HOW DOES INTERNAL MIGRATION BY EDUCATIONAL ATTAINMENT REACT TO REGIONAL UNBALANCES? THE CASE OF ITALY

Romano Piras, University of Cagliari - Department of Economics
Viale S. Ignazio, 17 - 09123 Cagliari (Italy)
Tel. +39 070 6753314, Fax. +39 070 660929, e-mail: pirasr@unica.it

Abstract
We conduct an empirical investigation of the determinants of migration flows across Italian regions taking explicitly into account the educational attainment of migrants. We follow the standard macroeconomic migration theory and test what a role has been played by income and unemployment differentials. With the exception of graduate migration, for all educational levels relative per worker GDP is the most important variable in explaining internal migration during the 1995-2005 period.
JEL Codes: J61, I21, R23.
Key words: Italy; labour migration; human capital; income and unemployment differentials.

1. Introduction

The literature on migration has shown a quite strong direct link between educational attainment and migration propensity. One reason of such a phenomenon is that, since returns to human capital are higher the higher the level of education, the opportunity cost of staying in an area (country or region inside a country) increases with the education attainment. Another reason that leads the more educated to migrate is that individuals who have a higher educational degree usually also have higher chances to find a job with respect to those who have only limited schooling. In the context of international mobility, Solimano (2006) claims that more talented people are more mobile than unskilled workers and that they find more favourable immigration policies in the host countries in which they typically try to migrate. As regards, for example, scientists and academics he claims that "... they leave their home countries attracted by higher salaries abroad, by the possibility of increasing their knowledge base and to transmit their own, to interact with peers of international recognition, and to pursue a successful career" (p. 5). A similar reasoning applies also for experts in science and technology, professionals in the health sector, entrepreneurs and managers and the like. We think that many of these factors affect internal migration across regions as well. As regards internal migration across Italian regions such a pattern has been recently found by Svimez (2009).

In this respect, Italy has some peculiarities that make it an interesting case study to be investigated. During the 1950s and the 1960s, millions of individuals moved from the backward Southern and, at that time, North-eastern regions towards the Central and North-western ones. Overall, it emerges from various papers[1] that during the 1960s and the 1970s, housing market conditions and real wages have played the main role in shaping migration across regions, whereas unemployment differentials have been less important. More recently, Basile and Causi (2007) find that from 1991 to 1995, when migration flows were generally decreasing, the effect of economic variables (unemployment and per capita income) on inter-provincial migration flows were negligible or nil; on the contrary, during the 1996-2000 period characterised by an increase of internal mobility, migration flows have reacted more rapidly to unemployment and per capita income. Fachin (2007) concentrates his analysis on male migration during the period from 1970 to 1996. His empirical analysis points out income in the sending regions as the main factor in the migration decision, whereas unemployment and income differentials are definitively less important. Etzo (2009) results broadly confirm those of Fachin (2007). In his analysis during the 1995-2002 time period per

[1] Salvatore (1977), Attanasio and Padoa-Schioppa (1991), Brunello *et Al*. (2001).

capita GDP in both sending and receiving regions has played a strong role in explaining internal migration. On the contrary, unemployment has a stronger impact for the sending regions but a minor role for the receiving regions.

Although important given their findings, none of these works has paid attention at the skill level of migrants and, in the light of what previously said, this is an important question to be investigated. Very recently Ciriaci (2010) has partially tackled this issue. She uses individual level data from a survey on labour market entry condition and focuses on graduate migration at provincial level. The main result, obtained with backward stepwise probit regressions, is that "instead of acting as a mechanism to equalise growth … the phenomenon of human capital flight acts as a disequilibrating mechanism" between the Central-northern and Southern regions (p. 28).

In this paper, differently from all previous studies, we investigate the determinants of migration flows across Italian regions taking explicitly into account all the educational level of migrants: primary school, lower- and upper-secondary school and university level. In order to tackle possible endogeneity problems we apply a GMM estimator.

The paper is organised as follows. In Section 2 we sketch out the empirical model. In Section 3 we discuss the data and the empirical strategy. In section 4 we present and discuss our empirical results. Finally, Section 5 concludes.

2. Model specification

A very general macroeconomic migration function can be written as,

$$m_{ij} = F(Y_i, Y_j, U_i, U_j, Z_i, Z_j)$$

where migration from country j to country i, m_{ij}, is "explained" by per capita incomes, Y_i and Y_j, and unemployment rates U_i and U_j, in both countries and by other economic and non economic push and pull factors $\mathbf{Z}_i$ and $\mathbf{Z}_j$. As to the choice of the migration variable, various definitions are available. In order to take into account the size effect - that is the fact that an increased population in a country leads *per se* to an increase in the level of migration - migration flows usually are specified as migration *rates*, namely as the ratio of migrants to resident population either in the sending or in the receiving country.

