September 2016
Volume: 13, **No**: 3, pp. 443 – 454
ISSN: 1741-8984
e-ISSN: 1741-8992
www.migrationletters.com

Article history: Received 22 November 2014; accepted 31 August 2015

CASE STUDY
Internal migration in Italy. Long-run analysis of push and pull factors across regions and macro-areas of the country

Romano Piras [ϒ]

Abstract

The great majority of empirical studies on internal migration across Italian regions either ignore the long-run perspective of the phenomenon or do not consider push and pull factors separately. In addition, Centre-North to South flows, intra South and intra Centre-North migration have been neglected. We aim to fill this gap and study interregional migration flows from different geographical perspectives. We apply four panel data estimators with different statistical assumptions and theoretical properties and show that long-run migration flows from the Mezzogiorno towards Centre-Northern regions are well explained by a gravity model in which per capita GDP, unemployment along with population at both origin and destination play a major role. On the contrary, migration flows from Centre-North to South has probably much to do with other social and demographic factors. Finally, intra Centre-North and intra South migration flows roughly obey to the gravity model, though not all explicative variables are relevant.

Keywords: Internal migration; gravity model; Italy; panel data methods

Introduction

Empirical studies on internal migration across Italian regions have been pursued in recent years with the great majority of them interested in the recent upsurge of internal mobility started from the middle of 1990s (Basile and Causi, 2007; Etzo, 2011; Mocetti and Porello, 2012; Lamonica and Zagaglia, 2013) and ignore the long-run perspective of the phenomenon. Others consider the long-run perspective, but do not consider push and pull factors separately (Attanasio and Padoa-Schioppa, 1991; Daveri and Faini, 1999; Cannari *et al.*, 2000; Brunello *et al.*, 2001; Fachin, 2007; Bonasia and Napolitano, 2012; Piras, 2012a, 2012b). A survey of empirical literature on Italian internal migration can be found in Piras (2016) who also study long-run migration trends of all Italian regions, but does not investigate migration across macro-areas of the country.

[ϒ] Romano Piras is Professor of Economics at the Department of Social and Institutional Sciences, University of Cagliari, Viale S. Ignazio, 17, 09123, Cagliari, Italy.
E-mail: pirasr@unica.it.

Briefly, almost all these works find that the fundamental role in explaining the internal migration pattern is played by macroeconomic variables. Conversely, other social and demographic variables have not been found as robust as the macroeconomic ones. These results confirm that internal migration in Italy has different determinants with respect to internal migration in other countries and international migration across countries in which, as well demonstrated by the current research on migration, networks and diaspora effects play a significant role. In addition, the empirical literature on Italy has typically investigated migration from Southern towards Centre-Northern regions. On the contrary, Centre-North to South flows, not to mention *intra* South and *intra* Centre-North migration, have been largely ignored. In this paper we aim to fill this gap and tackle interregional migration flows from different geographical perspectives: for the country as a whole, from the South to the Centre-North; from the Centre-North to the South; across Southern regions and, finally, across Centre-Northern regions. To accomplish these tasks, we use an extended gravity model of migration and apply it to the twenty Italian regions for the 1970-2005 period.

Empirical specification

Let us consider an extended gravity model of migration such as the following:

$$F_{ijt} = \frac{A_{ij}^{\gamma_1} P_{it}^{\alpha_1} P_{jt}^{\alpha_2}}{D_{ij}^{\gamma_2}} \prod_{s=1}^{m} X_{s,it}^{\beta_{s,i}} X_{s,jt}^{\beta_{s,j}} \qquad (1)$$

where F_{ijt} measures the number of individuals that moves from the region of origin *i* to the region of destination *j* at time *t*; obviously $F_{ijt} \neq F_{jit}$. Along with standard gravity variables, namely population at origin P_{it}, at destination P_{jt}, adjacency A_{ij} and distance D_{ij}, we assume that other variables affect bilateral migration flows. These variables are condensed in the two vectors $X_{s,i}^{\beta_{s,i}}$ and $X_{s,j}^{\beta_{s,j}}$, respectively. More precisely, we consider per capita GDP, pcy_{it} and pcy_{jt}, and unemployment rate, u_{it} and u_{jt}, at both origin and destination. Notice that the γ's, α's and β's are parameters to be estimated.

