40TH EUROPEAN CONGRESS

FURTHER EVIDENCE ON SPANISH REGIONAL CONVERGENCE (PER CAPITA INCOME VERSUS WELFARE, ARE THEY SO DIFFERENT?)

Francisco J. Goerlich and Matilde Mas

Universitat de València and Instituto Valenciano de Investigaciones Económicas

Address for correspondence

Francisco J. Goerlich

Departamento de Análisis Económico

Universitat de València

Avda. de los Naranjos, s/n. 46022 Valencia

Tel.: +34 96 382 82 53 / e-mail: Francisco.J.Goerlich@uv.es

A B S T R A C T

The paper offers new evidence on the convergence of the fifty provinces of Spain. Two variables are considered: per capita income, which is the one habitually used, and a cardinal index of welfare based on Sen (1974), which corrects the former by taking into account inequality in the personal distribution of income. It is centred on the cross-section evolution of distribution, increasing the statistical armoury by considering, as well as the simple variation coefficient, its weighted version, "box plots" representation, density functions (simple and weighted), and Lorenz curves.

There is currently an abundance of empirical evidence on the convergence of the regions and provinces of Spain. The origin of the new wave of studies is to be found in the contributions of Barro and Sala-i-Martin (1991, 1992, 1995) applied to data on the American economy; its member states ; samples of countries; and also of regions within each country or geographical area. The application of the concepts of β and σ convergence to the regions of Spain has generated a whole avalanche of studies, many of which can be found referenced in Goerlich and Mas (1998), and Cuadrado, Mancha and Garrido (1998). The great majority of them use simple dispersion statistics (the standard deviation of the logarithm, or the coefficient of variation) as a measurement of σ convergence, the only exceptions being Rabadán and Salas (1996), and Goerlich and Mas (1998), who also use weighted statistics, from the literature on inequality.

Dispersion statistics summarise in a single value all the information contained in cross-section data, and therefore do not consider the external form of the distribution, nor the changes that it undergoes across time. Quah (1993a, 1993b, 1996, 1997) has proposed the use of complementary statistics, which do capture the characteristics of all the distribution. Among them, we will consider two: the representation by "box plots", and density functions. Furthermore, borrowing concepts from the literature on inequality and welfare, we have also estimated Lorenz (1905) curves, closely linked to the measurement of the reduction in inequality that we are analysing.

The literature on convergence from a macroeconomic viewpoint has centred on the use of simple statistics, in which all observations have the same weight, irrespective of the size of the region or country they represent. The origin of this choice lies in the interest in the behaviour of geographical units, countries, regions, or provinces, which are assumed to share the same technology and fundamental parameters. However, when the size of the units under analysis is very unequal, it is also of interest to consider the information provided by weighted statistics (Goerlich and Mas (1998), Goerlich (2000)). For this reason, in this study we will consider both versions, simple and weighted, of one of the statistics most habitually used by the literature, the coefficient of variation.

The objections to using per capita income to measure the welfare enjoyed by a society are well known. Proposals for indices of welfare that take into account the personal distribution of income also abound, and there is much discussion of the properties that they must fulfil (Sen (1974)). For the national aggregate, and for the regions and provinces of Spain, there are currently available inequality indices (Gini (1912), Theil (1967), Atkinson (1970), and Mean Absolute Deviation) of the personal distribution of income, obtained on the basis of the information provided by the three Household Budget Surveys (*Encuestas de Presupuestos Familiares* - EPF) elaborated by the National Statistical Institute (*Instituto Nacional de Estadística* - INE) for our country.

The methodology for drawing up the indices can be found in Goerlich and Mas (1999), and is available on the Internet (http://www.ivie.es). The surveys were made during the periods 1973/74, 1980/81, and 1990/91, and are representative at the level of the fifty provinces of Spain[1]. This enables welfare indices to be constructed for the observations corresponding to the years 1973, 1981, and 1991. From the wide range available, we have selected the welfare index constructed on the basis of the Gini index, both for its popularity and for its relationship with the Lorenz curve, and for its ease of interpretation.

The paper is structured as follows. Section I summarises some methodological aspects of interest. Section II presents the descriptive statistics for the per capita GDP variable, covering the period 1973-1998. Section III compares the profiles followed by this variable and the welfare index of Sen (1974), the main conclusions being presented in section IV.

I. Methodological Aspects

Let us suppose that we have available n groupings of individuals whose per capita income we designate by x_i, $x_i = Y_i/N_i$, Y_i being the income and N_i the population of group $i = 1,2,...,n$. Also let p_i be the relative frequency, i.e., the percentage of population per grouping, $p_i = N_i/N$, $N = \Sigma_{i=1}^{n} N_i$; and y_i the proportion of income of grouping i, $y_i = Y_i/Y$, $Y = \Sigma_{i=1}^{n} Y_i$. The average per capita income for the aggregate can be expressed as a weighted arithmetic mean, $\mu = \frac{Y}{N} = \Sigma_{i=1}^{n} p_i x_i$.

It is well known that for this set of data, we can calculate a large number of dispersion statistics (Cowell (1995), Goerlich (1998, 2000)), each with different normative properties. These statistics will indicate the dispersion of the observations around a central value, normally the value of per capita income for the aggregate, μ. We consider that the indicator of dispersion must have certain basic properties: (i) it must be independent of the scale, i.e. it must remain unaltered if the income of each individual in the population (or the per capita income of each grouping) is altered in the same proportion (zero degree homogeneity of income); (ii) it must be independent of the size of the population, i.e. it must remain unaltered if the number of individuals at each level of income is altered in the same proportion (zero degree homogeneity of population); and (iii) it must satisfy the principle of transfers of Pigou (1912)-Dalton (1920), i.e. any transfer from a rich individual or province to a poorer one that does not invert their relative ranking must reduce the value of the index (Sen (1973)).

From among those indices that satisfy the above properties, we have selected the coefficient of variation, which is no more than the standard deviation divided by the average of the observations,

$$CV_{\omega}(x) = \frac{SD_{\omega}(x)}{\mu} \tag{1}$$

where $SD_{\omega}(x) = \sqrt{\Sigma_{i=1}^{n} p_i (x_i - \mu)^2}$. This statistic satisfies the three basic properties mentioned above[2].

Two observations are of interest. Firstly, the coefficient of variation defined in (1) is not the one habitually used in the literature, as it is a weighted statistic that takes into account the proportions of the population. However, the literature on macroeconomic convergence allocates the same weight to all observations, calculating the statistics in simple terms. The unweighted version of the indicator would be given by:

$$CV(x) = \frac{SD(x)}{\bar{x}} \tag{2}$$

where $\bar{x} = \frac{\Sigma_{i=1}^{n} x_i}{n}$ and $SD(x) = \sqrt{\frac{\Sigma_{i=1}^{n} (x_i - \bar{x})^2}{n}}$. Observe that this statistic centres the dispersion around an "incorrect" point, the simple arithmetic mean, $\bar{x}$, which does not coincide with the mean for the aggregate, μ, which is observable. Both statistics, (1) and (2), will be calculated to examine the influence of the weightings on the results.

Secondly, the interest in having the dispersion statistic satisfy the Pigou-Dalton principle of transfers makes it recommendable to rule out the most habitual dispersion statistic in the literature on growth and convergence, the standard deviation of logarithms, whether in its simple version, $SD(\log x) = \sqrt{\frac{\Sigma_{i=1}^{n}(\log x_i - \log \tilde{x})^2}{n}}$ or weighted version, $SD_{\omega}(\log x) = \sqrt{\Sigma_{i=1}^{n} p_i (\log x_i - \log \tilde{\mu})^2}$, where $\log \tilde{x} = \frac{\Sigma_{i=1}^{n} \log x_i}{n}$ and $\log \tilde{\mu} = \Sigma_{i=1}^{n} p_i \log x_i$. As accurately noted by Foster and Ok (1999), it is possible to find cases of practical relevance in which a reduction in the overall dispersion of the distribution, in the sense of Lorenz's (1905) dominance, is accompanied by increases in $SD(\log x)$, which calls into question the definition of σ-convergence in terms of this statistic (Barro and Sala-i-Martín (1995).

The coefficient of variation, like any other index, is a scalar numerical representation of a complex and multidimensional phenomenon. For this reason it is advisable to widen the horizon by examining additional characteristics of the data. Of the instruments available, we consider of interest the so-called "Box Plots", as they are a simple and graphic form of describing some characteristics of the data; the Lorenz curve; and density functions, as instruments that illustrate, from different points of view, the whole of the distribution. We describe these instruments below.

Box plots

A box-plot is merely a flat representation of some of the most outstanding characteristics of a set of data. Its main advantage is that, because it is a two-dimensional representation, several box-plots can be observed simultaneously on the same graph, allowing dynamic study of the evolution of some important characteristics of the distribution, e.g. existence, appearance or disappearance of outliers, dispersion or concentration of the data, or the symmetry or otherwise of the distribution.

