Working Paper 97-42
Economics Series 19
June 1997

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-98-75

AN INEQUALITY DECOMPOSITION METHOD WHICH MINIMIZES EQUIVALENCE SCALES 'CONTAMINATION' PROBLEMS

Coral del Río and Javier Ruiz-Castillo *

Abstract

Decomposable measures are a useful tool to analyze the impact of households characteristics on income or expenditure inequality. However, the results are sensitive to the choice of equivalence scales in a heterogenous population. In this paper, we assume that equivalence scales depend only on the number of persons in the household. In this context, we suggest a method to free the decomposition analysis from the possible 'contamination' that will arise if we use an inappropiate equivalence scale. The method is applied to the evolution of the standard of living in Spain during the 80's. We study the structure of Spanish inequality in 1980-81 and 1990-91, as well as the trend in overall inequality over time in terms of three factors: i) the change in within-group inequality, ii) the change in between-group inequality, and iii) the demographic change across partition subgroups.

Keywords: Inequality decomposition; Inequality structure; Inequality trend; Equivalence scales.

* del Rio, Departamento de Economía Aplicada, Universidad de Vigo; Ruiz-Castillo, Departamento de Economía, Universidad Carlos III de Madrid, E-mail: jrc@eco.

"An inequality decomposition method which minimizes equivalence scales 'contamination' problems"

Coral del Río* and Javier Ruiz-Castillo**

*Departamento de Economía Aplicada, Universidad de Vigo

**Departamento de Economía, Universidad Carlos III de Madrid

June 1997

This work has been completed under the research programme of the Cátedra *Gumersindo de Azcárate*, financed by the Ministry of Labour and Social Affairs. It has been presented in the 1996 meeting of the network *"Distribution and Redistribution of Income"*, in Bordeaux, financed by the European Communities (Contract #ERBCHRXCT940647). Financial help from Project PB93-0230 of the Spanish DGICYT is also acknowledged.

Abstract

Decomposable measures are a useful tool to analyze the impact of households characteristics on income or expenditure inequality. However, the results are sensitive to the choice of equivalence scales in a heterogenous population. In this paper, we assume that equivalence scales depend only on the number of persons in the household. In this context, we suggest a method to free the decomposition analysis from the possible 'contamination' that will arise if we use an inappropiate equivalence scale. The method is applied to the evolution of the standard of living in Spain during the 80's. We study the structure of Spanish inequality in 1980-81 and 1990-91, as well as the trend in overall inequality over time in terms of three factors: i) the change in within-group inequality, ii) the change in between-group inequality, and iii) the demographic change across partition subgroups.

KEY WORDS: Inequality decomposition; Inequality structure; Inequality trend; Equivalence scales

Abstract

Decomposable measures are a useful tool to analyze the impact of households characteristics on income or expenditure inequality. However, the results are sensitive to the choice of equivalence scales in a heterogenous population. In this paper, we assume that equivalence scales depend only on the number of persons in the household. In this context, we suggest a method to free the decomposition analysis from the possible 'contamination' that will arise if we use an inappropiate equivalence scale. The method is applied to the evolution of the standard of living in Spain during the 80's. We study the structure of Spanish inequality in 1980-81 and 1990-91, as well as the trend in overall inequality over time in terms of three factors: i) the change in within-group inequality, ii) the change in between-group inequality, and iii) the demographic change across partition subgroups.

INTRODUCCION

In applied work in the income distribution field, researchers have often available a rich information on household demographic, geographic and socioeconomic characteristics. Since Mookherjee and Shorrocks (1982), decomposable inequality measures have proved to be a useful tool to analize this information. For every population partition, decomposable measures of inequality allow us to express overall inequality in a cross-section as the sum of two terms: a weighted sum of *within-group* inequalities, plus a *between-group* inequality component. In this paper, the between-group component is calculated as if each person within a given group received the group's mean income(1).

Using decomposable measures, two interesting questions can be asked. In the first place, one can probe into the structure of a country's inequality in a single year by ranking all household characteristics in terms of their explanatory power of overall inequality, measured by the importance of the between-group term. In the second place, one can attempt to explain the trend in overall inequality over time in terms of three factors: i) the change in within-group inequality (due to changes in subgroup inequality values), ii) the change in between-group inequality (due to the relative variations in subgroup means), and iii) the demographic change across partition subgroups (due to shifts in subgroup population shares). A further advantage of using decomposition methods, is that the results for the total population can be easily related to the analysis at the specific subgroup level.

To simplify the analysis, these questions are usually answered for the distribution of equivalent income according to some reasonable equivalence scale(2). However, as Coulter et al (1992a) conclude, there is no single 'correct' equivalence scale for adjusting incomes. Thus, a range of scale relativities is both justifiable and inevitable. The problem, of course, is that overall inequality

measurement -and therefore the decomposition analysis- is known to be sensitive to scale choice.

In this paper, to make the analysis tractable we assume that equivalence scales depend only on the number of persons in the household. In this context, we suggest a method to free the decomposition analysis from the possible 'contamination' that will arise if we use an inappropiate equivalence scale. The method is applied to the evolution of the standard of living in Spain during the 80's, using the 1980-81 and 1990-91 *Encuestas de Presupuestos Familiares* (EPF for short), the Spanish budget surveys collected by the INE (*Instituto Nacional de Estadística*) with the main purpose of estimating the weights of the official consumer price index.

There are several reasons why the Spanish case is an interesting one. Spain gave itself a democratic regime during the mid 70's, and became full member of the European Community in 1986. During the last two decades, Spain has been involved in a complex process of economic modernization and liberalization, while striving at the same time to catch up in the construction of a Welfare State comparable to the one existing in other Western societies. Like in Portugal, who has gone through similar deep political and economic reforms during this period, in Spain there has been a decrease in both income and expenditure inequality, a different trend from most OCDE countries(3). From this perspective, the Iberian peninsula's experience could be of some interest to some other economies in transition, both in Latin America and in Eastern Europe.

The paper is organised in four Sections. In Section I we present the decomposition method. In Section II we describe the data and apply the method to discuss the sources of Spanish expenditure inequality in 1980-81 and 1990-91. Section III contains a brief description of the macroeconomic evolution of Spain during the 80's, and attempts an explanation of the trend in expenditure inequality in terms of the basic partition by household size and seven other household characteristics. The concluding Section IV summarizes what we have learned from the Spanish data. An statistical Appendix contains a summary of

results in the related literature, as well as information on population shares and subgroup means, and the definition of some variables.

I. THE DECOMPOSITION METHOD

I.1. Interpersonal comparisons of welfare

Assume we have a population of h = 1,...,H households whose living standards can be adequately represented by a one-dimensional variable we call income, x^h. Households can differ in income and/or a vector of household characteristics. As indicated in the Introduction, we assume that equivalence scales depend only on the number of persons in the household. Households of the same size are assumed to have the same needs and, therefore, their incomes are directly comparable. Larger households have greater needs, but also greater opportunities to achieve economies of scale in consumption. Assume that there are m = 1,...,M household sizes. Following Buhmann *et al* (1988) and Coulter *et al* (1992a, 1992b), for each household h of size m we define adjusted income by

$$z^h(\Theta) = x^h / m^\Theta, \ \Theta \in [0,1].$$

When $\Theta = 0$, adjusted income coincides with unadjusted household income, while if $\Theta = 1$, it becomes *per capita* household income. Taking a single adult as the reference type, the expression m^Θ can be interpreted as the number of equivalent adults in a household of size m. Thus, the greater is the equivalence elasticity Θ, the smaller are the economies of scale in consumption or, in other words, the larger is the number of equivalent adults(4).

Let x^m and $z^m(\Theta)$ be, respectively, the vector of original and adjusted incomes for households of size m. Notice that, if I is any index of relative inequality, then for each m

$$I(z^m(\Theta)) = I(x^m / (m^\Theta) = I(x^m).$$

Thus, within each subgroup with the same needs, we assume that the inequality of adjusted income is equal to the inequality of original income, independently of individual income and prices.

I. 2. The decomposition of overall inequality for the partition by household size

We say that an inequality index is decomposable by population subgroup, if the decomposition procedure of overall inequality into a within-group and a between-group term is valid for any arbitrary population partition. It is well known that the Generalized Entropy (GE) family of inequality indices are the only measures of relative inequality that satisfy the usual normative properties required from any inequality index and, in addition, are decomposable by population subgroup. (See, for example, Shorrocks (1984)). The family can be described by means of the following convenient cardinalization:

$$I_c(z(\Theta)) = (1/H)\,(1/c^2 - c)\,\Sigma_h\,\{(z^h(\Theta)/\mu(z^h(\Theta))^c - 1\},\ c \neq 0,1;$$

$$I_0(z(\Theta)) = (1/H)\,\Sigma_h\,\log\{\mu(z^h(\Theta)/z^h(\Theta)\};$$

$$I_1(z(\Theta)) = (1/H)\,\Sigma_h\,\{(z^h(\Theta)/\mu(z^h(\Theta)\}\log\{z^h(\Theta)/\mu(z^h(\Theta)\},$$

where $\mu(\cdot)$ is the mean of the distribution. The parameter c summarizes the sensitivity of I_c in different parts of the income distribution: the more positive (negative) c is, the more sensitive I_c is to differences at the top (bottom) of the distribution (Cowell and Kuga (1881)). I_1 is the original Theil index, while I_0 is the mean logarithmic deviation.

