Scale and composition effects of human capital on Spanish regional migration

J. Clemente[a], G. Larramona[a] and V. Montuenga[a,*]

Department of Economic Analysis, University of Zaragoza, 50005, Spain

The aim of this paper is to test the effects on internal migration in Spain of human capital, decomposed into two components: the size, and the composition of the labour force. Our results indicate that those Spanish regions that experienced increases in the ratio of skilled workers to unskilled workers attracted less immigration during the period 1965-1984. Between 1985 and 2000, when employment rates turned non-significant in determining migration, the opposite was observed.

JEL classification: J61, J31, O15, R11

Keywords: Migration, human capital, scale effect, composition effect.

---

[*] Corresponding author. E-mail: victor.montuenga@unizar.es

1. Introduction

Exogenous growth models see migration as a transitory phenomenon, since differences in wages across areas disappear as workers move from low-paid to high-paid areas. This notion has been continuously challenged by empirical facts. While the US has always been a receiving migration, the UK sent workers abroad in the nineteenth century, but it was a receiving country in the following century. More recently, Spain, a traditional country of emigrants, became a receiving country at the beginning of the new century.

Endogenous growth models allow for unidirectional movements in the long run with growth in wages proportional to the growth of the externality considered, fitting well the US case. Reichlin and Rustichini (1998) have pointed out that some of the driving force behind migration comes from the size and composition of the labour force, and that changes in the proportion of skilled over unskilled workers (the composition effect) may act in the opposite way to a mere increase in population, giving rise to inverse migration.

We investigate the determinants of net inter-regional migration flows in Spain, where a structural change was seen during the 1970s and 1980s. With data from 1964 to 2000, we show that both scale and composition effects have a significant effect, and are important elements in the current migration debate over inverse migration. Section II summarizes the literature and presents the model, Section III presents our empirical results, and Section IV concludes.

II. The model

Economically, wages and unemployment are the fundamental variables in explaining migration. However, Antolín and Bover (1997) found a reduced or null influence of both on inter-regional flows in Spain. More recently, Mulhern and Watson (2010) found that, in the second half of the 1990s, the responsiveness of migration to wages and unemployment differences increased due to labour market reforms and the massive influx of immigrants to Spain. Until the mid-1970s, elevated gross and net unidirectional movements from underdeveloped South-Western regions to more developed North-Eastern regions (and Madrid), were observed. Between the mid-1970s and the mid-1980s, flows decreased and then recovered in gross terms (Bover and Velilla, 2005; Maza and Villaverde, 2004). After that, the massive influx of foreigners altered the pattern of inter-regional migration (Hierro, 2009; Mulhern and Watson, 2010), what makes us to select 2000 as the last year in our sample.

We base our investigation in the model of Reichlin and Rustichini (1998) using non-stationary heterogeneous panel data techniques. Migration is driven by differences between expected earnings in two regions (*i*): the host (*1*) and the home region (*2*),[1] which are computed as the product of the employment rate and the wage. Regions differ in their initial stocks of human and physical capital. Additionally, there are two types of workers, skilled and unskilled, with L_i and N_i being, respectively, the number of skilled and unskilled workers in region *i*. Unskilled migrants take into account differences in unskilled expected earnings, $(1-U_i^N)\ v_i$, and skilled migrants the differences in skilled expected earnings, $(1-U_i^L)\ w_i$. The migration rate to region 1, m_1, is the net amount of workers received in the region over the total amount of native workers (L_1+N_1). Accordingly, m_1 can be expressed as the sum of the rate of skilled and unskilled migrants ($m_1^N + m_1^L$). These rates are then assumed to be represented by a logarithmic function, so that

$$m_1 = m_1^L + m_1^N = \log\left[\frac{1-U_1^L}{1-U_2^L}\frac{w_1}{w_2}\right]^{\gamma_1} + \log\left[\frac{1-U_1^N}{1-U_2^N}\frac{v_1}{v_2}\right]^{\gamma_2} \tag{1}$$

