CONVERGENCE AMONG THE SPANISH REGIONS: 
AN INference-BASED STOCHASTIC DOMINANCE APPROACH

Ismael Ahamdanech Zarco
Instituto Nacional de Estadística (I.N.E.)

Carmelo García Pérez
Universidad de Alcalá

Mercedes Prieto Alaiz
Universidad de Valladolid

Abstract

The traditional analysis of economic convergence between countries or regions is usually performed by comparing distribution means, such as per capita income. However, this kind of analysis, which is intimately related to the economic welfare of a society, presents only a partial approach to the economic convergence phenomena, given that the disparities within regions or countries are not considered. The empirical methodology used in this paper complements the traditional convergence approach, adding the aspects of efficiency and inequality in income distribution. Using first and second stochastic dominance, together with Lorenz dominance, the convergence between Spanish Regions from 1990 to 2003 is studied by means of per capita income data taken from the Spanish Household Budget Survey (EPF) of 1990-1991 and the Survey on Income and Living Conditions (SILC) of 2003.

Key Words: Convergence, Welfare, Stochastic dominance, Spain.

JEL Classification: D31, D63, I32
1. Introduction

The reduction of economic welfare disparities between regions is one of the most important objectives of the European Union. The difficulty of constructing, not only a common market, but a future political union between regions with huge differences in living standards, is widely accepted. According to this, a large share of the European Union budget is dedicated to trying to reduce the income gap between regions by means of different funds, such as the structural funds. In this context, a question arises: are really effective the structural funds in reducing economic welfare disparities between regions? This is the question we try to answer in this paper. To this end, there are some important issues that deserve to be remarked upon.

First of all, most studies that analyze economic convergence are based on the study of certain macroeconomic variables, such as average income, on the basis that average income evolution is a good approach to studying the trends of the level of economic welfare in the society (Sala i Martin, 1992; Quah, 1986). However, it is commonly accepted that a reduction of income inequality increases economic welfare: consider, for instance, the Sen Index of welfare (Sen, 1976) or the generalized Lorenz curve and how it relates inequality to welfare (Shorrocks, 1983). According to this, if we try to study the convergence process between regions we should take into account not only the income level evolution, but also the income inequality over time. This is an important fact which we consider in this paper by using the stochastic dominance approach: while rank dominance (first order dominance) only takes into account efficiency aspects, second order dominance, or Generalized Lorenz Dominance, also considers inequality aspects. Since we study both first and second order dominance, we introduce inequality into the analysis.

On the other hand, to study stochastic dominance it is necessary to use micro data from a sample survey. In our case, we employ income data from the Spanish Household Budget Survey (EPF) of 1990-1991 and the Survey on Income and Living Conditions (SILC) of 2003. However, using sample data we can commit sampling errors. We take into account this possibility and overcome the problem by using inference techniques developed by Beach and Davidson (1983), Bishop et al. (1992) and Davidson and Duclos (2000), among others. These techniques allow us to make more precise comparisons between the distributions analysed.
The paper is organised as follows. In section 2, we study the literature referred to the convergence process in Spain over the last three decades. Section 3 is dedicated to reviewing the methodology we use, while in section 4 we explain the data and some methodological decisions. In section five we present the main results obtained, and the conclusions are left to the last section.

2. Review of the Literature on Economic Convergence in Spain

There are many analyses on convergence in Spain, most of them focused on specific macroeconomic variables such as per capita income. Fuente (1996) and Goerlich and Mas (2001) have reviewed this literature. We can distinguish between two approaches regarding the methodology used to study the convergence hypothesis: on the one hand, there are works focused on the Barro and Sala i Martín (1991, 1992) approach; on the other hand, we find analyses that use the methodology proposed by Quah (1996).

