.12)***		(14.42)***
Pop. at dest		0.061		0.082		0.056
		(3.45)***		(4.58)***		(3.75)***
Wage at dest.		0.044		0.045		0.035
		(9.86)***		(10.47)***		(9.78)***
Constant	-1.109	0.002	-0.734	0.257	-1.250	-0.563
	(1.16)	(0.00)	(1.32)	(0.34)	(2.54)**	(0.87)
Dest dum	Yes	No	Yes	No	Yes	No
Orig dum	Yes	Yes	Yes	Yes	Yes	Yes
Method	OLS	OLS	Heckman	Heckman	Heckman	Heckman
Mills			-0.380	-0.446	-0.10	-0.99
			(6.86)***	(7.37)***	(0.22)	(1.88)*
Obs	3604	3084	5760	4992	5760	4992
R-squared	0.60	0.45				

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

Table 4. **Determinants of relative concentration.**

	(1)	(2)	(3)	(4)	(5)	(6)
	Rel conc	Rel conc	Rel conc	Rel conc	Change RC	Change RC
Lagged diasp	-0.502	-0.294	-0.514	-0.347	-0.008	-0.008
	(5.87)***	(3.54)***	(9.67)***	(5.73)***	(16.05)***	(15.45)***
Col. links	-4.635	-7.085	-4.619	-7.008	-0.040	-0.043
	(4.68)***	(6.41)***	(10.69)***	(14.75)***	(9.93)***	(10.45)***
Linguage	0.338	0.373	0.321	0.369	-0.004	-0.005
	(0.84)	(0.78)	(1.09)	(1.02)	(1.58)	(1.75)*
Log(dist)	0.266	0.628	0.269	0.615	0.006	0.006
	(1.24)	(3.73)***	(1.69)*	(3.91)***	(3.78)***	(4.26)***
Shengen	-0.193	-0.076	-0.180	-0.068	0.002	0.001
	(0.50)	(0.16)	(0.36)	(0.11)	(0.49)	(0.26)
Pop. at dest		0.956		0.988		0.003
		(7.13)***		(9.33)***		(3.50)***
Immig pol		-0.014		-0.013		0.000
		(1.31)		(0.84)		(1.51)
Soc exp		-1.509		-1.573		0.002
		(4.38)***		(4.44)***		(0.52)
Wage at dest		0.217		0.217		0.001
		(7.69)***		(8.57)***		(4.68)***
Constant	5.607	-18.397	-3.240	-10.824	-0.037	-0.111
	(0.29)	(4.70)***	(1.19)	(2.77)***	(1.60)	(3.33)***
Dest dum	Yes	No	Yes	No	Yes	No
Orig dum	Yes	Yes	Yes	Yes	Yes	Yes
Method	OLS	OLS	Heckman	Heckman	Heckman	Heckman
Mills			-0.405	-0.680	-0.873	-1.684
			(1.07)	(1.94)**	(2.44)**	(6.12)***
Observations	3920	3367	5730	4966	5730	4966
R-squared	0.29	0.17				

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

Extensions

- Is the impact non linear ? rolling regressions ; rolling regressions to eliminate the impact of small (or big) diasporas on educational composition.
- We get in general a decreasing marginal impact of migrants'network : consistent with double log specification impact on selection
- But no evidence of threshold : small diasporas matter !!!!

Extensions

- Is the impact non linear ? rolling regressions ; rolling regressions to eliminate the impact of small (or big) diasporas on educational composition.
- We get in general a decreasing marginal impact of migrants'network : consistent with double log specification impact on selection
- But no evidence of threshold : small diasporas matter !!!!