Coulter *et al* (1992a, 1992b) have shown how the inequality estimates provided by the GE family vary sistematically with the parameter Θ which captures the generosity of the scale. They illustrate their analysis with UK data(5). However, they suggest that using the GE family in its decomposable form restricts the 'contamination' of the inequality orderings that will arise if there is incomplete or incorrect information about the equivalence scales. In this way, *some* robust conclusions, independent of the equivalence scale, can be obtained.

To see this, consider the formula for the GE index when written in decomposable form for the partition by household size:

$$I_c(z(\Theta)) = \Sigma_m \ (v^m(\Theta))^c (p^m)^{1-c} I_c(z^m(\Theta) + I_c(\mu^1(\Theta),...,\mu^M(\Theta)),$$

where $v^m(\Theta)$ is the share of total adjusted income held by households of size m for each Θ; p^m is group m's population share, and $I_c(\mu^1(\Theta),...,\mu^M(\Theta))$ is the between-group inequality calculated as if each household of a given size m received that group's mean adjusted income $\mu^m(\Theta)$[6]. Taking into account that, for each m, $I_c(z^m(\Theta) = I_c(x^m)$, Coulter *et al* (1992a) indicate that using the 'wrong' equivalence scale contaminates only the between group component. Notice that, since $v^m(\Theta)$ depends on Θ for every m, this property only holds for c = 0. In this case, denoting by U and $C(\Theta)$ the uncontaminated and the contaminated terms, we have:

$$I_0(z(\Theta)) = U + C(\Theta), \qquad (1)$$

where

$$U = \Sigma_m \ p^m \ I_0(x^m) \qquad (2)$$

is the weighted average of the inequality within each household size with weights equal to population shares, and

$$C(\Theta) = I_0(\mu^1(\Theta),...,\mu^M(\Theta))$$

is the between-group inequality which depends on Θ.

I. 3. The decomposition of overall inequality for other partitions

For any other partition of the population into k = 1,..., K subgroups, the direct decomposition of overall inequality according to the index I_0 is:

$$I_0(z(\Theta)) = \Sigma_k \ p^k I_0(z^k(\Theta)) + I_0(\mu^1(\Theta),...,\mu^K(\Theta)).$$

Unfortunately, in this case both the within- and between-group terms are "contaminated" by Θ. Therefore, one cannot disentangle the impact of

characteristic k on overall inequality independently of the role played by the parameter Θ.

To overcome this shortcoming, we suggest to apply the decomposability property to the terms $I_0(x^m)$ in expression (2), which are independent of Θ. In other words, we propose to consider the partition induced by characteristic k into, say, the K geographical regions in a country within each homogeneous subgroup of equal size. Let x^{mk} be the unadjusted income vector of households of size m in region k, let μ^{mk} be the mean of that distribution, and let r^{mk} be the proportion of households of type k among households of size m. For each m, $I_0(x^m)$ is seen to be equal to the sum of two terms: i) the weighted sum of the inequality within each subgroup of the new partition, $I_0(x^{mk})$, with weights equal to the demographic shares, r^{mk}; and ii) the between-group inequality induced by characteristic k among households of size m, $I_0(\mu^{m1},...,\mu^{mK})$, calculated as if each household of size m receives the mean of the region where she lives in, μ^{mk}. That is,

$$I_0(x^m) = \Sigma_k\, r^{mk}\, I_0(x^{mk}) + I_0(\mu^{m1},...,\mu^{mK}).$$

If we now let p^{mk} be the proportion of households of size m in region k in the population as a whole, then we have

$$\begin{aligned} U &= \Sigma_m\, p^m\, [\Sigma_k r^{mk}\, I_0(x^{mk}) + I_0(\mu^{m1}(\Theta),...,\mu^{mK}(\Theta))] \\ &= \Sigma_{mk}\, p^{mk}\, I_0(x^{mk}) + \Sigma_m\, p^m\, I_0(\mu^{m1}(\Theta),...,\mu^{mK}(\Theta)). \end{aligned}$$

Thus,

$$U = W^k + B^k,$$

where

$$W^k = [\Sigma_{mk}\, p^{mk}\, I_0(x^{mk})], \tag{3}$$

and

$$B^k = [\Sigma_m\, p^m\, I_0(\mu^{m1},...,\mu^{mK})]. \tag{4}$$

Equation (3) is the within-group inequality in the partition by household size *and* characteristic k. On the other hand, we have seen that $I_0(\mu^{m1},...,\mu^{mK})$ is the between-group inequality induced by characteristic k among households of size

m. Thus, equation (4) is the between-group inequality attributable to characteristic k. Collecting terms, we have

$$I_0(z(\Theta)) = U + C(\Theta) = W^k + B^k + C(\Theta). \quad (5)$$

The implication for empirical work is clear. On the one hand, one can inmediately check how much of the overall adjusted income inequality is attributable to inequality within the partition by household size, independently of the value we assign to the equivalence elasticity Θ. On the other hand, to learn about the sources of inequality in a given year, we can rank the different partitions according to their explanatory power measured by the relative size of the 'uncontaminated' between-group term, B^k, which is independent of the generosity of the scale. If for two partitions k and j we find that B^k/U is much greater than B^j/U, then it is reasonable to say that the population characteristic k is more important as a determinant of overall inequality than is characteristic j.

II. THE STRUCTURE OF SPANISH INEQUALITY

II.1. The data

We use data from the Spanish EPF's for the periods April 1980-March 1981 and April 1990-March 1991. These are large budget surveys of 23,972 and 21,155 observations, representative a population of approximately 10 or 11 million households and 37 or 38 million persons in 1980-81 and 1990-91, respectively, occupying residential housing in all of Spain including the northern African cities of Ceuta and Melilla.

Household welfare is approximated by a measure of current consumption, namely, household total current expenditure on private goods and services, net of expenditures on the acquisition of certain durables, but inclusive of imputations for self-consumption, wages in kind, meals subsidised at work, and the rental value for owner-occupied and other non rental housing(7). We express household expenditures at constant prices of the Winter of 1991 by means

of household specific statistical price indices. Since we are interested in personal rather than household welfare, we follow the usual practice of studying the personal distribution in which each person is assigned the adjusted expenditures of the household to which she belongs (cf Danziger and Taussing (1979), Cowell (1984) and Shorrocks (1995)). All of our results use the weighting factors provided by the INE which make the sample representative of the entire population.

II. 2. Household characteristics

We study eight characteristics, classified into three groups: Demographic characteristics: 1) Household size (HSIZE); 2) Household type (HTYPE); 3) Age of household head (AGE). Geographic characteristics: 4) Size of the municipality (MUNS); 5) Autonomous Community of residence (AUTON). Socioeconomic characteristics: 6) Educational level attained by the household head (EDC); 7) Socioeconomic status of the household head (SOCIO); 8) Number of earners of labor and non-labor income in the household (NEARN).

Table 1 presents the frequency distributions of the 1990-91 population for households (H) and persons (P) classified by each of these characteristics. Borooah and Collins (1995) point out that if two characteristics k and j are related, the inequality decomposition based upon the disaggregation of the population into subgroups on the basis of a single characteristic, k or j, might be misleading. Therefore, we have included some new ones. TYPAGE combines HTYPE and AGE into 16 subgroups. GEO distinguishes urban from rural municipalities in each Autonomous Community, giving rise to 36 subgroups. AGEDC combines two human capital variables, the age and the educational level of the household head, giving rise to 14 subgroups. SOCEDC combines the two socioeconomic characteristics SOCIO and EDC into 16 subgroups. Finally, NEADC combines NEARN and EDC into 17 categories. (See the Appendix for a detailed definition of these five variables). For later reference, notice that the finer the partition, the greater the between-group term.

