Regional Convergence in Europe
and the Industry-Mix:
a Shift-Share Analysis

J. Esteban
Institut d'Anàlisi Econòmica, CSIC

This version: December 1999

Mailing address: Institut d'Anàlisi Econòmica, CSIC
Campus de la UAB
08193 Bellaterra. Barcelona
Tel 34 (9)3 5806612 Fax 5801452
e-mail: iiae2@cc.uab.es

SUMMARY: In this paper we try to elucidate the extent to which existing interregional inequality in aggregate productivities per worker within the European Union can be attributed to differences in the sectoral composition of activities, rather than to productivity gaps that are uniform across sectors. To this effect we use the *shift-share analysis* and show that regional specialization has a very minor role and that interregional differences can essentially be explained by uniform productivity gaps only. Our empirical results turn out to be statistically very significant and robust to different definitions of Gross Value Added (market prices and factor costs), different degrees of sectoral breakdowns, dates and alternative sets of countries. Our findings thus provide support to regional development policies focusing on actions producing uniform increases in regional productivities, such as infrastructures and human capital.

CSJA: O40, R11, R12 and R58

ACKNOWLEDGMENTS: I have benefited from useful comments by K. Stahl, A. de la Fuente, T. Garcia-Milà, R. Marimon, D. Ray and two anonymous referees. Responsibility for errors remains my own. I wish to thank most specially the very competent research assistance by J-A Duro. Financial support from the research grant CICYT PB96-0897 and the Fundació d'Economia Analítica is gratefully acknowledged.

1. Introduction

The work of Barro and Sala-i-Martín [1991] on regional convergence has triggered a number papers investigating the dynamics of regional income within the European Union. The results seem to be quite inconclusive. Abraham and van Rompuy [1995], Armstrong [1995], de la Fuente [1994], and Molle and Boeckhout [1995] have conducted different tests for the existence of *beta* convergence across the EU regions and all have obtained similar results: they have been able to identify a negative –but very small- convergence factor. Marcet [1994] and Canova and Marcet [1995], working with panel data, allow for regional dummies and obtain that the present relative income of the EU regions is very close to their steady income. Finally, Quah [1993] and [1996] uses an estimated transition matrix to estimate the evolution of the distribution of regional income. His results[1] do not identify a tendency to converge towards a common per capita income.

A possible explanation for such a seemingly weak convergence process is that most of the empirical work has dealt with *per capita* income, rather than productivity per worker. *Per capita* income can be thought of as the product of three factors: productivity per worker, employment rate and participation rate. Indeed, the regional variability in unemployment or participation rates, rather than dispersion in productivities, could be a possible explanation for the dispersion of the regional per capita incomes. Carlino [1992] and Browne [1989] show that for the United States[2] most of the interregional inequality in per capita income can be attributed to differences in employment rates. As for Europe, Esteban [1994] shows that interregional inequality in productivity per worker has significantly decreased during the 80s, but that by 1989 it still accounted for two thirds of the recorded inequality in per capita income.[3]

The persistence of significant interregional differences in aggregate productivity per worker may be compatible with a process of regional equalization of productivity, sector by sector.[4] Suppose that for each single sector productivity is equal across all the regions. As long as the (average) productivity per worker differs

---

[1] See also Neven and Gouyette [1995].

[2] In Esteban [1994] it is shown that interregional inequality in Europe (with 140 regions) is more than twice than inter-state inequality in the US (with 48 states).

[3] Esteban [1994] and Duro and Esteban [1998] show that the inequality of income measured with the Theil index can be decomposed into the sum of the inequalities of productivity per worker, activity rates and participation rates.

[4] Paci [1997] finds that during the 80s there has been a significant regional convergence in productivity per worker in the EU in industry and services, but not in agriculture.

across sectors, differences in the industry-mix can give rise to variations in aggregate productivity per worker at the regional level. Alternatively, it can be the case that the interregional differences in aggregate productivity affect all industrial sectors uniformly, possibly produced by the differential endowment of region-specific factors such as infrastructures or communications. The elucidation of the role of the industry-mix and of the region-specific productivity factors is relevant on at least three grounds. Firstly, it may help to clarify whether, below the persistence of interregional inequalities in aggregate productivity per worker, there has been a convergence in productivity sector by sector. Secondly, a critical role for the industry-mix would cast doubts on the relevance of the aggregative one-sector growth models in explaining the regional differences in per capita income. Finally, and from a policy point of view, the results would validate/question the appropriateness of the EU regional policy, essentially based on instruments geared to generate uniform productivity increases in backward regions (infrastructure and human capital).

