Unemployment and Inflation as economic determinants to explain Spanish income inequality: Are redistributive measures in Spain effective in reducing inequality?

Author: Juan Manuel Prior Clavero

Supervised by Javier Ordoñez Monfort

Email: al363344@gmail.com

Degree in Economics. Academic year: 2020/2021

Abstract:

In this paper, we will analyse the influence of the unemployment and inflation rate on Spanish income inequality. At the same time, we will evaluate how the main redistributive measures to reduce inequality behaved. All this, in a period with different economic cycles, which affected differently the evolution of income distribution. We will employ the Ordinary Least Squares (OLS) method, using time series, to carry out the study. The resulting methodology after applying differencing will allow us to show that both indicators have a negative influence on inequality, showing some signs of persistence. However, it will not adequately capture the effectiveness of redistributive measures. We will not even find evidence on the tax effect. Nevertheless, when analysing the evolution of the indicator used to measure the redistributive effect of the tax, we will see its effectiveness in reducing inequality has been gradually increasing over time.

Keywords: Inequality, Unemployment, Inflation, Redistributive Measures

JEL Classification: C32, E24, E31, D63, H23, H51, H52

Index

Contents **Pages**

Figures and tables

1. Introduction

Over the last 40 years, the concern to investigate the phenomenon of inequality has been gaining more and more interest, due to the global increase in the worsening distribution of income and wealth. The OECD report (2014) makes it clear that the problems of inequality are an issue to be addressed both in developed countries and in emerging countries. For this reason, several economists and researchers from around the world have tried to explain the origin of this phenomenon. In the case of Spain, it has not been possible to avoid this problem either. In fact, there is a clear consensus in the literature that, in most of OECD countries, income inequality has persisted over time.

However, Spain is considered, by several studies, as a particular case study when considering the analysis of income inequality. As specified in the article by Sarmiento (2017), the growth of inequality in Spain in the last three decades prior to the 2008 crisis, was not as marked compared to the United States or other European countries, with a more negative and abrupt evolution of inequality. In fact, taking into account the different economic cycles in Spain during this period, and their different effects on the distribution of wealth, a greater interest has emerged in investigating the evolution of inequality in the Spanish case. Research has been appearing, trying to identify which factors determine inequality, and other studies, where the redistributive capacity of the Spanish measures is evaluated. For instance, considering the study by Pena Trapero et al. (1996), he found that, during the 1980s, inflation increased inequality in the first income deciles. However, he did not find any relationship with unemployment, as did other economists who used a similar methodology. Later, in the article by Adiego and Ayala (2013), it was specified that, in the periods between 1980-1990 and 1995-2007, there was a favourable economic growth cycle. The first cycle was characterised by maintaining a positive evolution in income distribution, due to relatively effective redistributive measures. The second cycle, despite of being also characterised by a favourable economic growth, did not record substantial improvements in the distributive process. According to (Ayala, L., Ruiz-Huerta, J. and Martínez, R., 1993), since a recessive phase in unemployment happened in the early 1990s, negative changes in income distribution began to take place. They concluded that unemployment indeed have a negative influence on inequality. However, although the recession of unemployment began to recover at the end of the 1990s, other studies, such as the Adiego and Ayala (2013), have shown the gradual loss of effectiveness of redistributive measures during the second expansionary cycle.

On the other hand, concerning the report from the Bank of Spain by Anghel, et al (2018), it shows the 2008 real estate crisis had a negative impact mainly on the Spanish employment rate, increasing the inequality of capita income. Furthermore, in the report by García-Altés and Ortún (2014) they state that, after the outbreak of the financial and real estate crisis, Spain's problems with the effectiveness of its redistributive policies and labour market structure were dismantled and brought to light. He made a clear criticism of the Spanish economic and distributive model, according to his conviction that an urgent improvement is required.

Given the variety of results on the influence of unemployment and inflation on Spanish inequality, and the evaluation of its redistributive policies, it does not seem to be entirely clear how these factors have behaved in the literature. While some researchers defend the importance of these economic indicators in explaining inequality, others argue against this hypothesis. On the other hand, neither the role of redistributive measures is entirely clear.