TABLE 1. Household (H) and personal (P) 1990-91 frequency distributions -in %- for different household characteristics

HSIZE	H	P	HTYPE	H	P	AGE	H	P	AUTON	H	P
1member	10.0	2.9	1	10.0	2.9	Up to 24	1.5	1.3	Andalucía	16.6	17.8
2	22.3	13.1	2	17.0	10.0	25-34	13.4	13.2	Aragón	3.4	3.1
3	20.8	18.3	3	5.2	3.9	35-44	19.8	24.2	Asturias	3.0	2.9
4	25.0	29.3	4	16.2	17.4	45-54	19.1	24.3	Baleares	1.9	1.7
5	13.2	19.4	5	24.7	28.5	55-64, active	13.3	13.5	Canarias	3.5	3.8
6	5.4	9.6	6	10.6	16.0	55-64, retired	7.8	7.1	Cantabria	1.3	1.4
7 or +	3.3	7.4	7	7.2	11.3	65-74, male	11.8	9.3	C.-León	7.1	6.7
			8	6.5	5.9	65-74, fem.	4.4	2.1	C.-La M.	4.5	4.4
			9	2.7	4.1	75 or +, male	5.1	3.4	Cataluña	16.0	15.4
						75 or +, fem.	3.8	1.6	Co. Valen.	10.1	9.8
SOCIO	H	P	EDC	H	P	NEARN	H	P	Extremad.	2.9	2.9
1	2.8	3.3	1	4.4	3.5	1	1.5	1.5	Galicia	6.9	7.2
2	2.7	3.2	2	21.4	19.7	2	18.6	10.2	Madrid	12.6	12.6
3	33.5	38.5	3	38.3	39.4	3	9.6	7.5	R. Murcia	2.5	2.6
4	7.2	8.4	4	14.0	15.1	4	0.6	0.5	Navarra	1.3	1.3
5	3.0	3.5	5	7.4	7.4	5	0.5	0.6	País Vasco	5.4	5.5
6	8.2	9.1	6	5.2	5.4	6	24.2	25.3	La Rioja	0.7	0.7
7	0.3	0.4	7	4.6	4.7	7	14.0	14.8	Ceuta y	0.3	0.3
8	1.2	1.2	8	4.6	4.8	8	14.4	15.9	Melilla		
9	5.1	5.5	MUNS	H	P	9	3.8	4.7			
10	10.1	9.6	< 2,000	7.3	6.5	10	5.0	6.7			
11	16.1	12.1	2-10,000	19.4	19.4	11	2.6	3.6			
12	7.4	3.4	10-50,000	22.0	22.8	12	5.2	8.5			
13	0.2	0.2	50-500,000	31.8	32.7						
14	2.1	1.7	> 500,000	19.5	18.6						

HSIZE (household size). AGE (household head's age). MUNS (municipality size). AUTON (Autonomous Community). HTYPE (household type): 1. Single person and minors-only (up to 17 years old), 2. Couples, 3. Three or more adults without children, 4. Couples with young (18-30 years old) descendants, 5. Couples with minors, 6. Couples with minors and young descendants, 7. Couples with minors and/or young descendants, as well as older people (65 or more years), 8. Single parent households, 9. Households without couples with unrelated minors. EDC (household head's educational level): 1. Illiterate, 2. Without formal studies, 3. Grade school, 4. Primary school, 5. Secondary school, 6. Vocational school, 7. Three years College degree, 8. More than three years - College degree. SOCIO (household head socioeconomic category): 1. Workers, agricultural, 2. Self employed, agricultural, 3. Non-agricultural workers, 4. Non-agricultural self employed, 5. Supervisors, and Armed Forces, 6. Upper classes, 7. Unclassifiable persons in the labor force, 8. Part-time (less than 1/3 working day), 9. Unemployed, 10. Early retired (below 65 years of age), 11. Retired, male, 12. Retired, female, 13. Living off property income, 14. Other category outside of the labor force. NEARN (Number of earners of labor and non-labor income in the household): 1. No income earners, 2. One non-labor income earner, 3. Two or more non-labor income earners, 4. Only one earner from part time work, 5. Two or more earners, none of them from full time work, 6. One full time labor income earner, 7. Two earners, one of them full time, 8. Two full time labor income earners, 9. Three earners, one of them full time, 10. Three earners, two of them full time, 11. Three full time income earners, 12. Four or more earners of any kind.

II.1. The sources of Spanish inequality

Starting with the partition by household size, we saw in equation (5) that overall inequality is the sum of two terms: within-group or uncontaminated

inequality, U, which is independent of the equivalence scale parameter Θ, and between-group or contaminated inequality, C(Θ). Table 2 offers the evidence on this decomposition for 1980-81 and 1990-91. In both years, the importance of C(Θ) as an explanatory factor of overall inequality follows a non-linear pattern with Θ. When no allowance is made for household size, i.e. Θ = 0, between-group inequality accounts for a sizable percentage of overall inequality, 16-18 per cent. As Θ increases, reflecting the decreasing importance of economies of scale in consumption within the household, the ratio $C(\Theta)/I_0(z(\Theta))$ rapidly declines, increasing again as Θ approaches 1 and adjusted total expenditure becomes per capita total expenditure[8]. Therefore, in the analysis of the structure of inequality in a given year the risk of 'contamination' is non-negligible: different values for Θ imply a rather different explanatory role for the inequality between household sizes.

TABLE 2. The importance of between-group inequality in the partition by household size as a function of the equivalence scales parameter Θ

	Inequality in 1980-81				Inequality in 1990-91			
Θ =	(1) U	(2) C(Θ)	(3) $I_0(\Theta)$	(2)/(3) in %	(1) U	(2) C(Θ)	(3) $I_0(\Theta)$	(2)/(3) in %
0.0	0.1518	0.0294	0.1811	16.2	0.1361	0.0303	0.1664	18.2
0.2	0.1518	0.0132	0.1650	8.0	0.1361	0.0146	0.1507	9.7
0.4	0.1518	0.0039	0.1556	2.5	0.1361	0.0052	0.1412	3.7
0.7	0.1518	0.0033	0.1551	2.2	0.1361	0.0036	0.1396	2.6
1.0	0.1518	0.0202	0.1720	11.7	0.1361	0.0184	0.1545	11.9

U = uncontaminated inequality; C(Θ) = contaminated inequality; $I_0(\Theta)$ = overall inequality

Turning now towards the uncontaminated term, we study two questions: the internal structure of within-group inequality in the partition by HSIZE; and the decomposition $U = W^k + B^k$ for every partition k (equation (5)). For the first question, recall that $U = \Sigma_m U^m$, where $U^m = p^m I_0(x^m)$ and p^m is the demographic share of the household size in question. Table 3 presents the percentage distribution of U. In both years, single person and two member households (plus large households of 8 or more members in 1980-81), contribute

to within-group inequality well above what can be expected from their demographic importance. The opposite is the case for the important group of three to six member households which represent 75 per cent of the people.

TABLE 3. The individual contribution of each household size to within-group inequality

	Household size								
	1member	2	3	4	5	6	7	8 or more	All
1980-81	4.2	15.7	15.6	21.7	18.2	10.6	6.4	7.6	100.0
1990-91	5.2	17.0	17.3	27.5	17.4	9.0	4.0	2.5	100.0

In the second place, we study which characteristics account for a greater percentage of U. Table 4 contains the relevant information on the expression $(B^k/U)100$ for every k, where B^k is the between-group inequality induced by characteristic k.

TABLE 4. The importance of the between-group inequality induced by the different characteristics, B^k, relative to the uncontaminated inequality term, U, measured in percentage terms: $[B^k/U]100$ for each household characteristic k

	Demographic		Geographic		Socioeconomic		
	HTYPE	AGE	MUNS	AUTON	EDC	SOCIO	NEARN
1980-81	2.0	6.8	10.6	9.4	25.2	20.5	10.4
1990-91	3.6	7.3	9.5	11.3	22.1	15.0	10.7

	Combined characteristics				
	TYPAGE (HTYPE+AGE)	GEO (MUNS+AUTON)	AGEDC (AGE+EDEC)	SOCEDC (SOCIO+EDC)	NEADC (NEARN+EDC)
1980-81	5.3	15.7	21.7	27.6	25.1
1990-91	6.9	15.9	19.3	20.3	21.1

i) The demographic variables HTYPE and AGE contribute very little to the explanation of uncontaminated inequality. Even when they are combined into TYPAGE they explain less than 7 per cent of it.

ii) Separately, the geographic variables explain about 10 per cent, and jointly about 16 per cent. This should give some pause to (some) policy makers cronically worried by the improvement of interregional inequities. Although they are not talked about as much, urban/rural inequalities in Spain are as important as interregional ones. From a geographical perspective, people

interested in reducing Spain's inequality should look into the inequality within Autonomous Communities and within urban/rural subgroups.

iii) Not surprisingly, inequality is a phenomenon better explained by socioeconomic factors. SOCIO explains about 15-20 per cent and EDC about 22-25 per cent of U. Their combined effect, ranging from 20 to 27 per cent, adds very little to the explanation. The human capital variable AGEDC, surely because of the EDC component, explains about 20 per cent of U.

iv) The variable NEARN, reflecting the situation of household members in relation to economic activity, explain as little as the geographical ones. Only in combination with EDC, NEADC jumps to 21-25 per cent of the within-group inequality in the partition by household size.

These results are in agreement with those obtained for other countries using the same index I_0 and a specific set of equivalence scales. (See the summary results presented in Table A in the Appendix). Thus, like on other countries, in the Spanish case the capacity of a variety of household characteristics to explain overall inequality in a single cross-section is rather limited. Most inequality takes place within subgroups of the different partitions. Nevertheless, EDC and SOCIO -the educational level attained by the household head and its socioeconomic category, respectively- are the ones which account for the greatest percentage of the term uncontaminated by Θ.

