DISCUSSION PAPER SERIES

IZA DP No. 2512

Macroeconomic Conditions and the Distribution of Income in Spain

Lídia Farré-Olalla
Francis Vella

December 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Macroeconomic Conditions and the Distribution of Income in Spain

Lídia Farré-Olalla
University of Alicante

Francis Vella
Georgetown University
and IZA Bonn

Discussion Paper No. 2512
December 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 2512
December 2006

ABSTRACT

Macroeconomic Conditions and the Distribution of Income in Spain*

This paper analyzes the impact of changes in macroeconomic conditions on the income distribution in Spain. Using household data from the Encuesta Continuada de Presupuestos Familiares (ECPF) from 1985 to 1996, we disentangle the effect of aggregate variables on the income distribution by estimating counterfactual densities conditional on different macroeconomic scenarios. Our empirical approach allows for a flexible relationship between the income level and two constructed indices. The first index captures the influence of individual characteristics while the second captures the role of macroeconomic variables. The contribution of each of these variables to their respective indices is estimated by a semiparametric least squares procedure. We find that although inequality displays a decreasing trend over the earlier part of the period examined, the poor performance of the Spanish economy during the early 1990's appears to have reversed this trend. We also conclude that while inflation appears to have no impact on the distribution of income for the period examined, there were important redistributive roles for unemployment, government expenditure and the level of GDP.

JEL Classification: D31, E32, C14

Keywords: income distribution, aggregate fluctuations, semiparametric estimation

Corresponding author:

Francis Vella
Department of Economics
Georgetown University
Washington, DC 20057
USA
E-mail: fgv@georgetown.edu

* We are very grateful to Roger Klein for his input and many conversations regarding estimation of the empirical model. We have also benefited from the comments of Omar Licandro, Lisa Lynch and Pedro Mira. This paper was written while the authors were members of the Economics Department of the European University Institute.

1 Introduction

The manner in which aggregate shocks and macroeconomic policies are associated with changes in the income distribution is an important economic question as the shape and location of the income distribution affect many aspects of economic behavior. These include taxation revenues, welfare benefit payments and individual and household consumption and saving patterns. Accordingly, quantifying how changes in the income distribution systematically covary with movements in the "economic environment" is an important objective for economic policy. However, just as the macroeconomic environment is relevant for the income distribution (Blank et al. 1993; Cutler et al. 1991; Blank and Blinder 1985), it is well known that the characteristics of the individuals comprising the economy also have implications for the distribution of income (Juhn, Murphy and Pierce 1993). While previous studies have generally examined either the relationship between the income distribution and macroeconomic conditions, or that between the income distribution and the characteristics of the individuals in the economy, it is likely that the manner in which the individuals' characteristics interact with the economic environment also influences the distribution. For example, if certain worker types fare better (worse) under certain macroeconomic scenarios, then there is scope for changes in aggregate conditions to affect the distribution of income. This paper addresses this issue empirically by treating the level of income as a function of macroeconomic as well as individual factors. We employ a semiparametric procedure which models the income level as a function of both a weighted combination of macro variables and a weighted combination of individual characteristics. Our procedure also allows for a flexible interaction between these two weighted combinations. Having estimated the relationship between the individual in-

come level and these macro and individual variables, we can produce estimates of the income distribution conditional on the values of these variables. Moreover, we can also generate hypothetical distributions by varying the values of the macro variables or the individual characteristics.

An exercise of this kind would be most purposefully conducted over a period of substantial economic change. Accordingly, this paper uses Spanish household data taken from the Encuesta Continuada de Presupuestos Familiares (ECPF), over the period 1985-1996. The latter 1980's was a period of rapid economic growth, following Spain's entry into the European Community in 1986. However, the expansionary phase ended with a severe recession in the early 1990's which was accompanied by high levels of unemployment. During this period there were also notable changes in the income distribution. While the mean income level has increased since the first half of the 1980's, the level of income inequality decreased during the 1980's before increasing with the onset of the recession at the beginning of the 1990's. These movements suggest a link between macroeconomic activity and the income distribution.

Previous studies have examined the implications of macroeconomic conditions for the income distribution. One branch of research has focused on the extent to which different socioeconomic groups benefit from economic growth. Despite these efforts, it has proved difficult to establish a theoretical link between the distribution of income and economic well-being (Alesina and Rodrick 1994 and Galor and Zeira 1993). However, the empirical evidence has suggested some regularity. For example, Barro (2000) finds that the Kuznets curve emerges in a broad panel of countries. Dollar and Kraay (2002) also emphasize the importance of economic growth for poverty reduction, as they find that economic growth, more generally, relatively benefits the

poor. Other authors have examined the cyclical behavior of the income distribution. Among these, Blinder and Esaki (1978) and Blank and Blinder (1985) studied the effects of inflation and unemployment on the US income distribution after the 1950's. They conclude that while inflation has little effect on the distribution, unemployment exerts large and negative effects on the lower tail. As a result, the distribution of income widens in recessions and narrows during expansions. The theoretical work by Castañeda et al. (1998) also emphasizes the importance of unemployment for the cyclical behavior of income. They show that allowing for different unemployment spells across income groups is the most important factor to replicate the dynamics of the income distribution over the business cycle. Previous findings thus indicate that, as a result of different individuals' sensitivities to economy-wide shocks, macroeconomic activity affects not only the mean but also higher moments of the income distribution.

Despite the importance of heterogeneity in individuals' responses for the transmission of macroeconomic shocks, previous empirical work has taken an aggregate approach. Most of this research examines the reduced-form relationship between key economic indicators (i.e. unemployment, inflation, government expenditure or GDP growth) and some statistics of the distribution (i.e. the Gini coefficient or the income shares of the different quintiles). In this paper, we propose a more disaggregated procedure to quantify the effect of macroeconomic activity on the income distribution. The novelty is that we estimate the level of income conditional on different macroeconomic scenarios, allowing the effect of aggregate variables to vary across socioeconomic groups. Accordingly, we model the conditional expectation of an individual's income as an unspecified function of two indices. The first index is a linear

combination of individual characteristics, while the second is a linear combination of macroeconomic features. We then estimate the level of income as an unrestricted function of these two indices by adapting the semiparametric multiple index least squares estimator of Ichimura and Lee (1991). This approach provides estimates of the weighting coefficients that measure the contribution of each explanatory variable to its respective index. It also estimates the function which maps the two indices into individual income.

To disentangle the contribution of each aggregate variable to changes in the distribution of income we use counterfactual distributions. More precisely, we compare the actual income distribution with the hypothetical distribution that would have resulted if any particular variable had taken a different value while holding the others constant. The results indicate that macroeconomic activity has an important distributive impact. More explicitly, we find that the increase in the level of government expenditure and GDP during the 1980's and the 1990's had a significant positive association with the mean of the distribution. In contrast, the high unemployment rates in the first half of the 1990's increased the standard deviation and other measures of income dispersion. We do not find any statistically significant distributional effect for inflation, although this may partially reflect the absence of variability in inflation rates over the period.

We highlight that this paper does not aim to identify a causal relationship between macroeconomic variables and the level of individual income. That is, we treat the various macroeconomic scenarios that occur in the period studied as exogenous to the income distribution. This is a restrictive assumption as one can easily foresee situations where the income distribution and macroeconomic activity are jointly

determined. While incorporating such a relationship and exploring its implications for our results would be a worthwhile and very challenging project, we feel that our more limited approach provides useful insights and it is an important first step in addressing the more complicated scenario.

The rest of the paper is organized as follows. Section 2 summarizes the main changes that characterized the Spanish economy in the 1980's and the 1990's. Section 3 introduces the empirical model and the methodology and Section 4 describes the data. The results are presented in Section 5, followed by some concluding remarks in Section 6.

2 Macroeconomics and the Distribution in Spain

The substantial changes in the Spanish economy during the 1980's and 1990's provide an interesting scenario to analyze the empirical relation between macroeconomic activity and the income distribution. Compared to other OECD countries, the beginning of the 1980's were years of an economic slowdown in Spain, where the negative effects of the oil shocks (1973, 1979) were particularly severe. Spain had a heavily energy-dependent industry and this was exacerbated by the political instability after Franco's regime which made it difficult to implement appropriate economic measures.

After Spain joined the European Community in 1986 the economy started to recover. Figure 1 presents the evolution of the level and the growth rate of GDP and provides a clear picture of the main economic features of the period. The admission to the European Community coincided with an expansionary phase of the international economy that contributed to the positive economic performance of Spain. From 1986 to 1990, GDP grew at an average rate of 3.7 per cent. Positive short-term

expectations, reinforced by a higher degree of political stability, attracted foreign investors. Moreover, since there was a wide technological gap between Spain and the other industrialized countries in Europe, foreign investment also brought technological innovation. This transfer had an immediate positive effect on the labor market, and the unemployment rate fell from 21.45% in 1985 to 16.23% in 1989 (see Figure 2). These favorable conditions, together with controlled inflation rate (see Figure 2), resulted in an over performance of the Spanish economy in the context of the OECD.

Nevertheless, the international recession in the early 1990's strongly reduced economic activity in Spain. To gain credibility after joining the European Exchange Rate Mechanism (ERM), the Spanish authorities implemented a restrictive monetary policy aimed at decreasing the national level of prices. The strategy was successful for the tradable goods sector, but not for the non-tradable sector that lost competitiveness due to the low productivity of the national firms operating with obsolete technology. Low labor mobility and high wage indexation were major obstacles to the labor market's ability to adjust to the decreasing demand, and this resulted in a large increase in the unemployment rate. Slackening GDP growth, a substantial public and external deficit, and an unemployment rate above 24% were the main features of the Spanish economy in the first half of the 1990's.