Thus, the aim of this paper is to study the effects of inflation and unemployment on income inequality in Spain. We will also consider how the most important redistributive measures in the literature behaved. All this in a period characterised by the presence of different expansionary and recessionary phases, including the crisis of the 1990s and 2008. We will carry out the study from 1980 to 2014, so as to collect (as accurately as possible) the impact of the different cycles mentioned above, on our variables of interest. In this way, we will see if this work can be useful to corroborate some of the studies that confirm the relationship between inflation, unemployment and inequality, or to support those studies which defend the effectiveness of redistributive measures.

Following this introduction, we will perform in section 2 a more detailed review of the literature, which has been devoted to studying the phenomenon of inequality, using different approaches and methodologies. The description and analysis of the variables, and the data source will be explained in sections 3 and 4, respectively. In section 5, we will explain the methodology that it will be used to perform the study. Basically, an OLS regression will be carried out using time series, after first differences have been applied. This paper employs a simpler analysis compared to other studies; nevertheless, I believe that it can be useful to get a general idea of how the Spanish economic and distributional model is organised. In section 6, we will show both the results needed to employ our methodology, and the results extracted from the regression, explaining their economic implications. Finally, in the conclusion, we will summarise the empirical results obtained with their respective limitations.

2. Literature review

Income inequality has been one of the least explored phenomena by economists compared to other areas of study. However, there is a growing body of research that attempts to investigate the income gap between rich and poor. Over the last two centuries, numerous studies, which have used different approaches, have appeared in order to analyse income inequality.

First, we have those which have analysed the persistence of income inequality. This approach consists of considering whether this phenomenon is transitory or permanent in the long term. Several economists, have used this technique to analyse the persistence of inequality in different countries, and have simultaneously tried to identify the shocks or determinants that cause it to persist. In the literature, globalisation and technological change have been the most important determinants when analysing the determinants. For example, it was shown the external financial, economic and social openness of most countries in the world, worsened the distribution of income globally, between and within countries. There was some controversy about this theory, since research, such as Dollar (2005), Dollar and Kroay (2002), showed (in addition to the clear persistence of inequality in a sample of 137 countries) that globalisation had a little significant effect on the distribution of world income. However, they did not exclude the existence of a link between globalisation and economic growth, with poverty reduction. In fact, the positive effects of globalisation on inequality were also discussed. The article by Milanovic (2016) comments that, in developed countries, those who benefit from the economic phenomenon of globalisation are the middle and lower classes. However, studies which supported the theory that globalisation increased the inequality gap between countries were predominant. For example, in the case of Galbraith's (2011) research, there was strong evidence of the negative effects of globalisation on the distribution of world income between and within countries. While other economists, such as Bourguignon (2015), finally demonstrated that globalisation can behave as an exogenous determinant of inequality across countries. Within each country, in particular, there would be greater or lesser inequality, depending on the set of political and economic decisions made in each of the nations.

Taking the latter as a starting point, papers analysing the persistence of inequality, which took into account the main policy variables, were better placed. For example, research by Mocan (1999) found that both inflation and unemployment had a significant effect on income distribution in the US economy. In the work of Christopoulos and McAdam (2017), the persistence of inequality in OECD and additional countries was shown. They demonstrated the existence of unit roots and persistence in the gross and net income gini indices, corroborating the permanent effects of shocks, such as those derived from innovation and financial shocks, on inequality across countries. They showed that the fiscal redistributive effect was not enough to reduce persistence. Most of these papers used the technique of differentiation to carry out the analysis of persistence. There are few papers which analyse persistence, by using fractional integration and cointegration techniques. In particular, Gil-Ana et al (2019) investigate both the degree of persistence of inequality and the influence of economic growth, inflation and unemployment on inequality in OECD countries. They conclude that there is significant persistence in virtually all countries, including Spain, and that the only significant variable which explains (decreasing) inequality over time is economic growth. This is odd, because it is one of the few studies which goes against the widespread theory that economic growth worsens inequality. Importantly, this latest paper also analyses the persistence of other macroeconomic variables. They show that, in the case of Spain, inflation, unemployment and economic growth are persistent over time.

In addition, as mentioned above, the influence of technological change has also been extensively studied for its relevance in determining persistence. Griliches (1969) discussed his theory that technological advances required a greater demand for capital accumulation, and therefore required greater specialisation in the human labour factor. According to his theory, this led to an increase in the wage gap. Because the specialised human factor accumulated higher earnings, compared to those with less specialisation. However, there are several who go against this theory, claiming that the higher the specialisation and productivity, the smaller the inequality gap in a country. For example, Card and Dinardo (2002) confirmed this claim in the United States, because they found little evidence for Griliches (1969). They found that it was changes in the minimum wage, the rate of unionisation, and labour market reallocation which were actually behind the increase in inequality in the United States.