In this paper we wish to examine in a unified framework the role of the differences in sectoral productivity and in the industry-mix in explaining the interregional inequality in productivity per worker. To this effect in Section 2 we develop a reformulation of the *shift-share analysis* to decompose regional aggregate productivity per worker into three components: industry-mix, productivity differential and allocative. In Section 3 we describe the data and test for the role played by each component. We also test whether interregional differences can be attributed to one *shift-share* component alone. Our main finding is that differences in productivity per worker can be fully explained by the existence of region-specific productivity differentials, uniform across sectors. This result is robust to alternative definitions of gross value added (market prices and factor costs) and sets of countries.

2. A multisectoral analysis of interregional inequality: the *shift-share* approach.

The question we address can be most usefully approached by means of the *shift-share* analysis. This was originally proposed by Dunn [1960] as a forecasting technique for regional growth (usually, employment). The essential idea is to analyze the extent to which the difference in growth between each region and the national average is due to the region performing uniformly better than average on all industries or to the fact that the region happens to be specialized in fast growing sectors.[5]

---

[5] See recent applications to the US by García-Milà and McGuire [1993] and to Spain by García-Milà and Marimon [1996].

Esteban (1972) modified the standard two factor decomposition and extended it to the sum of three components: structural, differential, and allocative. The first component indicates the growth share attributable to the particular industry-mix of each specific region. The second one measures the part due to the region growing faster at the sectoral level (possibly because of a higher productivity). Finally, the third component measures the covariance between the two previous components. This can be interpreted as the contribution to regional growth deriving from its specialization in those activities where the region is most competitive.

Even though *shift-share analysis* was originally conceived as a technique to analyze regional employment dynamics, it is straightforward to extend it to the decomposition of interregional aggregate productivity differentials. Aggregate average productivity per worker is the weighted sum of the productivities at the sectoral level. Thus, a particular region can have an aggregate productivity per worker above the mean because of two reasons (or a combination of both). On the one hand, it can be that in all, or most, sectors this region has a productivity per worker above the mean. On the other hand, it can be the case that sectoral productivities are not different from the mean, but that this region is specialized in those sectors with higher productivity per worker. For instance, the average productivity in agriculture, in industry or in the service sector could be identical across the EU regions. Yet, the regions specialized in services would have an aggregate productivity per worker higher than those specialized in agriculture.

In order to establish the role of region-specific productivity differentials and/or industry-mix we contrast each region against a prototype, benchmark region endowed with sectoral productivities and industry-mix equal to the Europe-wide average. The singularity of any particular region is captured by the differences between the actual and prototype industry-mix and between actual and prototype sectoral productivities. Therefore, the shift between the actual aggregate productivity of a given region and the European average productivity per worker can be imputed to each type of difference with respect to the prototype or to both. To this end we compute for each region the fictitious aggregate productivity that this region would have had had it differed from the benchmark standards in one respect only. The difference with respect to the prototype, European productivity per worker gives the contribution of each of these two factors in explaining the gap between actual and benchmark regional aggregate productivity per worker. We shall now show that this gap is the sum of the contribution of these two factors, together with a third component collecting the covariance between the two.

Let p_i^j be sector j's employment share in region i. Therefore $\sum_j p_i^j = 1$, for all regions i. We shall denote by p^j sector j's employment at the European level. Here we shall also have that $\sum_j p^j = 1$. Similarly, we shall use x_i^j and x^j to denote the productivity per worker in sector j and region i, and at the European level, respectively. Therefore, we shall have the following equalities,

$$x = \sum_j p^j x^j \text{ and } x_i = \sum_j p_i^j x_i^j .$$

We can now define the three components of the regional deviation in productivity.