Another remarkable fact of the period is that government spending and, in particular, social expenditure programs rose faster in Spain than in the European Community as a whole. This increase was mainly financed by a rise in taxation, which itself had a redistributive impact because 52% of all income tax revenues were paid by the top 10% of the income-earners, and 20% by the top 1% (Maravall 1997). To evaluate the redistributive effects of the increase in social expenditure programs Table 1 presents

the percentage of GDP assigned to each program and its annual growth rate. From this Table, the most relevant feature is the significant change in the internal structure of social spending. The late 1980's were characterized by an increase in the amount of resources aimed at improving the economic well-being of the most disadvantaged, with a large increase in public spending assigned to sickness and survivor benefits. However, at the beginning of the 1990's public spending on these social programs was substantially reduced in favor of an increase in unemployment subsidies and pensions programs which do not necessarily benefit the poorest members in the society.

The structural change in the Spanish economy, resulting from joining the European Community, combined with the changes in the internal composition of government expenditure and the high unemployment rates which characterized the early years of the 1990's, may have influenced the relative position of individuals in the income distribution. Table 2 shows some descriptive statistics to examine this possibility and reveals that the log income mean substantially increased between 1985 and 1992 (from 12.81 to 13.12). However, with the advent of the recession at the beginning of the 1990's, the mean failed to grow for the remainder of the period. Thus, the evolution of this variable seems to be strongly related to the growth path of the economy.

Table 2 also reports the time series of both the standard deviation and Gini coefficients for log income.[1] These two measures indicate that the dispersion in income decreased during the 1980's. This trend, which goes in the opposite direction to the evolution of inequality in most other OECD countries, might be the result of the increase in economic resources and the substantial redistributive role of the public

[1] See Bourguignon (1979) for a detailed survey of inequality measures.

sector in the 1980's. However, Figure 3 plots the time series of these measures and indicates that the decreasing trend in income inequality was reversed when the economic activity started to decelerate. The high unemployment rates that characterized the economic crisis, and the variations in the composition of public spending during the first half of the 1990's, may therefore be responsible for the change in the trend of income dispersion.

3 Econometric Strategy

Previous work on the distributional dimension of macroeconomic activity indicates that different individuals may have different responses to changes in the economic environment. However, the manner in which aggregate variables combine with individual characteristics in determining income levels is likely to be complicated, and economic theory provides limited guidance to the empirical link between them. Accordingly, we model the level of individual income as an unconstrained function of two indices. The first index captures the effect of individual characteristics (i.e. age, gender, education), while the second captures the role of aggregate shocks and policies (i.e. the unemployment rate, the level of public expenditure). In estimation, we employ the semiparametric least squares estimator of multiple index models by Ichimura and Lee (1991). The level of income is modeled as follows:

$$y_{it} = h(I_{1it}, I_{2t}) + e_{it}, \tag{1}$$

where y_{it} denotes the level of income of individual i in period t, I_{1it} is the index of individual characteristics and I_{2t} is the index of aggregate variables which indicates

the state of the economy. The function $h(.,.)$ is unknown and captures how the two arguments of the function map into the individual level of income. The e_{it} is a zero-mean error term which is assumed to be uncorrelated with the variables in the two indices. The index I_{1it} can be represented as a linear combination of individual characteristics in the vector X_{it}; while the index for the state of the economy, I_{2t}, can be expressed as a linear combination of aggregate variables capturing the trend of the economy and the deviations from this trend, contained in the vector Z_t.

That is:

$$I_{1it} = \beta' X_{it}, \; I_{2t} = \alpha' Z_t, \tag{2}$$

where β and α are vectors of unknown coefficients.

If we assume a linear functional form for $h(.,.)$ and a strictly additive relation between the two indices, the α and β coefficients could be estimated by Ordinary Least Squares.[2] The semiparametric multiple index estimator does not specify a functional form for $h(.,.)$ and estimates α and β by replacing $E[y_{it}|\beta' X_{it}, \alpha' Z_t]$ with its nonparametric estimate when minimizing the following objective function:

$$Q(\beta, \alpha) = \sum_{t=1}^{T} \sum_{i=1}^{N} \tau_{it} \{y_{it} - E[y_{it}|\beta' X_{it}, \alpha' Z_t]\}^2 \tag{3}$$

where τ_{it} is a trimming function.[3]

[2] This approach was adopted by Keane and Moffitt (1988), who added to a traditional Mincerian equation the aggregate unemployment rate to analyze the cyclical properties of wages.

[3] In estimating semiparametric models trimming is required when the density of the data is low at these observations. Following the requirements outlined by Klein and Vella (2006), we introduce a trimming function, τ_{it}, that places zero weight on observations below 5% and above 95% on the basis of the distribution of X and Z.

The conditional expectation $E[y_{it}|\beta' X_{it}, \alpha' Z_t]$ is estimated using kernel functions. That is:

$$E[y_{it}|\beta' X_{it}, \alpha' Z_t] = \frac{\sum_{l=1}^{T}\sum_{j=1}^{N} h_{jl}^{-2} K\left(\frac{\beta' X_{it}-\beta' X_{jl}}{h_{jl}}, \frac{\alpha' Z_t-\alpha' Z_l}{h_{jl}}\right) y_{jl}}{\sum_{l=1}^{T}\sum_{j=1}^{N} h_{jl}^{-2} K\left(\frac{\beta' X_{it}-\beta' X_{jl}}{h_{jl}}, \frac{\alpha' Z_t-\alpha' Z_l}{h_{jl}}\right)}. \tag{4}$$

where $K(.)$ is a bivariate normal kernel and h_{jl} is the window width. While Ichimura and Lee (1991) use higher order kernels to reduce the bias associated with the kernel estimator, we follow Klein and Vella (2006) who find that, in the case of the double index model, the finite sample performance of the estimator is significantly improved if local smoothing is used as a bias reduction technique.[4]

With the estimated parameters, $\hat{\beta}$ and $\hat{\alpha}$, one can compute the expected level of income for each individual as a function of his/her characteristics and the state of the economy. This is given by the conditional expectation:

$$\hat{E}[y_{it}|\hat{\beta}' X_{it}, \hat{\alpha}' Z_t] = \hat{E}[y_{it}|\hat{I}_{1it}, \hat{I}_{2t}] = \hat{y}_{it} \tag{5}$$

where $\hat{E}[.]$ denotes the kernel estimate. With the $\hat{y}'_{it}s$ an estimate of the income distribution can be produced as,

$$\hat{f}_t(\hat{y}_{it}) = \frac{1}{N}\sum_{j=1}^{N} K(\frac{\hat{y}_{it}-\hat{y}_{jt}}{h}). \tag{6}$$

[4]With local smoothing, the window width in the final estimation of $E[.]$ varies in each observation and depends on a pilot density estimator. Klein and Vella (2006) employ bivariate kernels that depend on an estimated sample covariance matrix. They orthogonalize the column vectors I_1 and I_2 and then estimate their joint density estimator as the product of two independent kernels. The same methodology is employed in this paper.

Using the predicted coefficients and the estimated relationship between the indices, the distributional effect of macroeconomic activity can be evaluated by comparing the actual distribution of income with the hypothetical distribution that would have prevailed at some alternative value for I_2. Accordingly, we compute first the empirical income distribution at the estimated values of $\hat{I}_{1it}$ and $\hat{I}_{2t}$. Then, keeping $\hat{I}_{1it}$ constant, we replace any element in $\hat{I}_{2t}$ by its value of interest to obtain the alternative index $\hat{I}^c_{2t}$. Next, the predicted level of individual income in the counterfactual scenario is estimated using the original empirical distribution. More explicitly:

$$\hat{E}^c[y_{it}|\hat{I}_{1it},\hat{I}^c_{2t}] = \frac{\sum_{l=1}^{T}\sum_{j=1}^{N} h_{jl}^{-2} K\left(\frac{\hat{I}_{1it}-\hat{I}_{1jl}}{h_{jl}}, \frac{\hat{I}^c_{2t}-\hat{I}_{2l}}{h_{jl}}\right) y_{jl}}{\sum_{l=1}^{T}\sum_{j=1}^{N} h_{jl}^{-2} K\left(\frac{\hat{I}_{1it}-\hat{I}_{1jl}}{h_{jl}}, \frac{\hat{I}^c_{2t}-\hat{I}_{2l}}{h_{jl}}\right)} = \hat{y}^c_{it}. \tag{7}$$

On the basis of these estimated distributions one can then employ the usual techniques to evaluate changes in the distribution of income.

4 The Data

The Spanish household survey (ECPF, 1985-1996) collects information on income, consumption and socioeconomic characteristics for a sample of 2,000 households each year, interviewed on a quarterly basis. This is a rotating panel where households are interviewed a maximum of eight quarters. To analyze the distributional aspects of macroeconomic activity, it would be desirable to have a data set with a longer time series dimension to observe more variation in the aggregate variables. However, as noted above, the performance of the Spanish economy changed considerably over the period covered by the ECPF.

The ECPF was originally constructed to collect information on household expen-

ditures in order to compute the Consumer Price Index. It also contains information on household income from which we construct a measure for quarterly net disposable income, which includes labor market earnings (i.e. wages and salaries net of taxes) and various transfer payments (for example, unemployment insurance and workers' compensation). In the survey, income and related questions refer to previous quarter's information, and income reported in the first and third quarter of the year includes extra salaries.[5] Hence, to prevent extra salaries from distorting the results we only use the information reported during the second quarter. The information on income in nominal pesetas is deflated to year 1992 prices using the CPI.[6]

Excluded from the sample are observations where the household head is retired or out of the labor force. Self-employed workers are also excluded, as these incomes are found to be systematically misreported.[7] Our final sample contains 21.097 observations over the period 1985-1996, which represents approximately 1.750 observations each year. Table 3 presents summary statistics for the variables used in the empirical analysis. Note that the unit of observation is the household head and the dependent variable we employ is equivalent income, which results from dividing total household income by an equivalence scale that takes into account the number of member in the household.[8] In the empirical analysis, male and partner income recipient are dummy variables set equal to 1, respectively, if the household head is a male and his/her partner receives any source of labor income. Education is measured as the highest

[5] Most employment contracts in Spain include a bonus payment equivalent to a month's salary paid twice a year (at Christmas and in July before the summer vacation).