Stemming from a regional production function with increasing returns to scale, see Reichlin and Rustichini (1998), gaps of skilled and unskilled (w_1/w_2) and (v_1/v_2) depend on three factors; first, the capital-labour ratio (k_1/k_2), where $k_i = \frac{K_i}{N_i}$ is the capital-unskilled labour ratio, capturing the absence of perfect movement of capital;[2] second, the relative size of the unskilled labour force (N_1/N_2), capturing the scale effect; third, the relative composition of the labour force (π_1/π_2), where $\pi_i = \frac{L_i}{N_i}$ is the skilled labour per unit of unskilled labour, capturing the composition effect. Therefore, the specification to be estimated is

$$m_1 = \gamma_0^{'} + \gamma_1^{'}\log(\frac{1-U_1^L}{1-U_2^L}) + \gamma_2^{'}\log(\frac{1-U_1^N}{1-U_2^N}) + \gamma_3^{'}\log\left(\frac{k_1}{k_2}\right) + \gamma_4^{'}\log\left(\frac{\pi_1}{\pi_2}\right) + \gamma_5^{'}\log\left(\frac{N_1}{N_2}\right) + e \tag{2}$$

[1] There are two regions with index *i=1,2* respectively, where 1 is the specific region and 2 the whole set of the remaining regions.

[2] This implies that marginal product of capital may diverge across regions (Yamamoto, 2008). We thus relax the assumption of perfect mobility of capital imposed in Reichlin and Rustichini (1998), and test it explicitly with our data.

III Data and Empirical Results

We use annual observations from 1964 to 2000, for the 17 Spanish NUTS2 regions. Net migration flows between each region and the rest of the country are obtained from the Spanish National Statistics Office. Data on physical capital come from the BBVA Foundation, and data on employment by qualifications from Valencia Institute of Economic Research.[3]

The empirical analysis is carried out in three stages: first, assessing the non-stationarity of series; second, testing for co-integration; and, third, estimating the panel co-integration relationship. Additionally, heterogeneity in the panel must be assessed and dealt with. In a first stage, we apply unit root tests for panel data (Levin et al. 2002; Im et al. 2003) finding that the hypothesis I(1) is non-rejected, whereas the hypothesis of series being I(2) is strongly rejected (Table 1, upper and bottom panels, respectively). In a second stage, we apply the cointegration tests (Pedroni, 1999). Table 2 shows clear evidence that variables exhibit a long term relationship.

Prior evidence suggests modelling Equation (2) in a dynamic way to capture the marked inertia observed in the Spanish migratory flows (Maza and Villaverde, 2004; Hierro, 2009). Consequently, we can think of equation (2) as a long-run relationship in which any deviation of the equilibrium is recovered through an error correction model. We allow for heterogeneity across sectional units and re-parameterise (2) as

$$\Delta m_{1t} = \phi_i (m_{1t} - X_{1t}'\gamma) + \lambda_1 + \delta_1 \Delta \log(\frac{1-U_{1t}^L}{1-U_{2t}^L}) + \delta_2 \Delta \log(\frac{1-U_{1t}^N}{1-U_{2t}^N}) + \delta_3 \Delta \log(\frac{k_{1t}}{k_{2t}}) +$$
$$+ \delta_4 \Delta \log(\frac{\pi_{1t}}{\pi_{2t}}) + \delta_3 \Delta \log(\frac{N_{1t}}{N_{2t}}) + \mathrm{e}_{1t} \qquad (3)$$

and estimate (3) through the PMG model.[4] X_1 is the vector of all explanatory variables, γ's are the long-run coefficients, δ's are the short-run coefficients, and ϕ is the error correcting speed of the adjustment term. Regional intercepts (λ) are included.

We present estimates for the whole period and two sub-periods, Table 3.[5] Co-integration is confirmed, since ϕ is negative in all cases. Thus, any deviation from the long-run is

[3] Skilled are those with secondary non-compulsory education or higher, with unskilled otherwise.

[4] The PMG (Pooled Mean Groups, Pesaran *et al.*, 1999) constrains the long-run coefficient vector to be equal across panels while allowing for region-specific short-run adjustment coefficients. A similar approach has been followed by Fachin (2007) for estimating internal migration in Italy.