III. THE INEQUALITY TREND DURING THE 80'S

III. 1. The background

From a macroeconomic point of view, the recession started after the first oil crisis in 1973 did not finish in Spain until 1986. According to the National Accounts figures, from 1980 until 1985 real GDP at prices of 1986 grew only at an average rate of 1.4 per cent, while from that date until 1991 GDP grew at an average rate of 4.1 per cent. It is well known that Spain exhibits the highest official unemployment rate in the EU. From 1980 to 1991, it rose first up to 21 per

cent of the labor force in 1986, to be reduced later down to a 15.6 per cent in 1991. During the 80's, consumer price inflation in Spain was 3 percentage points above the EU average.

The household budget surveys which provide the data source for this paper, inform us about the rate of growth of consumption during the 80's. Let us denote by $\mathbf{z}_1(\Theta)$ and $\mathbf{z}_2(\Theta)$ the 1980-81 and 1990-91 distributions of adjusted household expenditures at Winter of 1991 prices. Mean adjusted household expenditures raised in real terms at an average annual rate in the range 2.3 to 3.3 per cent, depending on the value we give to the parameter Θ. What do we know of the change in overall inequality in Spain during this period? In Del Río and Ruiz-Castillo (1996a) we found that $\mathbf{z}_2(\Theta)$ Lorenz dominates $\mathbf{z}_1(\Theta)$ for all Θ, so that there is an unambigous improvement in relative inequality in real terms[9].

Using the Lorenz consistent inequality index I_0, we find that real inequality decreased from 8 to 10 per cent for different values of Θ (See Table 5 below). Decomposition analysis is a useful tool in the search for an explanation of this trend. On the one hand, both in Spain and in other countries, shifts in subgroup population shares are often quoted as potential determinants of inequality change. On the other hand, relative variation in subgroup means may give rise to important changes in between-group inequality. In what follows, we first present the evidence on the potential explanatory factors. Then we use the decomposition of the inequality trend to verify whether or not those factors are important. Finally, since the inequality trend unexplained through these two avenues must be necesarily accounted for by changes in specific subgroup inequality values, we check which subgroups have influenced the most (and the least) the reduction in uncontaminated inequality.

Table B in the Appendix presents the shifts in demographic shares for the eight variables for which we have information. The most important features can be summarized as follows.

1. Taking 5 person households -whose numbers remain essentially constant- as a reference, the larger the household the larger the population

decline, and the smaller the household the larger the increase. As a result, average household size decreases from 3.70 to 3.41 members.

2. Fertility rates are declining, so that households with minors (HTYPE equal to 5, 6 or 7) are losing importance. Young persons (between 18 and 30 years old), heavily hitted by unemployment, remain at their parents homes rather than forming independent households of their own (HTYPE = 4). As a result of a delay in marriage commitments -the other side of this coin- the subgroup of households headed by a person 35-44 years of age is increasing. Single person households and couples without children, as well as single parent households (HTYPE = 1, 2, and 8, respectively), are increasing more rapidly than the population as a whole.

3. The Spanish population is aging. Taking households headed by an early retired person in the 55-64 age bracket as a reference, all older households are increasing (more so if headed by a female), while younger households are loosing importance.

4. Relatively large sized households are leaving small municipalities, while more people are moving into medium sized cities and provincial capitals between 50,000 and 500,000 inhabitants.

5. Only Madrid and three less developed Autonomous Communities (the Canary Islands, Andalucía and Extremadura) are growing above the national rate. The País Vasco and Navarra, the East of the country (Cataluña, Comunidad Valenciana y Murcia), and Castilla-La Mancha grow close to the average rate, while the rest of the North (Cantabria, Asturias, Galicia) Aragón and La Rioja, plus Castilla-León and the Baleares Islands grow at a smaller rate or even loose some population.

6. Households headed by persons whith a secondary school degree and beyond, are increasing, while the opposite is the case for illiterates, people without formal studies or with only grade school.

7. Households headed by someone in the agricultural sector, part-time employment, and the small subgroup living off property income are loosing

demographic importance. The mixed subgroup headed by supervisors or Armed Forces personnel, the upper socioeconomic category and, above all, retired people of all kinds, are increasing above national rates.

8. The small groups with no income earners or persons engaged only in part time employment (**NEARN** = 1, 4, 5), and the important group with only one person with full time employment (**NEARN** = 6) are loosing population. Households consisting of people outside of the labor force (**NEARN** = 2, 3), and households with two or more persons with a full time job (**NEARN** = 7-12) are growing above the national average.

For all partitions, Table C in the Appendix presents the subgroup mean household expenditures in index form for an intermediate value of the equivalence scales parameter $\Theta = 0.4$. The only unambigous pattern we observe is the decline in the range of the index numbers for the two socioeconomic variables **SOCIO** and **EDC**. This leads us to expect an important positive contribution by these variables to the reduction in the uncontaminated inequality during this period.

In the remainder of this section we analyze the explanatory power of demographic shifts and changes in subgroup means, starting with the partition by household size.

III. 2. The partition by household size

Let us denote by $\Delta I(\Theta)$ the change in inequality between the two situations, i.e., $\Delta I(\Theta) = I_0(z_2(\Theta)) - I_0(z_1(\Theta)$. This magnitude can be expressed as

$$\Delta I(\Theta) = \Delta U + \Delta C(\Theta), \tag{6}$$

where:

$$\Delta U = U_2 - U_1 = \Delta W + \Delta D, \tag{7}$$

$$\Delta W = \Sigma_m \, p_1^m \, [I_0(x_2^m) - I_0(x_1^m)], \tag{8}$$

$$\Delta D = \Sigma_m \, [p_2^m - p_1^m] \, I_0(x_2^m), \tag{9}$$

$$\Delta C(\Theta) = I_0(\mu_2^1(\Theta),..., \mu_2^M(\Theta)) - I_0(\mu_1^1(\Theta),..., \mu_1^M(\Theta)). \qquad (10)$$

Equation (7) is the change in uncontaminated inequality, which is seen to be the sum of two terms: equation (8), which is the weighted sum of inequality changes within each household size, and equation (9) which captures the impact on uncontaminated inequality of demographic changes across the partition by household size. Both are independent of Θ, which only affects equation (10), namely, the change in between-group inequality in the partition by household size. Taking into account the information in Table 2, we present in Table 5 the estimates of the trend in uncontaminated, contaminated and overall inequality in percentage terms.

TABLE 5. The trend in uncontaminated, contaminated and overall inequality during the 80's

Rates of change, in %, of:	Equivalence scales parameter Θ				
	0.0	0.2	0.4	0.7	1.0
Overall inequality = $[(I_2(.) - I_1(.))/I_1((.)]100$	- 8.1	- 8.7	- 9.3	- 10.0	- 10.2
Contaminated relative to overall inequality = $[(C_2(.) - C_1(.))/I_1(.)]100$	0.5	0.8	0.8	0.1	- 1.0
Uncontaminated relative to overall inequality = $[(U_2 - U_1)/I_1(.)]100$	- 8.6	- 9.5	- 10.1	- 10.1	- 9.1
Uncontaminated inequality = $[(U_2 - U_1) / U_1]100$	- 10.3	- 10.3	- 10.3	- 10.3	- 10.3

The first row in Table 5 informs about a reduction in real inequality of about 8/10 per cent, which is rather robust to changes in the parameter Θ. How do we explain this change in term of changes inside the partition by household size? In the first place, we observe a non-linear pattern in the change in contaminated inequality. As Θ grows from 0 to 1, first there is a gradual increase in between-group inequality, which eventually becomes a decrease explaining 10 per cent of the overall change. However, changes in between-group inequality are relatively unimportant as an explanatory factor of overall inequality. In the second place, as indicated in the fourth row of Table 5, uncontaminated inequality decreases by approximately 10 per cent. This means that most of the reduction in overall

inequality -at least 90 per cent- must be attributed to the change in the uncontaminated term.

The next question is to find out how much of the change in uncontaminated inequality, ΔU, is due to shifts in the population frequency distribution by household size, ΔD. According to equation (7), $\Delta U = \Delta W + \Delta D$. We have already seen that $(\Delta U/U_1)100 = -10.3$ We estimate that $(\Delta W/U_1)100 = -11.7$, while $(\Delta D/U_1)100 = 1.4$. The ratio of these two magnitudes is, approximately, equal to - 13. Therefore, the switch towards smaller households during the 80's, documented in Table B in the Appendix, gives rise to an increase in inequality which represents about 13 per cent of the overall change in the uncontaminated term. Taking into account that smaller households are more unequal (see Table 3), this is of course to be expected.

We have seen that about 113 per cent of the reduction in the uncontaminated term must be explained within the partition by HSIZE. To investigate the contribution of each subgroup to the decrease in W, let us write equation (8) as $\Delta W = \Sigma_m \Delta W^m$, where $\Delta W^m = p_1^m [I_0(x_2^m) - I_0(x_1^m)]$ and p_1^m is the share in the population in 1980-81 of persons in households of size m. In Table 6 we look at the expression $(\Delta W^m/\Delta W)100$ for each m.

TABLE 6. Contribution by each subgroup to within-group inequality in the partition by household size

	Household size									
	1member	2	3	4	5	6	7	8	9 and more	All
In %:	6.8	20.9	23.9	2.2	17.3	0.2	8.2	12.0	9.7	100.0

We observe that, relative to their demographic importance, the contribution to the reduction in inequality by single person and 7 or more person households is very large. These two groups explain one third of the change but they represent 10 per cent of all persons. On the contrary, 4 to 6 person households, which contain almost 60 per cent of the population, contribute only about 20 per cent of the change.