[6] Note that in January 1999 the peseta was set such that one euro was equal to 166.386 pesetas.

[7] See Oliver, Ramos and Raymond-Bara (2001).

[8] We use the equivalence scale estimated by the OECD for Spain during the period of study (Oliver et al. 2001).

grade completed and converted to four dummy variables: less than high school, high school, some college and college or more. The size of the city in the ECPF takes four different values and is converted to the following dummy variables: town_1 (less than 1,000 inhabitants), town_2 (between 1,000 and 10,000), town_3 (between 10,000 and 50,000) and town_4 (more than 50,000). In estimating the model we employ town_4 and less than high school graduates as the excluded groups.

To examine the distributional effect of macroeconomic activity we compare the distribution of income in 1989 and 1994. The consequences of the Spanish expansion and recession are clearly visible in these two years. After joining the European Community in 1986, the economy started growing and in 1989 it possessed the highest rate of GDP growth and the lowest level of unemployment over the period. By contrast, the unemployment rate in 1994 was above 24% as a result of the crisis that hit the economy in the first half of the 1990's. The density of log income for these two years is plotted in Figure 4. When comparing these distributions, two major issues deserve attention. First, the mean shifted to the right over the period. Second, the distribution in 1994 appears more dispersed. Both of these movements are confirmed by the descriptive statistics in Table 4.

We investigate the distributional effect of four main aggregate variables. The unemployment and inflation rates are widely used in studies of earnings as indicators of the economic cycle. An important redistributive role is found for unemployment as it is unequally distributed across the population, with higher unemployment rates among lower-wage workers (Blank 1989). In contrast, inflation has generally not been found to exert an insignificant effect on the distribution (Blinder and Esaki 1978, and Blank and Blinder 1985). Accordingly, we include in the model the quarterly rate of

civilian unemployment, which increased 7 percentage points between 1989 and 1994, and the quarterly rate of inflation which remained almost constant over the period. Table 5 presents the change between 1989 and 1994 in all the aggregate variables included in the empirical investigation.

Political economy models also suggest that redistributive policies are a powerful instrument to affect the distribution of income (Perotti 1993). However, the effects depend on the nature of the spending programs and the tax financing system (Blejer and Guerrero 1991). During our sample period, government spending increased and its composition substantially changed. To investigate the distributional effects associated to changes in this aggregate we include the log of government expenditure in the empirical model. Since the level of GDP increased significantly over the period, the log of GDP is also included in the estimation to control for a possible trend in the evolution of individual income.

Movements in the income distribution can also be related to variations in the characteristics of individuals comprising the economy. However, the period of time considered here is relatively short and demographic shifts are not expected to be important. Table 5 displays the average of individual socioeconomic characteristics for 1989 and 1994. These statistics indicate small changes between the two years. However, the possibility that these factors influenced the distribution of income is also considered in the next section.

5 Model Specification and Results

5.1 Specification of the Individual and Aggregate Indices

In estimating the coefficients corresponding to the model in (1) the variables in the respective indices are the following:

$$I_{1it} = \beta_1 * gender + \beta_2 * age + \sum_{k=3}^{5} \beta_k * education\ level + \sum_{k=6}^{8} \beta_k * city\ size + \beta_9 * partner\ income\ recipient$$

$$\begin{aligned} I_{2t} &= \alpha_1 * unemployment + \alpha_2 * inflation + \alpha_3 * \log(government\ expenditure) + \\ &\quad \alpha_4 * \log(GDP). \end{aligned}$$

Semiparametric estimation of (1) requires normalizations of location and scale to identify the model coefficients. Location is fixed by excluding the constant term in both indices. The scale is normalized by fixing the coefficient for "gender" in the first index and that for "unemployment" in the second to 1.[9] Ichimura and Lee (1991) note that identification requires that each index contains at least one variable not included in the other index. Here, the level of income is a function of an "individual characteristics" index and an "aggregate variables" index and the exclusion restrictions are then satisfied. For the sake of computation, we standardize all the variables in the model to have mean zero and standard deviation 1.

Table 6 presents the estimated coefficients obtained by the semiparametric procedure discussed in Section 3.[10] The estimated standard errors appear in parentheses.

[9] This applies to any semiparametric procedure that involves kernel estimators.

[10] In estimation, we use as starting values for the maximization problem the estimated OLS coef-

These coefficients capture the weight or the contribution of each explanatory variable to the value of the index. Given the employed normalizations, the coefficients can only be interpreted in relative terms. That is, if the unemployment rate and the log of government expenditure increased by one standard deviation (noting that after standardizing the variables all have mean zero and standard deviation equal to unity) the contribution of the log government expenditure to the change in the value of the second index would only be 30% of the contribution of the unemployment rate (noting that it would also have the opposite sign). The absolute value of the contribution of the inflation rate would be even smaller, with only 15% of the contribution of the unemployment rate. Finally, the log of GDP has a weight in the second index similar to the unemployment rate, but with the opposite sign. The same type of analysis applies for the index of individual characteristics.

It is important to note that while the sign of the estimated coefficients indicates the direction of the contribution of each explanatory variable to the value of the index, the sign is not informative about the direction of the effect on the level of individual income. Statements concerning the sign of the coefficients are only possible if one is willing to make some assumptions on the functional form of $h(.)$. Without predetermined assumptions on the form of $h(.)$, the effect of each explanatory variable on the level of individual income can still be quantified using counterfactual distributions. Accordingly, we employ the coefficients and the estimated relationship between the two indices and the level of individual income, to obtain the distribution of income conditional on different values of the explanatory variables. By comparing

ficients. The programs used for estimation are adapted from those used in the simulation evidence and empirical example in Klein and Vella (2006). We acknowledge the role of Roger Klein in writing these programs.

the resulting distributions, we can calculate the effect of changes in aggregate and individual variables on the level of income as well as assess where in the distribution these changes exert the largest impact.

In the next section we examine the contribution of different aggregate variables to the observed differences in the income distribution between 1989 and 1994. Table 7 presents the standardized values of the explanatory variables for these two years. Over the period the unemployment rate changed the most dramatically. The increase in this variable more than doubled the increase in log of GDP. Moreover, it was 1.6 times larger than the increase in log of government expenditure and 5 times larger than the change in the inflation rate. The large contribution of the unemployment rate, and to a lesser extent that of government expenditure and GDP, on changing the value of the index suggests that these variables are likely to be responsible for most of the observed variations in the level of income. In contrast, given the small weight associated with the inflation rate, this variable would have had to decrease about 7 percentage points to change the value of the second index to the same extent as the change resulting from the observed increase in the unemployment rate. Table 7 also indicates that individual characteristics remained almost constant between 1989 and 1994, and therefore they do not seem to be responsible for the observed changes in the income distribution.

5.2 Quantifying the Effect of Macroeconomic Activity

To assess the distributional effects of macroeconomic activity we quantify the contribution of different aggregate indicators to the observed changes in the income distribution between 1989 and 1994. Figure 5 plots the kernel density estimates of

predicted log income in these two years and shows that, between 1989 and 1994, the mean shifted to the right and the distribution became more disperse. The increase in the mean might be the result of the positive growth path that characterized the Spanish economy after joining the European Community. However, the higher dispersion observed in 1994 suggests that the positive shift was not homogeneous over the domain of the distribution.

To measure these changes, the first and second columns in Table 8 report some descriptive statistics of the predicted log income distribution in 1989 and 1994. Over the period, the mean of the distribution rose from 13.02 to 13.06, representing an increase of 110 euros in real terms (base 1992). At the same time, the standard deviation increased by 13% (from 0.26 in 1989 to 0.30 in 1994). Table 9 presents the log income level at different percentiles of the distribution. The first and second columns of this table indicate that the higher income dispersion in 1994 resulted from the increase in the difference between the $10^{th} - 50^{th}$ percentile of the distribution. Figure 5 suggests that this increase was mainly due to an improvement in the economic situation of individuals located near the mean of the distribution that did not benefit those in the lower tail, who were relatively worse off in 1994 than in 1989.

To quantify the contribution of different economic variables to these changes we now compare the predicted density of log income in 1994, based on 1994 values for the macroeconomic variables, to a series of hypothetical densities based on the macroeconomic variables being held at their 1989 level. The use of counterfactual distributions is frequent in labor economics dating back to Oaxaca's (1973) study of the gender wage gap. Like the familiar Oaxaca decomposition our procedure ignores general equilibrium effects. Despite this limitation our approach indicates the potential im-

portance of the various variables considered in explaining the observed changes in the income distribution.

First consider the distributional effects associated with inflation. In Figure 6 the solid line represents the predicted density in 1994 and the overlapping line, given by a dashed line, is the counterfactual distribution at the 1989 inflation level. The figure clearly indicates that inflation had no distributional impact. This, however, probably reflects the small variation in the inflation rate over the period (from 0.998 percent in 1989 to 0.796 percent in 1994). After joining the European Community the Spanish authorities implemented a deflationary policy to gain credibility and this kept inflation low during the 1990's. In addition, the estimated coefficient for inflation is very small and changes in inflation thereby contribute little to the macroeconomic index. This result is consistent with previous empirical work that finds an economically insignificant redistributive effect for inflation (Blank and Blinder 1985).