In the second stage, we have those studies which have analysed inequality, considering the influence of economic growth and the opening of financial markets, using a smaller temporal focus. We have already commented previously on the uncertainty about the effect of economic growth. If we focus on the Spanish case, a study carried out by Ochando (2010) tried to explain how economic growth directly affects income distribution. He searched the keys to explain why the last long period of economic growth in Spain, prior to the 2008 crisis, did not mean an improvement in income distribution. He concludes that, although there was growth in GDP, changes in the labour market, the impact of immigration, and less effective and redistributive redistributive policies, were the factors that caused the worsening of income distribution.

Another issue to be debated in the literature is whether the opening of financial markets improves income distribution. For example, in the case of the study by Buman and Lensink (2015), they discuss that financial liberalisation would improve distribution in those countries that are in favour of maintaining a large financial debt with the exterior. They talk over economic agents have a greater capacity for consumption and investment, thanks to foreign indebtedness. However, other studies show the opposite. For example, De Haan and Sturm (2017) argue that there are some countries where, by financially liberalising, increase inequality. They explain it is the quality of political institutions in a given country which determines the impact of financial liberalisation on inequality. In developing countries with favourable institutions, both financial openness and financial development would be positive for inequality. Considering their argument, financial development is favourable for economic development, because it would benefit those individuals with lower incomes.

Next, we will discuss studies which are more focused on the analysis of the case of inequality in Spain over the last 30 years. Using different methodologies, some economists have also tried to identify the determinants of inequality in the Spanish case, assessing the redistributive capacity of measures and policies. On the one hand, there are those who, using the methodological approach formulated by Fields (2003), have tried to identify the contribution of each explanatory factor to inequality, disaggregating the variance of the logarithm of income into the different explanatory contributions to inequality. In the study by Adiego and Ayala (2013), they used it to identify which factors influenced the truncation of the downward trend in inequality in Spain, maintained since the 1970s. They conclude that the educational level of households and its relationship with economic activity were relevant in explaining the increase in inequality. Wage differentials were generated between the different educational categories. Furthermore, they state that the growing risk of the relative loss of income of certain types of households, due to the inactivity of all household members, is an important cause of the increase in inequality (especially since the beginning of the 2008 crisis). Other papers, such as Pijoan-Mas and Sanchez-Marcos (2010), used a similar methodology. They managed to demonstrate and convince the literature that the decrease in inequality in the 1980s was due to the expansive economic phase of that time, and to the good performance of the redistributive process.

On the other hand, another recurrent methodology used in the Spanish literature is known as the SURE system (Zellner, 1962). It is used to consider the influence of economic cycles on income inequality. To do so, this methodology consists of regressing the different quintile shares of total (primary or net) income on the business cycle, inflation and a long-term trend. It was first applied by Blinder and Esaki (1978) to the United States. It has been updated by various economists to additional countries, most of whom approximated the business cycle using unemployment and inflation rates. Most of them found that unemployment and inflation have a negative effect on inequality (increasing it). In contrast, other economists found that a change in the price level has a positive effect for low income groups and a negative effect for high income groups. This might go against economic logic.

In addition, a study by Ayala et al. (1995), which uses an inequality index as a dependent variable instead of an income variable, failed to find any direct relationship between unemployment and the Spanish inequality index. In the study by Doncel et al. of the Universidad Rey Juan Carlos, they use the time series approach using the same methodology, to look at the cyclical influence on primary and disposable income, from 1985-1996, covering the cycle where inequality remained stable. They use proxys variables of the business cycle, such as the unemployment rate, the rate of change of GDP, and the Hodick-Prescott filter (1980). Basically, they analyse the stabilising role played by the public sector in stabilising the flow of income, through taxes and transfers, after the impact of the cycle. They conclude that both unemployment and inflation behave regressively, to the disadvantage of those with lower primary incomes. However, they argue that the public sector at the time managed to effectively mitigate these effects on the unemployment rate to a significant extent, but failed to fully compensate the regressivity in the case of the inflation rate.

It is clear there is a broad consensus that unemployment and inflation factors influence inequality in Spain, regardless of the methodology used. On the other hand, it is not entirely clear what the role of Spanish redistributive measures has been. We will try to see what impact the determinants of inflation and unemployment had on income inequality, and what were the effects of the redistributive measures applied. In this way, we may be able to corroborate some of the studies discussed above.