III. 3. The explanation of the inequality trend in terms of the remaining variables

For any partition k, the following decomposition of the overall inequality trend is useful:

$$\Delta I(\Theta) = \Delta U + \Delta C(\Theta) = \Delta W^k + \Delta B^k + \Delta D^k + \Delta C(\Theta), \qquad (11)$$

where:

$$\Delta W^k = \Sigma_{mk}\, p_1^{mk}\, [I_0(x_2^{mk}) - I_0(x_1^{mk})], \qquad (12)$$

$$\Delta B^k = \Sigma_m\, p_1^m\, [I_0(\mu_2^{m1},\ldots,\mu_2^{mK}) - I_0(\mu_1^{m1},\ldots,\mu_1^{mK})], \qquad (13)$$

$$\Delta D^k = \alpha^k + \beta^k, \qquad (14)$$

$$\alpha^k = \Sigma_m\, [p_2^m - p_1^m] I_0(\mu_2^{m1},\ldots,\mu_2^{mK}), \qquad (15)$$

and

$$\beta^k = \Sigma_{mk}\, [p_2^{mk} - p_1^{mk}] I_0(x_2^{mk}). \qquad (16)$$

Contrary to the approximation originally suggested by Mookherjee and Shorrocks (1982) and used also by Jenkins (1995), equation (11) provides an exact decomposition of the overall inequality change[10]. Equations (12), (13) and (14) capture the contribution to the trend in uncontaminated inequality attributable to, respectively:

- the change in inequality within the double partition by household size and characteristic k;

- the change in between-group inequality induced by characteristic k;

- the demographic change across subgroups. This term captures the impact of demographic changes in two steps. On the one hand, equation (15) captures the shifts in population shares in the partition by household size weighted by the between-group inequality induced by partition k among households of each size. On the other hand, equation (16) reflects the impact of changes across the k-subgroups within households of the same size, weighted by the inequality within each subgroup k of size m.

Of course, for each partition k we have that

$$\Delta U = \Delta W^k + E^k$$

where $$E^k = \Delta B^k + \Delta D^k \qquad (17)$$

is the total effect attributable to characteristic k. If for two partitions k and j we find that $E^k/\Delta U$ is much greater than $E^j/\Delta U$, then it is reasonable to say that the characteristic k is more important as an explanatory factor of the change in the uncontaminated term of overall inequality than is characteristic j. Thus, our next question is to study the impact of both the demographic change across subgroups in the different partitions, ΔD^k, and the changes in subgroup means which give rise to changes in ΔB^k. The evidence on $(\alpha^k/\Delta U)100$, $(\beta^k/\Delta U)100$, $(\Delta D^k/\Delta U)100$, $(\Delta B^k/\Delta U)100$, and $(E^k/\Delta U)100$ for each k is, respectively, in columns (1) to (5) of Table 7.

TABLE 7. The impact of demographic changes across partition subgroups and the change in between-group inequality induced by each partition, in percentage terms relative to the reduction in uncontaminated inequality

	(1)	(2)	(3)	(4)	(5)		(1)	(2)	(3)	(4)	(5)
AUTON	-27.5	4.7	-22.8	-12.2	-35.1	HTYPAGE*	-23.1	3.6	-19.4	-12.7	-32.1
HTYPE	-23.0	7.1	-15.9	-18.9	-34.8	GEO*	-33.5	8.6	-24.9	5.6	-19.3
AGE	-22.3	2.4	-19.9	-0.3	-20.2	AGEDC*	-32.8	2.0	-30.8	40.6	9.7
NEARN	-2.5	3.0	0.5	4.4	4.9	NEADC*	-24.6	4.1	-20.5	55.6	35.1
MUNS	-12.7	-2.2	-14.9	25.8	10.9	SOCEDC*	-43.1	2.5	-41.6	89.3	47.7
SOCIO	-39.3	4.3	-35.0	64.0	29.0						
EDC	-4.0	1.1	-2.9	54.9	52.0						

(1) = $\alpha^k/\Delta U$, equation (15); (2) = $\beta^k/\Delta U$, equation (16); (3) = (1) + (2) = $\Delta D^k/\Delta U$ = percentage change in the uncontaminated term due to shifts in population shares in partition k; (4) = $\Delta B^k/\Delta U$ = change in the uncontaminated term due to the change in between-group inequality induced by characteristic k; (5) = (3) + (4) = E^k = total effect attributable to characteristic k

AUTON = Autonomous Community; HTYPE = household type; MUNS = municipality size; AGE = household head's age; NEARN = number of earners; SOCIO = household head's socioeconomic category; EDC = household head's educatioonal level

(*) See the Appendix for definitions

---

i) We have already seen that the shift towards smaller households captured by ΔD amounts to an inequality increase which is equal to

approximately 13 per cent of the reduction in uncontaminated inequality. By comparing equations (9) and (15), we see that the difference between ΔD and α^k is that changes in population shares are now weighted by the between-group inequality induced by each k in the partition by **HSIZE**. According to column (1) in Table 7, it appears that for the variables **AUTON, HTYPE, AGE** and, above all, **SOCIO**, the between-group inequality term $I_0(\mu^{m1},...,\mu^{mK})$ for $m \leq 4$ is relatively large. (The opposite is the case for the partitions **NEARN** and **EDC**). Since we saw in Table 4 that B^k is particularly large for **SOCIO**, that α^k has a large absolute value for this variable comes as no surprise.

ii) As far as β^k is concerned, we observe in column (2) of Table 7 that, except for **MUNS**, shifts in population shares within the double partition by **HSIZE** and the remaining characteristics have had a positive contribution, albeit small, to the reduction in uncontaminated inequality. This does not offset the effect captured in α^k. Therefore, as shown in column (3), the net demographic effect on ΔU attributable to ΔD^k is negative in all cases except **NEARN**. It is of some importance for **SOCIO** (35 per cent), negligible for **EDC**, and between 15 and 22 per cent of ΔU for the remaining variables.

iii) Turning now towards between-group inequality effects, column (4) suggests two comments. The socioeconomic variables **EDC** and **SOCIO** which exhibited the greatest explanatory capacity of overall inequality in both years (see Table 4), make a large positive contribution to ΔU. On the other hand, only the variables **AUTON** and **HTYPE** have a significantly negative contribution in this score.

iv) The consequences of all of the above is reflected in column (5) which presents the total effect attributable to each k. There are two types of variables from this point of view. Those for which the increase in between-group inequality lead to a *negative* total effect -**AUTON, HTYPE** and **AGE**- and those for which the opposite is the case -**EDC** and **SOCIO**. For the first type, we must go down to the contributions by individual subgroups to understand the inequality change. On the contrary, the second group of variables provide a partial but

important explanation of the overall trend which amounts to one third or one half, respectively, of the reduction in uncontaminated inequality.

v) The five combined characteristics reported about in the right-hand side of Table 7 do not add up anything new to the analysis. Only the role of EDC, when combined with AGE and NEARN in AGEDC and NEADC, is worth pointing out.

The final question is: which individual subgroups play a particular large role in the explanation of the change in the within-group term ΔW^k? We can rewrite equation (12) as $\Delta W^k = \Sigma_k \Delta I^k$, where $\Delta I^k = \Sigma_k p_1^{mk} [I_0(x_2^{mk}) - I_0(x_1^{mk})]$. The information about the expressions $\Delta I^k/\Delta W^k$ for the more important partitions -SOCIO and EDC- is in Table 8 (for the remaining partitions the information is available upon request).

TABLE 8. Individual contributions to the change in within-group inequality in selected partitions

SOCIO	Contribution to the reduction in within-group ineq.	Personal frequency distribution	EDC	Contribution to the reduction in within-group ineq.	Personal frequency distribution
1	12.1	3.3	1	11.7	3.5
2	9.8	3.2	2	45.9	19.7
3	52.1	38.5	3	51.9	39.4
4	-10.6	8.4	4	2.5	15.1
5	0.4	3.5	5	-5.7	7.4
6	0.2	9.1	6	1.8	5.4
7	-0.1	0.4	7	5.2	4.7
8	7.4	1.2	8	-13.3	4.8
9	-7.7	5.5		100.0	100.0
10	7.5	9.6			
11	17.2	12.1			
12	5.2	3.4			
13	1.9	0.2			
14	4.4	1.7			
	100.0	100.0			

SOCIO (household head's socioeconomic category): 1. Workers, agricultural, 2. Self employed, agricultural, 3. Non-agricultural workers, 4. Non-agricultural self employed, 5. Supervisors, and Armed Forces, 6. Upper classes, 7. Unclassifiable persons in the labor force, 8. Part-time (less than 1/3 working day), 9. Unemployed, 10. Early retired (below 65 years of age), 11. Retired, male, 12. Retired, female, 13. Living off property income, 14. Other category outside of the labor force. EDC (household head's educational level): 1. Illiterate, 2. Without formal studies, 3. Grade school, 4. Primary school, 5. Secondary school, 6. Vocational school, 7. Three years College degree, 8. More than three years College degree.