In contrast, however, the movements in the unemployment rate had an important distributional effect. Figure 7 plots the distribution of predicted log income in 1994, denoted by a solid line, and the one that would have resulted had the unemployment rate remained at its 1989 level, denoted by a dashed line, noting that the unemployment rate increased by 7 percentage points, from 17% in 1989 to 24% in 1994. The figure shows that a lower level of unemployment in 1994 would have shifted the distribution to the right. The second and third columns of Table 8 reveals that the increase in unemployment between 1989 and 1994 reduced the mean of log income from 13.09 to 13.06 and increased its standard deviation by 7% (from 0.28 to 0.30). Table 9 also reports the log income level at different percentiles for the various hypothetical distributions. This table indicates that a lower level of unemployment in 1994 would have

narrowed the distance between the tails of the distribution, $(25^{th} - 75^{th})$ percentiles, and in particular the differences in income levels for those located below the median, $(10^{th} - 50^{th})$ percentiles. These findings suggest that the high unemployment rate in the early 1990's was responsible for the observed increased in income dispersion, as unemployed workers were unable to benefit from the positive economic effects of joining the European Community.

The integration of Spain into the European Community contributed to a rapid economic growth which only slowed down with the recession at the beginning of the 1990's. However, the log of GDP displayed a positive trend over this period and this appears to partially explain the increase in income mean between 1989 and 1994. Figure 8 plots the predicted log income density for 1994, given by the solid line, and that which would have prevailed had the level of GDP remained at its lower level in 1989 (denoted by the dashed line). Figure 8 indicates that the higher level of GDP in 1994, 9% above that in 1989, shifted the entire distribution to the right. The sixth column in Table 8 displays the descriptive statistics for the counterfactual distribution of log income at the 1989 level of GDP. A comparison across columns indicates that the change in the level of GDP is the main factor in explaining the increase in log income mean over the period (from 12.92 to 13.06). Moreover, the increase in the level of GDP narrowed the distribution as the highest gains, in terms of income associated with the increase in the level of GDP, were experienced in the lowest fifth percentile of the distribution.

Finally we consider the distributional effects of government expenditure. In Figure 9 the dashed line denotes the log income density that would have resulted if government expenditure had remained at its 1989 level, noting that the 1989 level of

public expenditures was 8% lower than in 1994. While less dramatic that the change associated with the increase in GDP, the rise in government expenditure also shifted the log income distribution to the right. This suggests that had the government reduced public expenditure as a result of the economic crisis in the early 1990's, the individual cost of the recession would have been larger. Indeed, an important percentage of government expenditure in 1994 was due to social welfare expenditures to soften the effects of the recession (i.e. unemployment benefits, health and education services and pension systems). The fifth column in Table 8 reports the descriptive statistics for the counterfactual log income distribution at the 1989 level of public expenditure. A comparison with the second column indicates that the increase in the level of government expenditure raised the mean of predicted log income from 12.98 to 13.06. However, it also contributed to an increase in the difference between the $10^{th} - 50^{th}$ percentile. This negative distributional effect might be due to the changes in the internal composition of public expenditure over the period. Thus the beginning of the 1990's were characterized by an increase in social expenditure to finance health, education and, in particular, pensions and unemployment compensation; while expenditure on sickness and survivors benefit programs, with a stronger effect on the lower tail of the distribution, was reduced.

To identify where in the distribution changes in aggregate variables exert the largest impact, Figure 10 plots the difference between the predicted density in 1994 and the hypothetical densities based on 1989 values for the macroeconomic variables. The vertical line marks the mean of the distribution in 1994. In this figure, a "flat line" indicates that the variable considered does not have any distributional effect. As expected, the line with long dashes, which plots the differences due to changes in the

level of inflation, is almost flat. However, there are important distributional effects for all the other aggregate variables. Changes in the distribution due to variations in the level of GDP are marked by the line with dots and dashes. The redistributive effects in this case are clear. The increase in the level of GDP reduced the mass below the mean of the distribution and shifted it to points around and above it. The increase in government expenditure, represented by a line with short dashes, had a similar positive effect. However, its magnitude is smaller and it seems to be more concentrated around the mean of the distribution. Accordingly, the poorest did not benefit from the increase in government expenditure as much as they did from the increase in the level of GDP. The solid line represents the distributional effect of unemployment. As a result of the increase in unemployment in 1994, workers at different points of the distribution slid back towards lower parts of the distribution. This contributed to the fattening of the lower tail of the distribution and the increase in income dispersion.

The results from the empirical analysis allow us to conclude that, unlike inflation, the remaining aggregate variables had a significant effect on the distribution of income. In particular, the increase in the level of GDP and government expenditure, with a positive effect on most of the distribution, reduced the negative effect of the high unemployment rates in the early 1990's and moderated the economic costs of the recession.

Although this paper focuses on the distributional aspects of macroeconomic activity the characteristics of individuals in the economy also affect the distribution of income. As noted above, the distribution of individual characteristics remained relatively stable over the period and, accordingly, we do not expect them to have

an important contribution to the observed changes in the income distribution. To examine this possibility Figure 11 reports the difference between the log income distribution in 1994 and the one that would have resulted if individual characteristics had remained as they had in 1989, denoted by a dotted line. Changes in the individual characteristics appear to have shifted the distribution to the right. This result is consistent with the aging of the Spanish population and the upgrading of workers' skills. However, the distributional effect of individual characteristics is smaller than that of aggregate variables. This is clear from Figure 11 which highlights that the difference between the actual density in 1994 and the hypothetical density with individual characteristics at the 1989 level, is smaller than the difference between the actual density and the one with aggregate variables at their 1989 level. These results suggest that the intense macroeconomic activity in the 1980's and 1990's was primarily responsible for most of the observed changes in the Spanish income distribution over the period.

6 Conclusions

This paper analyzes the effect of various economic variables on changes in the income distribution for a sample of Spanish households over the period 1985-1996. We employ a semiparametric double index based procedure that estimates the level of individual income as a flexible function of individual characteristics and aggregate economic variables. This characterization of the income generating process allows us to predict the level of individual incomes under different macroeconomic scenarios.

Using counterfactual income distributions we conclude that the macroeconomic activity over the period has an important redistributive effect. The high unemploy-

ment rate, which results from the recession in the early 1990's, fattens the lower part of the income distribution and explains a substantial proportion of the increase in income dispersion observed during the 1990's. The rise in the level of GDP and government expenditure, however, appears to shift the income distribution to the right, and reduces the individual cost of the economic crisis at the beginning of the 1990's.

References

[1] Alesina, A. and D. Rodrik (1994): "Distribution Politics and Economic Growth", *Quarterly Journal of Economics,* 109, 465-490.

[2] Barro, R. J. (2000): "Inequality and Growth in a Panel of Countries", *Journal of Economic Growth*, 5, 5-32.

[3] Blank, R. (1989): "Disaggregating the effect of business cycles on the distribution of income", *Economica*, Vol 56, No 222, 141-163.

[4] Blank, R., D. Card, F. Levy and L.J. Medoff (1993): "Poverty, Income Distribution, and Growth: Are They Still Connected?", *Brooking Papers on Economic Activity*, Vol. 1993, N^o 2, 285-339.

[5] Blank R. and A. Blinder (1985): "Macroeconomics, Income Distribution and Poverty", *NBER Woking Paper 1567.*

[6] Blejer, M. and I. Guerrero (1990): "The impact of Macroeconomic policies on income distribution: An empirical study of the Philippines", *The Review of Economics and Statistics*, 72, Vol. 3, 414-423.

[7] Blinder, A. and H. Y. Esaki (1978): "Macroeconomic Activity and Income Distribution in the Postwar United States", *The Review of Economics and Statistics*, Vol. 60, No. 4 (Nov. 1978), 604-609.

[8] Bourguignon, F. (1979): "Decomposable Income Inequality Measures ", *Econometrica*, 47, 901-920.

[9] Castañeda, A., J. Díaz-Giménez and J. Ríos-Rull (1998): "Exploring the income distribution business cycle dynamics", *Journal of Monetary Economics*, 42, 93-130.

[10] Cutler D.M., F.K. Lawrence, D. Card and R. Hall (1991): "Macroeconomic Performance and the Disadvantaged", *Brooking Papers on Economic Activity*, Vol. 1991, N^o 2, 1-74.

[11] Dollar, D. and A. Kraay (2002): "Growth is Good for the Poor", *Journal of Economic Growth* 7, 195-225

[12] Galor, O. and J. Zeira (1993): "Income Distribution in Macroeconomics", *Review of Economic Studies,* 60, 35-52

[13] Ichimura, H. and L.F. Lee (1991): "Semiparametric least squares (SLS) and weighted SLS estimation of multiple index models: Single equation estimation", in Nonparametric and Semiparametric Methods in Econometrics and Statistics, ed. W. Barnett, J. Powell and G. Tauchen, Cambridge University Press.

[14] Juhn, C., K. Murphy, and B. Pierce, (1993) "Wage Inequality and the Rise in Returns to Skill", *The Journal of Political Economy*, Vol 101, No. 3, 410-442.

[15] Keane, M. and R. Moffitt (1988): "Real Wage over the Business Cycle: Estimating the Impact of Heterogeneity with Micro Data", *Journal of Political Economy*, 1988, Vol 96, No 6.

[16] Klein, R. and F. Vella (2006): "A Semiparametric Model for Binary Response and Continuous Outcomes Under Index Heteroscedasticity", *IZA working paper 2383.*

[17] Maravall, J. M. (1997): "Regimes, Politics, and Markets", *Oxford University Press.*

[18] Oaxaca, R. (1973): "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review*, 14, 193-709.

[19] OECD, (1996): "Social Expenditure Statistics of OECD Member Countries", *Unpublished Manuscript.*

[20] Oliver, J., X. Ramos, and J.L. Raymond-Bara (2001): "Recent Trends in Spanish Income Distribution: A Robust Picture of Falling Income Inequality", *UAB Working Papers* 107.