Analyses based on the Barro and Sala i Martín methodology try to compare, from different econometric models, whether the poor regions have a higher growth than the rich ones. Moreover, these studies analyze whether the per capita income of the different regions tends to the same value in the long term (absolute beta convergence) or to unequal steady states (conditional beta convergence). Generally, this kind of analysis also studies whether the relative dispersion of average regional income distribution decreases over time (sigma convergence). Raymond and García (1994), Dolado, González-Paramo and Roldán (1994) and Mas et al. (1994) reach similar conclusions using different variables in the analysis: while there is convergence until the end of the 1970s, from then on this process stopped. More recently, María-Dolores and Solanes (2002) research the convergence in the Gross Added Value per employee during the period 1955-1997. Their results are in line with other works, although they estimate a slower convergence velocity (1.78% per year). Goerlich and Mas (2002) study the evolution of income in the Spanish regions between 1995-2000 from three different points of view. On the one hand, they analyze the geographic location of the economic activity, reaching the conclusion that most of the activity is located in the North-East, Madrid and the islands (Balearic and Canaries). On the other hand, they study convergence in per capita income and productivity, confirming the reductions of the differences until the late seventies. Finally, the paper shows that inequality has decreased.
Another approach for studying the convergence between regions is that proposed by Quah (1996). The author remarks on the importance of analyzing, on the one hand, the outside shape of the density function, and on the other, the regions’ movements inside the distribution. The first branch of analysis would imply the estimation of density functions using parametric or non-parametric methods. Most of the works in this approach use a methodology based on the kernel estimators of the density functions. The hypothesis of convergence would imply a density function with a dispersion decreasing over time until it converges at a point with probability 1. Gardeazabal (1996), Lamo (2000) and Tortosa et al. (2005) analyze the convergence between the provinces and Spanish regions with the Quah methodology. All these works confirm that there has been convergence from the middle of the last century to the end of 1990s, although more intense until the middle seventies.

The works we have reviewed are based only on efficiency criteria, but there is another research approach based not only on efficiency criteria but also on equality, using microeconomic data from household surveys.

One of the earliest works that considers both efficiency and inequality criteria is that by Callealta, Casas and Núñez (1996). The authors carry out an exhaustive analysis of the inequality level with different indexes, and they analyze the welfare level through generalized Lorenz curves. One of the most remarkable conclusions reached in this work is that there were almost no variations in the regions’ relative positions with respect to the general welfare level of the nation as a whole.

In this line, we can situate the work of Goerlich and Mas (2002) who try to complement the analysis of some macroeconomic variables with the study of the inequality and economic welfare level of the different regions. They achieve a negative correlation between the inequality level and per capita income, i.e., richer regions show less inequality. Another conclusion they reach is that the economic welfare level is, above all, determined by per capita income.

Ayala, Jurado and Pedraja (2006) analyze inequality with the Gini index and different indexes of the Atkinson and generalized entropy family. They reach similar conclusions to those we have seen before: while some convergence would appear to exist, there are few changes in the relative positions of the Spanish regions.
Villar (2006) bases the economic welfare analysis on the welfare index proposed by Sen (1976). In this work, one of the main conclusions is that some regions that were not originally in a very good position (Castile and Leon, Extremadura, Andalusia and Murcia) have increased their welfare level; thus, regional differences have decreased, but not in a big way.

3. Theoretical Approaches for Studying Convergence

As we have mentioned, research on economic convergence has usually been focused on some macroeconomic variables, such as per capita income. This approach is based on the relationship between income and economic welfare. If it is true, convergence will be obtained when all the regions have the same per capita income, which would imply that all of them had reached the same economic welfare.

The distribution of this variable, which we will call “income” from now on, can be represented in a population with \( n \) individuals, divided in \( H \) regions, as:

\[
\frac{x_{11}, \ldots, x_{n_1}}{\text{Region 1}}, \ldots, \frac{x_{1h}, \ldots, x_{n_h}}{\text{Region } h}, \ldots, \frac{x_{1H}, \ldots, x_{n_H}}{\text{Region } H}
\]

where \( x_{ih} \) is the income of the \( i \)-th individual in the region, \( h \)-th and \( n_h \) is the population of region \( h \)-th. Following Milanovic (2006) we can analyze the differences between regions or countries using three different points of view.

On the one hand, we can study the dispersion of the average income of the different regions, i.e., we can analyze the dispersion of the distribution:

\[ \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_H \]

Beta and sigma convergence with non-weighted indicators or the Quah (1996) methodology uses this approach from the national accounts data.

On the other hand, we can analyze disparities on the assumption that all the individuals from the same region have the same income, which would be the average income of the region. This implies the analysis of the dispersion of the distribution:

\[ \frac{\bar{x}_1}{n_1}, \ldots, \frac{\bar{x}_j}{n_j}, \ldots, \frac{\bar{x}_h}{n_h}, \ldots, \frac{\bar{x}_H}{n_H} \]

This approach is equivalent to the sigma convergence analysis with weighted indicators, using, as above, data from the national accounts.

Both approaches described have the same problem: they do not take into consideration the disparities inside the region. For this reason, the third way, proposed
by Milanovic, for studying convergence is based on the analysis of the disparities not only between regions but also inside them. The works based on this approach use household surveys instead of national accounts data.

The methodology we apply in this paper is in accordance with the third kind of works described by Milanovic, and it is based on the stochastic dominance concept, which we review in the next section.