A box-plot, with all its elements, can be examined in graph 1. The vertical axis represents the scale of the variable, in our case the normalised per capita income. In order to abstract from growth, it is best to normalise the variable for the average of the aggregate, so the variable in question is now $\frac{x_i}{\mu}$. This is equivalent to normalisation by $CV_{\omega}(x)$. The square, or box, represents the inter-quartile range. The 0.75 quartile, $\xi_{.75}$, forms the top of the square and the 0.25 quartile, $\xi_{.25}$, the bottom. By construction, 50% of the mass of probability of distribution is contained in the box. The height of the box therefore represents the inter-quartile range, which is a habitual measurement of dispersion in statistics. A greater inter-quartile range will be displayed as a box of greater height, indicating that 50% of the density of *x* is relatively scattered. On the other hand, a smaller inter-quartile range will be seen as a shorter box, and indicates that 50% of the density of *x* is relatively concentrated.

The horizontal line inside the box is the median, or 0.50 quartile, a measurement of position of the distribution of the variable. The positioning of this line in relation to the top or bottom of the box provides graphical information on the form of the distribution. If the median is not in the middle of the box, the distribution is asymmetrical. In the case of Graph 1 there is evidence of asymmetry to the left, i.e. towards the lower part of the distribution.

Two vertical lines appear at the top and bottom of the box. The ends of these lines, drawn horizontally, are known respectively as the upper and lower adjacent values. On the basis of the inter-quartile range $R(\xi_{.25})$, the upper adjacent value is defined as the observed value of the variable represented not greater than $\xi_{.75} + 1.5 \text{x} R(\xi_{.25})$, and the lower adjacent value as the observed value of the variable represented not less than $\xi_{.25} - 1.5 \text{x} R(\xi_{.25})$. The maximum longitude between adjacent values will be given by the interval, $[\xi_{.25} - 1.5 \text{x} R(\xi_{.25}) , \xi_{.75} + 1.5 \text{x} R(\xi_{.25})]$, but in general will be shorter. The reason is that, within this interval, the extreme observations are selected. The adjacent values are therefore order statistics that correspond to current observations of the variable, and cover the range of observations that are not considered to be outliers.

Finally, the observations situated beyond the adjacent values are the outliers, upper if they are greater than the upper adjacent value, and lower if they are lower than the lower adjacent value. These values are represented individually by small horizontal lines. Thus, in the example of graph 1 we can observe 3 upper outliers and 2 lower ones.

Adjacent values, then, play a dual role. On the one hand, they demarcate the range of observations that are not considered atypical. On the other, they indicate the distance between the extreme values of these observations and the outliers, enabling us to observe their distance or closeness to the majority of the distribution. It is possible that there are no outliers, so that the adjacent values will in fact be the extreme values of the set of observations, the maximum and/or the minimum of the distribution.

Box plots therefore summarise a great deal of information, and are useful fundamentally for two reasons: (i) for determination and evolution of outliers, and (ii) in the analysis of the dispersion or concentration of the distribution; more precisely, of the 50% of the density of probability associated with $R(\xi_{.25})$.

Lorenz Curves

The popularity of these curves makes detailed description of them unnecessary. Let us consider, without loss of generality, that the per capita income of the various provinces has been ordered non-decreasingly, $x_1 \leq x_2 \leq \ldots \leq x_n$, and let us call this ordered vector of per capita incomes $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$. Let us define Fs as the proportion of population that receives a per capita income equal to or less than x_s, $F_s = \Sigma_{i=1}^{s} p_i$, and Φ_s as the cumulative proportion of income corresponding to the level of per capita income x_s, $\Phi_s = \Sigma_{i=1}^{s} y_i$. That is to say, Φ_s is the proportion of income received by the groupings with a level of per capita income equal to or less than x_s. Defining $F_0 = \Phi_0 = 0$, the relationship between Φ_s and F_s is what is called the Lorenz curve.

Numerous indices attempt to summarise the graphic information provided by the Lorenz curve in a quantitative measurement that shows the divergence between this curve and the situation of perfect equality, defined by the diagonal. The most popular of these measurements, which will be used later, is the Gini index, G (Gini (1912)). From the geometrical point of view the Gini index is defined as twice the area between the Lorenz curve and the line of perfect equality. The Gini index varies, for continuous distributions, between 0, perfect equality, and 1, maximum inequality, and will be greater the further the Lorenz curve is from the line of perfect equality. However, from the computational viewpoint it is better to have available a formula to give us the above result. We can therefore define the Gini index as an average of the mean relative difference (Kendall and Stuart (1963)).

$$G = \frac{1}{2\mu} \sum_i \sum_j p_i p_j \left| x_i - x_j \right| \tag{3}$$

where $\sum_i$ is to be understood as $\Sigma_{i=1}^{n}$.

Observe that, as they have been defined, both the Gini index and the Lorenz curve are weighted statistics.

Density functions

The most suitable procedure for illustrating the characteristics of a set of data is, probably, to represent the distribution from which the observations come, $f(x)$. For example, if we know for certain that the data come from a normal distribution, $f(x) = \phi(x)$, it would be enough to know their average and their variance, drawing the normal density for the corresponding values of these parameters. In practice, however, the form of the distribution is not known for certain, but it can be estimated without imposing *a priori* any functional form. The procedure consists, intuitively, of smoothing a histogram in order to obtain a continuous representation (Silverman (1986)). For this, we assume that each observation, x_i, provides certain information about the density underlying the observations in a certain interval, such that at each point the estimated density is given by

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h} \right) \tag{4}$$

This estimator is known in the statistical literature as a kernel estimator, with the kernel represented by $K(\bullet)$ (Silverman (1986)).The construction of this estimator requires two choices: the kernel function, $K(\bullet)$, and the smoothing parameter, h, also called "window" or band width by some authors. We will not enter into technical details which can be found in Silverman (1986), Scott (1992) or Simonoff (1996). Suffice it to say that in practice the gaussian kernel is used.

$$K(s)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{s^2}{2}\right\} \quad (5)$$

which implies distributing the information content of x_i in accordance with this density of probability, and that the smoothing parameter was chosen optimally for this kernel (Silverman (1986)). Graph 2 shows graphically the construction of the estimator.

Thus the ordinates of the function, $y = f(x)$, are obtained as the sum of the heights of the superimposed individual densities.

The estimator defined in (4) constitutes a simple estimation, since it gives the same importance to each observation. The corresponding weighted estimator is obtained directly by replacing the simple sum by the corresponding weighted sum (DiNardo, Fortin and Lemieux (1996)),

$$\hat{f}_{\omega}(x)=\frac{1}{h}\sum_{i=1}^{n}p_i K\left(\frac{x-x_i}{h}\right) \quad (6)$$

which provides information not only on the distribution of the observations, but on the distribution in terms of the population underlying them.

In the empirical application carried out in the two following sections the density functions were estimated for the normalised variable, $\frac{x_i}{\mu}$.

Welfare

If we accept the assumptions necessary for making comparisons of cardinal welfare functions (Sen (1974)), it is then possible to derive indices of welfare from the average value of the per capita income, a measurement of central position of the distribution, and an index of inequality ranging from 0 to 1. Comparisons of welfare usually involve two characteristics of distribution: position and dispersion (Shorrocks (1983)).

Let us define the index of per capita welfare for a province as

$$w=\mu(1-G) \quad (7)$$

G being the Gini index. In this way, $W = Nw = Y(1-G)$ provides the value of the real aggregate income which, equally divided among the population, would generate the same level of well-being as the current real income, when inequality is measured by the Gini index. Clearly, since $0 \leq G \leq 1$, $0 \leq W \leq Y$.

To sum up, the above welfare indicator adjusts the per capita income of a province by an index that reflects the dispersion, or inequality, in the personal distribution of income in that province. Indices of quality other than G could potentially generate other results.

II. Convergence in per capita GDP

The habitual presentation of σ-convergence uses the simple coefficient of variation, or the standard deviation of the logarithm of the variable. For the reasons given above, which can be found in greater detail in Goerlich (2000), it has been chosen to consider the first of them in the presentation made in graph 3, which also shows the weighted version of this indicator. Although information is available from 1955 (Fundación BBV (various years)) and Sophinet (http://bancoreg.fbbv.es)), 1973 was chosen as the starting year in order to facilitate comparison with the welfare indices, as seen in section III.

The majority of authors who have analysed the convergence of the Spanish provinces (e.g. Mas, Maudos, Pérez and Uriel (1995)) have pointed out that convergence was especially intense until the mid-1970s, slowing down from then onwards. However, the visual impression given by graph 3 is that convergence continued in subsequent years, stagnating after 1993[3]. Only by observing the range of variation of the two variables can it be seen that the reduction was not very intense, above all if it is compared with that of the 1950s and 1960s. Consequently, as was easy to anticipate, the choice of period affects the conclusions that derive from the mere observation of the statistics of dispersion.

Graph 3 also indicates that, in this case, the use of simple or weighted dispersion statistics has no practical consequences, since both present similar profiles, especially up to the year 1993. We will return to this point later.

The coefficient of variation summarises in a single value the distance between the provinces, and the movements that are produced within the distribution across time. A first step in the analysis of the external form of the distribution is the presentation by means of box plots as described in section I. Graph 4 shows the box plots corresponding to four observations: 1973, 1981, 1991, and 1998. In this graph it can be seen that the reduction of inequality was very modest during these years. But the following aspects of interest can also be observed.