Equation (1) conjectures that bilateral migration flows from origin *i* to destination *j* are spurred by the "mass" of the two regions given by population P_{it} and P_{jt}: the bigger their mass, the bigger the bilateral flows. Furthermore, higher level of per capita income in the sending (receiving) region leads to lower (higher) migration flows from *i* to *j*. On the contrary, higher level of unemployment rates in the sending (receiving) region leads to higher (lower) migration flows from *i* to *j*.

Based upon equation (1), we estimate the following empirical specification for bilateral migration flows across Italian regions:

$$\ln F_{ijt} = \beta_1 \ln pcy_{it-1} + \beta_2 \ln pcy_{jt-1} + \beta_3 \ln u_{it-1} + \beta_4 \ln u_{jt-1} + \beta_5 \ln n_{it-1} + \beta_6 \ln n_{jt-1} + \\ + \beta_7 \ln D_{ij} + \beta_8 \ln A_{ij} + \lambda_t + \nu_{ij} + \varepsilon_{ijt} \quad (2)$$

where pcy_{it-1} is per capita GDP in the sending region; pcy_{jt-1} represents per capita GDP in the receiving region; u_{it-1} is the unemployment rate in region *i*; u_{jt-1} is the unemployment rate in region *j*; λ_t is a time effect; ν_{ij} is the specific time-invariant effect associated with the region-pair *i* and *j*; finally ε_{ijt} is an additive error term uncorrelated with the covariates.

Leaving all technicalities aside (for more details see Rose and van Wincoop, 2001; Egger and Pfaffermayr, 2003), one can show that the origin-destination effects ν_{ij} account for any time invariant historical, geographical, political, cultural and other bilateral effects. Since most of these influences usually remain unobserved, including bilateral origin-destination effects is the natural way of controlling them. Moreover, also time specific effects λ_t are usually introduced to capture cyclical influences that commonly affect all regions. D_{ij} is the kilometric distance between regional capitals and is expected to discourage migration, since longer distance migration implies higher physical and psychological costs associated with moving. A_{ij} is a dummy variable which equals one if the origin and destination region share a common border and zero otherwise (as for the two main Italian island, Sicily and Sardinia, we have considered them as adjacent to those regions towards which they are connected by regular ferry boat services). The assumption here is that if two regions are adjacent, it is easier to migrate from *i* to *j*.

We apply four panel data estimators with different statistical assumptions and theoretical properties and, for the sake of simplicity, we refrain from dealing with any issue concerning stationarity of the series and heterogeneity of units.[1] We show all the results and comment on standard diagnostic tests and performances of the different estimators. Yet, we do not stick to a particular estimator, rather we provide a wider discussion based on overall results. The random effects (RE) model posits independence between ν_{ij} and ε_{ijt} and assumes that the explanatory variables are independent from both ν_{ij} and ε_{ijt}. If these conditions are met, the RE model is consistent and more efficient than the fixed effects (FE) model. The latter, on the contrary, is always consistent also in the absence of endogeneity or errors in variables. The trouble with the FE model is that by using a within transformation of the data to erase the unobserved region-pair specific effects, it wipes out all time-invariant variables, such as distance. Furthermore, some explanatory variables might be endogenous, namely possibly correlated with ν_{ij}, thought they are orthogonal to ε_{ijt}. In such a situation, the RE estimator cannot be used since it posits

[1] These issues are fully investigated in Piras (2016).

independence between both ν_{ij} and ε_{ijt}. In order to alleviate the potential endogeneity and/or the possibility of reverse causality of some explanatory variables and in addition to Hausman and Taylor (1981) instrumental variable estimation technique described below, we follow the standard practice (Mayda, 2010; Mocetti and Porello, 2012) of lagging by one year all time-variant variables. By so doing we also account for previous period information available to potential migrants.