Figure 4.a. Estimating ρ with rolling regressions
Dependent = change in diaspora size

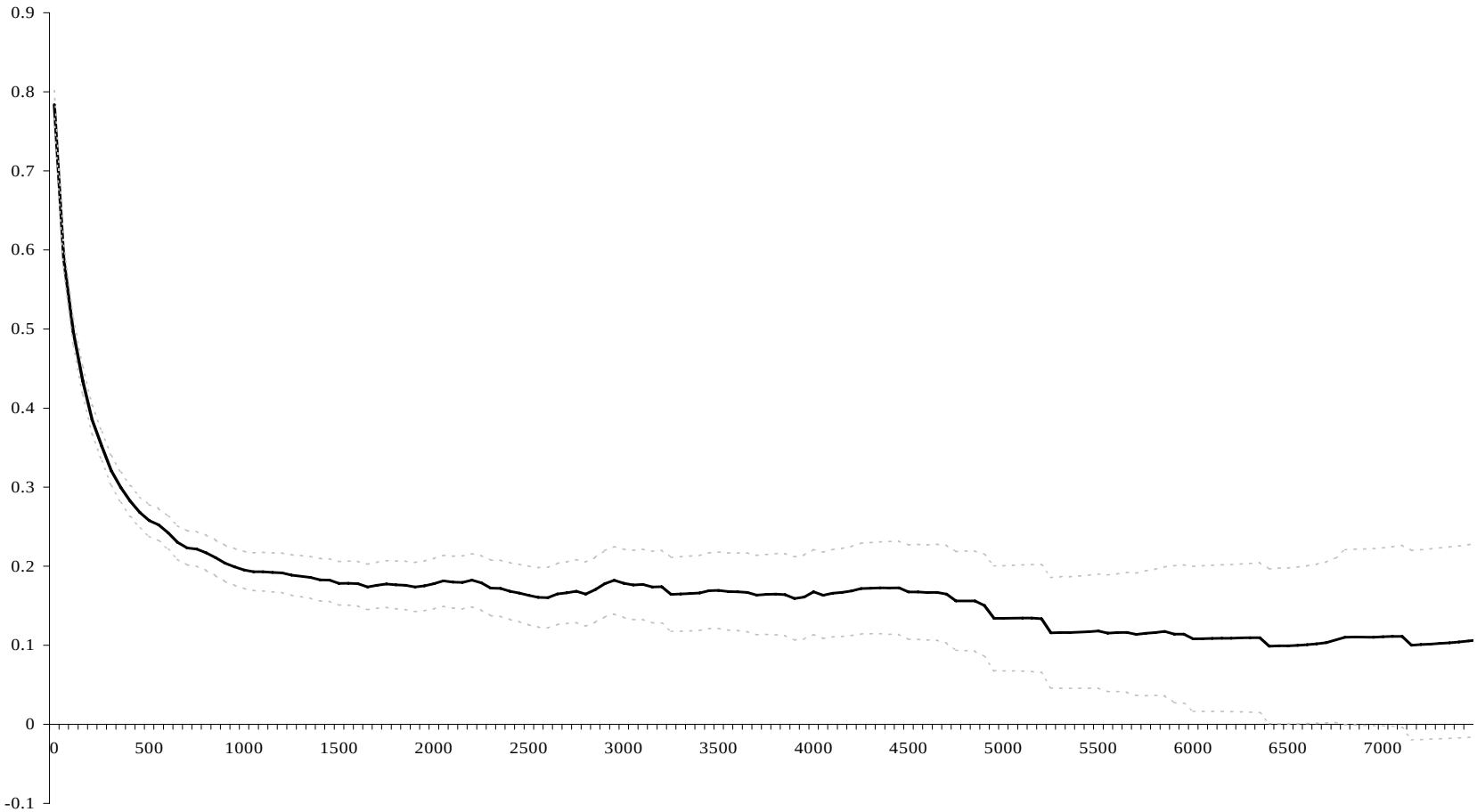
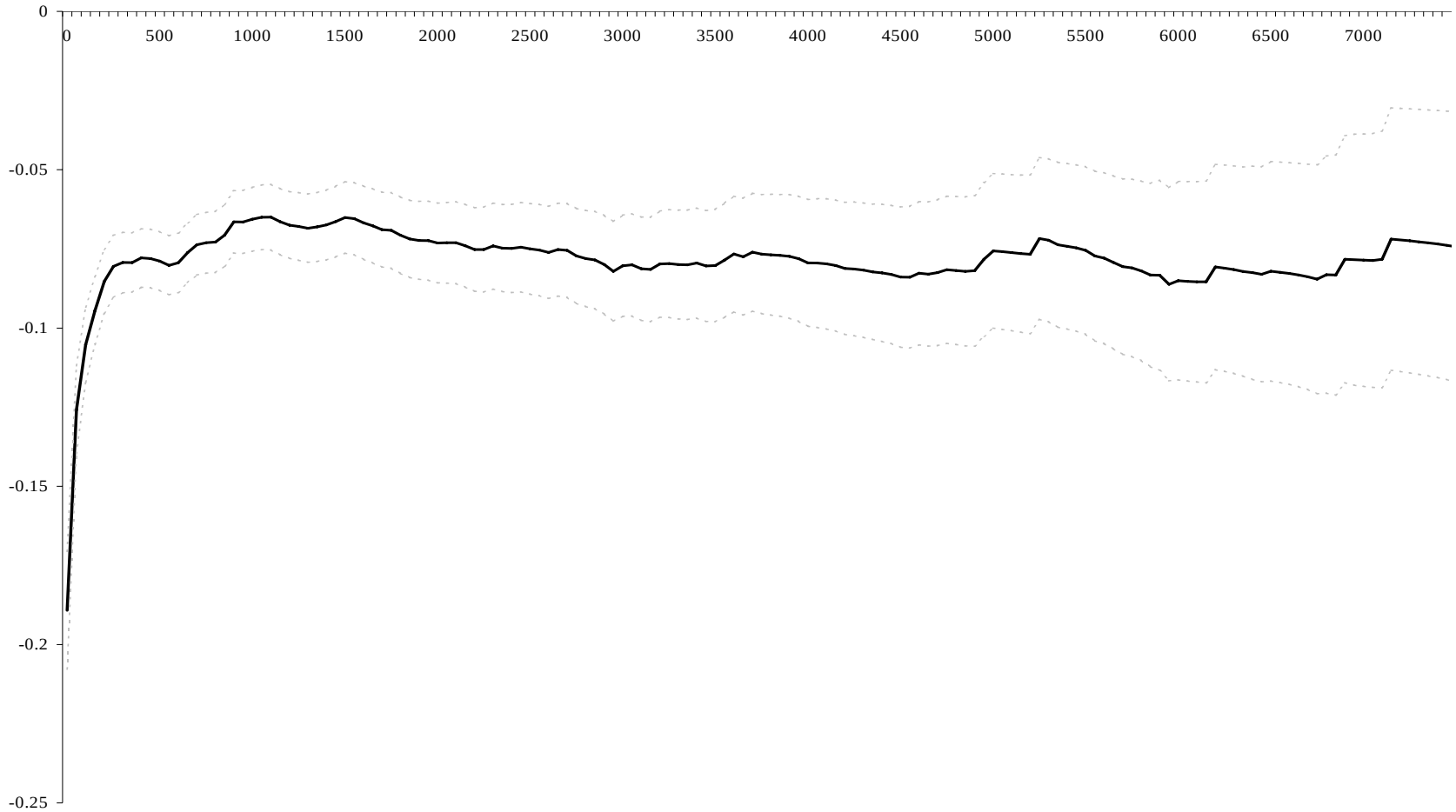