3.1 First and Second Order Stochastic Dominance and Lorenz Dominance

Let $F^a$ denote the cumulative income distribution function of region $A$ and $X^a(p) = \inf \{ x : F^a(x) \geq p \}$, $p \in [0,1]$ its inverse function or quantile function. Thus, region $A$ rank dominates region $B$ ($X^a >_r X^b$) if and only if $X^a(p) \geq X^b(p)$ for every $p \in [0,1]$ with at least one strict inequality.

Saposnik (1981, 1983) establishes the relationship between rank dominance and economic welfare, proving the following theorem:

$$X^a >_r X^b \iff W(X^a) > W(X^b), \forall w \in W_p,$$

where $W_p$ denotes the class of anonymous and increasing welfare functions, i.e., the class of functions according to the assumptions of the Pareto principle and anonymity.

We can define the Generalized Lorenz curve (Shorrocks, 1983) as the Lorenz curve scaled by the income mean. Then, we can say that the region $A$ second order dominates region $B$ ($X^a >_g X^b$) if and only if $G^a(p) \geq G^b(p)$ with at least one strict inequality, where $G^a(p)$ represents the Generalized Lorenz curve ordinates for $A$ (equivalently for $B$). In this context, Shorrocks (1983) proves that:

$$X^a >_g X^b \iff S(X^a) \geq S(X^b), \forall w \in W_s,$$

where $W_s$ is the class of welfare functions increasing and S-concave$^1$.

Bishop, Formby and Thistle (1991) contribute to clarifying the relationship between first and second order dominance, proving that: $X^a >_g X^b$ implies $X^a >_r X^b$.

Finally, as is well known, Lorenz dominance has inequality implications. If we note the ordinates of the Lorenz curve as $L^a(p)$ and $L^b(p)$ for regions $A$ and $B$,

---

$^1$ Dasgupta, Sen and Starret (1973) show that S-concavity is enough to introduce the Pigou-Dalton principle.
respectively, then region \( A \) Lorenz dominates region \( B \) if and only if \( L^*(p) \geq L^*(p) \) with at least one inequality prevailing. If this happens, inequality will be higher in region \( B \).

### 3.2 Stochastic Dominance and Economic Convergence

We have seen that if region \( A \) rank dominates (or second order dominates) region \( B \) we can say that the economic welfare is higher in \( A \). And even more important, we can say that by assuming a few, widely, though not universally\(^2\), accepted assumptions (only the anonymous Pareto principle in first order dominance; anonymous Pareto principle and Pigou-Dalton principle in second order dominance).

In this context, following Bishop, Formby and Thistle (1992) we can say that there has been complete economic convergence when there are no significant differences in ordinates of inverse function or generalized Lorenz curves for the regions analyzed. On the other hand, two regions are not converging when the number of statistically different ordinates decreases over time. With the terminology we have adopted, we can say that there is complete convergence if, from a situation of rank dominance, we reach another one where \( X^*(p) = X^*(p) \) for every \( p \in [0,1] \). If there is no convergence using first order dominance, we can analyze convergence by adding the Pigou-Dalton principle and there will be complete convergence if \( G^*(p) = G^*(p) \) for every \( p \) (in both cases, the regions will be converging if the number of inequalities decreases over time).

### 3.3 Stochastic Dominance and Statistical Inference

When we work with sample data we are exposed to committing sampling errors, and some crossing between curves or some differences between ordinates may not be statistically significant, so, using inference, we can rank some distributions which we could not by using only a descriptive statistics approach.

In this context, some statistical tests can be used to study whether the differences between the ordinates of the curves are significant. Beach and Davidson (1983) derive the joint variance-covariance structure of generalized Lorenz ordinates (\( \Pi \)). These authors show that the Generalized Lorenz ordinates vector \( \hat{G} = (\hat{G}_1, \hat{G}_2, ..., \hat{\mu}) \) is asymptotically normal because \( \sqrt{n}(\hat{G} - G) \) has a limiting \( K+1 \) variable normal distribution, where \( n \) is the sample size.

\(^2\) Consider, for example, Rawl’s approach.
Since the Lorenz curve ordinates can be written as a linear transformation of the ordinates of the generalized Lorenz curve, $\hat{\mathcal{L}}_i = \hat{G}_i / \mu$, where $\mu$ is the income mean, using $\Pi$, it is possible to obtain the joint variance-covariance matrix of the Lorenz curve ordinates. In this case, $\sqrt{n}(\hat{L} - L)$ has a limiting normal multivariate distribution, too. Using these distributions, Bishop, Formby and Thistle (1989) introduce a pair-wise statistical inference test to compare two generalised Lorenz or Lorenz ordinates.