Following Hatton (1995), in recent studies, the empirical methodology on migration across countries or regions applies a logarithmic or a semi-logarithmic equation approach. Since we investigate net regional migration rates, we cannot apply the former given that net migration can obviously take negative values. The latter, does not suffer from this drawback and has the desirable characteristic of making the rising of net migration rate not linear with respect to the independent variables, thus implying that migration follows some form of saturation pattern.

The above general macroeconomic migration function must be tailored to the case under investigation. Firstly, as claimed by Smith and Swanson (1998), net migration rates are very useful in many circumstances in that, for instance, they provide a summary measure of one component of population change. Secondly, whenever the concern for the migration impact is on labour markets, the spotlight of many migration studies on net migration is a sound and reasonable choice. Thirdly, in Italy the flow of migrants has almost always been unidirectional from the South to the Centre-North and, as argued by Bentolila and Dolado (1991), it does not make a difference whether net rather than gross migration rates are used. It is then convenient to assume that region's i net migration rate depends on relative per capita income and unemployment rate. In so doing, we introduce a sort of source region all-other-destinations comparison into the analysis, overcoming the need to introduce a bilateral comparison of these variables from each couple of regions and specifying a much more

parsimonious empirical specification. Frequently, because of the direct and indirect transportation costs of migration, an element taken explicitly into account in determining migration is the distance between the sending and the receiving region. In our empirical approach, given the definition of net migration with respect to all other regions, this is precluded.2 Furthermore, because it takes time for information to be acquired, migration is likely to respond with a lag to changed circumstances and a dynamic formulation is required. Finally, if one assumes that in a decade the push and pull factors captured by $\mathbf{Z}_i$ and $\mathbf{Z}_j$ do not change significantly over time, then they can be modelled as constants specific to each region.

For these reasons we study internal migration flows by educational attainment across the 20 Italian regions according to the following dynamic model:

$$(1) \qquad m_{iITAt}^{edu(h)} = \beta_1 m_{iITAt-1}^{edu(h)} + \beta_2 \ln\left(\frac{Y_i}{Y_{ITA}}\right)_t + \beta_3 \ln\left(\frac{U_i}{U_{ITA}}\right)_t + \nu_{it}$$

where $m_{iITAt}^{edu(h)}$ is the net migration rate of region's i population with educational level h with respect to all other regions, that is:

$$(2) \qquad m_{iITAt}^{edu(h)} = \frac{(\text{inflows} - \text{outflows})_{iITAt}^{edu(h)}}{\text{population}_{it}^{edu(h)}} \times 1000$$

and $edu(h)$ corresponds to educational level h, namely primary, lower-secondary, upper-secondary and university level.[3] $\nu_{it} = \mu_i + \varepsilon_{it}$, where μ_i are regional effects and ε_{it} is an additive error term. As can be seen from equation (2), net migration rates by educational level are computed with respect to population with the same educational attainment. In addition, $(Y_i/Y_{ITA})_t$ is region's i per worker[4] GDP relative to national average, $(U_i/U_{ITA})_t$ measures region's i relative unemployment rate. To check for the robustness of our empirical results, we use three different variables: unemployment rate, youth unemployment rate and unemployment rate by educational level h.

Regional per worker GDP differential is a proxy for differing wages and, more generally, wealth expectations for region i vis-à-vis all other regions; thus a positive link with the net migration rate is expected. Ideally, it would have been better to have a measure of wages or incomes for the different educational levels we are studying, but unfortunately such a measure is not available at regional level for the time period we are analysing. Higher relative regional unemployment rates discourage people from moving in and spur residents to move out, consequently we expect a negative relation between relative unemployment and

2 For similar reasons we cannot take into account the past periods stock of individuals who migrated as a proxy for network effects. However, the inclusion of the lagged dependent variable in the regression can be seen as a crude approximation of a migration chain variable.

3 In Italy the schooling system is articulated basically into four levels: primary school (5 years), lower-secondary school (3 years), upper-secondary education (from 3 to 5 years) and university degree (from 3 to 6 years according to the field of study). In recent years, various reforms have changed and are still changing the Italian educational system, particularly the university one. Upper-secondary education, which is under reform as well, is differentiated into six categories: classical, scientific, artistic, primary teacher training, technical schools and vocational education. With the exception of vocational education that lasts up to three years and does not allow entering university, all the other categories do consent it.