[21] Perotti, R. (1993): "Political Equilibrium, Income Distribution and Growth", *Review of Economic Studies* 60, 755-776.

Appendix

Figure 1: GDP (level and growth rate)

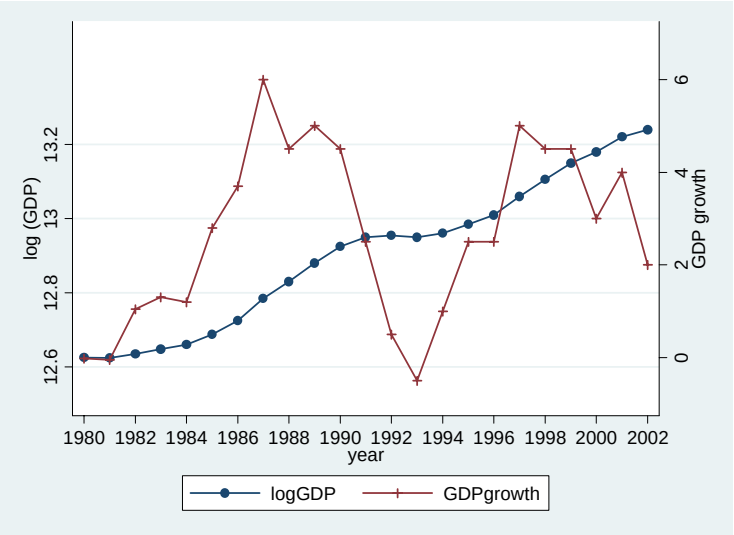

Figure 2: Unemployment and Inflation

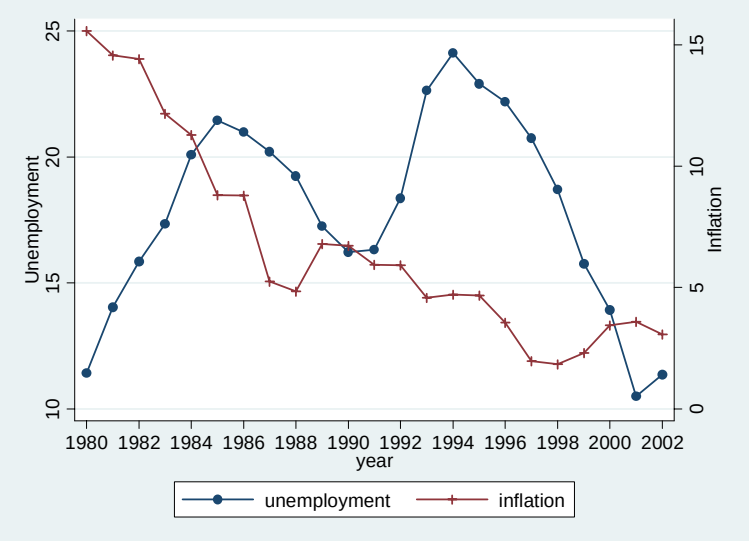

Figure 3: Inequality Measures

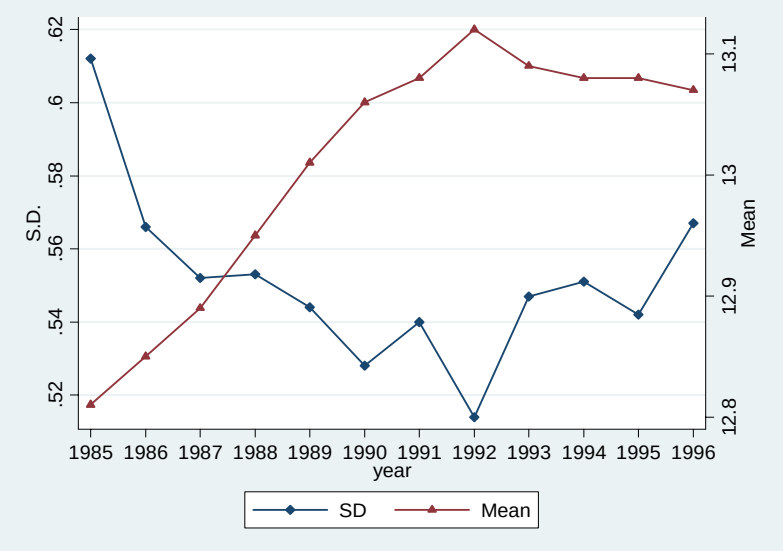

Figure 4: Density of log income: 1989 and 1994