3. Starting hypotheses and analysis of variables

3.1 Hypotheses

Following the methodology that will be better explained in section 5, we will try to answer the following hypotheses related to the evolution of inequality in Spain:

- Hypothesis 1: Macroeconomic indicators of unemployment and inflation have a negative influence on inequality in Spain, increasing it.
- Hypothesis 2: The most important measures implemented by the Spanish government in order to reduce the problem of inequality are effective.

That is, using income inequality as the dependent variable, we will consider if the respective unemployment and inflation rates are important to explain the evolution of Spanish inequality. Additionally, we will check the effectiveness of Spanish redistributive measures. It has been decided to choose both indicators, because I consider them to be important economic variables to consider, due to the macroeconomic imbalances they can cause in the economy. We could also have chosen GDP. However, there was a risk of incurring a perfect correlation with other variables. The aim is not to make an accurate approximation of cycles by using these two variables, as some of the studies cited above. Our aim is just to see whether they really have a linear relationship with inequality that is consistent with economic logic. On the other hand, we have chosen those variables related to reducing inequality. These are considered, in the literature, to be some of the most important for lowing the problem of income distribution. Both in OECD countries, and in developing countries. We shall see if this really happens in Spain.

3.2 Description of variables

For a better understanding of the variables to be used in the study, and the hypotheses to be tested, we are illustrating in a summary table the description of each variable, and its expected economic relationship with the variable to be explained. Secondly, we will analyse the evolution of each variable, briefly explaining the political and economic environment over the period studied.

It should be noted that inequality will be measured by the gini index, which is an economic measure that will be close to 0 in the case of perfect equality, and 100 in the case of maximum inequality between the different groups of disposable income in Spain.

Table 1: description of the variable to be explained and of the macroeconomic variables

Dependent variable	Description	Source data	Hypothesis fulfillment
Gini index	It measures the inequality index in the distribution of disposable income in net terms. The higher the index, the greater the inequality among the social groups of the population.	Standardized World Income Inequality Database (SWIID).	
Macroeconomic explanatory variable	**Description**	**Source data**	**HP fulfillment**
Unemployment rate	Represents the percentage of unemployed out of the total labor force (active population).	World Economic Outlook Database	The more unemployment, the higher income inequality.
Inflation rate	Represents the interannual variations of average prices on consumption, expressed in percentage variations.	World Economic Outlook Database	The more inflation, the higher income inequality in Spain.

Source: own elaboration based on data obtained from several sources

Table 2: description of the redistributive variables

Redistributive explanatory variable	Description	Source data	HP fulfillment
Public spending on education	Expressed as the percentage used by the State towards the field of Education, of the total GDP.	Data from Espuelas (2013) and Databank	The higher spending on education, the less inequality.
Investment in Health	Both private and State investments will be considered. Expressed as the percentage of the total of GDP towards the field of Health.	OECD Main Economic Indicators Database.	The higher investment in Health, the less income inequality.
Reynolds-Smolensky index	It measures the redistributive ability of income tax in Spain.	Standardized World Income Inequality Database (SWIID).	The more redistributive the income tax in Spain, the less inequality.

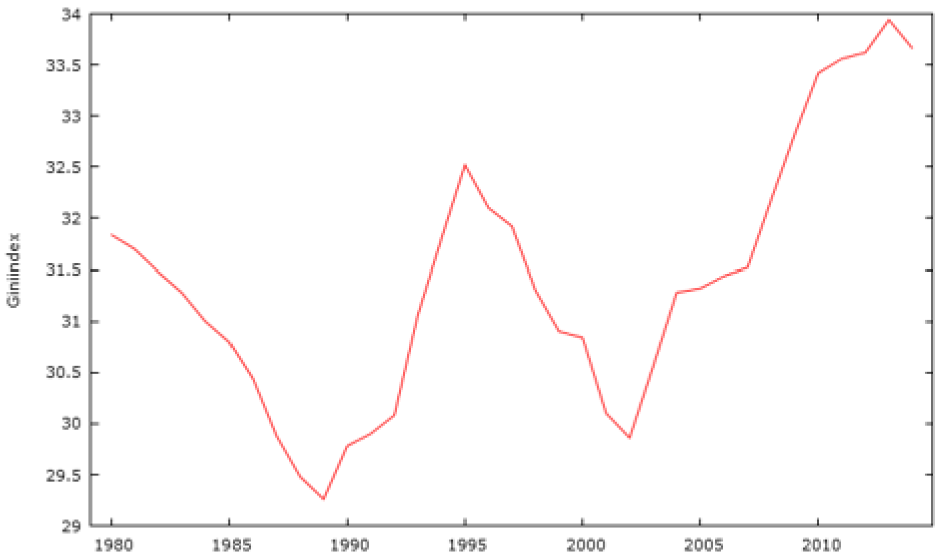
Source: own elaboration based on data from various sources