Firstly, among the Spanish provinces only one outlier is detected in these four years, the Balearic islands in the year 1973. Secondly, the distance between the upper adjacent value and the 0.75 quartile, which contains the thirteen provinces with highest per capita GDP, is appreciably greater than the distance corresponding to the lower adjacent value, where the thirteen poorest provinces are situated. This indicates that, as well as the outlier, some rich provinces are far ahead of the rest. Indeed, if we observe the upper limit of the inter-quartile range, we can see that in practice only these thirteen provinces, out of fifty, present values higher than the national average. The next twenty-four, contained in the range defined by the box, present a per capita GDP from 73% to 109% of the average depending on the years, while that of the thirteen poorest ranged from 54% to 77%.

As a consequence of the above, the median that divides the box is always lower than the national average, about 90%, though there was an increase between 1973 and 1998. This information indicates that a process of reduction of inequality, though not very intense, has taken place during these years, and this also appears to be confirmed by the reduction of the adjacent values, and in spite of the slight widening of the inter-quartile range. Also, the position of the median, always situated below the average value, points to a right-inclining asymmetrical distribution which will be analysed later.

The geographical location of the provinces belonging to the inter-quartile range, and the adjacent values, appear in maps 1 and 2. In the starting year, all the provinces adjoining the French frontier belonged to the group of the thirteen richest. They were joined by the outlier Balearics, Madrid, with the second highest level of per capita GDP and therefore defining the extreme of the upper adjacent value, Valencia and Valladolid. Within the box, which contains the twenty-four provinces contained between the first and third quartile, are provinces adjoining those above. The exceptions are the western provinces of Galicia and Andalusia which also enter into this group. Finally, the thirteen provinces contained in the lower adjacent value were situated in the south and west of the Peninsula.

Twenty-five years later some changes had occurred, but in general the thirteen poorest were still in the south and west of the peninsula, including all the Andalusian provinces with the exception of Almeria. Also the thirteen richest continued to be those closest to the French frontier (except Huesca), the Balearics and Madrid. In general, the impression given by these maps is one of mobility of the provinces with highest income levels towards the north and the Mediterranean east, leaving the poorest in the south and north-west.

While box plots present some interesting aspects of the distribution, density functions offer a complete characterisation. Graph 5 shows, in panel a), the simple density functions, and in panel b) the weighted ones, corresponding to the two extreme years, 1973 and 1998. The first impression is that now the distinction between simple and weighted statistics is indeed important, as the profiles of the two functions are very different.

The simple density functions that appear in panel a) of this graph confirm some of the information already provided by the box-plots. Indeed, the extremes of the distribution indicate its maximum and minimum values, and its shortening between those years is a first measurement of the slight reduction of inequalities that has occurred. This result is confirmed by the narrowing of the function around the average value. Also confirmed is the intuition provided by the box-plots that the distribution was asymmetrical towards the right, since it falls less steeply at that side.

However, the simple density function offers further information of interest. In 1973 there were three local maxima, two of them in the middle of the distribution, and the third around the extreme of the upper adjacent value. The latter is much lower than the others, indicating that a very small number of the thirteen provinces at the upper extreme of the distribution belonged to this "club", whereas in that of the poorest the number was higher, as indicated by the ordinate axis of graph 5. Twenty five years later there was no "club" of provinces, as the three "humps" had practically disappeared, leaving a single maximum, skewed slightly to the left of the median, in the area of 80% of the national per capita income. The distribution shows, however, an evident asymmetry.

The profile offered by panel b) of graph 3 is very different. The weighted density functions, which take into account the size of the population of the various provinces, indicate that in these years there were two clearly differentiated "clubs". The origin of this result lies in the fact that some of the richest provinces, fundamentally Madrid and Barcelona, are also those with the highest numbers of inhabitants. Though the existence of these "clubs" has been maintained over the twenty five years, the graph indicates that there has been a reduction of the gap between the two groups, as the weighted density functions have narrowed/moved closer together and the local maximum corresponding to the provinces with highest per capita GDP has moved to the left. However, the maximum corresponding to the "club" of the poorest has also moved to the left, indicating that in 1998 there were more inhabitants in provinces with per capita GDP below 80% of the national average than in 1973.

An overall view of what has happened between the two "clubs" is provided by the Lorenz curves that appear in graph 6. The usual presentation of these curves has been modified in order to make observation easier. Now, a greater proximity to the abscissa axis indicates less inequality. From 1973 to 1998 a reduction of inequality occurred in all divisions of the distribution, but not between 1981 and 1998, as the inhabitants of the provinces situated in the middle part of the distribution worsened during these years. Seen from another point of view, we could argue that there is a growing polarisation of the population into two groups of provinces, above average (about 130% of the national per capita income value) and below average (about 80%), with a gradual disappearance of the "middle class", understood as the population inhabiting the provinces with intermediate per capita income. In any case, the need to resort to this presentation also indicates that the reduction of inequality was very modest.

III. Per capita GDP and Welfare

If it is accepted that welfare is positively related to the level of income, and negatively related to inequality in the personal distribution of income, it is of interest to use an indicator that summarises the behaviour of both variables. As indicated in section I we use an index of welfare based on the Gini index (Sen (1974)). Considering this indicator, and referring to the variable total per capita expenditure, maps 3 and 4 locate the most unequal provinces in the west and south of the peninsula, geographically opposed to per capita GDP. Therefore, the poorest provinces are also those that present highest levels of inequality in the personal distribution of income.

The inter-province differences in the Gini index are less pronounced than in per capita GDP. For this reason, the modification of the last variable made by our indicator does not introduce significant differences. Thus, maps 5 and 6 offer a very similar geographical vision of welfare, though not identical, to that offered by maps 1 and 2 referring to per capita GDP. The provinces with highest indicators of welfare continue to be found in the north and east of the Peninsula.

Table 1 summarises the statistics of dispersion, coefficient of variation and max/min corresponding to the two variables, confirming the two results that we had already anticipated. Firstly, that in neither of two cases is the distinction between the simple and the weighted coefficient of variation important. And secondly, that the inter-province differences in per capita income are smaller than the differences in welfare, due to the negative correlation between inequality and level of per capita income. This result applies whether we use the coefficient of variation or the max/min range statistic.

However, while the coefficient of variation, weighted and simple, show a sustained reduction from 1973 to 1998, the max/min statistic indicates that, although per capita income continued the process of convergence throughout the period, welfare behaved conversely from 1981 to 1991. The origin of the divergence in those years must be found in the extremes of the distribution. Graph 7 confirms that the Lorenz curves cross at the top, reinforcing the picture already given by the Balearics' loss of leadership in 1981 as indicated by the table. This province experienced an atypical increase in the Gini index, so great that it placed it in second place in the ranking, behind the province of Lugo.

The box plot representations corresponding to the welfare variable (graph 8) and its comparison with that of per capita GDP (graph 4), offer similar views except in the following aspects. Firstly, in 1973, the Balearics disappear as an outlier. This is because of the downward correction experienced by this province when the personal distribution of income is taken into account. Secondly, the size of the box plots is greater in terms of welfare than of per capita income, confirming that the inter-province differences in the first variable are greater that those in the second. And thirdly, from 1973 to 1981 we can see a clear reduction of the dispersion of welfare, practically unnoticeable before, which again expands between 1981 and 1991. As we have already mentioned, this result has its origin in the province of the Balearics.

The relative similarity of the information offered by both variables is also reflected in the density functions that appear in graph 9. They all offer similar profiles to each other, but also show additional information. The greatest differences are found among the provinces contained in the inter-quartile range, while in the adjacent values they are practically superimposed. It is therefore the provinces with per capita income between 70% and the national average value that are most affected by the unequal distribution of income.

As was to be expected, the comparison of the Lorenz curves appearing in graph 10 confirms this result. The provinces of Spain are more unequal in terms of welfare than in terms of per capita income, especially in the centre of the distribution.

IV. Conclusions

The statistics which, like the coefficient of variation or the standard deviation of the logarithm, summarise in a single parameter all the information contained in the distribution, do not allow us to extract a large part of the wealth of information that it contains. This study has proposed expanding it with other statistics, box plots, density functions, and Lorenz curves, to reflect the external shape of the distribution. They all confirm that a slight reduction of inequality among provinces in per capita income occurred between 1973 and 1998. But the existence of convergent "clubs" is also detected.

In general, the poorest provinces are also the most unequal in terms of personal distribution of income, which helps to widen the inter-province differences in welfare. However, although the inequality of distribution is not the same in all the Spanish provinces, the differences in per capita income are more pronounced. For this reason, the profiles shown by the two variables are very similar, the greatest differences being observed in the middle reaches of the distribution.

Acknowledgements

This study was carried out in the framework of the DGICYT's project SEC98-0895. Francisco J. Goerlich wishes to acknowledge the financial assistance of the *Instituto Valenciano de Investigaciones Económicas* (IVIE).

Footnotes

---

[1] In the third quarter of 1997, the INE set in motion the *Encuesta Continua de Presupuestos Familiares (*ECPF$_{1997}$) [Continuous Survey of Household Budgets]. The new survey replaces the ECPF$_{1985}$ and also the EPF. It is representative at the scale of the seventeen regions, but not that of the provinces. For this reason, the latest information that will be available for the provinces is from 1991.

[2] The squared coefficient of variation is cardinally equivalent to the Theil index with parameter $\beta = 2$ (Cowell (1995)).

[3] It is possible that the increase in dispersion that took place in 1995 originated in statistical problems. The detailed inspection of the series of the Fundación BBV indicates that, in this year, some provinces presented growth rates much higher than the general trend.

References

Atkinson, A. B. (1970): "On the measurement of inequality", *Journal of Economic Theory*, 3, 244-263.