The *industry-mix* component μ_i of region i measures the differential productivity accruing from region i's specific sectoral composition, once we assume that sectoral productivities in each region are equal to the European averages. We thus write

$$\mu_i = \sum_j (p_i^j - p^j) x^j . \tag{1}$$

μ_i takes on positive values if the region is specialized ($p_i^j > p^j$) in sectors with high productivity at the European level and de-specialized ($p_i^j < p^j$) in sectors of low productivity. μ_i is maximal if the region were specialized in the most productive sector Europe wide. Conversely, μ_i would attain its minimum if the region were specialized in the least productive sector.

Note that (1) can be rewritten as

$$\sum_j p_i^j x^j = x + \mu_i . \tag{2}$$

The LHS is the average productivity per worker in region i if regional and national productivities would coincide sector by sector. Expression (2) then says that region i's average productivity would be equal to the national average plus the industry-mix component.

The *productivity differential* component π_i focuses on the contribution of sectoral productivity differences to the shift between regional and national average productivities, on the assumption that the region's industry-mix coincides with the national one. We then define π_i as

$$\pi_i = \sum_j p^j (x_i^j - x^j). \tag{3}$$

π_i takes on positive values if the region has sectoral productivities above the European average. Furthermore, for a given sectoral productivity differential, π_i is increasing in the share of this sector at the European level.

Note again that (3) can be rewritten as

$$\sum_j p^j x_i^j = x + \pi_i . \qquad (4)$$

The LHS stands for the average productivity of region i when its industry mix equals the national one and hence any differential in average productivity must be caused by sectoral productivity differences. Region i's average productivity could thus be expressed as the sum of the national average plus the productivity differential component.

The *allocative* component α_i is defined as

$$\alpha_i = \sum_j (p_i^j - p^j)(x_i^j - x^j) . \qquad (5)$$

This component is positive if the region is specialized, relative to the European average, in sectors whose productivity is above the European average and negative if below.[6] α_i attains its maximum if the region is completely specialized in the sector with the largest productivity differential with respect to the European average. This component is an indicator of the efficiency of each region in allocating its resources over the different industrial sectors. The allocative component can also be viewed as measuring the co-variance between sectoral specialization and productivity advantages.

We can now bring together the three components and write

$$x_i - x = \mu_i + \pi_i + \alpha_i . \qquad (6)$$

In (6) we have the gap between regional and national average productivities additively decomposed into the three components. Each component aggregates one source of regional differential productivity.

Before going into the empirical exercise, let us give an interpretation of the sign of the different components. Suppose that the European economy had not undergone productivity advances for a long period, so that sectoral technologies are

[6] Observe that accordingly with the original shift-share decomposition the allocative component was integrated into the productivity differential component obtaining the following productivity differential component

$$\tilde{\pi}_i = \sum_j p_i^j (x_i^j - x^j).$$

uniform across regions and factor mobility has fully exploited any local advantage. Marginal and average productivities have become equated through the regions, sector by sector. The size of a specific sector in the different regions might depend on history: the past existence of locational advantages that will have faded away by the end of the reallocation process. In that state of affairs, regional differences in per capita aggregate productivity might still subsist due to the fact that some regions have ended up being specialized in sectors with particularly high (or low) average productivity. This situation will be recorded in our decomposition by positive and negative values for the industry mix component, while the other two components would record zeroes in every region.

Suppose now that a productivity shock has taken place in some specific sector and at a particular location. The national average productivity in that sector has increased because of the shock and hence the productivity differential component will be negative in all but the benefited region, where it will take on positive values. Regions specialized in that sector will record an increase in the industry-mix component. If the increase takes place in a non-specialized region, then the allocation effect is negative until it has attracted more than average resources or the differential advantage has been caught up by other regions.

3. Decomposing regional inequalities

3.1. The data

We have used the data published at the regional level NUTS 2[7] by Eurostat in the data set REGIO 1995. Sectoral data on regional employment and Gross Value Added (at factor costs and at market prices) are provided for a breakdown of 6 and 17 sectors, respectively. Thus, it is in principle possible to compute sectoral average productivities at the regional level. However, the data are far from complete. This has imposed severe restrictions on the countries and period covered by our exercise.