For the generalized Lorenz ordinates, the null and alternative hypotheses are:

$$H_{0,i} : G_i^a = G_i^b \quad \text{and} \quad H_{A,i} : G_i^a \neq G_i^b \quad \text{for each} \quad i = 1, 2, \ldots, K+1 \quad (1)$$

where $G_i^a$ and $G_i^b$ are the generalized Lorenz ordinates for each $i$ of the income vectors $A$ and $B$.

The statistical test for equality of the $i^{th}$ elements of the vectors $G^a$ and $G^b$ will be:

$$T_{GL,i} = \frac{\hat{G}_i^b - \hat{G}_i^a}{\left( \hat{\sigma}_ii^a \right)^{1/2} + \left( \hat{\sigma}_ii^b \right)^{1/2}} \quad \text{for} \quad i=1,2,\ldots,K.$$ 

Where $\hat{\sigma}$ is the estimation of the elements of $\Pi$, whose formula is given by Beach and Davidson (1983). Under the null hypothesis $T_{GL,i}$ is asymptotically normal. The critical values for this test are determined by the Student Maximum Modulus distribution, which accounts for the correlation between the variables. If the null hypothesis is not rejected, we cannot rank. But if we reject the overall null hypothesis, there are three possible outcomes:

a) weak generalised Lorenz dominance: if for some quantiles $G_i^a > G_i^b$ and for other quantiles $G_i^a = G_i^b$.

b) strong generalised Lorenz dominance: if for all $i$ $G_i^a > G_i^b$.

c) the Lorenz generalised curves cross if for some quantiles $G_i^a > G_i^b$ and for other quantiles $G_i^a < G_i^b$. In this case, we cannot compare the welfare associated with the distributions $X^a$ and $X^b$ using the second-order dominance approach.

---

3 Obtained from the tables for the percentiles provided by Stoline and Ury (1979).
To extend these results to Lorenz dominance is relatively easy since the asymptotic distribution is, as we have said, similar. The null and alternative hypotheses are the same, merely changing $G_i$ by $L_i$. The statistical test will be:

$$T_{Li} = \frac{\hat{L}_i^a - \hat{L}_i^b}{\left[ \left( \hat{\sigma}^a_i \right) + \left( \hat{\sigma}^b_i \right) \right]^{1/2}}$$

for $i=1,2,\ldots,K$.

where $\hat{\nu}$ is the estimation of the elements of the joint variance-covariance matrix of the Lorenz ordinates distribution. The critical values are obtained as in the previous test, and the implications are the same.

Finally, the statistical test for rank dominance can be seen as an extension of the test for second order dominance. Beach et al. (1994) develop the joint variance-covariance matrix of quantile ordinates. The vector of $K$ sample quantile means $\hat{\mu} = (\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_K)'$ is a linear transformation of $\hat{G}$, then $\hat{\mu}$ is also asymptotically normal (Rao, 1973). Beach et al. (1994) illustrated the procedure for a specific choice of quantiles such as deciles ($K=10$). Given the variances, we can test the $k$ sub-hypotheses using the statistic:

$$T_{GLi} = \frac{\hat{\mu}_i^a - \hat{\mu}_i^b}{\left( \text{Var} (\hat{\mu}_i^a) \right)^{1/2}}$$

for $i = 1,2,\ldots,K$.

The critical values for this test are also determined by the Student Maximum Modulus distribution.

To analyze the robustness of the results we extend the study by using the Davidson and Duclos (2000) test. This test fixes some points $x_1, x_2, \ldots, x_k$ (the same for the two distributions compared) and considers the following statistics:

$$\hat{D}_y(x) = \frac{1}{N(s-1)!} \sum_{i=1}^{N} (x - y_i)^{s-1}$$

$$\hat{D}_z(x) = \frac{1}{N(s-1)!} \sum_{i=1}^{N} (x - z_i)^{s-1}$$

---

4 The formula is also obtained in Beach and Davidson (1983).
where \( y \) and \( z \) represent the distributions compared, \( N = n_y + n_z \) and the \( u_+ = \max(u,0) \). Notice that for rank dominance \( s = 1 \) and for generalized Lorenz dominance \( s = 2 \).