The Hausman and Taylor (1981) estimator provides an appealing approach to estimate a gravity model in a panel data framework in that it makes use of the several dimensions of the panel data in order to overcome the possible correlation of some explanatory variables with the time-invariant effects and, at the same time, it is able to estimate time-invariant variables as well. Baltagi *et al.* (2003) propose a test procedure to compare the random effects, fixed effects and the Hausman and Taylor estimators. As it is well known, the choice between the random effects and the fixed effects estimators is based upon the standard Hausman test: rejection of the null implies that the fixed effects estimator has to be preferred, otherwise the random effects should be chosen. Baltagi *et al.* (2003) suggest that if the null is rejected, a second Hausman test should be carried out to discriminate between the fixed effects and the Hausman and Taylor estimators. If, in this second Hausman test, the null is not rejected, then the Hausman and Taylor estimator is chosen; otherwise if it is rejected, this testing procedure leads to the fixed effects estimator as the one to be preferred.

A final question to be tackled regards the presence of cross-sectional dependence in the errors due, for example, to common shocks or unobserved components. Such dependence is likely to be present in our sample of Italian regions, given that they are deeply integrated among themselves. In our empirical analysis, in order to formally test whether the cross-sectional units are independent, we estimate equation (2) and use Frees (1995) test on the residuals. Once cross-sectional dependence is possibly detected, we may apply the Driscoll and Kraay (1998) approach (FEDK) in order to obtain a covariance matrix estimator that yields heteroskedasticity consistent standard errors that are also robust to general forms of both temporal and spatial dependence. Thus, the level of significance of the FE and the FEDK estimated coefficients can be compared to check the robustness of the empirical results.

Data and empirical results

Data on unemployment and interregional migration flows are from ISTAT (various years, b) and ISTAT (various years, a), regional per capita GDP and population comes from Svimez (2011).

Let us start with Table 1 that contains the regressions regarding equation (2). The overall picture is good as the majority of variables in all regressions are statistically significant and correctly signed. The testing procedure advocated by

Baltagi *et Al.* (2003) suggests that the FE estimator is to be preferred with respect to the RE and the HT estimator, whereas the Frees (1995) test signals the presence of cross-sectional dependence in both the RE and the FE estimator. Thus, our preferred estimator is the FEDK one in which migration flows are negatively associated with per capita GDP at origin with an estimated elasticity of -0.367 and unemployment at destination with an estimated elasticity of -0.169. Total population, both at origin and at destination exert a positive role on total interregional migration flows and its impact is very similar (1.976 at origin and 2.191 at destination). Finally, and somehow surprisingly, neither per capita GDP at destination, nor unemployment at origin affect the migration decision.

Table 1. Determinants of bilateral migration flows (Total).

Variables	(1) RE	(2) FE	(3) HT	(4) FEDK
Per capita GDP (origin)	-0.505***	-0.367***	-0.475***	-0.367***
	(0.000)	(0.000)	(0.000)	(0.001)
Per capita GDP (dest.)	0.017	0.191***	0.060	0.191
	(0.887)	(0.009)	(0.276)	(0.269)
Unemp. Rate (origin)	0.033	0.049***	0.037***	0.049
	(0.359)	(0.001)	(0.004)	(0.123)
Unemp. Rate (dest.)	-0.191***	-0.169***	-0.184***	-0.169***
	(0.000)	(0.000)	(0.000)	(0.000)
Population (origin)	1.029***	1.976***	1.382***	1.976***
	(0.000)	(0.000)	(0.000)	(0.000)
Population (dest.)	1.042***	2.191***	1.470***	2.191***
	(0.000)	(0.000)	(0.000)	(0.000)
Distance	0.012		-0.119	
	(0.879)		(0.533)	
Proximity	0.929***		0.564*	
	(0.000)		(0.056)	
		Diagnostics and tests statistics		
R-squared	0.383	0.393		0.393
F- and Wald χ^2 test	4509.66	70.09	8609.46	479.91
	(0.000)	(0.000)	(0.000)	(0.000)
Frees (1995) test	38.002	39.626		
	(0.000)	(0.000)		
Hausman test FE vs RE	175.81			
	(0.000)			
Hausman test FE vs HT		123.86		
		(0.000)		