Because of the lack of data, we have performed our exercise measuring productivity by the gross value added (GVA) at market prices for a breakdown of 17 sectors. The countries for which this information is available (for 1986 only) are

[7] There are around 150 such regions in the European Union as a whole. Their size varies quite substantially, ranging from regions not exceeding the 200,000 inhabitants to regions like the Southeast (UK) which includes the greater London. The number of NUTS 2 regions per country also varies quite substantially. Ireland, for instance, is considered one single region whilst Germany contains 31 NUTS 2 regions.

Belgium, France, Italy, Portugal and Spain. In order to test the robustness of our results with respect to a larger set of countries, to the use of GVA at factor costs, and to different years, we had to restrict ourselves to a breakdown into 6 sectors. But, even with this sacrifice we could add Germany only. Further, when comparing 1986 with 1989 we lose Belgium. Finally, data on GVA at factor costs are available for 1986 only and for the five largest EU countries. In view of the impossibility of defining a core set of countries with reasonable significance, we have opted to carry out the robustness analysis with four different sets of countries. These sets are:

Set 1: Belgium, France, Italy, Portugal and Spain. We have information on the GVA at market prices for 1986 and a breakdown of 17 sectors. Notice that Germany and the UK are not included.

Set 2: Belgium, France, Germany, Italy, Portugal and Spain. This is the largest set of countries possible. We have the same information as in set 1, but only for 6 sectors.

Set 3: France, Germany, Italy, Portugal and Spain. For this set of countries we have the same information as for set 2 referred to 1989.

Set 4: France, Germany, Italy, Spain and United Kingdom. It contains the five largest countries in the EU. This is the only set for which we have information on the GVA at factor costs, for 1986, and for 6 sectors.

The intersection of the four sets contains France, Italy and Spain, only.

3.2. The exercise

One way of measuring the role played by each of the *shift-share* components in explaining interregional differences in productivity per worker is to compute the relative weight of the variance of each component in the overall observed variance, together with a term collecting the covariances.

Since regional average productivity shifts $x_i - x = y_i$ can be decomposed into the sum of the three factors, as in (6), it is easy to obtain that

$$\text{var}(y) = \text{var}(\mu)+\text{var}(\pi)+\text{var}(\alpha)+2[\text{cov}(\mu,\pi)+\text{cov}(\mu,\alpha)+\text{cov}(\pi,\alpha)]. \quad (7)$$

In order to have a sharper appreciation of the role played by each component, we shall also test whether interregional differences in aggregate productivity per worker can be explained by a model including one single component of the *shift-share* decomposition presented in Section 2. To this effect we shall estimate the following three models:

$$x_i - x = a_\mu + b_\mu \mu_i + \varepsilon_\mu, \quad (8a)$$
$$x_i - x = a_\pi + b_\pi \pi_i + \varepsilon_\pi, \quad \text{and} \quad (8b)$$
$$x_i - x = a_\alpha + b_\alpha \alpha_i + \varepsilon_\alpha. \quad (8c)$$

3.3. The main result

We start by computing (7) using the data set 1, with the GVA at market prices for 1986 and for 17 sectors. The countries included are Belgium, France, Italy and Spain. The result obtained is given in Table 1.

Table 1
Share on total variance by components, 17 sectors. 1986

$\frac{var(\mu_i)}{var(x_i)}$	$\frac{var(\pi_i)}{var(x_i)}$	$\frac{var(\alpha_i)}{var(x_i)}$	$\frac{2\sum cov}{var(x_i)}$
0,1416	0,7037	0,0889	0,0658

The results are quite unambiguous. Most of the observed interregional variance in aggregate productivity per worker is attributable to pure productivity differentials. It is also worth noting the modest weight of the covariances.

In order to have a sharper view of the role played by the different components, we have tested the model (8) using again the data set 1. The results obtained are given in Table 2.

Table 2: Parameter estimates for 17 sectors, data set 1

	$\hat{a}$	$\hat{b}$	R^2
Model 8a (μ)	-881.86	2.0123	0.5735
Model 8b (π)	-958.11	1.1264	0.8928
Model 8c (α)	-2400.71	-0.8732	0.0678

It is immediate that the fit given by the productivity differential component is much better than when using any other component separately. The value of R^2 is substantially higher. Moreover, the value of the corresponding *b* parameter is remarkably close to unity.