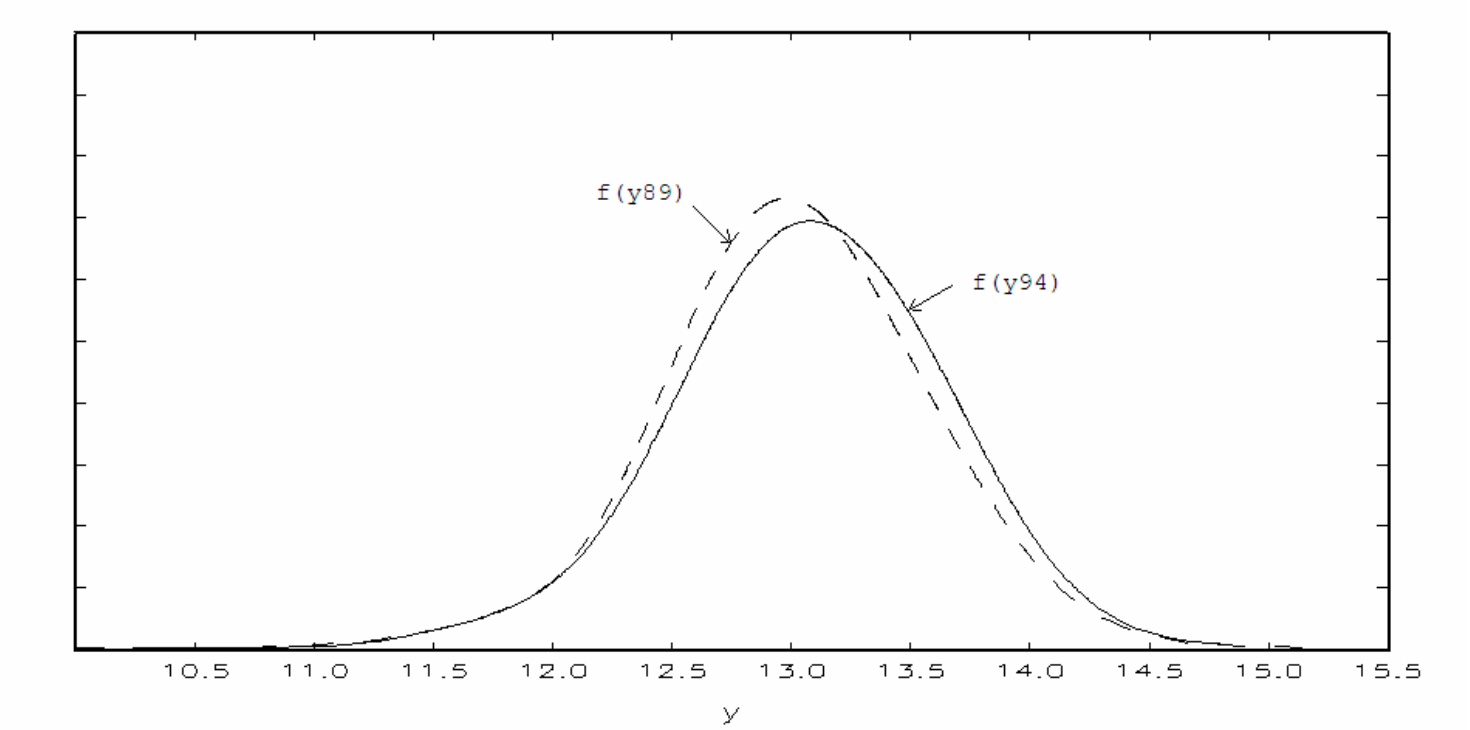

Figure 5: Density of log predicted income: 1989 and 1994

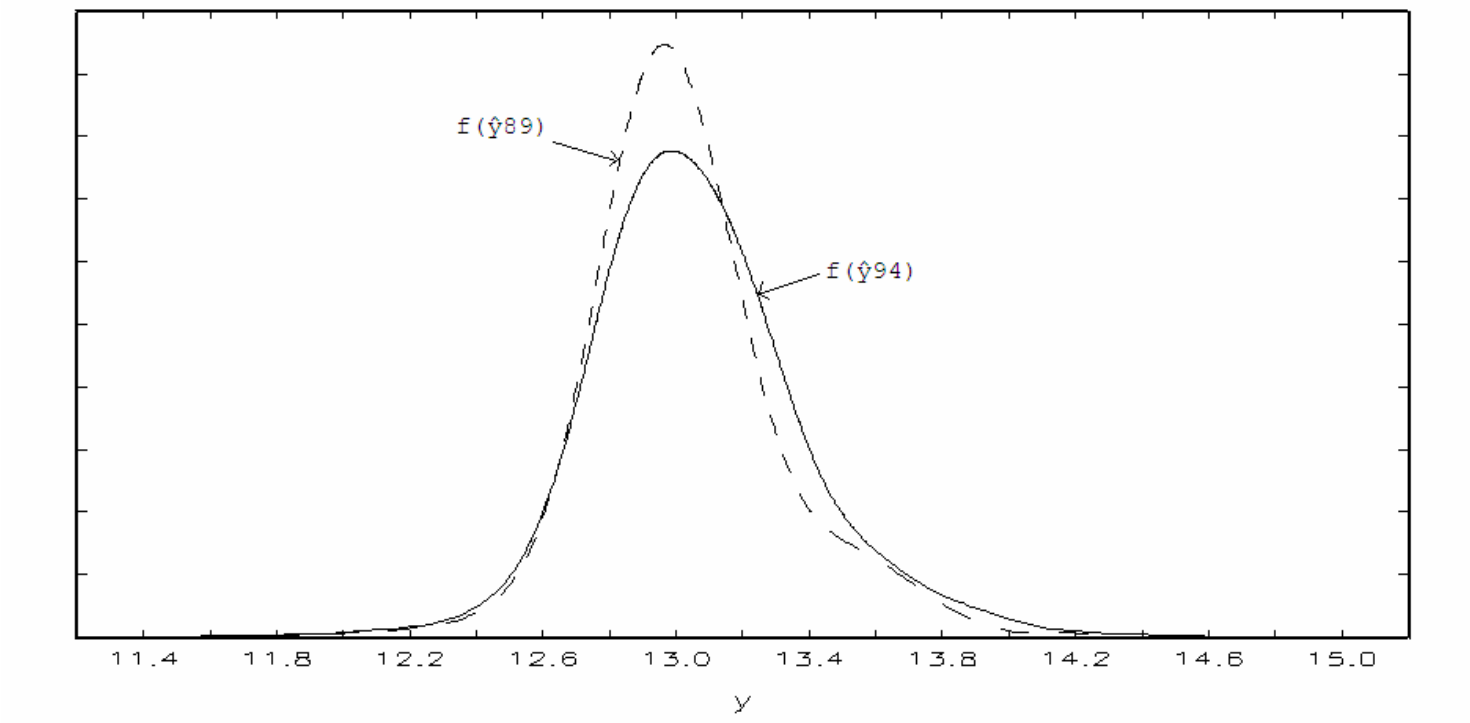

Figure 6: Counterfactual Density of log income, (Inflation)

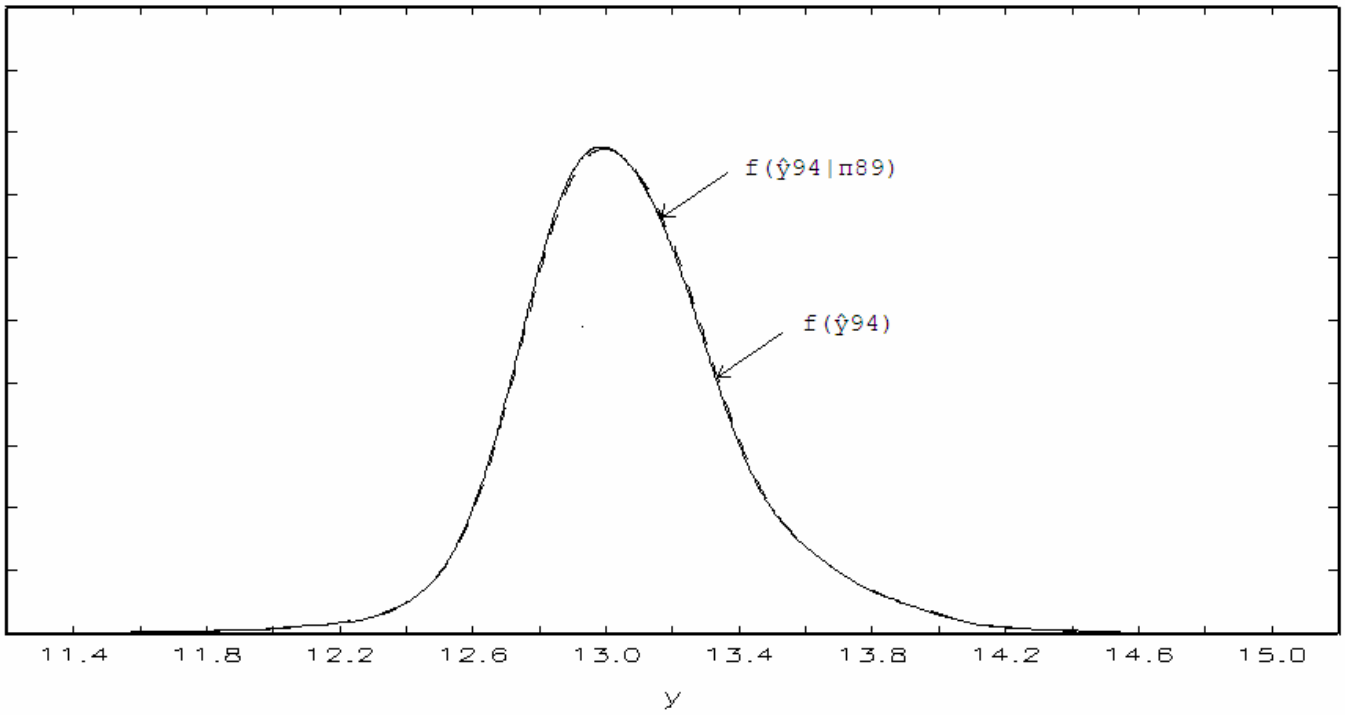

Figure 7: Counterfactual Density of log income, (Unemployment)

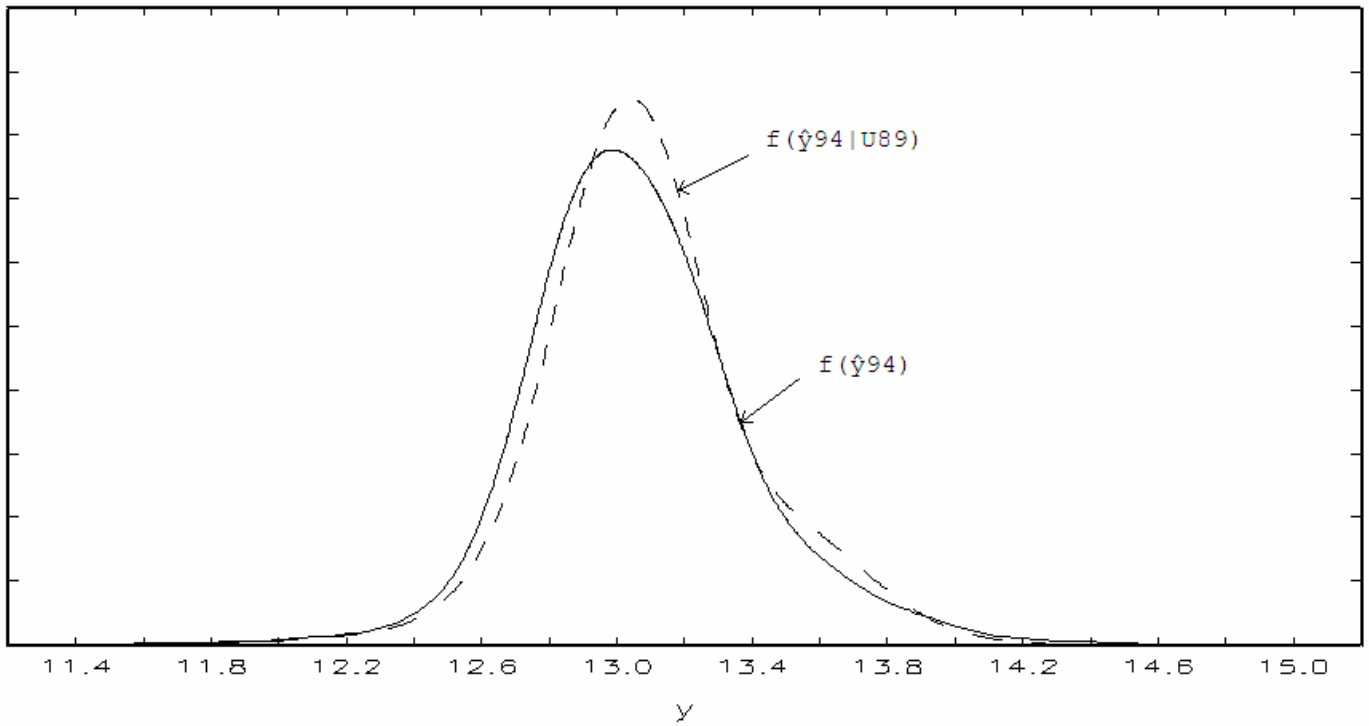

Figure 8: Counterfactual Density of log income, (GDP)