[5] Bover and Velilla (2005); Maza and Villaverde (2004) and Hierro (2009) concur that in the mid-1980s there was a shift in migration behaviour due to the increase in both intra-regional flow movements and return migration, thus contributing to an increase in gross inter-regional flows, but not in net flows.

corrected with short-run adjustments. The relative employment rates of skilled and unskilled are both statistically significant for the whole period. Whereas an increase in the relative employment rate of unskilled stimulates immigration into region 1, an increase in the case of skilled deters it. This may explain the lack of significance observed in the prior literature, since these works make no distinction between high and low qualifications of the workforce. Over time, these effects appear to vanish.

The relative capital to unskilled labour ratio is statistically significant and positively estimated, capturing the existence of barriers to the free movements of capital and the attractive characteristics of better-endowed regions. The decentralized political system in Spain permits numbers of public investment projects at regional level, especially after the 1980s, when the process of regional decentralisation became more prominent.

An increase in the relative number of unskilled workers in region 1 leads immigration into region 1, confirming the scale effect. By contrast, the composition effect is found to be negative, but insignificant. There is, however, a different behaviour in the two sub-periods considered. Between 1962 and 1984, the parameter is negative and statistically significant, while it is also significant, but positive, between 1985 and 2000. These results may be due to the change in the productive structure of the Spanish economy of the 1980s, when the government implemented new industrial and education policies that modified labour demand in favour of skilled workers. Thus, the composition effect may be behind the inverse migration in Spain, as Reichlin and Rustichini (1998) point out.

IV. Conclusion

We have investigated the phenomenon of inverse migration by studying inter-regional net migration flows in Spain during the last part of the 20th century through non-stationary heterogeneous panel data techniques. Following Reichlin and Rustichini (1998), wages are decomposed into scale and composition effects.

Our results show that, except for the scale effect, there is a change in the mid-1980s. Relative employment rates lose significance in explaining migration, being substituted by the relative capital to unskilled labour ratio and the composition effect. As a consequence, those regions experiencing increases in the ratio of skilled workers-to-unskilled workers induced more immigration, while prior to 1985, the opposite was observed.

References

Antolin, P. and Bover, O. (1997) Regional migration in Spain: the effect of personal characteristics and of unemployment, wage, and house price differentials using pooled cross-sections, *Oxford Bulletin of Economics and Statistics*, **59**, 215-35.

Bover, O. and Velilla, P. (2005) Migration in Spain: Historical Background and Current Trends in *European Migration: What Do We Know?*, (Ed.) K. Zimmermann (Ed.), CEPR and Oxford University Press, Oxford, pp. 389-414.

Fachin, S. (2007) Long-run trends in internal migrations in Italy: a study in panel cointegration with dependent variables, *Journal of Applied Econometrics*, **22**, 401-28

Hierro, M. (2009) Modelling the dynamics of internal migration flows in Spain, *Papers in Regional Science*, **88**, 683-92.

Im, KS., Pesaran, MH. and Shin, Y. (2003) Testing for unit roots in heterogeneous panels, *Journal of Econometrics*, **115**, 53-74.

Levin, A., Lin, CF. and Chu, C. (2002) Unit root tests in panel data: asymptotic and finite sample properties, *Journal of Econometrics*, **108**, 1-24.

Maza, A. and Villaverde, J. (2004) Internal migration in Spain: a semiparametric analysis, *The Review of Regional Studies*, **34**, 156-71.

Mulhern, A. and J. Watson (2010) Spanish inter-regional migration: an enigma resolved. *Applied Economic Letters*, **17**, 1355-59.

Pedroni, P. (1999) Critical values for cointegration tests in heterogeneous panels with multiple regressors, *Oxford Bulletin of Economics and Statistics*, **61**, 653-70.

Pesaran, M.H., Shin, Y. and Smith, R. P. (1999). Pooled mean group estimation of dynamic heterogeneous panels, *Journal of the American Statistical Association*, **94**, 621-34.

Reichlin, P., Rustichini, A. (1998) Diverging patterns with endogenous labor migration. *Journal of Economics Dynamics and Control*, **22**, 703–28.