This result tells us that we can have a very accurate prediction of the differences in aggregate productivity per worker between any two European regions on the basis of the differential component only. It follows that the observed interregional differences are essentially due to productivity shifts that are uniform across sectors.

3.4. Robustness of the result

The data set we have used in the previous exercise contains countries with a strong Southern European bias. Towards testing for the inclusion of other countries we use data set 2 which includes Germany, but at the cost of drastically reducing the number of sectors from 17 to 6.

Table 3: Parameter estimates: 17 vs. 6 sectors and data set 1 vs. 2

	Model 8a (μ)			Model 8b (π)			Model 8c (α)		
	$\hat{a}$	$\hat{b}$	R^2	$\hat{a}$	$\hat{b}$	R^2	$\hat{a}$	$\hat{b}$	R^2
Data set 1 17 sectors	-881.86	2.0123	0.5735	-958.11	1.1264	0.8928	-2400.71	-0.8732	0.0678
Data set 1 6 sectors	-998.34	2.2888	0.4872	-356.56	1.1155	0.9753	-2135.93	-1.8673	0.1628
Data set 2 6 sectors	-1345.46	2.3690	0.5320	-392.91	1.1184	0.9751	-2849.40	-2.2095	0.2131

From Table 3 we can conclude that, for the same data set 1, the reduction from 17 to 6 sectors (row 1 and 2) decreases slightly the quality of fit of the industry-mix component, while it increases it further for the productivity differential component. This was to be expected. If we now compare rows 2 and 3 we can see that the inclusion of Germany does not reduce the explanatory role of the productivity differential component. Furthermore, the value of the parameter b_π is almost identical in all the estimations.

Our data refer to one year only. Towards testing for the robustness of our results in time, we use data set 3. In this, we have information for 1989, excluding

Belgium. In Table 4 we compare the results for 1986 and 1989 using data set 3. We observe that the impact of the different components as well as the estimated parameter values remain unchanged.

Table 4: Parameter estimates: 1989 and 1986, 6 sectors (data set 3)

	Model 8a (μ)			Model 8b (π)			Model 8c (α)		
	$\hat{a}$	$\hat{b}$	R^2	$\hat{a}$	$\hat{b}$	R^2	$\hat{a}$	$\hat{b}$	R^2
1986	-1191.97	2.5306	0.5488	-334.22	1.1193	0.9746	-2768.60	-2.6423	0.2314
1989	-868.12	2.5768	0.6169	-516.53	1.1491	0.9689	-3017.60	-2.4313	0.2009

Our last test concerns the use of value added at market prices instead of at factor costs. We would like to exclude the role played by taxes and transfers. In Table 5 we present the results of fitting (8) using Gross Value Added at factor costs. Unfortunately this forces us to use a different data set, namely data set 4. It contains France, Germany, Italy, Spain and the United Kingdom, the five largest EU countries. It can be readily verified that the change in the concept of Value Added does not affect the results concerning the explanatory power of productivity differentials. Note, however, the sharp fall in the R^2 corresponding to the industry-mix component.

Table 5: Parameter estimates for 6 sectors, data set 4
Gross Value Added at factor costs (1986)

	$\hat{a}$	$\hat{b}$	R^2
Model 8a (μ)	-434.14	1.6879	0.1696
Model 8b (π)	-659.73	1.0779	0.9570
Model 8c (α)	-1227.79	0.9843	0.0115

Summarizing, it seems remarkable that in all the tests we have obtained the same result. It is that region-specific productivity differentials account for virtually all interregional differences in aggregate productivity per worker.

4. Summary and policy implications

In this paper we have offered a method to asses which part of interregional differences in average productivity is attributable to the region-specific productivity differentials and which to the particular sectoral composition. In order to single out the role of these factors we have used the standard shift-share analysis to decompose regional income distances with respect to the European mean into three components: industry-mix, productivity differentials and allocation. The empirical results obtained are remarkably neat and robust: interregional differences can be almost fully explained by region-specific productivity differentials. Evidence suggests that regions which lag behind suffer from a uniform productivity gap, with very little role for the specific sectoral specialization into activities with high or low productivity per worker. These results are robust to various tests and in particular to changes in the number of sectors, the set of countries used, and the chosen years.