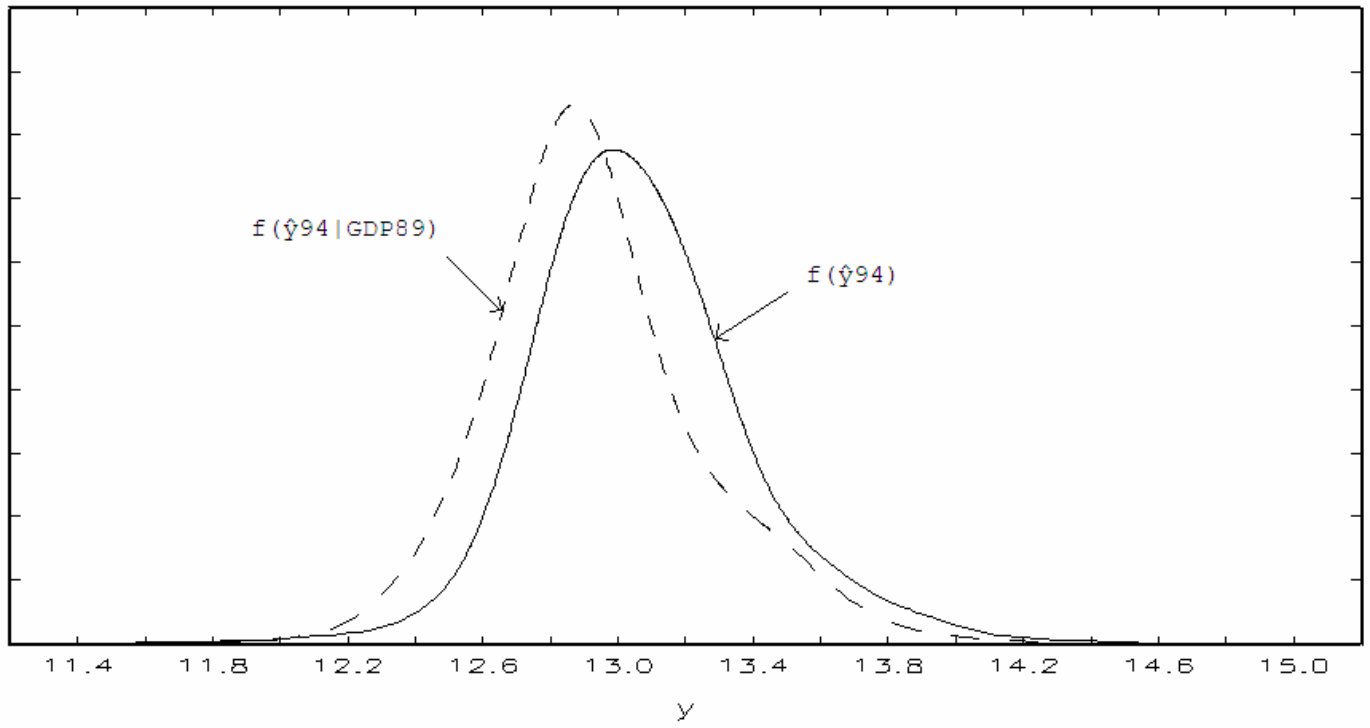

Figure 9: Counterfactual Density of log income, (Public Spending)

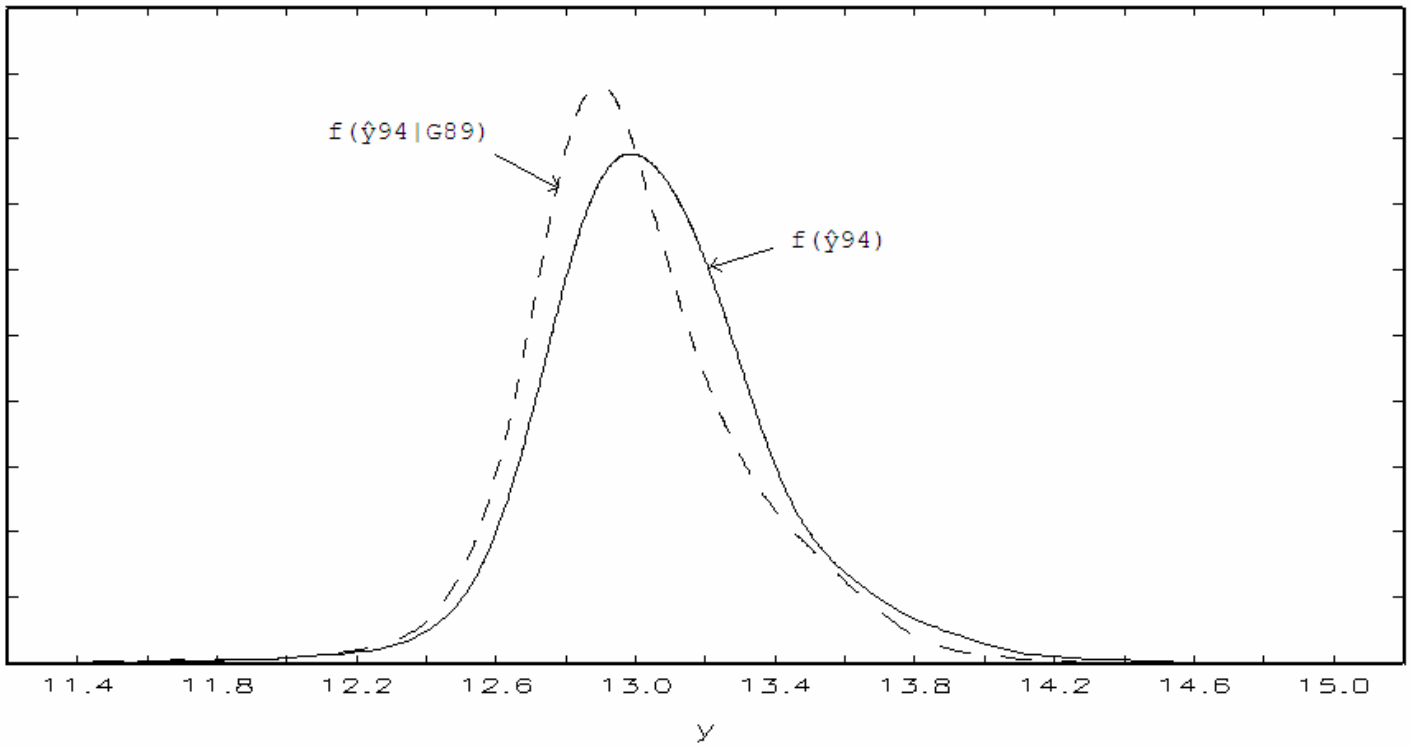

Figure 10: Difference between the Density in 1994 and Counterfactuals

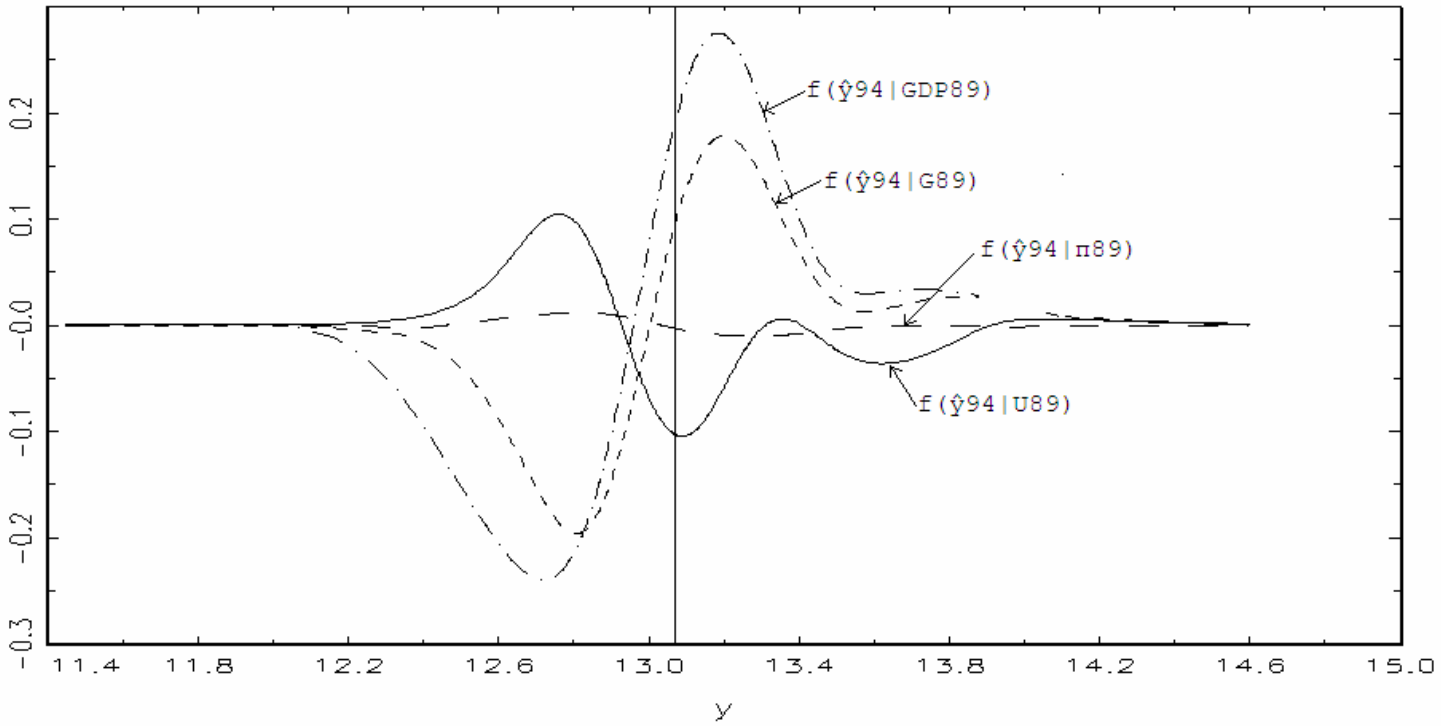

Figure 11: Difference between the Density in 1994 and Counterfactuals (Aggregate and Individual variables)