Barro, R. and X. Sala y Martín (1991): "Convergence across States and Regions", *Brooking Papers on Economic Activity*, 1 107-182.

__________ (1992): "Convergence", *Journal of Political Economy*, 100, 223-251.

__________ (1995): *Economic Growth*, McGraw-Hill.

Cowell, F. (1995): *Measuring Inequality*, 2nd Edition, LSE Handbooks in Economics, Prentice Hall, London. (1st. Edition 1977, Phillip Alan Publiserhs Limited, London).

Cuadrado Roura, J.R., Mancha, T. and R. Garrido: *Convergencia Regional en España. Hechos, Tendencias y Perspectivas*, Fundación Argentaria, Visor.

Dalton, H. (1920): "The measurement of inequality of income", *Economic Journal*, 30, 348-361.

DiNardo, J., Fortin, N. M. and Lemieux, T. (1996): "Labor market institutions and the distribution of wages, 1973-1992: A semi-parametric approach", *Econometrica*, 64, 5, September, 1001-1044.

Foster, J. E. and Ok, E. A. (1999): "Lorenz dominance and the variance of logarithms", *Econometrica*, 67, 4, July, 901-907.

Gini, C. (1912): "Variabilità e mutabilità, contributo allo studio delle distribuzioni e relazioni statistiche", *Studi Economico-Giuridici dell' Universiti di Cagliari*, 3, part 2, 1-158.

Goerlich, F.J. (1998): "Desigualdad, Diversidad y Convergencia: (Algunos) Instrumentos de Medida", Monograph, Instituto Valenciano de Investigaciones Económicas, December.

Goerlich, F.J. (2000): "Desigualdad, Diversidad y Convergencia: (Más) Instrumentos de Medida –Estadística Descriptiva-", Monograph, Instituto Valenciano de Investigaciones Económicas, March.

Goerlich, F.J. and M. Mas (1998): "Medición de las desigualdades: variables, indicadores y resultados", *Moneda y Crédito*, Segunda Época, 207, 59-86.

Goerlich, F.J. and M. Mas (1999): "Medición de la desigualdad: Contribución a una base de datos regional", Monograph, Instituto Valenciano de Investigaciones Económicas.

Kendall, M. G. and Stuart, A. (1963): *The Advanced Theory of Statistics. Volume 1: Distribution Theory.* 2ª Ed. Griffin, London.

Lorenz, M. C. (1905): "Methods for measuring the concentration of wealth", *Publications of the American Statistical Association*, 9, 209-219.

Mas, M., Maudos, J., Pérez, F., and E. Uriel (1995): "Growth and Convergence in the Spanish Provinces" in Amstrong, H. and R. Vickerman (eds), *Convergence and Divergence Among European Regions*, Pion.

Pigou, A. C. (1912): *The Economic of Welfare*, London. (Edited by MacMillan, New York in 1952).

Quah, D. (1993a): "Galton´s fallacy and test of the convergence hypothesis", *The Scandinavian Journal of Economics*, 95, 4, December, 427-443.

Quah, D. (1993b): "Empirical crosss-section dynamics in economic growth", *European Economic Review*, 37, 2/3, April, 426-434.

Quah, D. (1996): “Convergence empirics across economies with (some) capital mobility”, *Journal of Economic Growth*, 1, March, 95-124.

Quah, D. (1997): “Empirics for growth and distribution: Stratification, polarization, and convergence clubs”, *Journal of Economic Growth*, 2, March, 27-59.

Rabadán, I. and R. Salas (1996): “Convergencia y redistribución intertemporal en España: Efectos de los impuestos directos, cotizaciones sociales y transferencias”, *Economía Pública*, September, Fundación BBV.

Scott, D. W. (1992): *Multivariate Density Estimation: Theory, Practice, and Visualization*, John Wiley & Sons, New York.

Sen, A. (1973): *On Economic Inequality*, Oxford University Press, Oxford.

Sen, A. (1974): “Informational Bases of Alternative Welfare Approaches”, *Journal of Public Economics*, 3, 387-403, 1974.

Shorrocks, A. F. (1983): "Ranking income distributions", *Economica*, 50, 3-17.

Silverman, B. W. (1986): *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, Monographs on Statistics and Applied Probability nº 26, London.

Simonoff, J. S. (1996): *Smoothing Methods in Statistics*, Springer-Verlag, Berlin.

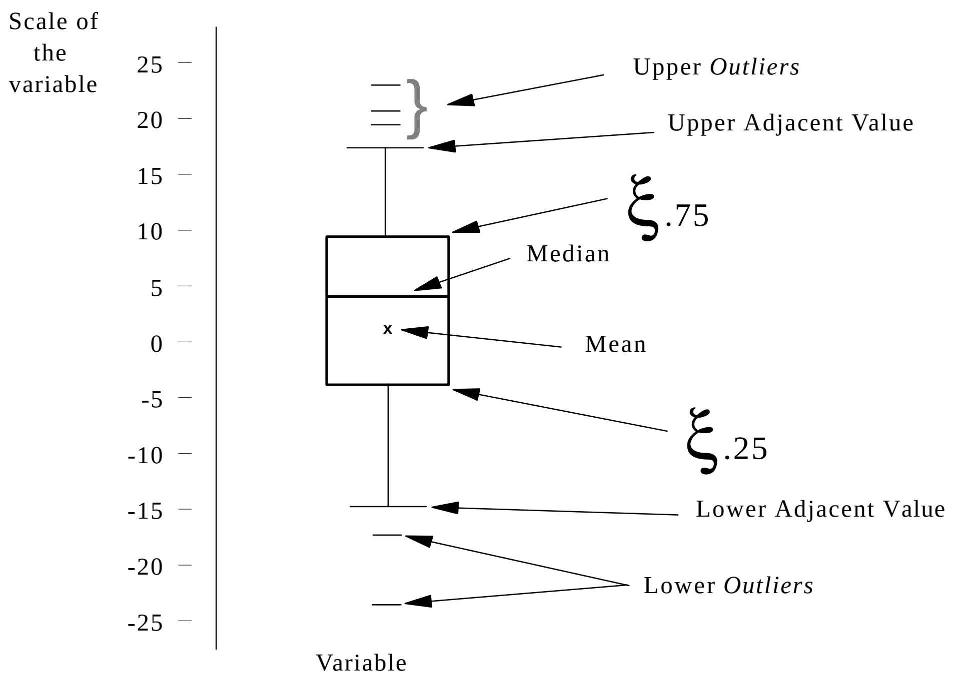
Theil, H. (1967): *Economics and Information Theory*, Amsterdam, North-Holland.

Table 1. PIB per capita and welfare. Dispersion statistics

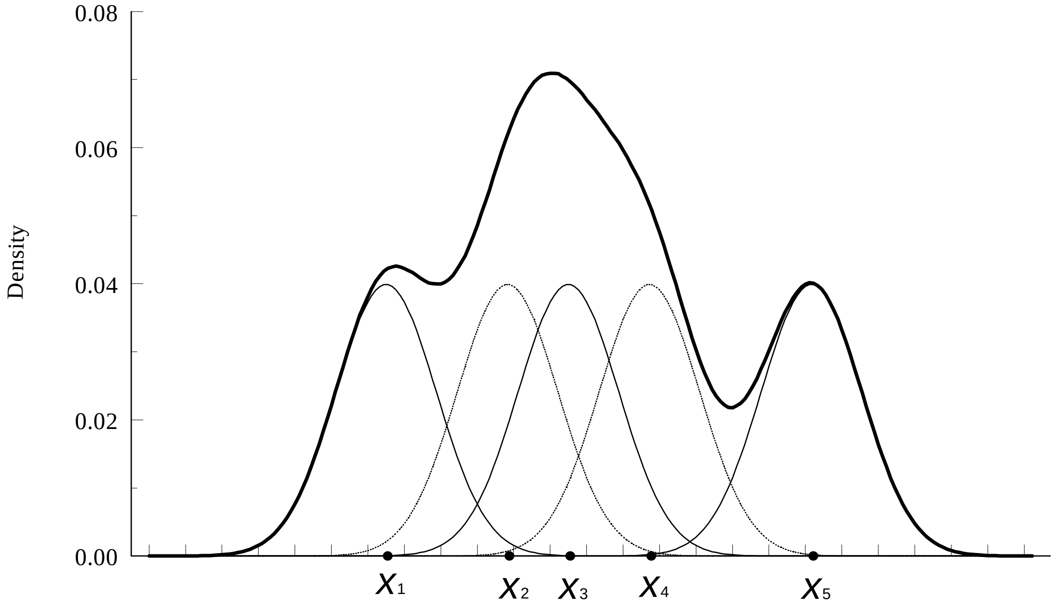
	PIB per cápita						Bienestar					
	Spain (thousand of constant pesetas)	Provincial Indices					Spain (thousand of constant pesetas)	Provincial Indices				
		Coefficient of variation	Weighted coefficient of variation	Maximum (normalized)*	Minimum (normalized)	Max/Min		Coefficient of variation	Weighted coefficient of variation	Maximum (normalized)*	Minimum (normalized)	Max/Min
1973	750.8	0.270	0.275	1.47 (Balears)	0.54 (Badajoz)	2.72	512,6	0.291	0.291	1.52 (Balears)	0.50 (Badajoz)	3.04
1981	807.1	0.242	0.235	1.48 (Balears)	0.58 (Badajoz)	2.55	550,3	0.259	0.250	1.40 (Balears)	0.55 (Badajoz)	2.54
1991	1,108.4	0.223	0.224	1.46 (Balears)	0.61 (Badajoz)	2.39	773,5	0.240	0.234	1.52 (Balears)	0.58 (Badajoz)	2.62