4 We have conducted all regressions also with per capita instead of per worker GDP. In general the estimates are slightly better with the latter variable. All results are available from the author upon request.

net migration.

3. Empirical analysis

3.1 Data

All data comes from the national institute of statistics. Interregional migration flows by educational attainment (ISTAT, various years, a) are classified into five groups: *laurea* (university degree), *diploma di scuola media superiore* (upper-secondary school), *licenza media inferiore* (lower-secondary school), *licenza elementare* (primary school) and *nessun titolo* (no schooling). It is worth noticing that in Italy changing residence is not compulsory and it could be the case that individuals move from one region to another without going to municipal register in order to record it; thus, recorded data under-report actual migration flows.

As regards net migration rates by educational level, we compute them with respect to population with the same educational attainment. In order to do that, we use resident population by educational level as classified by ISTAT according to a representative sampling (ISTAT, various years, b). Unfortunately, starting from 1993 ISTAT brings together resident population with primary school and with no schooling and, as a consequence, we are forced to compute regional net migration rates pooling them into a single variable. For the sake of simplicity, however, in the rest of the paper we will refer only at primary school level net migration rate, but it should be clear that it has to be understood as net migration rate of people with no schooling or with primary school educational attainment. Finally, regional per worker GDP comes from (ISTAT, 2009).

3.2 Estimation strategy

As it is well known, in dynamic panel data estimation, both random and fixed effects models are biased. In order to overcome such a bias various alternatives are available, ranging from instrumental variables techniques, to GMM estimators. These estimators are particular useful in our case since it could be the case that regional per worker GDP and unemployment rate are endogenously determined. In fact, on the one hand migration influences regional per capita GDP by changing regional population, on the other it has an effect on the labour force and on employment, thus influencing regional unemployment (Bentivogli and Pagano, 1999). In our empirical analysis, to tackle this issue, we adopt the system GMM estimator put forward by Arellano and Bover (1995) and Blundell and Bond (1998).

All estimation results are reported with robust *p*-values, consistent in the presence of any pattern of heteroskedasticity and autocorrelation within panels. To determine whether instruments are valid, the Hansen test of overidentifying restriction is reported. Finally, Arellano-Bond first and second order serial correlation tests for the idiosyncratic error term are also reported. The GMM estimator is consistent if there is not second order serial correlation in the error term of the first-differenced equation. Thus, failure to reject the null hypothesis of no second order serial correlation gives support to the validity of instruments. Overall the diagnostic tests of the regressions we report below suggest that instrument are valid and that the GMM estimator is consistent.

4. Empirical results and discussion

Table 1 reports the results for total net migration rate. Per worker GDP and unemployment rate are all significant and have the expected signs, the lagged dependent variable and the youth unemployment rate are significant up to 1%. These results confirm what recently found by, among others, Basile and Causi (2007) and Etzo (2009). As regards primary school net migration rate (Table 2) relative per worker GDP is significant, though sometimes only at 10%. The problem here is that the unemployment rate and the

unemployment rate by educational level are statistically significant and have the wrong positive sign. This result could be explained in terms of return migration. In fact, those individuals who moved from the Southern towards the Centre-northern regions during the 1950s and the 1960s were very low educated, with primary education only or no schooling at all. Many of them, after they retired forty years later, come back to their regions of origin, namely to Southern ones where relative unemployment rates are higher than in the rest of the country. At lower-secondary school educational level (Table 3), together with the lagged dependent variable the relative per worker GDP is almost always significant. On the contrary, the unemployment rate variables are never significant. Thus, analogously to primary school net migration these results suggest that unemployment, however defined, is not crucial for the migration decision of lower-secondary school migrants. In Table 4 the picture that emerges for upper-secondary school net migration is quite different since together with the lagged dependent variable and the relative per worker GDP also the three unemployment rate variables are always negative and highly significant. Anyhow, migrants react more promptly to income rather than unemployment differentials. Finally, university level net migration rate (Table 5). Somehow surprisingly, this component of internal migration does not respond to per worker GDP differentials. Nevertheless graduate migrants strongly react to unemployment rate by educational attainment and, to a small extent, to youth unemployment rate. Their behaviour is fairly different from that of the other migrants and our results suggest that job opportunities rather than income differentials are the key elements in their migration decision.