*Observations: 13,300. Units: 380. Robust p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Year dummies not shown. In the HT estimate per capita GDP and unemployment rates at both origin and destination have been treated as endogenous.*

These results can only be compared with those of Etzo (2011) and with Piras (2016). Our point estimate of per capita GDP at origin conforms with the former who, depending on the estimation procedure, reports a range varying from -0.328 to -0.679, and with the latter who reports various estimates in the interval between -0.137 and -0.569. As just said, the estimated coefficient of per capita GDP at destination is not significant in our FEDK estimate, whereas Etzo (2011) finds a strong and statistically significant result ranging from 0.368 up to 0.770. Similarly, Piras (2016) finds that per capita GDP at destination pulls migrations with an estimated elasticity that varies between 0.345 and 0.958. Unemployment rate at destination (-0.169) is strongly statistically significant with an estimated elasticity higher with respect to Etzo (2011) and Piras (2016) who both find similar results around -0.10 or lower. As for population, again our results confirm those of Etzo (2011) and Piras (2016) who report a strong positive effect at origin and at destination, with an array of estimated coefficients varying between 1.004 and 3.897 at origin and between 0.959 and 1.758 at destination. To sum up, our results suggest that, in the long-run, per capita GDP operates more as a push rather than as a pull factor and that population exerts a positive and symmetric role in pulling and pushing individuals across Italian regions.

When we restrict our attention at South-North migration flows, these results are strengthened: as one can see from Table 2, all estimation methods show that virtually all variables included in equation (2) have the expected sign and are highly significant. The testing procedures imply that the HT estimator is to be preferred and that cross-sectional dependence is still present in the RE and the FE estimators. However, the estimated parameters are very similar between the HT and the FEDK estimators, therefore in our opinion both of them should be considered good ones. On the one hand, per capita GDP at origin (with an estimate around 0.7 in both regressions) and unemployment rate at destination (with an estimate from -0.477 in the FEDK estimate to -0.517 in the HT estimate) refrain people from moving from the South to Centre-Northern regions. On the other hand, per capita GDP at destination, with an estimated elasticity above one, and unemployment rate at origin, which reports a point estimate virtually identical in both regressions and equals to 0.33, spur migration flows from the *Mezzogiorno* towards the rest of the country. The regional mass measured by population is a strong force in driving migration patterns. It is worth stressing that the pulling force of population at origin (from 1.303 to 1.466 in the HT and FEDK regression, respectively) is lower than the attractive role played by the same variable at destination where its estimated elasticity is well above four in the HT estimate and close to five in the FEDK estimate.

In contrast with this strong evidence, once we look at the other way round, a completely different image emerges. Table 3 reports the results for the North-South migration flows and it is evident that the estimated equations offer a very poor explicative power. Only population at origin affects the migration decision. On the contrary, the other variables are either not statistically significant or, even worst, when they turn out to be so, they display the wrong sign.

Table 2. Determinants of bilateral migration flows (from South to Centre-North).

Variables	(1) RE	(2) FE	(3) HT	(4) FEDK
Per capita GDP (origin)	-0.811***	-0.725***	-0.744***	-0.725***
	(0.001)	(0.000)	(0.000)	(0.000)
Per capita GDP (dest.)	0.673**	1.144***	1.052***	1.144***
	(0.019)	(0.000)	(0.000)	(0.006)
Unemp. Rate (origin)	0.330***	0.332***	0.330***	0.332***
	(0.000)	(0.000)	(0.000)	(0.000)
Unemp. Rate (dest.)	-0.664***	-0.477***	-0.517***	-0.477***
	(0.000)	(0.000)	(0.000)	(0.000)
Population (origin)	1.038***	1.446***	1.303***	1.446***
	(0.000)	(0.000)	(0.000)	(0.000)
Population (dest.)	1.301***	4.748***	4.092***	4.748***
	(0.000)	(0.000)	(0.000)	(0.000)
Distance	-0.197		0.390	
	(0.253)		(0.704)	
Proximity	0.914***		-0.278	
	(0.000)		(0.879)	
		Diagnostics and tests statistics		
R-squared	0.417	0.477		0.477
F- and Wald χ^2 test	4620.39	87.19	2894.23	88.93
	(0.000)	(0.000)	(0.000)	(0.000)
Frees (1995) test	12.214	11.221		
	(0.000)	(0.000)		
Hausman test FE vs RE	413.16			
	(0.000)			
Hausman test FE vs HT			1.99	
			(1.000)	