** Spain = 1*

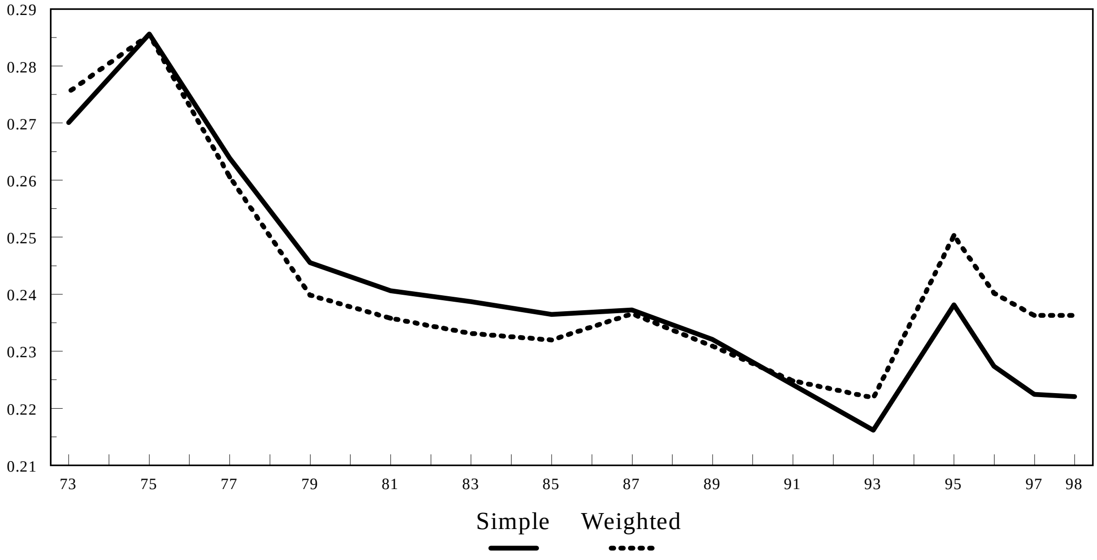
GRAPH 1. *Box-Plot*



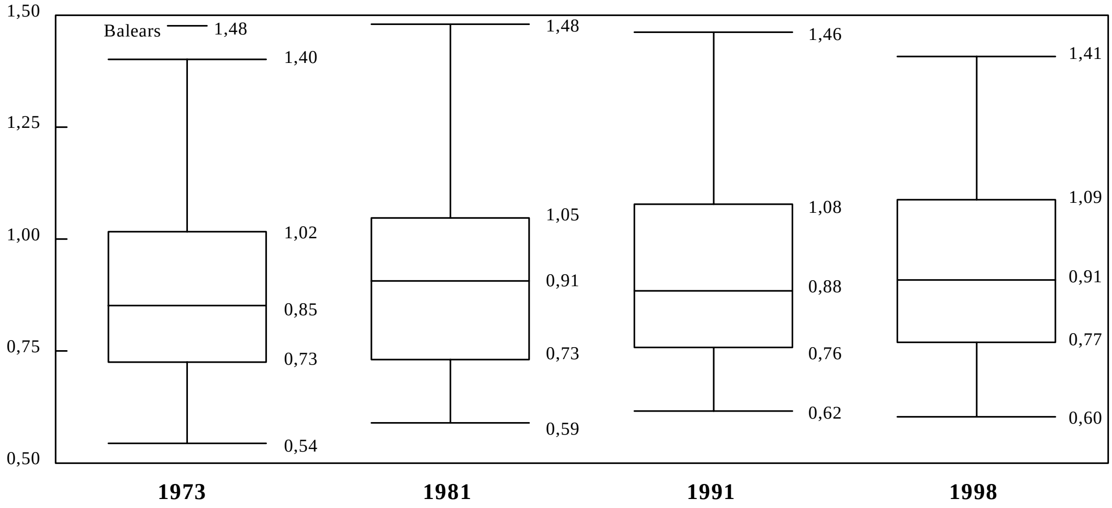
GRAPH 2. Density Estimation



GRAPH 3.
GDP per capita dispersion.
Coefficient of variation. Simple and weighted.

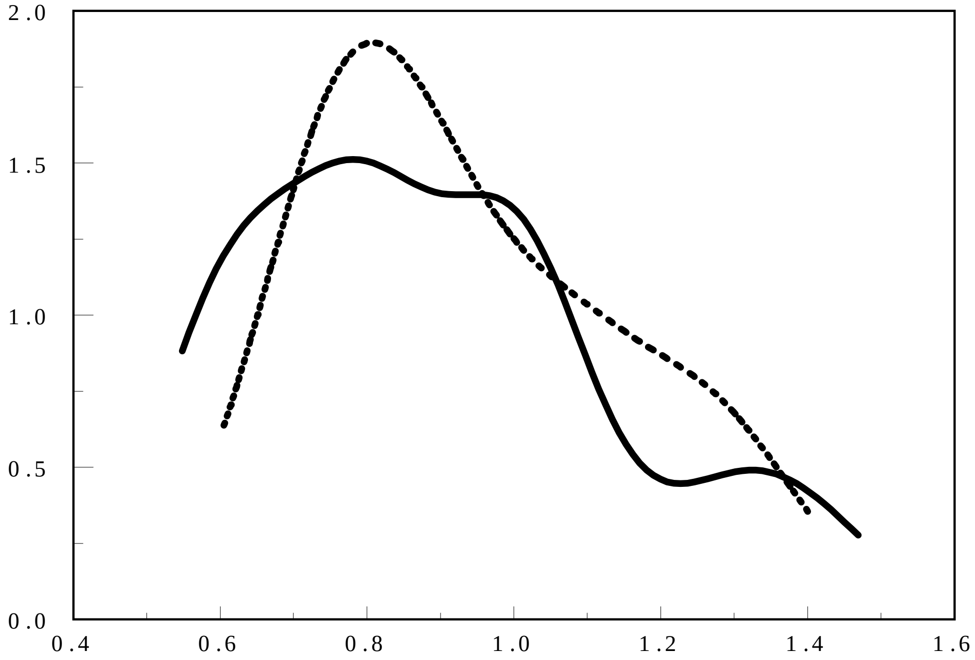


GRAPH 4.
GDP per capita. Box-Plots.

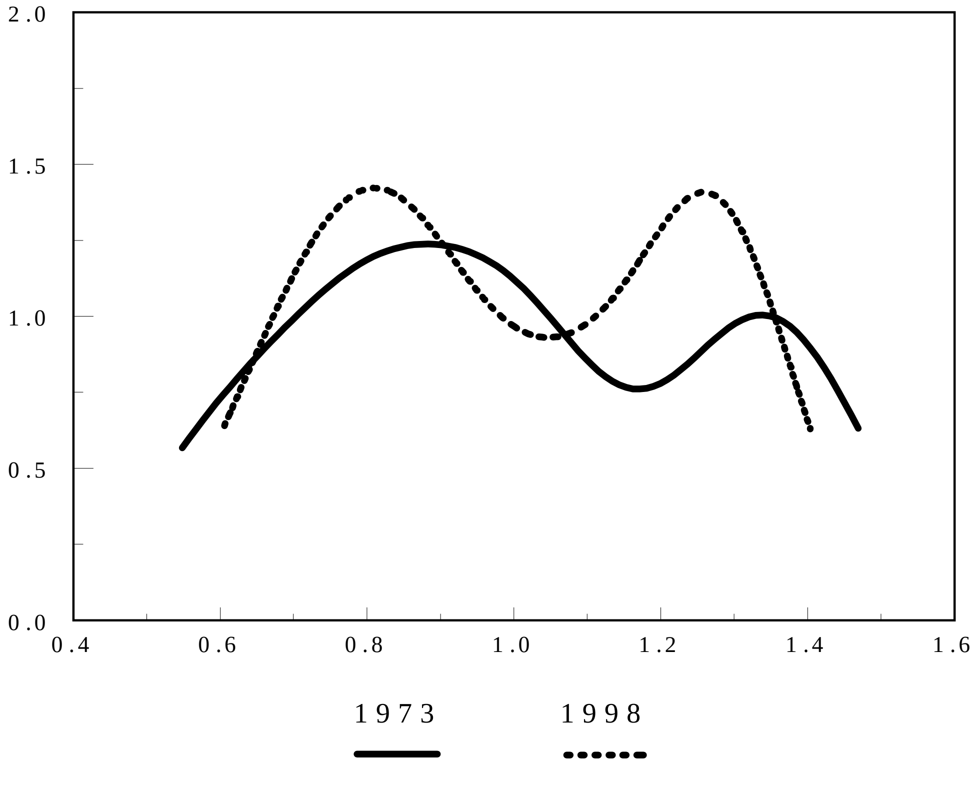


GRAPH 5.
Density functions.
GDP per capita.