The variances are:

\[
\hat{V}_y^s(x) = \left( \frac{1}{N((s-1)!^2)} \right) \sum_{i=1}^{N} (x - y_i)^{2(s-1)} - \hat{D}_y^s(x)^2
\]

\[
\hat{V}_z^s(x) = \left( \frac{1}{N((s-1)!^2)} \right) \sum_{i=1}^{N} (x - z_i)^{2(s-1)} - \hat{D}_z^s(x)^2
\]

Davidson and Duclos (2000) propose the following normalized statistic:

\[
T^s(x) = \frac{\hat{D}_y^s(x) - \hat{D}_z^s(x)}{\sqrt{\hat{V}_y^s(x)}}
\]

where, if we have unpaired observations independently drawn from two populations:

\[
\hat{V}_y^s(x) = \hat{V}_y^s(x) + \hat{V}_z^s(x)
\]

Davidson and Duclos show that under the null hypothesis \( H_0 : D_y^s(x) = D_z^s(x) \), \( T^s(x) \) is asymptotically distributed as a standard normal variate.

### 4. Empirical Analysis

#### 4.1. Data and Methodological Decisions

As we have observed, regional convergence analysis tries to study the discrepancies in the welfare level between different regions. In this context, the use of microeconomic data becomes very important, especially in the Spanish case, where there was a lack of disaggregated information that allowed us to obtain reliable regional results. In this sense, the appearance of the SIL in 2004 was an important advance since it is an annual survey and it will let us carry out regional analysis. In this work, we have used data from the EPF 1990-1991, with data referred to 1990, and the SILC for 2004, with data referred to 2003.

The interest variable is the disposable income per household, including total household income, adding transfers and deducing taxes and Social Security contributions. Our approach to this variable has been the household total income corrected by the total number of individuals of the household (we will call this variable
income or per capita income). The analysis in per capita terms seeks to introduce the analysis in the literature on convergence.

Income data proceeding from the files of the two surveys has been weighted using the weights given by the sample design and the number of individuals in the households.

Finally, 2003 data has been deflated by the price indexes of each region and expressed in 1990 constant euros.

4.2. Results: Is there Convergence?

We have considered three levels of analysis. On the one hand, we have compared those regions which have been receiving structural funds from the EU (that is, those with less than the 75% of the mean per capita income of the EU) with the regions that did not receive those funds. On the other hand, we have compared three groups of regions: those that will continue to receive structural funds after 2007, those that never received funds and those that will stop receiving funds gradually (phasing regions). Finally, to be more exhaustive, we have studied the convergence between regions individually.

4.2.1 Convergence between Regions Objective 1 and non-Objective 1.

Since we have studied the convergence for the period 1991-2003 we have selected the regions that were objective 1 in 1990. These regions were Andalusia, Asturias, Cantabria, Canary Islands, Castile-Leon, Castile-La Mancha, the Community of Valencia, Extremadura, Galicia and the Region of Murcia. Thus, we have two groups to compare: the regions that received structural funds and those that did not.

The first conclusion that should be remarked upon is that there has not been convergence, at least statistically significant, using both rank dominance and second order dominance. Figure 1 plots the quantile functions by decile for rich regions (those that did not perceive structural funds) and objective 1 regions for the period 1990 and 2003. As we can see, in both periods the function for rich regions is above the one for objective 1 regions. This is confirmed by looking at Table 1, where it is shown that the differences between rich and objective 1 regions remain statistically significant at a 1% level in 2003, exactly the same as in 1990, for all the deciles, using the Beach et al. (1994) test.
Table 1. Rich regions and Objective 1 regions: Conditional mean income, by decile, 1990 and 2003

<table>
<thead>
<tr>
<th>Decile</th>
<th>1990</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 1</td>
<td>Group 2</td>
</tr>
<tr>
<td>1</td>
<td>1609,837</td>
<td>1080,384</td>
</tr>
<tr>
<td></td>
<td>(23,937)</td>
<td>(12,339)</td>
</tr>
<tr>
<td>2</td>
<td>2349,752</td>
<td>1702,137</td>
</tr>
<tr>
<td></td>
<td>(21,358)</td>
<td>(11,692)</td>
</tr>
<tr>
<td>3</td>
<td>2830,026</td>
<td>2085,290</td>
</tr>
<tr>
<td></td>
<td>(25,255)</td>
<td>(11,897)</td>
</tr>
<tr>
<td>4</td>
<td>3292,488</td>
<td>2428,559</td>
</tr>
<tr>
<td></td>
<td>(26,965)</td>
<td>(12,747)</td>
</tr>
<tr>
<td>5</td>
<td>3717,289</td>
<td>2773,551</td>
</tr>
<tr>
<td></td>
<td>(25,970)</td>
<td>(14,878)</td>
</tr>
<tr>
<td>6</td>
<td>4186,097</td>
<td>3164,219</td>
</tr>
<tr>
<td></td>
<td>(31,512)</td>
<td>(15,387)</td>
</tr>
<tr>
<td>7</td>
<td>4742,434</td>
<td>3604,826</td>
</tr>
<tr>
<td></td>
<td>(38,875)</td>
<td>(18,818)</td>
</tr>
<tr>
<td>8</td>
<td>5516,887</td>
<td>4186,640</td>
</tr>
<tr>
<td></td>
<td>(49,855)</td>
<td>(23,140)</td>
</tr>
<tr>
<td>9</td>
<td>6821,899</td>
<td>5093,475</td>
</tr>
<tr>
<td></td>
<td>(72,915)</td>
<td>(33,956)</td>
</tr>
<tr>
<td>10</td>
<td>11306,260</td>
<td>8433,287</td>
</tr>
<tr>
<td></td>
<td>(337,191)</td>
<td>(146,567)</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses.