*Observations: 3,360. Units: 96. Robust p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Year dummies not shown. In the HT estimate per capita GDP and unemployment rates at both origin and destination have been treated as endogenous.*

These results reaffirm that internal migration in Italy has very different motivations and causes, depending on whether South to North or North to South flows are involved. In the long-run, interregional migration has been unidirectional from the South to the Centre-North and has been primarily driven by economic factors and population. On the contrary, North-South migration has nothing to do with economic forces, but has probably much to do with other social and demographic factors that our model is not able to capture. As pointed out by previous works (Mocetti and Porello, 2012; Piras, 2012b; ISTAT, 2012; Etzo, 2011) a possible explanation of these results could be given in terms of return migration. Many of those individuals who emigrated from the *Mezzogiorno* to the Centre-North decide to come back towards their regions of origin after they get retired, thus partially fuelling the North-South flows. Given the long-lasting economic divide between the Southern and the

Centre-Northern regions, with the former having permanently higher unemployment rates and lower per capita GDP levels with respect to the latter, the results of Table 3 are perfectly in line with the Italian internal migration experience.

Table 3. Determinants of bilateral migration flows (from Centre-North to South).

Variables	(1) RE	(2) FE	(3) HT	(4) FEDK
Per capita GDP (origin)	-0.167	0.123	0.009	0.123
	(0.475)	(0.450)	(0.943)	(0.535)
Per capita GDP (dest.)	-0.317	-0.476***	-0.423***	-0.476***
	(0.109)	(0.001)	(0.000)	(0.002)
Unemp. Rate (origin)	-0.325***	-0.179***	-0.227***	-0.179*
	(0.000)	(0.000)	(0.000)	(0.063)
Unemp. Rate (dest.)	0.112*	0.106***	0.110***	0.106
	(0.057)	(0.001)	(0.001)	(0.173)
Population (origin)	1.218***	3.950***	3.149***	3.950***
	(0.000)	(0.000)	(0.000)	(0.000)
Population (dest.)	0.939***	-0.082	0.304	-0.082
	(0.000)	(0.823)	(0.196)	(0.822)
Distance	0.134		0.983	
	(0.486)		(0.243)	
Proximity	1.290***		0.664	
	(0.000)		(0.657)	
Diagnostics and tests statistics				
R-squared	0.359	0.398		0.398
F- and Wald χ^2 test	233.16	41.71	2102.27	87.74
	(0.000)	(0.000)	(0.000)	(0.000)
Frees (1995) test	8.922	8.092		
	(0.000)	(0.000)		
Hausman test FE vs RE	236.77			
	(0.000)			
Hausman test FE vs HT		134.99		
		(0.000)		

*Observations: 3,360. Units: 96. Robust p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Year dummies not shown. In the HT estimate per capita GDP and unemployment rates at both origin and destination have been treated as endogenous.*

Table 4 reports the results for migration flows across Centre-Northern regions. The standard testing procedure point to the FEDK as the preferred one, though HT estimated elasticities are very close to FEDK. Per capita GDP at origin (which reports a point estimate of -0.792 in the HT regression and of -0.840 in the FEDK regression) and unemployment at destination (estimates around -0.12) refrain migration, population exerts a positive role on it, while neither per capita GDP at destination, nor unemployment at origin are statistically significant. Overall, though less strong than South to Centre-North, the explicative power of the gravity model is quite good also for intra Centre-North flows.

Table 4. Determinants of bilateral migration flows (across Centre-Northern regions).