For the socioeconomic category, the following subgroups are characterized by a 60 per cent contribution to the reduction in within-group inequality which is large relative to their demographic importance which is only one third of the population: households headed by one person in the agricultural categories or part-time employment (**SOCIO** = 1, 2, 8), as well as persons outside the labor force (**SOCIO** = 10-14). Which subgroups contribute negatively or very little to the improvement in within-group inequality is equally revealing. These are households headed by someone in the upper classes, the unemployed and the self employed (**SOCIO** = 6, 9, 4), which have a combined negative contribution of 18 per cent and approximately that same percentage of the population.

For the educational level, the subgroups with a large contribution are households headed by an illiterate or a person without formal studies (**EDC** = 1, 2), which contribute more than half of the reduction in within-group inequality but represent less than a quarter of the population. At the opposite extreme, we have households headed by a person who has attained a secondary school or a College degree (**EDC** = 5, 8), with a 20 per cent combined negative contribution and 12 per cent of the population.

IV. CONCLUSIONS

Inequality decomposition methods provide a useful approach to the investigation of the structure of inequality in a given year, the explanation of the inequality trend over time, and the conection between the inequality of the population as a whole and the inequality within specific subgroups.

In this paper we have adopted a simple model in which equivalence scales depend only on household size through a parameter Θ which captures the importance we may want to give to economies of scale in consumption within the household. In this context, we have proposed a decomposition method which minimizes equivalence scales contamination problems. In the empirical

part of the paper we apply this method to household expenditures data in Spain for 1980-81 and 1990-91.

As far as the structural aspects of inequality in a given year, the main lessons obtained can be summarized as follows.

1. As in previous studies in Spain and other countries, we have confirmed that, due to the contaminated term $C(\Theta)$, the overall inequality in a given year varies substantially with the equivalence scales selected. If the equivalence elasticity Θ is chosen outside the range (0.4, 0.7), then the explanatory power of the between-group inequality term in the partition by household size could be unduly important. We find also that the trend in $C(\Theta)$ changes signs as a function of Θ, although its importance as an explanatory factor of the overall inequality trend is not large (between 1 and 9 per cent). Therefore, what requires explanation is the uncontaminated term and its changes over time.

2. It appears that, like in other countries, we have to live with the fact that the explanatory power of the between-group inequality attributable to household characteristics is, at most, between 20 to 28 per cent of uncontaminated inequality. We confirm also some patterns found elsewhere: i) the importance of demographic variables, like household type and the age of the household head, is negligible; ii) the importance of inter-regional inequities -which explains less than 10 per cent of uncontaminated inequality and about the same as the urban/rural variable- is smaller than the attention they receive in current debates on national inequality; iii) some socioeconomic variables, notably, the educational level and the socioeconomic category of the household head, provide the greatest explanatory power in both years.

The implication is clear: to understand the structure of inequality in any country we must look inside specific subgroups. Our method permits to select those groups whose contribution to within-group inequality is particularly large (or small), independently of the choice of the equivalence scales parameter Θ.

Turning now towards the inequality trend, the main conclusions are the following:

1. Overall inequality decreases in the 8 to 10 per cent range, depending on the value of the equivalence scales parameter Θ. Our decomposition method allows us to conclude that the non linear changes in the contaminated term have a low explanatory power. Most of the overall inequality trend is accounted for by a reduction of 10.3 per cent in the uncontaminated term.

2. Demographic factors in the partition by household size have a relatively small but negative effect on uncontaminated inequality. If the shift towards smaller households, also found in other Western countries, would have been the only change, then the uncontaminated inequality would have increased by a 1.4 per cent. This is offset by a 11 per cent reduction in within-group inequality. Single person and 7 or more person households contributed to this reduction well above their demographic importance in the population.

3. Once we apply the decomposition property of our inequality measure to each household size separately, in general the demographic factor associated to other partitions has also a negative effect on uncontaminated inequality. Sometimes the role of a new variable reinforces the demografic effect due to shifts in household sizes -as in the case of the socioeconomic category- while in other cases the demographic effect turns out to be negligible -as in the cases of the educational level and the number of income earners in the household.

4. It is clear that the key variables for explaining the reduction in uncontaminated inequality are the socioeconomic category and the educational level of the household head. As far as the first one is concerned, households headed by a retired person have mean expenditures below the population as a whole but have improved their relative positions regardless of their age and sex. Since they have increased their numbers, this must be the consequence of the workings of the public and universal Social Security retirement system(11). On the other hand, both agricultural categories, also below the average, have improved their lot too, although this could be partly the consequence of their diminished population share. These facts, together with the drastic fall of the upper classes, explain why this partition's total effect accounts for one third of the

change in uncontaminated inequality. This happens through the reduction of between-group inequality which more than offsets the inequality increase due to the demographic factor. The remaining two thirds of the reduction of uncontaminated inequality take place within the partition subgroups. In so far as households headed by a person retired or in the agricultural labor force contribute very heavily to this reduction of within-group inequality, it appears that the forces which lead to an increase of their mean household expenditures cause also a reduction of their inequality.

As far as the educational level is concerned, only the illiterates and those without formal studies maintain their relative positions in mean household expenditures. The remaining subgroups all loose relative ground. A fact for which, at present, we do not have a convincing explanation. Since the demographic effect does not substract much from the total effect, the reduction in the between-group inequality is the factor essentially responsable for one half of the reduction in uncontaminated inequality. The lowest educational categories are the subgroups which contribute disproportionally to the reduction in within-group inequality.

To understand fully the role of these two variables, we must recall for which subgroups inequality increases. These are, on the one hand, the subgroups headed by someone with a secondary school degree (whose numbers greatly increased over the period), and the highly educated, and on the other hand the self-employed, the upper classes, and those headed by an unemployed person.

5. How can we summarize these separate but interconnected pieces? In the first place, at the lower tail of the distribution (the retired and the agricultural categories, the aged, the illiterate and those without formal studies -all of which are well represented among one and two person households), there has been an increase in the mean household expenditures *and* an important decrease in inequality. The first fact causes an improvement in uncontaminated inequality through between-group effects; the second causes the same impact, but directly through the within-group term. In the second place, in some segments of the

upper tail of the distribution there has been a loss in relative positions -operating through the between-effects already mentioned- *and* a loss in inequality. More research on the income side is needed to know whether the final picture is one of an increase in labor and/or market earnings inequality, compensated by an increasingly generous system of public transfers which leads to the present reduction in uncontaminated inequality. That the improvement in old age and disability pensions is a key element of any explanation is already clear from the evidence presented in this paper.

6. Finally, one may ask what conection can we establish between high unemaployment rates and the reduction of inequaliyty. In this paper we have seen that the number of persons living in thouseholds headed by unemployed person has remained practically constant, amounting to 5.5 per cent of the population in 1990-91. The subgroup mean expenditures index has increased slightly from 72 to 78 during the 80's, contributing to a decrease in between-group inequality. However, this is one of the few subgroups for which inequality has increased.

On the other hand, we know that that unemployment is particularly large among the young. We have seen that households with young persons between 17 and 30 years old have experienced a considerable increase. This is in line with Revenga (1991)'s suggestion that this arrangement might have helped offset the negative effects that, otherwise, unemployment among the young could have caused in overall inequality. These households relative position -measured by the mean expenditures index- has remained stable, and their contribution to the improvement of uncontaminated inequality is not much different that their demographic weight. More about the impact of these living arrangements will be known when we study in another paper the population partition by individual, rather than household, characteristics.

NOTES

(1) For recent discussions of the issues involved in decomposition analysis, see Cowell and Jenkins (1994) and Ruiz-Castillo (1995a).

(2) See, for example, Jenkins (1995) and Borooah and Collins (1995) for the UK, Gouveia and Tavares (1995) and Rodrigues (1993) for Portugal, and Tsakloglou (1993) for Greece. In the seminal contribution to decomposition analysis, Mookherjee and Shorrocks (1982) study the evolution of household unadjusted income by age of household using published UK group data from the *Family Expenditure System*; therefore, no equivalence scale is considered at all.

(3) For Portugal, see Gouveia and Tavares (1995) and Rodrigues (1993); for Spain, Del Río and Ruiz-Castillo (1996a, 1996b); and for the international experience, see, for instance, Atkinson *et al* (1995). As a matter of fact, judging from a comparable data source, the inequality reduction in Spain is taking place since 1973-74 (cf Ruiz-Castillo (1995b).

(4) In particular, Borooah and Collins (1995) report that the McClements scales, used by Jenkins (1995) and many other authors in the UK, represent an equivalence elasticity of 0.6. According to Ruiz-Castillo (1996), the so-called OCDE scales, widely used internationally, represent an equivalence elasticity of 0.8.

(5) This has been confirmed in other countries. For Portugal, see Rodrigues (1993). For Spain, see Ruiz-Castillo (1995b) for the period 1973-74 to 1980-81, and Section II of this paper for the period 1980-81 to 1990-91.