Figure 1. Rich regions and Objective 1 regions: Conditional mean income, by decile, 1990 and 2003

The picture is very similar when we consider second order dominance. In this case, Figure 2 shows that the Generalized Lorenz curve for rich regions lies above the one for objective 1 regions at every decile, while in Table 2 we can see that the differences between ordinates remain significant at the 1% level in 2003 for all the deciles, using the Beach and Davidson (1983) test.
Table 2. Rich Regions and Objective 1 Regions Generalized Lorenz Ordinates, by Decile, 1990 and 2003

<table>
<thead>
<tr>
<th>Decile</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Test statistic</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>160.909</td>
<td>107.692</td>
<td>19.76***</td>
<td>168.027</td>
<td>114.284</td>
<td>13.05***</td>
</tr>
<tr>
<td></td>
<td>(2.394)</td>
<td>(1.234)</td>
<td></td>
<td>(3.427)</td>
<td>(2.284)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>395.884</td>
<td>277.906</td>
<td>25.30***</td>
<td>443.857</td>
<td>309.085</td>
<td>18.65***</td>
</tr>
<tr>
<td></td>
<td>(4.114)</td>
<td>(2.195)</td>
<td></td>
<td>(6.180)</td>
<td>(3.746)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>678.886</td>
<td>486.435</td>
<td>27.96***</td>
<td>784.586</td>
<td>553.926</td>
<td>22.34***</td>
</tr>
<tr>
<td></td>
<td>(6.125)</td>
<td>(3.142)</td>
<td></td>
<td>(8.767)</td>
<td>(5.454)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1008.135</td>
<td>729.291</td>
<td>30.17***</td>
<td>1181.027</td>
<td>847.008</td>
<td>25.12***</td>
</tr>
<tr>
<td></td>
<td>(8.263)</td>
<td>(4.139)</td>
<td></td>
<td>(11.034)</td>
<td>(7.416)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1379.864</td>
<td>1006.646</td>
<td>32.34***</td>
<td>1629.567</td>
<td>1189.198</td>
<td>26.48***</td>
</tr>
<tr>
<td></td>
<td>(10.261)</td>
<td>(5.282)</td>
<td></td>
<td>(13.828)</td>
<td>(9.244)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1798.474</td>
<td>1323.068</td>
<td>33.50***</td>
<td>2139.611</td>
<td>1578.078</td>
<td>27.62***</td>
</tr>
<tr>
<td>7</td>
<td>2272.717</td>
<td>1683.550</td>
<td>33.92***</td>
<td>2716.881</td>
<td>2024.870</td>
<td>28.62***</td>
</tr>
<tr>
<td></td>
<td>(15.505)</td>
<td>(7.834)</td>
<td></td>
<td>(20.000)</td>
<td>(13.589)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2824.406</td>
<td>2102.214</td>
<td>33.89***</td>
<td>3373.336</td>
<td>2542.968</td>
<td>28.87***</td>
</tr>
<tr>
<td></td>
<td>(19.084)</td>
<td>(9.487)</td>
<td></td>
<td>(23.735)</td>
<td>(16.254)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3506.596</td>
<td>2611.562</td>
<td>33.27***</td>
<td>4165.657</td>
<td>3171.219</td>
<td>28.51***</td>
</tr>
<tr>
<td>10</td>
<td>4637.222</td>
<td>3454.891</td>
<td>23.29***</td>
<td>5393.694</td>
<td>4175.087</td>
<td>23.83***</td>
</tr>
<tr>
<td></td>
<td>(46.157)</td>
<td>(21.118)</td>
<td></td>
<td>(41.777)</td>
<td>(29.498)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses.  
*** Statistically significant at the 1% level.