Variables	(1) RE	(2) FE	(3) HT	(4) FEDK
Per capita GDP (origin)	-0.655**	-0.840***	-0.792***	-0.840***
	(0.014)	(0.003)	(0.000)	(0.000)
Per capita GDP (dest.)	-0.119	-0.119	-0.123	-0.119
	(0.650)	(0.664)	(0.178)	(0.630)
Unemp. Rate (origin)	0.009	-0.023	-0.002	-0.023
	(0.866)	(0.637)	(0.929)	(0.606)
Unemp. Rate (dest.)	-0.107**	-0.119**	-0.121***	-0.119***
	(0.046)	(0.019)	(0.000)	(0.004)
Population (origin)	0.850***	0.471	0.818***	0.471**
	(0.000)	(0.165)	(0.000)	(0.015)
Population (dest.)	0.862***	0.888***	0.862***	0.888***
	(0.000)	(0.007)	(0.000)	(0.002)
Distance	-0.398***		-0.391**	
	(0.003)		(0.018)	
Proximity	0.790***		0.808***	
	(0.000)		(0.000)	
		Diagnostics and tests statistics		
R-squared	0.501	0.502		0.502
F- and Wald χ^2 test	6311.92	87.21	5303.43	42.33
	(0.000)	(0.000)	(0.000)	(0.000)
Frees (1995) test	6.794	6.549		
	(0.000)	(0.000)		
Hausman test FE vs RE	634.75			
	(0.000)			
Hausman test FE vs HT		46.77		
		(0.183)		

*Observations: 4,620. Units: 132. Robust p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Year dummies not shown. In the HT estimate per capita GDP and unemployment rates at both origin and destination have been treated as endogenous.*

Finally, Table 5 reports the results for migration flows across Southern regions. On the one hand, contrary to intra Centre-North flows, here the results clearly indicate that per capita GDP at origin is not statistically significant and that per capita GDP at destination is a magnet for individuals who decide to move. On the other hand, analogously with intra Centre-North flows, unemployment at destination and population at origin show the expected signs, whereas population at destination is statistically significant only in the RE and HT estimators.

Table 5. Determinants of bilateral migration flows (across Southern regions).

Variables	(1) RE	(2) FE	(3) HT	(4) FEDK
Per capita GDP (origin)	-0.109	0.193	0.128	0.193
	(0.516)	(0.245)	(0.233)	(0.159)
Per capita GDP (dest.)	0.389*	0.375*	0.385***	0.375**
	(0.058)	(0.100)	(0.000)	(0.028)
Unemp. Rate (origin)	-0.057	-0.051	-0.056*	-0.051
	(0.274)	(0.331)	(0.085)	(0.259)
Unemp. Rate (dest.)	-0.188**	-0.185**	-0.184***	-0.185***
	(0.014)	(0.017)	(0.000)	(0.008)
Population (origin)	0.962***	1.480**	0.992***	1.480***
	(0.000)	(0.041)	(0.000)	(0.000)
Population (dest.)	0.871***	0.796	0.867***	0.796
	(0.000)	(0.241)	(0.000)	(0.216)
Distance	-1.0101***		-1.022***	
	(0.000)		(0.000)	
Proximity	0.531***		0.523**	
	(0.009)		(0.050)	
	Diagnostics and tests statistics			
R-squared	0.630	0.631		0.631
F- and Wald χ^2 test	8805.94	156.36	3420.61	93.23
	(0.000)	(0.000)	(0.000)	(0.000)
Frees (1995) test	2.158	2.187		
	(0.000)	(0.000)		
Hausman test FE vs RE	249.83			
	(0.000)			
Hausman test FE vs HT		177.26		
		(0.000)		

Observations: 1,960. Units: 56. Robust p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Year dummies not shown. In the HT estimate per capita GDP and unemployment rates at both origin and destination have been treated as endogenous.

Conclusions

In this paper we have investigated bilateral migration flows across Italy during the 1970-2005 period. We have estimated an extended gravity model of internal migration in which migration flows have been regressed on per capita GDP, unemployment rate and population at origin and at destination. We have applied four panel estimators in order to better understand the overall picture of internal migration across Italian regions.