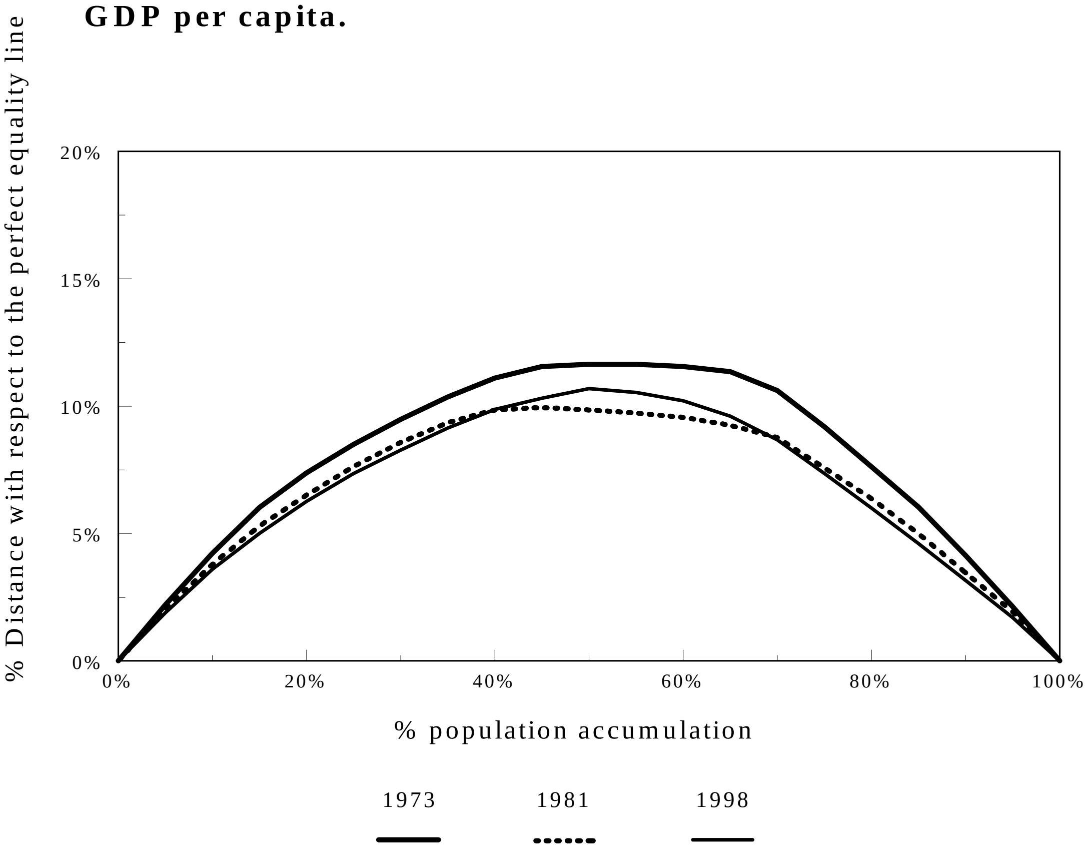
a) Simple



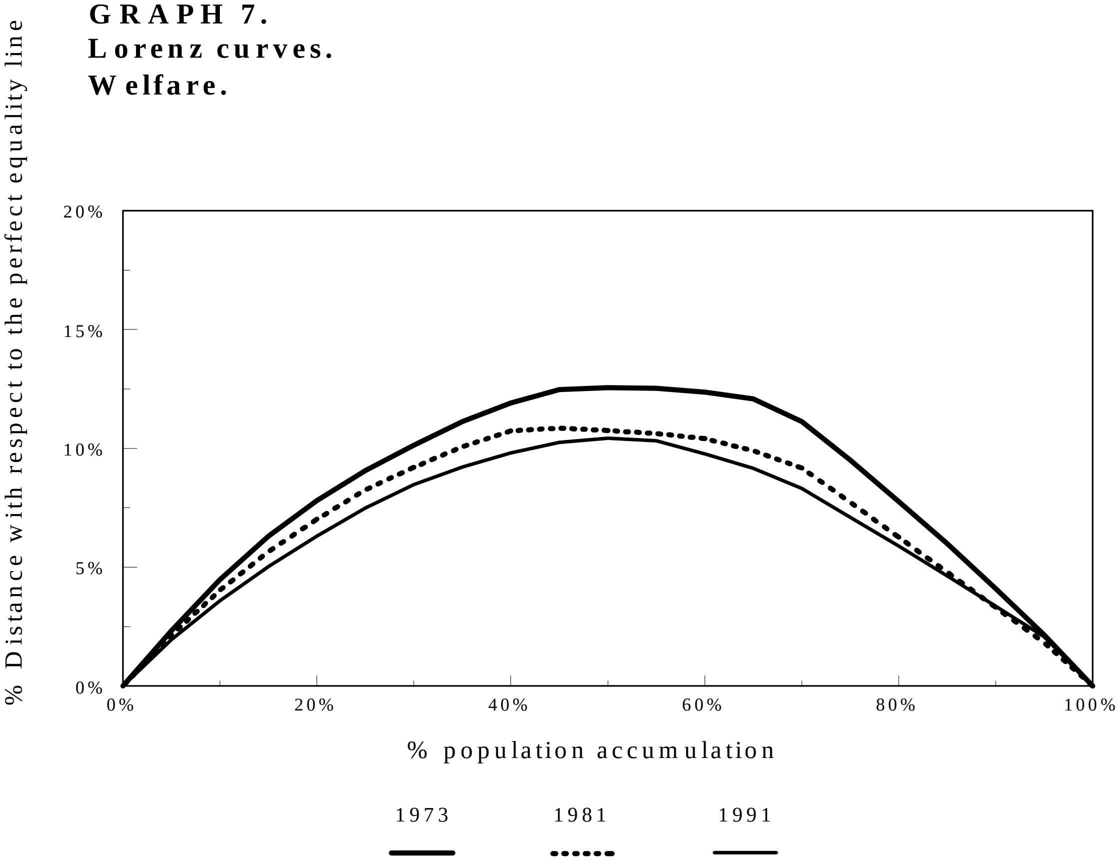
b) Weighted



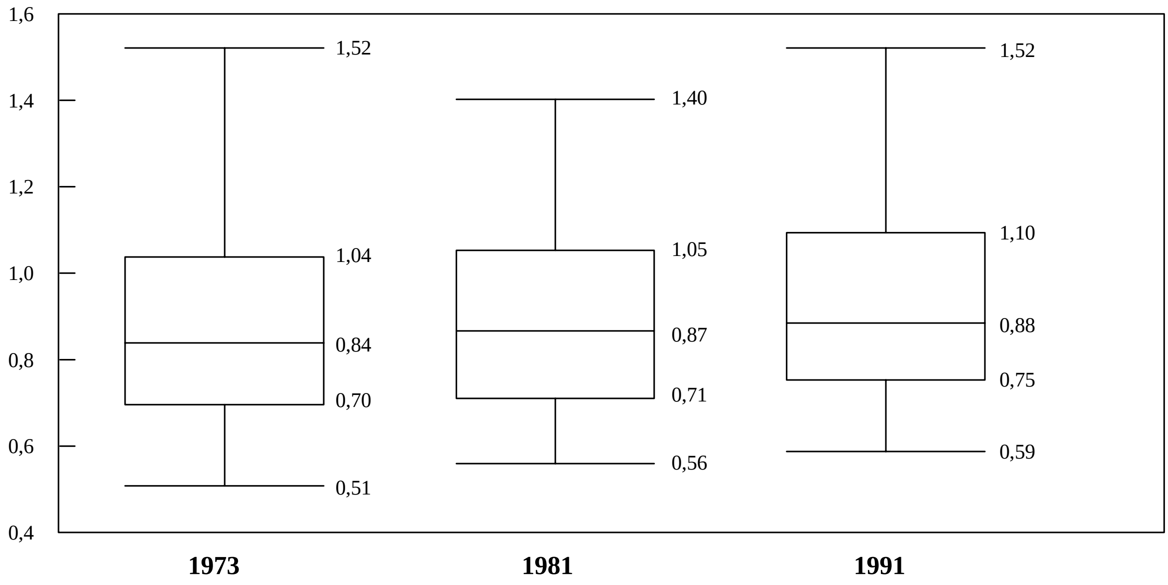
GRAPH 6.
Lorenz curves.
GDP per capita.



GRAPH 7.
Lorenz curves.
Welfare.