5. Conclusions

In this paper we have analysed interregional migration in Italy taking explicitly into account the educational level of migrants.

We have followed the recent empirical literature on migration and estimated how internal migration rates by educational attainment react to some fundamental economic variables such as relative per worker GDP and various measures of relative unemployment rates. We have computed migration rates with respect to population with the same educational level and applied a GMM dynamic panel data methodology to cope with endogeneity problems. As for unemployment we have used three different definitions: unemployment rate, youth unemployment rate and unemployment rate by educational attainment. Overall our empirical evidence can be summarised as follows. First, as regards total net migration, we have found that per worker GDP and unemployment rate are all significant and have the expected sign. Second, in the basic regressions for primary school level net migration (which also contains individuals with no schooling) relative per worker GDP is significant; surprisingly, though, unemployment rate and unemployment rate by educational level are also significant but have the wrong positive sign. We think that this result could be explained in terms of return migration. Third, at lower-secondary school educational level, relative per worker GDP is almost always significant, whereas the unemployment rate variables are never significant. Four, the picture that has emerged for upper-secondary school net migration is quite different since together with the lagged dependent variable and the relative per worker GDP also the three unemployment rate variables are always negative and highly significant. Five, university level net migration rate does not respond to per worker GDP differentials, rather it reacts to unemployment rate by educational attainment and to youth unemployment rate. Their behaviour is thus different from that of the other migrants and suggests that job opportunities rather than income differentials are the key elements in their migration decision. Six, for all educational levels but for university level, the relative per worker GDP variable is definitively the most important in explaining internal migration across Italian regions during the 1995-2005 period.

This result broadly agrees with some empirical works on Italian regions concerning the previous decades (Attanasio and Padoa-Schioppa, 1991; Brunello *et Al.* 2001) and more recent studies (Etzo, 2009; Basile and Causi, 2007).

Further research should be done in order to investigate whether other factors such as, for example, the industrial structure of Italian regions could help to better explain the recent interregional migration flows across Italy. Finally, in our econometric analysis regions are treated as a-spatial unities assuming implicitly that they are independent regardless of them to be contiguous or not. The latest advances in spatial econometrics literature could potentially enrich the understanding of Italian internal migration.

References.

Arellano M. and Bover O. (1995), "Another look at the instrumental variables estimation of error-components models". *Journal of Econometrics*, 68: 29-51.

Attanasio O. and Padoa-Schioppa F. (1991), "Regional inequalities, migration and mismatch in Italy, 1960-1986". In Padoa-Schioppa F. (ed), *Mismatch and Labour Mobility*. Cambridge University Press, Cambridge.

Basile R. and Causi M. (2007), "Le determinanti dei flussi migratori nelle province italiane: 1991-2001". *Economia & Lavoro*, XLI: 139-159.

Bentivogli C. and Pagano P. (1999), "Regional disparities and labour mobility: the Euro-11 versus the USA". *Labour*, 13: 737-760.

Bentolila S. and Dolado J. (1991), "Mismatch and internal migration in Spain, 1962-1986". In Padoa-Schioppa F. (ed), *Mismatch and Labour Mobility*. Cambridge University Press, Cambridge.

Blundell R. and Bond S. (1998), "Initial conditions and moment restrictions in dynamic panel data models". *Journal of Econometrics*, 87: 115-143.

Brunello G., Lupi C. and Ordine P. (2001), "Widening differences in Italian regional unemployment". *Labour Economics*, 8: 103-129.

Ciriaci D. (2010), "University quality, interegional brain drain and spatial inequality. The case of Italy". *LLEE Working Document* n. 87.

Etzo I. (2009), "Determinants of interegional migration in Italy: A panel data analysis". MPRA 8637.

Fachin S. (2007), "Long-run trends in internal migration in Italy: a study in panel cointegration with dependent units". *Journal of Applied Econometrics*, 22: 401-428.

Hatton T. (1995), "A model of UK migration 1870-1913". *Review of Economics and Statistics*, 77: 407-415.

ISTAT (2009), *Conti economici regionali. Anni 1995-2008*. ISTAT, Rome.

ISTAT (various years, a), *Movimento migratorio della popolazione residente. Iscrizioni e cancellazioni anagrafiche*. ISTAT, Rome.

ISTAT (various years, b), *Rilevazione trimestrale delle forze di lavoro*. ISTAT, Rome.

Salvatore D. (1977) "An Econometric Analysis of Internal Migration in Italy". *Journal of Regional Science*, 17: 395-408.