Our results can be briefly summarised as follow. Internal migration flows are inversely linked with per capita GDP at origin and unemployment rate at destination, but neither per capita GDP at destination, nor unemployment at origin seem to affect them. These two variables, however, are significantly linked with internal migration when we consider the flows from the *Mezzogiorno* regions towards the Centre-Northern ones. Also intra Centre-North and intra South migration flows roughly obey to the gravity model we have used to estimate the long run patterns of internal mobility. On the contrary, migration flows from the Centre-North to the South seem not to be driven by economic factors; rather they probably obey to other social and demographic factors connected with the end of the working age period and with the beginning of the retirement life cycle of Southerners previously employed in the Centre-North of the country. Such a claim, however, should be further investigated in future research using also other more general estimation techniques. We are aware that migration is a very complex and layered phenomenon which the gravity model alone cannot fully explain. Yet, the main purpose of this paper has been a long-run quantitative analysis for Italy and, in this perspective, we believe that it can give a contribution to the explanation of internal migration flows. Add also the fact that data unavailability in many dimensions that are worth to be investigated, such as for example class and gender, poses big problems whenever such a long-run quantitative analysis is undertaken. Further empirical research is called for in order to extend the number of variables used in applied analysis and also to improve the quality of the data.

References

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, Cambridge University Press.

Baltagi, B. H., Bresson, G. and Pirotte, A. (2003). Fixed effects, random effects or Hausman–Taylor? A pretest estimator. *Economics Letters*, 79: 361-369.

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

Bonasia, M. and Napolitano, O. (2012). Determinants of interregional migration flows: The role of environmental factors in the Italian case. *Manchester School*, 80: 525-544.

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

Cannari, L., Nucci, F. and Sestito, P. (2000). Geographic labour mobility and the cost of housing: evidence from Italy. *Applied Economics*, 32: 1899-1906.

Daveri, F. and Faini, R. (1999). Where do migrants go?. *Oxford Economic Papers*, 51: 595-622.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 80: 549-560.

Egger, P. and Pfaffermayr M. (2003). The proper panel econometric specification of the gravity equation: A three-way model with bilateral interaction effects. *Empirical Economics*, 28: 571-580.

Etzo, I. (2011). The determinants of the recent interregional migration flows in Italy: A panel data analysis. *Journal of Regional Science*, 51: 948-966.

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.

Frees, E. W. (1995). Assessing cross-sectional correlation in panel data, *Journal of Econometrics*, 69: 393-414.

Hausman, J. A. and Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica*, 49: 1377-1398.

ISTAT (2012). *Le Recente Mobilità Territoriale in Italia. Le Migrazioni dal Mezzogiorno al Centro-Nord nel Periodo 1995-2008*. Argomenti n. 44, ISTAT, Rome.

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

ISTAT (various years, b). *Forze di lavoro*. ISTAT, Rome.

Lamonica, G. and Zagaglia, B. (2013). The determinants of internal mobility in Italy, 1995-2006: A comparison of Italians and resident foreigners. *Demographic Research*, 29: 407-440.

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

Mocetti, S. and Porello, C. (2012). Le nuove migrazioni interne: tendenze nuove di un vecchio fenomeno. *Rivista di Politica Economica*: 275-310.

Piras, R. (2012a). Internal migration across Italian regions: macroeconomic determinants and accomodating potential for a dualistic economy. *Manchester School*, 80: 499-524.

Piras, R. (2012b). Disentangling the role of migrants' educational level in the long-run Italian internal migration trends. *Studies in Regional Science*, 42: 377-396.

Piras, R. (2016). A long-run analysis of push and pull factors of internal migration in Italy. Estimation of a gravity model with human capital using homogeneous and heterogeneous approaches. *Papers in Regional Science*, DOI: 10.1111/pirs.12211

SVIMEZ (2011). *150 Anni di Statistiche Italiane: Nord e Sud, 1861-2011*, Il Mulino, Bologna.