Figure 4.b. Estimating ρ with rolling regressions
Dependent = log skill-ratio



Policy implications

- Diaspora effects explain a significant part of the migration flows, their skill composition and the relative concentration
- Furthermore, a large part is also explained by factors totally exogenous for governments (distance, colonial links, language, wage differentials (Grogger and Hanson, 2008 : 55 pct by income maximisation)
- $\rightarrow$ Small residual part left to immigration policies $\rightarrow$ beyond a change in the family reunification programs (included in our diaspora effects), little room for policy actions.

Policy implications (cont'ed)

- Suppose that France would like to implement a point system (one way of selective immigration policy)
- If still family reunification program, likely that France will still attract a lot of unskilled migrants coming from African countries with large diasporas (Algeria, Morrocco, Mali, ...) $\rightarrow$ share of unskilled likely to increase
- If program of ethnic diversification $\rightarrow$ likely to fail in presence of large diasporas (see concentration results)

Conclusions

- First comprehensive statistical analysis of diasporas
- Diasporas exert very strong effect on size but also on skill composition of migrants
- Implications : very difficult to try to have heterogeneity in diasporas in the presence of very large migrants' network (France, US)
- Difficult to conduct selection immigration policies in the presence of large diasporas : strong attractors of unskilled legal (and of course illegal) migrants.