(6) For the connection between the two separate issues -the inequality index cardinalization choice, and the definition of between-group inequality- see Cowell and Jenkins (1994).

(7) This is not the place to argue for or against expenditure versus income (See Ruiz-Castillo (1996)). However, the following studies refered to below for comparison purposes use also household expenditure as a proxy for household welfare: Gouveia and Tavares (1995) and Rodrigues (1993) for Portugal, and Tsakloglou (1993) for Greece.

(8) This is indeed the patern reported by Coulter *et al* (1992a, b) for the UK, by Rodrigues (1993) for Portugal, and by Ruiz-Castillo (1995b, 1996) for Spain in 1973-74 and 1980-81 at Winter of 1981 prices.

(9) This result is robust to different measures of a household standard of living, including total household income.

(10) Of course, any linear combination of base and current inequality values (or demographic shares) could be used to weight the terms in equations (9), (15) and (16) (or equations (8), (12) and (13), respectively).

(11) For a recent analysis of the Spanish Sociual Security system, see Boldrin et al (1997).

APPENDIX

Summary of results in the literature

In the literature, the explanation of overall inequality provided by a given characteristic is measured by the percentage that (contaminated) between-group inequality represents relative to overall (income or expenditure) inequality. For each country, a range of estimates for different years is presented in Table A. For each characteristic, the number in parenthesis is the number of subgroups in the corresponding partition. On the other hand, using several members of the Atkinson family of inequality indices, Cowell and Jenkins (1994) find that the percentage of overall US 1986 adjusted family income inequality explained by sex, race, age and earner status, ranged from 10 to 30 per cent.

For a rigorous comparison with our estimates, the explanation of the uncontaminated inequality provided by B^k in Table 4 and the explanation of overall inequality $I_0(z(\Theta))$ provided by $C(\Theta)$ in Table 2 should be added up for some appropiate Θ to reach a value comparable to the estimates in Table A. Thus, for example, the (contaminated) explanation provided by characteristics **EDC** and **HSIZE** in the Spanish case could amount to a 30 per cent of $I_0(z(\Theta))$ for values of Θ outside of the range (0.4, 10.7).

TABLE A. Percentage of overall inequality explained by several characteristics in different countries

	Portugal[a]	Greece[b]	UK[c]
HSIZE	5.6/7.4 (6)	-	-
HTYPE	-	-	14.1/17.7 (10)
HAGE	2.0/3.3 (6)	2.9/4.6 (7)	4.4/8.0 (7)
REGION	2.8/4.6 (5)	8.7/12.4 (9)	2.1/4.3 (11)
MUNS	5.5/7.8 (4)	2.9/4.6 (2)	-
SOCIO	16.7/20.7 (8)	-	13.3/17.1 (5)
EDC	21.0/27.2 (4)	18.1/22.5 (4)	-
NEARN	6.0/8.6 (5)	-	15.0/20.5 (7)

a: 1980, 1990, in Rodrigues (1993) b: 1974, 1982, in Tsakloglou (1993) c: 1971-1986, in Jenkins (1995)

Combined household characteristics: definitions

There are five of these characteristics.

1) GEO combines MUNS and AUTO as indicated: distinguishing urban from rural municipalities in each Autonomous Community, giving rise to 36 subgroups.

2) TYPAGE combines HTYPE and AGE as follows:

TYPAGE = 1 if HTYPE = 1 and AGE > 65 years
2 if HTYPE = 1 and AGE < 65 years
3 if HTYPE = 2 and AGE > 65 years
4 if HTYPE = 2 and AGE < 65 years
5 if HTYPE = 3 and AGE > 65 years
6 if HTYPE = 3 and AGE < 65 years
7 if HTYPE = 4 and AGE > 65 years
8 if HTYPE = 4 and AGE <65 years
9 if HTYPE = 5 and AGE > 30 years
10 if HTYPE = 5 and AGE < 30 years
11 if HTYPE = 6
12 if HTYPE = 7
13 if HTYPE = 8 and AGE > 65 years
14 if HTYPE = 8 and AGE < 65 years

15 if HTYPE = 8 and AGE > 65 years
16 if HTYPE = 8 and AGE < 65 years.

3) AGEDC combines AGE and EDC as follows:

AGEDC = 1 if AGE s 30 and EDC = 1, 2, 3
2 if AGE s 30 and EDC > 3
3 if 30 < AGE s 40 and EDC = 1, 2, 3
4 if30< AGE s40 and EDC> 3
5 if 40 < AGE s 54 and EDC = 1, 2, 3
6 if 40 < AGE s54 and EDC> 3
7 if 54 < AGE s 64, active, and EDC = 1, 2, 3
8 if 54 < AGE s 64, active, and EDC > 3
9 if 54 < AGE s 64, retired, and EDC = 1, 2, 3
10 if 54 < AGE s 64, retired, and EDC > 3
11 if 64 < AGE, male, and EDC = 1, 2, 3
12 if 64 < AGE, male, and EDC > 3
13 if 64 < AGE, female, and EDC = 1, 2, 3
14 if 64 < AGE, female, and EDC > 3.

4) SOCEDC combines SOCIO and EDC as follows:

SOCEDC = 1 if SOCIO = 1 and EDC = 1, 2, 3
2 if SOCIO = 1 and EDC > 3
3 if SOCIO = 2 and EDC = 1, 2, 3
4 if SOCIO = 2 and EDC > 3
5 if SOCIO = 3, 5, 7, and EDC = 1, 2, 3
6 if SOCIO = 3, 5, 7 and EDC > 3
8 if SOCIO = 4 and EDC > 3
9 if SOCIO = 6 and EDC = 1, 2, 3
10 if SOCIO = 6 and EDC > 3
11 if SOCIO = 8, 9 and EDC = 1, 2, 3
12 if SOCIO = 8, 9 and EDC > 3
13 if SOCIO = 10 and EDC = 1, 2, 3
14 if SOCIO = 10 and EDC > 3
15 if SOCIO = 11, 12, 13, 14, and EDC = 1, 2, 3
16 if SOCIO = 11, 12, 13, 14, and EDC > 3.

5) NEADC combines NEARN and EDC as follows:

NEADC= 1 if NEARN = 1
2 if NEARN = 2 and = 1, 2, 3
3 if NEARN = 2 and EEDC > 3
4 if NEARN = 3 and EDC = 1, 2, 3
5 if NEARN = 3 and EDC > 3
6 if NEARN = 4, 5 and EDC = 1, 2, 3
7 if NEARN = 4, 5 and EDC > 3
8 if NEARN= 6 and EDC = 1, 2, 3
9 if NEARN = 6 and EDC > 3
10 if NEARN = 7 and EDC = 1, 2, 3
11 if NEARN = 7 and EDC > 3
12 if NEARN = 8 and EDC = 1, 2, 3
13 if NEARN = 8 and EDC > 3
14 ifNEARN = 9, 10, 11andEDC=1,2,3
15 if NEARN = 9, 10, 11 and EDC > 3
16 if NEARN = 12 and EDC = 1, 2, 3
17 if NEARN = 12 and EDC > 3

TABLE B. Household characteristics. Percentage change of households (H) and persons (P) in all subgroups from 1980-81 to 1990-91. In % relative to the 1980-81 distribution

Percentage changes for the population as a whole: 12.3 for households and 3.84 for persons

HSIZE	H	P	HTYPE	H	P	AGE	H	P	AUTON	H	P
1 member	44.9	44.9	1	44.8	44.7	Upto24	-34.3	-38.1	Andalucía	17.5	8.1
2	19.3	19.3	2	19.1	19.1	25-34	-2.3	-11.2	Aragón	9.6	2.1
3	25.8	25.8	3	-4.8	-5.1	35-44	8.4	-2.4	Asturias	4.3	-0.3
4	19.3	19.3	4	68.9	75.4	45-54	-8.5	-11.2	Baleares	3.9	3.6
5	0.2	0.2	5	-9.2	-15.9	55-64, act.	-9.2	-15.9	Canarias	22.7	8.4
6	-20.6	-20.6	6	5.2	-1.3	55-64, ret.	86.9	96.0	Cantabria	10.6	3.9
7 or +	-40.8	-42.0	7	-11.4	-15.6	65-74, males	35.9	37.6	C.-León	10.1	1.8
			8	37.8	27.7	65-74, females	41.9	42.7	C.-La M.	12.7	4.1
			9	9.1	9.1	75 or +, males	53.4	45.7	Cataluña	11.1	0.6
						75 or+,females	73.7	70.0	Co. Valen.	13.3	4.2
SOCIO	H	P	EDC	H	P	NEARN	H	P	Extremad.	15.5	6.4
1	-29.3	-35.3	1	-31.9	42.1	1	-17.6	-21.0	Galicia	7.4	0.4
2	-45.6	47.6	2	-3.1	-18.2	2	27.6	12.7	Madrid	16.1	5.4
3	-7.2	-10.5	3	-9.4	-15.8	3	121.5	113.3	R. Murcia	13.8	9.0
4	20.3	13.2	4	135.3	133.1	4	-55.7	-55.2	Navarra	12.8	2.7
5	54.8	44.6	5	70.2	59.4	5	-58.0	-58.6	PaísVasco	11.2	-0.1
6	46.0	33.7	6	290.2	260.1	6	-29.7	-35.1	La Rioja	4.6	1.3
7	6.9	-5.0	7	54.8	41.3	7	35.0	22.0	Ceuta y	10.1	8.3
8	-41.9	-38.1	8	46.7	38.8	8	27.9	18.8	Melilla		
9	15.6	1.7	MUNS	H	P	9	119.0	96.3			
10	64.6	61.0	< 2,000	-27.1	-34.0	10	65.5	52.3			
11	51.7	51.5	2-10,000	15.3	5.0	11	17.8	5.1			
12	71.5	65.9	10-50,000	17.8	6.9	12	115.5	92.8			
13	-35.0	-28.0	50-500,000	22.8	14.5						
14	12.2	10.6	> 500,000	12.7	2.9						