Smith S. and Swanson D. (1998), "In defence of the net migrant". *Journal of Economic and Social Measurement*, 24: 249-264.

Solimano, A. (2006), "The international mobility of talent and its impact on global development". *United Nations University – World Institute for Development Economics Research*, Working Paper n. 8.

SVIMEZ (2009), *Rapporto 2008 sull'economia del Mezzogiorno*. Il Mulino, Bologna.

Table 1. Total net migration rate.

	(1)	(2)
Lagged dep. Variable	0.492***	0.537***
	[0.001]	[0.000]
Per worker GDP	5.175**	4.702*
	[0.025]	[0.074]
Un. Rate	-1.010*	
	[0.074]	
Youth Un. Rate		-1.102***
		[0.005]
A-B AR(1)	-2.55	-2.50
	[0.011]	[0.013]
A-B AR(2)	-0.25	0.55
	[0.799]	[0.584]
Hansen	14.05	13.17
	[0.230]	[0.283]

Sample period 1995-2005. Total observations: 200. Robust *p*-values in brackets. Constant term not reported. Hansen tests is a test on the validity of over-identifying restrictions, the null hypothesis is that instruments are valid. A-B AR(1) and AR(2) tests are Arellano-Bond first and second order serial correlation tests, respectively. ***, ** and * denote significance at 1, 5 and 10 per cent level, respectively.

Table 2. Primary school net migration rate.

	(1)	(2)	(3)
Lagged dep. Variable	0.025	0.125	0.429**
	[0.922]	[0.673]	[0.029]
Per worker GDP	7.974*	5.539*	6.729**
	[0.058]	[0.068]	[0.036]
Un. Rate	1.064*		
	[0.051]		
Youth Un. Rate		0.355	
		[0.456]	
Un. Rate by ed. Level			0.797*
			[0.076]
A-B AR(1)	-1.87	-2.04	-2.41
	[0.062]	[0.042]	[0.016]
A-B AR(2)	1.70	2.08	2.17
	[0.089]	[0.038]	[0.030]
Hansen	11.97	13.66	10.75
	[0.366]	[0.253]	[0.378]

See Table 1 for more details.

Table 3. Lower secondary school net migration rate.

	(1)	(2)	(3)
Lagged dep. Variable	0.839***	0.907***	0.795***
	[0.000]	[0.000]	[0.000]
Per worker GDP	3.771***	4.791**	5.155**
	[0.009]	[0.029]	[0.041]
Un. Rate	0.202		
	[0.728]		
Youth Un. Rate		0.299	
		[0.479]	
Un. Rate by ed. Level			-0.189
			[0.822]
A-B AR(1)	-1.65	-1.54	-1.87
	[0.099]	[0.124]	[0.061]
A-B AR(2)	0.69	0.62	0.80
	[0.492]	[0.537]	[0.425]
Hansen	15.74	15.61	17.09
	[0.151]	[0.210]	[0.146]

See Table 1 for more details.

Table 4. Upper secondary school net migration rate.

	(1)	(2)	(3)
Lagged dep. Variable	0.519***	0.514***	0.488***
	[0.000]	[0.000]	[0.001]
Per worker GDP	11.013***	9.627***	7.414**
	[0.001]	[0.005]	[0.045]
Un. Rate	-1.161**		
	[0.011]		
Youth Un. Rate		-1.539***	
		[0.000]	
Un. Rate by ed. Level			-1.302***
			[0.007]
A-B AR(1)	-3.01	-2.42	-3.39
	[0.003]	[0.015]	[0.001]
A-B AR(2)	0.67	1.46	-0.41
	[0.501]	[0.145]	[0.679]
Hansen	15.10	14.64	13.70
	[0.178]	[0.199]	[0.250]

See Table 1 for more details.

Table 5. University level net migration rate.

	(1)	(2)	(3)
Lagged dep. Variable	0.773***	0.769***	0.726***
	[0.000]	[0.000]	[0.000]
Per worker GDP	2.590	1.863	1.844
	[0.804]	[0.838]	[0.750]
Un. Rate	-1.816		
	[0.259]		
Youth Un. Rate		-1.981**	
		[0.035]	
Un. Rate by ed. Level			-1.762***
			[0.004]
A-B AR(1)	-2.06	-2.56	-2.15
	[0.040]	[0.011]	[0.031]
A-B AR(2)	0.36	0.31	0.43
	[0.720]	[0.756]	[0.670]
Hansen	13.28	15.03	14.78
	[0.275]	[0.181]	[0.254]

See Table 1 for more details.