Yamamoto, K. (2008) Location of industry, market size, and imperfect international capital mobility *Regional Science and Urban Economics*, **38**, 518-32.

Table 1. Spatial corrected panel unit roots tests

VARIABLE		$Log(1-U^L_1/ 1-U^L_2)$	$Log(1-U^N_1/1- U^N_2)$	$Log(k_1/ k_2)$	$Log(\pi_1/ \pi_2)$	$Log(N_1/ N_2)$	m_1
a) Testing for the null hypothesis I(1) vs I(0)							
Levin, Lin & Chu t (Common unit root)	Statistic	0.54240	1.97010	-0.06389	2.53065	2.6965	3.38773
		(0.7062)	(0.9756)	(0.4745)	(0.9943)	(0.9965)	(0.9996)
Im, Pesaran and Shin W-stat (Individual unit root)	Statistic	0.06365	1.74673	-0.28368	0.71236	1.78004	-0.72091
		(0.5254)	(0.9597)	(0.3883)	(0.7619)	(0.9625)	(0.2355)
b) Testing for the null hypothesis I(2) vs I(1)							
Levin, Lin & Chu t (Common unit root)	Statistic	-6.3648	-9.15634	-2.00026	-4.23456	-2.82157	-4.05210
		(0.0000)*	(0.0000)*	(0.0227)*	(0.0000)*	(0.0024)*	(0.0000)
Im, Pesaran and Shin W-stat (Individual unit root)	Statistic	-12.0430	-14.7682	-8.3457	-11.6823	-9.24952	-10.9135
		(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)

* indicates significance at 1% level. P-values between parentheses.

Table 2. Pedroni co-integration test

Null Hypothesis: No co-integration				
Alternative hypothesis: common AR coefs.				
	Statistic	Prob.	Weighted Statistic	Prob.
Panel rho-Statistic	1.234204	0.1863***	1.738785	0.088**
Panel PP-Statistic	-3.987137	0.0001*	-4.842569	0.0000*
Panel ADF-Statistic	-6.713748	0.0000*	-6.946168	0.0000*
Alternative hypothesis: individual AR coefs.				
Group rho-Statistic	2.471305	0.0188*		
Group PP-Statistic	-6.514086	0.0000*		
Group ADF-Statistic	-7.933942	0.0000*		

* indicates

Table 3. Long- and short-run PMG estimates

	1962-2000		1962-1984		1985-2000	
	Coef.	*t-ratio*	*Coef.*	*t-ratio*	*Coef.*	*t-ratio*
Error correction coefficient (ϕ)	-0.358	(-8.59)	-0.515	(-8.26)	-0.432	(-4.77)
Long run coefficients (γ)						
$Log(1-U^L_1/1-U^L_2)$	-0.034	(-4.37)	0.043	(2.10)	-0.010	(-1.03)
$Log(1-U^N_1/1-U^N_2)$	0.034	(4.40)	0.081	(4.30)	0.008	(0.79)
$Log(k_1/k_2)$	0.021	(5.44)	0.004	(0.78)	0.033	(6.98)
$Log(\pi_1/\pi_2)$	-0.003	(-0.63)	-0.017	(-3.31)	0.005	(1.72)
$Log(N_1/N_2)$	0.008	(3.43)	0.013	(2.66)	0.019	(7.57)
Short-run coefficients (δ) (average)						
$dLog(1-U^L_1/1-U^L_2)$	0.022	(1.89)	-0.078	(-1.50)	0.005	(0.32)
$dLog(1-U^N_1/1-U^N_2)$	0.090	(0.43)	0.142	(1.39)	0.011	(0.55)
$dLog(k_1/k_2)$	0.093	(4.03)	0.159	(2.39)	0.006	(0.19)
$dLog(\pi_1/\pi_2)$	0.013	(1.98)	0.034	(1.61)	0.007	(0.60)
$dLog(N_1/N_2)$	0.118	(4.54)	0.175	(2.77)	0.018	(0.50)
Log likelihood	2707.446		1538.339		1442.234	
Number of obs.	612		340		272	