HSIZE (household size). AGE (household head's age). MUNS (municipality size). AUTON (Autonomous Community). HTYPE (household type): 1. Single person and minors-only (up to 17 years old), 2. Couples, 3. Three or more adults without children, 4. Couples with young (18-30 years old) descendants, 5. Couples with minors, 6. Couples with minors and young descendants, 7. Couples with minors and/or young descendants, as well as older people (65 or more years), 8. Single parent households, 9. Households without couples with unrelated minors. EDC (household head's educational level): 1. Illiterate, 2. Without formal studies, 3. Grade school, 4. Primary school, 5. Secondary school, 6. Vocational school, 7. Three years College degree, 8. More than three years - College degree. SOCIO (household head socioeconomic category): 1. Workers, agricultural, 2. Self employed, agricultural, 3. Non-agricultural workers, 4. Non-agricultural self employed, 5. Supervisors, and Armed Forces, 6. Upper classes, 7. Unclassifiable persons in the labor force, 8. Part-time (less than 1/3 working day), 9. Unemployed, 10. Early retired (below 65 years of age), 11. Retired, male, 12. Retired, female, 13. Living off property income, 14. Other category outside of the labor force. NEARN (Number of earners of labor and non-labor income in the household): 1. No income earners, 2. One non-labor income earner, 3. Two or more non-labor income earners, 4. Only one earner from part time work, 5. Two or more earners, none of them from full time work, 6. One full time labor income earner, 7. Two earners, one of them full time, 8. Two full time labor income earners, 9. Three earners, one of them full time, 10. Three earners, two of them full time, 11. Three full time income earners, 12. Four or more earners of any kind.

TABLE C. Mean adjusted household expenditure for Θ = 0.4 for the different subgroups in all partitions, in index form with a value of 100 for the population as a whole.

Column (1) = 1980-81, Column (2) = 1990-91

HSIZE	(1)	(2)	HTYPE	(1)	(2)	AGE	(1)	(2)	AUTON	(1)	(2)
1member	77	82	1	77	82	Up to 24	99	87	Andalucía	85	85
2	90	89	2	90	89	25-34	104	97	Aragón	101	91
3	102	101	3	96	95	35-44	99	100	Asturias	96	103
4	105	107	4	113	115	45-54	106	110	Baleares	105	109
5	102	101	5	100	97	55-64, act.	108	112	Canarias	90	93
6	98	98	6	106	104	55-64, ret.	86	92	Cantabria	113	94
7 or +	99	92	7	99	100	65-74, male	85	90	C.-León	92	90
			8	95	96	65-74, females	72	78	C.-La M.	75	84
			9	90	90	75 or +, males	69	72	Cataluña	110	119
						75 or+,females	59	65	Co. Valen.	97	89
SOCIO	(1)	(2)	EDC	(1)	(2)	NEARN	(1)	(2)	Extremad.	68	72
1	69	74	1	60	62	1	65	71	Galicia	96	91
2	76	81	2	76	76	2	69	72	Madrid	125	128
3	107	100	3	97	91	3	72	77	R. Murcia	93	87
4	103	109	4	123	102	4	76	88	Navarra	123	121
5	120	130	5	152	128	5	77	76	PaísVasco	120	113
6	175	154	6	126	114	6	103	99	La Rioja	97	96
7	106	108	7	160	144	7	97	98	C. y Meli.	93	75
8	83	101	8	197	188	8	122	123			
9	72	78	MUNS	(1)	(2)	9	93	94			
10	81	90	< 2,000	81	78	10	110	110			
11	76	84	2-10,000	80	84	11	114	121			
12	66	73	10-50,000	91	92	12	111	110			
13	104	131	50-500,000	109	104						
14	94	78	> 500,000	126	127						

REFERENCES

Atkinson, A. B., L. Rainwater and T. Smeeding (1995), *Income Distribution in OECD Countries: the Evidence from the Luxembourg Income Study (LIS)*. Paris: OECD.

Boldrin, M., S. Jimenez, F. Peracchi, and J Ruiz-Castillo,"Social Security and Retirement in Spain", forthcoming in D. Wise and J. Gruber (eds), *Social Security : An International Perspective*, NBER and University of Chicago Press.

Borooah, V. and G. Collins (1995), "Was There a Regional Dimension to Changes in Income Inequality in the UK over 1982-1992? An Analysis Based on a Joint Decomposition of Income by Region and Employment Status," The Microsimulation Unit, Department of Applied Economics, University of Cambridge, DAE Working Papers Number MU 9507.

Buhmann, B., L. Rainwater, G. Schmauss and T. Smeeding (1988), "Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study Database,"Review *of Income and* Wealth, **34**: 115-142.

Coulter, F., F. Cowell and S. Jenkins (1992a), "Differences in Needs and Assessment of Income Distributions," *Bulletin of Economic Research,* **44**: 77124.

Coulter, F., F. Cowell and S. Jenkins (1992b), "Equivalence Scale Relativities and the Extent of Inequality and Poverty," *Economic Journal,* **102**: 1067-1082.

Cowell, F. (1984), "The Structure of American Income Inequality," *Review of Income and Wealth,* **30**: 351-375.

Cowell, F. and S. Jenkins (1994), "How Much Inequality Can We Explain? A Methodology and an Application to the USA," Distributional Analysis Research Programme, London School of Economics, Discussion Paper, No. DARP-7.

Cowell, F. A. and K. Kuga (1985), "Inequality Measurement: An Axiomatic Approach," *European Economic Review,* **15**: 287-305.

Danziger, S. and M. Taussing (1979), "The Income Unit and the Anatomy of the Income Distribution," *Review of Income And Wealth,* **79**: 365-375.

Del Rio, C. and J. Ruiz-Castillo (1996a), "Ordenaciones de bienestar e inferencia estadística. El caso de las EPF de 1980-81 y 1990-91", in *"La desigualdad de recursos. Segundo Simposio sobre la distribución de la renta y la riqueza",* Fundación Argentaria, Colección Igualdad. Volumen VI, 9-44.

Del Rio, C. and J. Ruiz-Castillo (1996b), "Intermediate Inequality and Welfare. The Case of Spain, 1980-81 to 1990-91", Universidad Carlos III de Madrid, Working Paper 96-03, Economic Series 03.

Jenkins, S. P. (1995), "Accounting for Inequality Trends: Decomposition Analyses for the UK, 1971-86," *Economica*, **62**: 29-63.

Gouveia, M. and J. Tavares (1995), "The Distribution of Household Income and Expenditure in Portugal: 1980 and 1990," *Review of Income and Wealth*, **41, 1**: 1-17.

Mookherjee, D. and A. Shorrocks (1982), "A Decomposition Analysis of the Trend in UK Income Inequality," *Economic Journal*, **92**: 886-902.

Revenga, A. (1991), "La liberalización económica y la distribución de la renta: la experiencia española," *Moneda y Crédito*, **193**: 179-224.

Rodrigues, C. F. (1993), "Measurement and Decomposition of Inequality in Portugal, 1980/81-1990/91," Department of Applied Economics, University of Cambridge, E.S.R.C. Discussion Paper, MU 9302.

Ruiz-Castillo, J. (1995a), "Income Distribution and Social Welfare: A Review Essay," *Investigaciones Económicas*, **XIX**: 3-34.

Ruiz-Castillo, J. (1995b), "The Anatomy of Money and Real Inequality in Spain, 1973-74 *to* 1980-81," *Journal of Income Distribution*, **5**: 265-281.

Ruiz-Castillo (1996), "A Simplified Model for Social Welfare Analysis. An Application to Spain, 1973-74 to 1980-81," Universidad Carlos III, Working Paper, 96-04, Economic Series 04.

Shorrocks, A. F. (1984), "Inequality Decomposition by Population Subgroups," *Econometrica*, **52**: 1369-1388.

Shorrocks, A. F. (1995), "On Income Distribution Comparisons for Heterogeneous Populations," University of Essex, Discussion Paper Series, No. 447.

Tsakloglou, P. (1993), "Aspects of Inequality in Greece. Measurement, Decomposition and Intertemporal Change: 1974, 1982," *Journal of Development Economics*, **40**: 53-74.