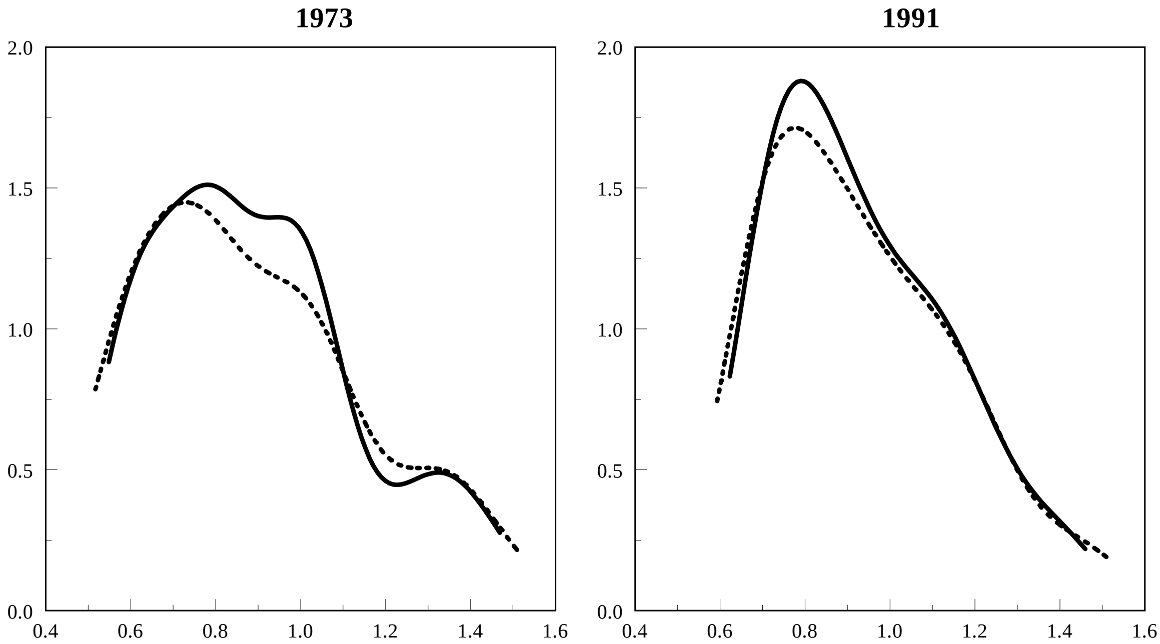


GRAPH 8.
Welfare. Box-Plots.

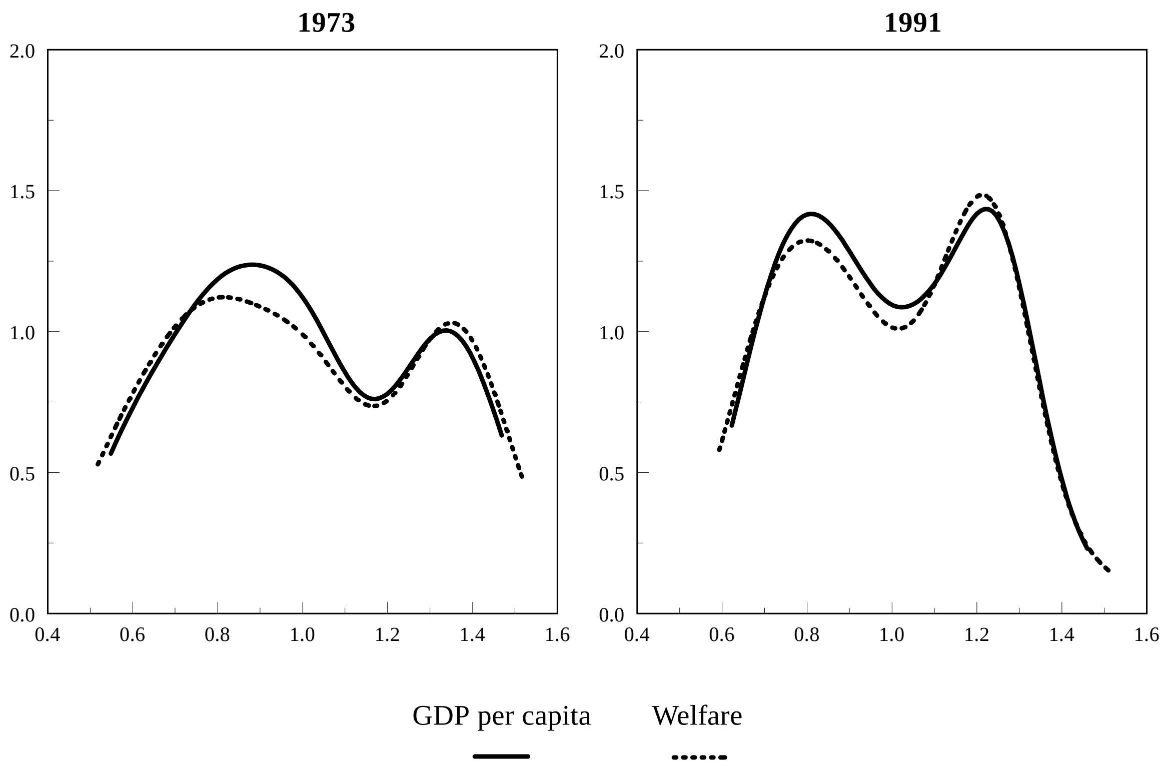


GRAPH 9.
Density functions.