Figure 2. Rich Regions and Objective 1 Regions: Generalized Lorenz Ordinates, by Decile, 1990 and 2003

The situation becomes different when we analyze inequality. In this case, we can see (Figure 3) that, while the Lorenz curve was almost equal for both groups of regions in 1990, in 2003 Lorenz curve for rich regions is above the one for objective 1 regions, which implies that inequality was higher in those regions in 2003. This is confirmed in Table 3. In this table we can see that in 1990 there was statistically significant Lorenz
dominance only for the first three first (and for the third only at 5% significance level). However, in 2003, the differences are statistically significant at a 1% level for all the deciles except the last one (5% level of significance). That is to say, inequality has increased in the period analyzed in objective 1 regions, with respect to rich regions.

### Table 3: Rich Regions and Objective 1 Regions Lorenz Ordinates, by Decile, 1990 and 2003

<table>
<thead>
<tr>
<th>Decile</th>
<th>1990 Group 1</th>
<th>1990 Group 2</th>
<th>Test statistic</th>
<th>2003 Group 1</th>
<th>2003 Group 2</th>
<th>Test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.035</td>
<td>0.031</td>
<td>5.37***</td>
<td>0.031</td>
<td>0.027</td>
<td>4.76***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.085</td>
<td>0.080</td>
<td>4.08***</td>
<td>0.082</td>
<td>0.074</td>
<td>6.23***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.146</td>
<td>0.141</td>
<td>3.08**</td>
<td>0.145</td>
<td>0.133</td>
<td>6.99***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.217</td>
<td>0.211</td>
<td>2.56</td>
<td>0.219</td>
<td>0.203</td>
<td>6.99***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.298</td>
<td>0.291</td>
<td>1.97</td>
<td>0.302</td>
<td>0.285</td>
<td>6.25***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.388</td>
<td>0.383</td>
<td>1.26</td>
<td>0.397</td>
<td>0.378</td>
<td>5.84***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.490</td>
<td>0.487</td>
<td>0.61</td>
<td>0.504</td>
<td>0.485</td>
<td>5.19***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.002)</td>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.609</td>
<td>0.608</td>
<td>0.11</td>
<td>0.625</td>
<td>0.609</td>
<td>4.17***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.003)</td>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.756</td>
<td>0.756</td>
<td>0.05</td>
<td>0.772</td>
<td>0.760</td>
<td>3.21**</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.003)</td>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* Standard errors are in parentheses.

***Statistically significant at the 1% level.

### Figure 3: Rich Regions and Objective 1 Regions Lorenz Ordinates, by Decile, 1990 and 2003

The results of first and second order dominance using the Davidson-Duclos test are almost the same. The fixed points $x_1, x_2, \ldots, x_K$ we have considered in order to
calculate statistical tests were the income deciles of Spanish general income distribution in 1990 and 2003. The differences are significant at 1% level in all the deciles in both first and second order dominance. Again, although the statistical significance remains, the test value decreases in all the ordinates except the very highest in second order dominance. This pattern could mean a smooth trend to convergence, and confirm the robustness of the results.

4.2.2 Convergence with Three Groups

As we have said, there is a group of regions which will stop receiving funds gradually from 2007, those we will call phasing regions. It would be interesting to study whether these regions have effectively converged if we analyze the whole distribution and not just a measure such as per capita income, and if we take into account other variables that affect economic welfare, such as inequality. Moreover, we will consider the evolution of these regions compared with what we have called rich regions (Madrid, Navarre, Basque Country, Catalonia, La Rioja, Aragon, Cantabria and Balearic Islands) and those that remain objective 1, or convergence objective regions using UE Cohesion Policy 2007-2013 terminology.

In this group we include statistical phasing-out regions and statistical phasing-in regions. Statistical phasing-out regions are those whose GDP per capita is above the 75% of the average EU GDP per capita as a result of the statistical effect of the last EU enlargement (Asturias and Murcia). Statistical phasing-in regions include those regions that would have moved from objective 1 by natural effect, that is to say, by surpassing the 75% of the average EU GDP per capita (Community of Valencia, Canary Islands and Castile and Leon).

Galicia, Castile-La Mancha, Andalusia and Extremadura, whose GDP per inhabitant is less than 75% of the Community average.
Figure 4 shows the quantile function for these three groups of regions. As we can see there are differences in both the 1990 and 2003 periods in these functions, but it seems that difference between rich and phasing regions decreased in 2003. However, Table 4 shows that differences remain statistically significant at the 1% level, although it is interesting to observe that the statistics for all deciles decreases, which would imply a (non-statistically significant) reduction of differences and some convergence.

The situation is similar when we compare rich regions with those that will remain objective 1 after 2007, although in this case the statistic decreases for all the deciles except the last one, which increases.