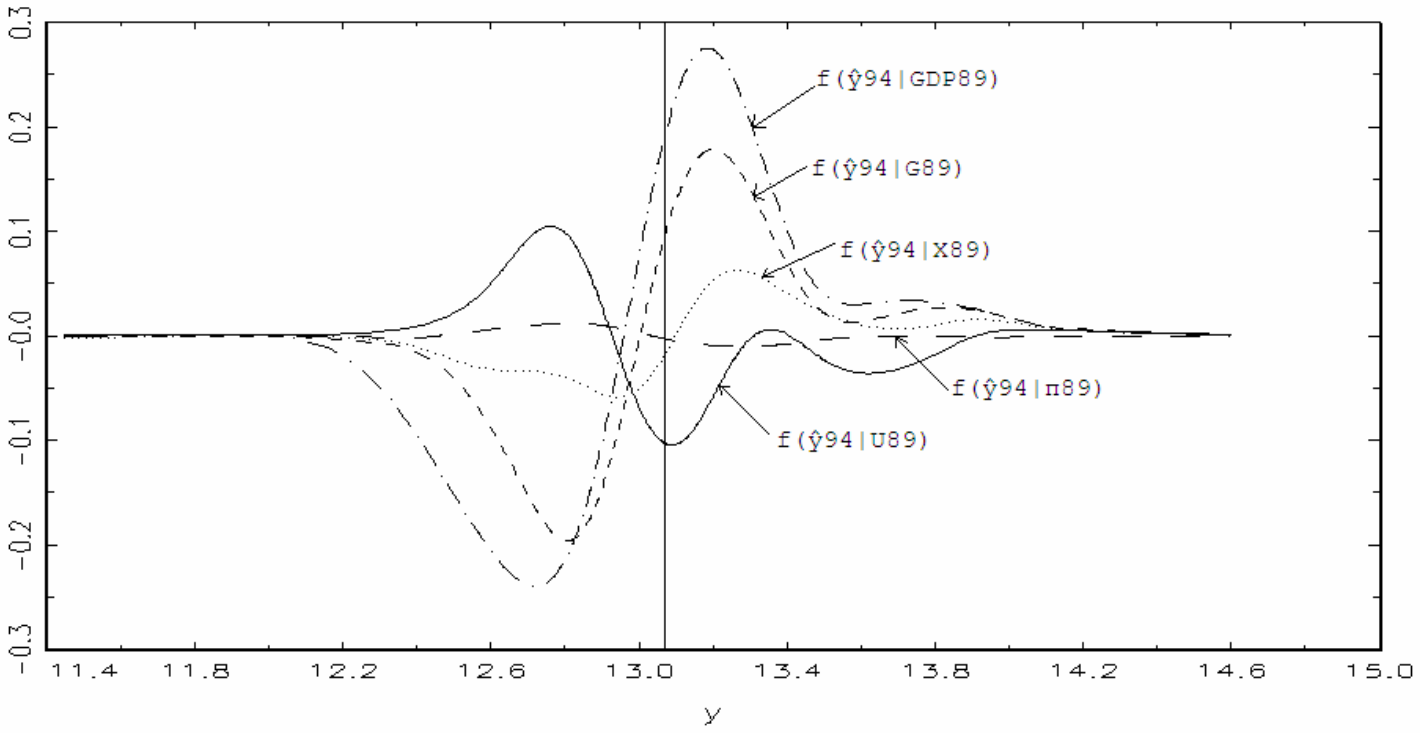

Table 1
Composition of Government Expenditure

Public Expenditure (Percentage of GDP)

	1985	*1986*	*1987*	*1988*	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*
Total Social Expenditure	18.71	18.38	18.28	18.81	19.14	19.82	20.67	21.41	22.52	23.54
Health	4.6	4.49	4.54	5.02	5.1	5.45	5.61	5.7	5.73	5.8
Disability Cash Benefits	1.35	1.33	1.3	1.32	1.36	1.27	1.33	1.38	1.47	1.54
Old-Age Cash Benefits	6.05	6	5.92	5.9	5.92	6.04	6.17	6.39	6.76	7.11
Occupational Injury and Disease	0.41	0.41	0.41	0.42	0.44	0.47	0.48	0.49	0.5	0.51
Sickness Benefits	0.74	0.71	0.7	0.73	0.78	0.85	0.92	0.99	1.04	1.08
Services for Eldery and Disable	0.09	0.08	0.09	0.1	0.11	0.14	0.13	0.14	0.16	0.17
Survivors Benefits	1.88	1.83	1.8	1.8	1.85	1.92	1.99	2.09	2.19	2.24
Family Cash Benefits	0.25	0.2	0.16	0.15	0.14	0.12	0.12	0.14	0.15	0.15
Unemployment Benefits	2.91	2.59	2.54	2.44	2.43	2.63	2.97	3.33	3.79	4.49

Real Annual Growth Rate

	1985	*1986*	*1987*	*1988*	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*
Total Social Expenditure	5.93	2.21	4.36	4.67	5.23	5.8	4.53	4.22	4.63	4.55
Health	-0.04	0.61	6.95	16.22	6.45	10.68	5.44	2.25	-0.7	1.25
Disability Cash Benefits	3.74	1.43	3.57	6.63	7.89	-2.91	6.63	4.83	4.9	4.8
Old-Age Cash Benefits	5.93	2.21	4.36	4.67	5.23	5.8	4.53	4.22	4.63	5.15
Occupational Injury and Disease	-0.32	1.11	6.42	8.67	8.31	11.82	4.55	2.27	0.9	1.12
Sickness Benefits	-0.99	-0.7	4.72	8.7	12.87	12.2	11.44	8.04	4.4	3.56
Services for Eldery and Disable	8.63	-1.49	17.71	11.61	18.16	26.85	-0.17	9.19	11.22	8.79
Survivors Benefits	4.37	0.88	3.82	4.74	7.96	7.56	6.01	5.87	3.55	2.3
Family Cash Benefits	-13.18	-17.77	-12.41	-5.06	-3.73	-6.25	-1.59	21.58	2.46	2.11
Unemployment Benefits	18.16	-8.17	3.7	1.01	4.36	11.99	15.67	12.8	12.55	18.52

Source: Social Expenditure Statistics of OECD Member Countries (1996)

Table 2: Income Distribution

Year	Mean*	S.D.	Gini
1985	12.81	0.612	0.315
1986	12.85	0.566	0.299
1987	12.89	0.552	0.292
1988	12.95	0.553	0.282
1989	13.01	0.544	0.296
1990	13.06	0.528	0.291
1991	13.08	0.54	0.279
1992	13.12	0.514	0.274
1993	13.09	0.547	0.29
1994	13.08	0.551	0.298
1995	13.08	0.542	0.287
1996	13.07	0.567	0.292

*log quarterly income in pesetas 1992

Table 3: Descriptive Statistics

	Mean	*S.D.*
Individual Variables		
Income	13.01	0.56
Male	0.92	
Partner with income	0.91	
Age	44.22	10.16
Education:		
Less than HS	0.14	
HS graduate	0.6	
Some college	0.14	
College or more	0.11	
City Size:		
<1,000 (1)	0.25	
1,000-10,000 (2)	0.2	
10,000-50,000 (3)	0.39	
>50,000 (4)	0.16	

Aggregate Variables		
Unemployment	20.07	2.73
Inflation	0.91	0.39
Log (G)	9.99	0.05
Log (GDP)	11.5	0.1

Table 4: Income Distribution (1989, 1994)

	1989	*1994*
Mean	13.01	13.08
10th-90th	1.29	1.32
10th-50th	0.62	0.68
50th-90th	0.67	0.65
25th-75th	0.64	0.69
5th-95th	1.7	1.75
S.D.	0.54	0.55
Gini	0.29	0.29

Table 5: Observed Changes in Explanatory Variables

	1989	*1994*
Individual Variables		
Male	0.95	0.93
Partner with income	0.94	0.92
Age	44.23	44.74
Education:		
Less than HS	0.18	0.18
HS graduate	0.61	0.62
Some college	0.1	0.11
College or more	0.1	0.09
City Size:		
<1,000 (1)	0.23	0.24
1,000-10,000 (2)	0.21	0.21
10,000-50,000 (3)	0.4	0.39
>50,000 (4)	0.16	0.16

Aggregate Variables		
Unemployment	17.26	24.22
Inflation	0.998	0.796
Log (G)	9.96	10.04
Log (GDP)	11.48	11.57

Table 6: Estimated Coefficients*

Variables	First Index	Variables	Second Index
gender	1	Unemployment	1
age	12.8489 (3.9549)	Inflation	-0.1459 (0.0736)
town 1	-16.0861 (4.9142)	log G	-0.2930 (0.1227)
town 2	-8.3570 (2.5902)	log GDP	-0.9899 (0.1712)
town 3	-2.6917 (0.9194)		
partner with income	18.1429 (5.6754)		
HS graduate	19.1098 (5.8501)		
Some College	28.4569 (8.7223)		
College or more	41.0209 (12.5976)		

*Town 4 and less than HS are ommited for identification purposes

Table 7: Observed Changes in the Normalized Explanatory Variables

Variables	1989	1994	Variables	1989	1994
gender	0.0295	-0.0867	HS graduate	0.0064	-0.0172
age	0.0098	-0.0148	Some College	-0.0624	0.1053
town 1	0.0095	-0.0107	College or more	-0.0236	0.0428
town 2	0.0003	0.0145	Unemployment	-1.0567	1.6018
town 3	-0.0052	-0.0197	Inflation	0.2274	-0.3117
town 4	-0.0046	0.0229	log G	-0.5228	1.1416
partner with income	0.0481	-0.029	log GDP	-0.1648	0.9219
less than HS	0.0756	-0.1212			

*For individual variables the change in mean is reported.

Table 8: Descriptive Statistics
(Predicted and Counterfactual Income Distributions)

	1989	1994	1994|U89	1994|89	1994|G89	1994|GDP89
Mean	13.02	13.06	13.09	13.07	12.98	12.92
10th-90th	0.62	0.69	0.71	0.67	0.66	0.76
10th-50th	0.2	0.31	0.28	0.27	0.21	0.28
50th-90th	0.42	0.42	0.43	0.4	0.45	0.48
25th-75th	0.25	0.35	0.3	0.36	0.31	0.33
5th-95th	0.87	0.96	0.98	0.95	0.95	1.02
S.D.	0.26	0.3	0.28	0.3	0.29	0.31
Gini	0.14	0.17	0.16	0.17	0.16	0.17

Table 9: Income Levels at Different Percentiles
(Predicted and Counterfactual Income Distributions)

	1989	*1994*	*1989|U89*	*1994|n89*	*1994|G89*	*1994|GDP89*
5th	12.69	12.68	12.69	12.68	12.58	12.47
10th	12.77	12.76	12.79	12.76	12.72	12.6
25th	12.88	12.87	12.91	12.87	12.82	12.77
50th	12.97	13.03	13.07	13.03	12.93	12.88
75th	13.13	13.22	13.21	13.23	13.13	13.07
90th	13.39	13.44	13.5	13.43	13.38	13.36
95th	13.56	13.63	13.66	13.63	13.53	13.49