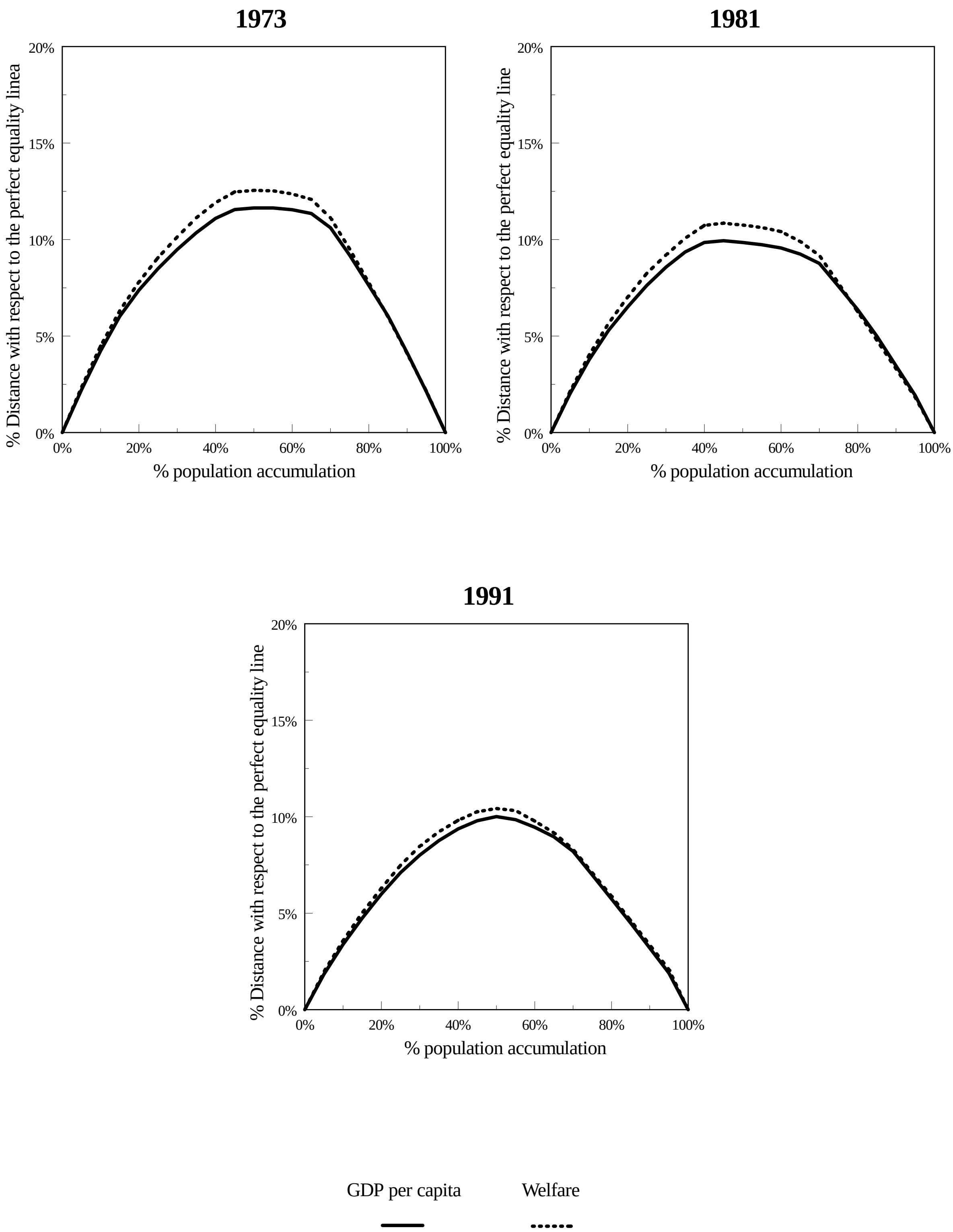
a) Simple



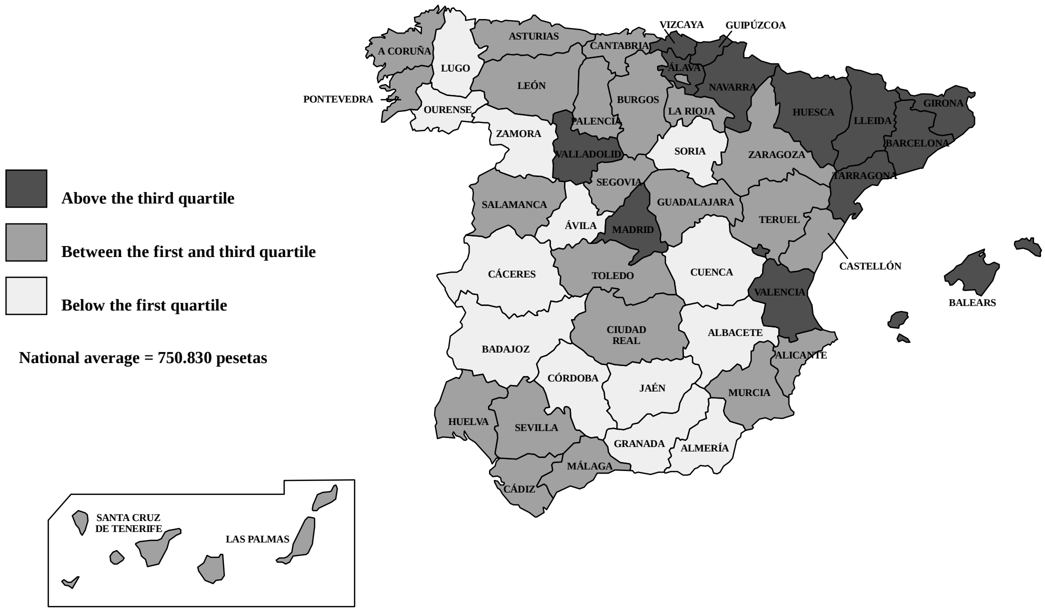
b) Weighted



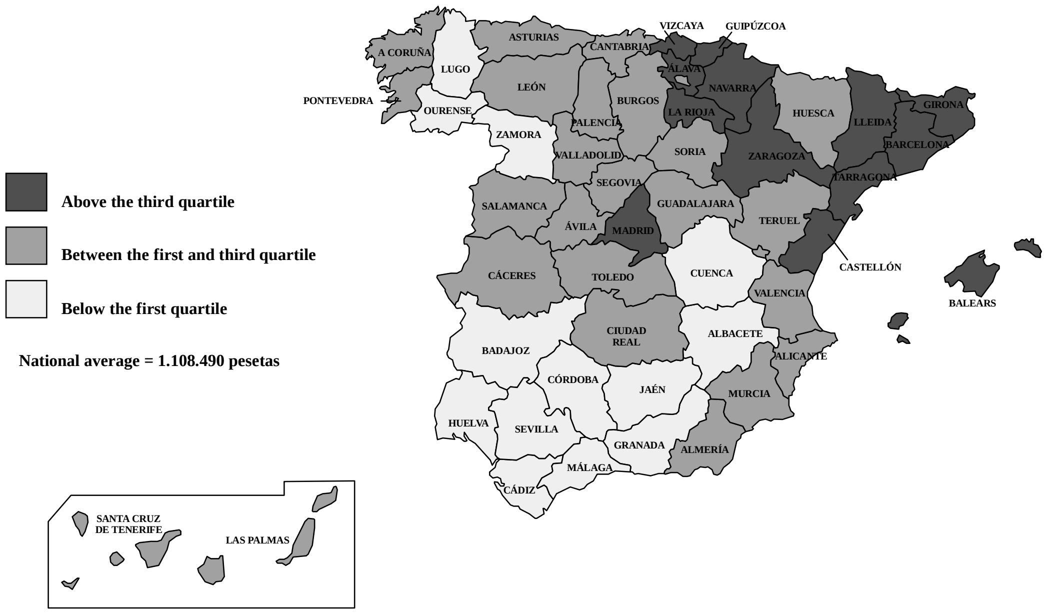
GRAPH 10.
Lorenz curves.



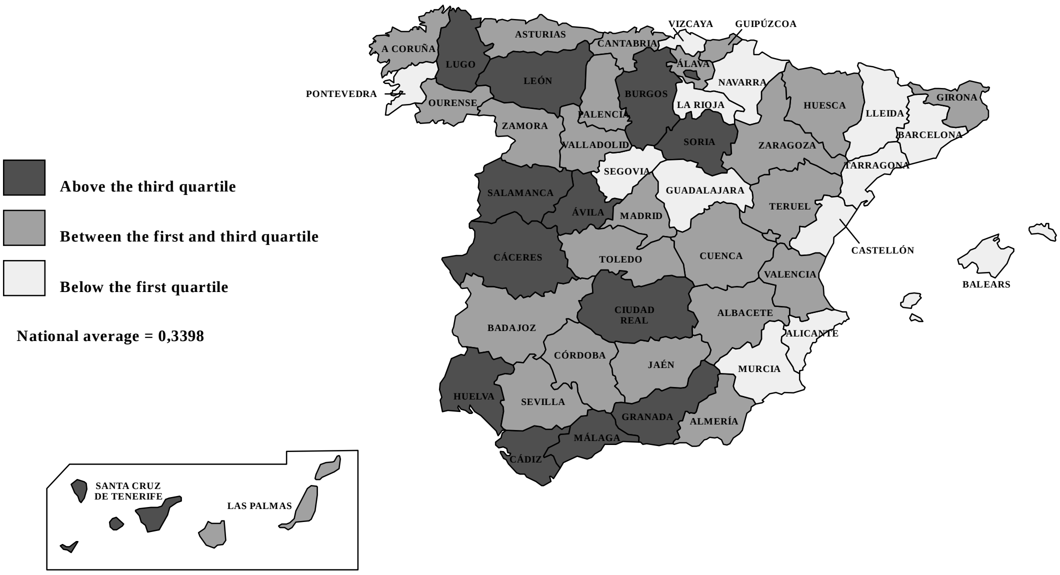
MAP 1.
GDP per capita. 1973.
Constant pesetas of 1986.



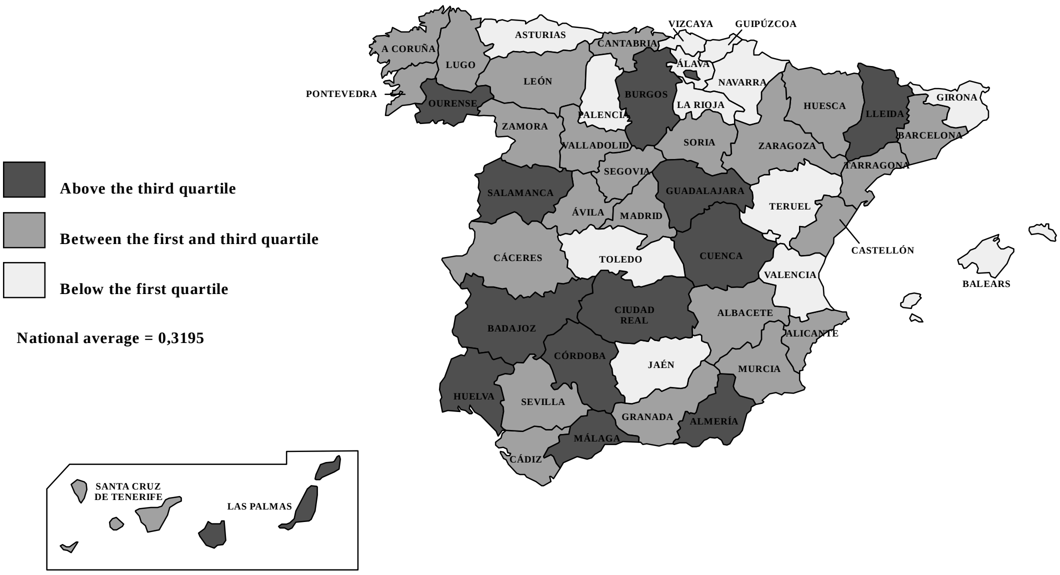
MAP 2.
GDP per capita. 1998.
Constant pesetas of 1986.



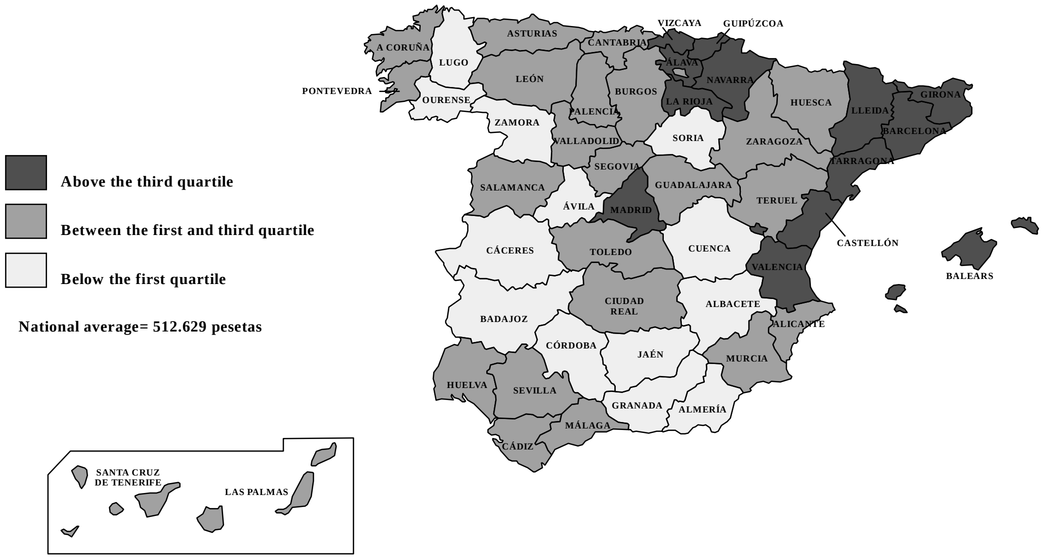
MAP 3.
GINI Index. 1973/74.
Total expenditure per capita.



MAP 4.
GINI Index. 1990/91.
Total expenditure per capita.



MAP 5.
Welfare. 1973.
Constant pesetas of 1986.



MAP 6.
Welfare. 1991.
Constant pesetas of 1986.