The results are almost the same if we compare phasing regions with those that will continue to period receive structural funds. The differences are statistically significant in every decile but in the very highest one in both periods. However, there is a reduction in the value of the statistics for all the deciles (again except the last one), although this reduction does not imply statistically significance changes.

The picture is quite similar when we use the second order dominance approach.

As we can see in Figure 5, differences in the Generalized Lorenz curve seems to have decreased between rich and phasing regions in the period analyzed. This is the same conclusion we reach when looking at Table 5, if we take into account the value of the statistics, which have decreased for all the deciles. However, the differences remain statistically significant at the 1% level for all the deciles.
If we compare phasing regions and objective 1 after 2007, we reach the same conclusions: there are some reductions in statistical values but it does not manage to change the statistical significance (the difference is significant at the 1% level in every decile in both periods).

Again the situation changes when we analyze inequality with the Lorenz dominance. Figure 6 shows that while the Lorenz curves for the three groups of regions were very similar in 1990, there has been an inverse process towards convergence in inequality in the period studied, since differences appear in some deciles.

![Figure 6. Rich Regions, Phasing Regions and Convergence Regions: Lorenz Ordinates, by Decile, 1990 and 2003](image)

The differences between Lorenz curves are confirmed by Table 6. In this Table we can see that while there were no statistically significant differences between the Lorenz curve deciles of rich and phasing regions in 1990, in 2003 there were statistically significant differences for the five lowest deciles. In accordance with this, we can say that inequality has increased in phasing regions with respect to rich regions, in what we can call a process of divergence.

Regarding differences between phasing regions and objective 1 regions after 2007, there were some changes in 2003 with respect to 1990. Phasing regions Lorenz dominates (weak dominance) in both periods. However, while in 1990 the statistically different deciles are the first five, in 2003 the significance changes to the deciles four to nine.
4.2.3 Convergence Considering the Regions Individually

The analysis of the Spanish regions individually shows the same pattern as the one we have seen above, i.e., although there is an approximation among some regions, the conclusion is a lack of convergence. Hesse diagrams (Figure 8) represent the ranking for second order dominance\(^7\) in 1990 and 2003. As we can see, there are two groups of regions in both periods: one that dominates Spain and another dominated by the whole country, and there have been few movements between these two groups.

If we study this pattern more carefully (See, Ahamdanech, Garcia and Prieto 2007) the conclusion is similar. The few changes that deserve to be mentioned are the following: Cantabria was dominated by Spain in 1990 and dominated Spain in 2003, improving its relative economic welfare. On the other hand, Castile and Leon worsened its relative economic welfare: there was no significant difference with Spain in 1990 while in 2003 it was dominated by the country.

![Figure 7: Second order dominance ranking (1990 and 2003).](image)

5. Conclusions

Throughout this work we have studied the convergence between the Spanish regions that have received structural funds and those that have not in the last 20 years.

\(^7\) Ranking for first order dominance is very similar in 1990 and identical in 2003.
In order to carry out this analysis, we have used stochastic dominance techniques, above all, for two reasons: on the one hand, it is a powerful tool in the sense that it incorporates few, widely accepted assumptions to rank the economic welfare. On the other hand, its use allows us to extend the convergence analysis to the whole distribution. Moreover, with the use of statistical inference, we extend the validity of the results since we take into account the possibility of the existence of sampling errors.

Indeed, we have used two different tests to extend the robustness of the results. And, as we have seen, the outcome is similar when using either of the two.

We have had the possibility of using two microeconomic data surveys, something that has allowed us to use the previous techniques. In this sense, it will be interesting to have more SILC waves at our disposal, since it will allow the researchers an analysis by regions. Nevertheless, we have used two periods of time sufficiently separated so as to study the effects of the structural funds over a long time.

The principal conclusion we have reached is a lack of convergence between the two groups of regions, at least if we just take into account the significance of the results. However, if we look at the value of the different tests, we observe that there seems to be some convergence. This is confirmed both by the Bishop, Formby and Thistle approach and the Davidson-Duclos one. Moreover, we can see an improvement of the phasing regions, i.e., of those regions that will stop receiving funds gradually from 2007. All this allows us to be optimistic.

Finally, although the conclusion is that statistically speaking there is no convergence, we have to point out that this does not mean a failure in the objectives of the structural fund. On the one hand, all the regions that have received these funds have improved their economic welfare (see Ahamdanech, Garcia and Prieto, 2007) if we compare distributions over time. On the other hand, it is necessary to remember that there were huge differences between regions in the first period studied, and that there seems to be a pattern of reduction of these differences, at least by comparing the tests statistics values for discrepancies between curve ordinates.

References