There are several implications of these results that we wish to underline. First a technical one: the decomposition of the traditional shift-share productivity component into a proper productivity differential component and an allocation component, as proposed by Esteban [1972], is empirically meaningful. We obtain an excellent fit for the net productivity differential component, while the allocation component appears as an uncorrelated perturbation. Second, our results suggest that nothing essential is lost if the analysis of regional convergence is explicitly or implicitly carried out with one-sector models, instead of multisector models. Finally, and from a policy point of view, our results indicate that in order to close the gap between advanced and retarded regions development policies should focus on factors affecting uniformly the productivity of backward regions. Our empirical evidence seems to back the pertinence of present EU regional policies based on structural funds essentially geared to improve on infrastructures and human capital.

References

Abraham, F. and P. van Rompuy (1995), Regional Convergence in the European Monetary Union, Papers in Regional Science 74, 125.142.

Armstrong, H. W. (1995), Convergence among Regions of the European Union, 1950-1990, Papers in Regional Science, 74, 143-152.

Armstrong, H.W. and R.W. Vickerman (eds.) (1995), Convergence and Divergence Among European Regions, Pion, London.

Barro, R.J. and X. Sala-i-Martin (1991), Convergence across States and Regions, Brookings Papers on Economic Activity 1:1991, 107-182.

Browne, L.E. [1989], Shifting Regional Fortunes: the Wheel Turns, New England Economic Review, Federal Reserve Bank of Boston.

Caminal, R., J. Esteban, A. de la Fuente and X. Vives, [1994] Crecimiento y Convergencia Regional en España y Europa, 2 vols, Fundación de Economía Analítica, Barcelona, 1994.

Canova, F. and A. Marcet [1995], The Poor Stay Poor: non-convergence across countries and regions, CEPR discussion paper # 1265.

Carlino, G.A. [1992], Are Regional Per Capita Earnings Diverging? Business Review, Federal Reserve Bank of Philadelphia, 3-12.

de la Fuente, A. and X. Vives [1995], Infrastructure and education as instruments of regional policy: evidence from Spain, Economic Policy 20, 13-51.

de la Fuente, A. [1994], Crecimiento y Convergencia, in Caminal et al. [1994] vol. 2, pp. 199-247.

Duro, J.A. and J. Esteban [1998], Factor Decomposition of Cross-Country Income Inequality, 1960-1990 Economics Letters, 60, 269-275.

Dunn, E.S. [1960], A Statistical and Analytical Technique for Regional Anlysis, Papers and Proceedings of the Regional Science Association 6, 97-112.

Esteban, J. [1972], A Reinterpretation of Shift-Share Analysis, Regional and Urban Economics 2, 249-261.

Esteban, J. [1994],"La Desigualdad Interregional en Europa y en España: descripción y análisis" in Caminal et al. [1994] vol 2, pp. 13-84.

García-Milà, T. and T. McGuire [1993], Industrial Mix as a Factor in the Growth and Variability of States' Economies, Regional Science and Urban Economics 23, 229-241.

García-Milà, T. and R. Marimon [1996], Integración Regional e Inversión Pública en España, in R. Marimon (ed), La Economía Española: una perspectiva diferente, Antoni Bosch editor, Barcelona, 197-256.

Marcet A. [1994], Los pobres siguen siendo pobres: convergencia entre regiones y países, un análisis bayesiano de datos de panel, in Caminal et al. [1994] vol.2 , pp. 249-270.

Molle, W. and S. Boeckhout (1995), Economic Disparity under Conditions of Integration. A Long Term View of the European Case, Papers in Regional Science 74, 105-123.

Neven, D. and C. Gouyette (1995), Regional Convergence in the European Community, Journal of Common Market Studies 33, 47-65.

Paci, R. (1997), More Similar and Less Equal. Economic Growth in the European Regions, Weltwirtschaftliches Archiv 133, 609-34.

Quah, D. (1993), Empirical Cross-Section Dynamics in Economic Growth, European Economic Review 37, 426-434.

Quah, D. (1994), Empirics for Economic Growth and Convergence, European Economic Review 40, 1353-75.