The importance of considering these three redistributive variables should be highlighted. According to the OECD report (2014), investing in the areas of education and health is fundamental to improve the well-being and quality of life of all people in society. Thus, those with lower incomes could have easier access to healthcare, and additionally receive better training in the field of education. In this way, they should contribute to the growth of social welfare, so that they will have a better access to work, and to income generation. Regarding the fiscal redistributive effect, several studies such as Argimon et al. (1986) or Onrubia and Picos (2013), claim the importance of considering taxes as a distributive tool, because of their progressivity. We would try to check whether income tax really fulfils its redistributive role for those with lower incomes, as with public spending on education and health investment.

3.3 Evolution of variables

To understand better our variables, we will briefly analyse how they evolved over the period analysed. In this way, it will help us to consider if we should take into account the existence of some kind of trend.

Figure 1: graph of the evolution of the gini index

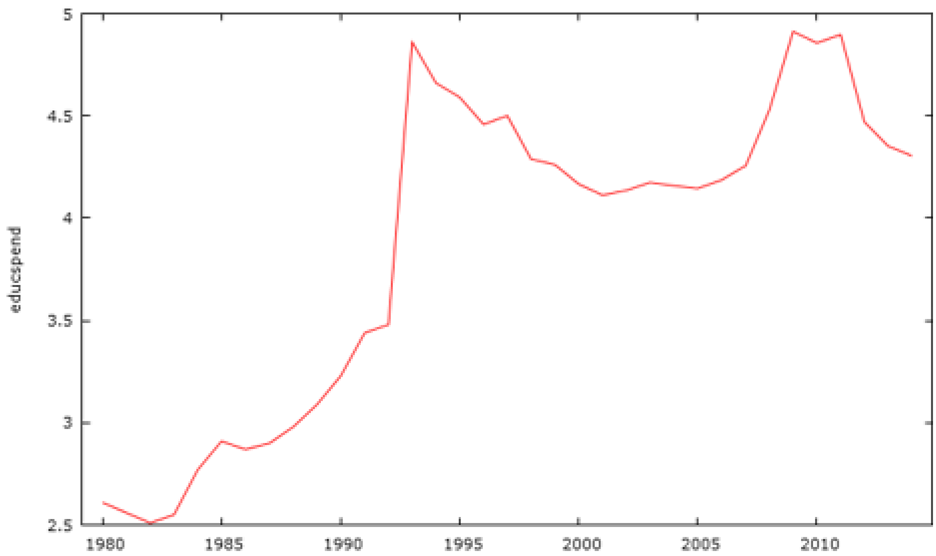


Source: own elaboration based on data from the SWIID

In view of the evolution of the gini index, it can be seen there is a relatively increasing trend. In a favourable economic context for inequality, the index decreased from the 1980s to the 1990s. Despite the recession at the beginning of the 1980s, this was a period where more productive and better paid jobs were available, especially after the start of the economic increasing from 1985 onwards. Moreover, throughout that decade, better policies were applied in the redistribution of wealth, according to Ayala et al. (1993)

and other studies. However, the economic and financial crisis of the 1990s, triggered by the bursting of the real estate bubble in Japan and by the oil price tensions, interrupted the downward trend in inequality. Pijoan-Mas and Sanchez-Marcos (2010) corroborated that inequality declined until 1991, only to grow moderately thereafter, after the impact of the crisis. In the mid-1990s, the Spanish economy entered an expansionary phase until 2007. From 1995 until the beginning of the new millennium, the gini index decreased slightly. From 2002 onwards, however, the index continued rising. This coincided with a massive influx of immigrants. After Spain joined the Monetary Union (1999), there was a favourable financial climate, where financial and mortgage credit began to be abused. After the bursting of the housing bubble in 2008, there was a major financial and economic crisis, which had a drastic negative impact on inequality.

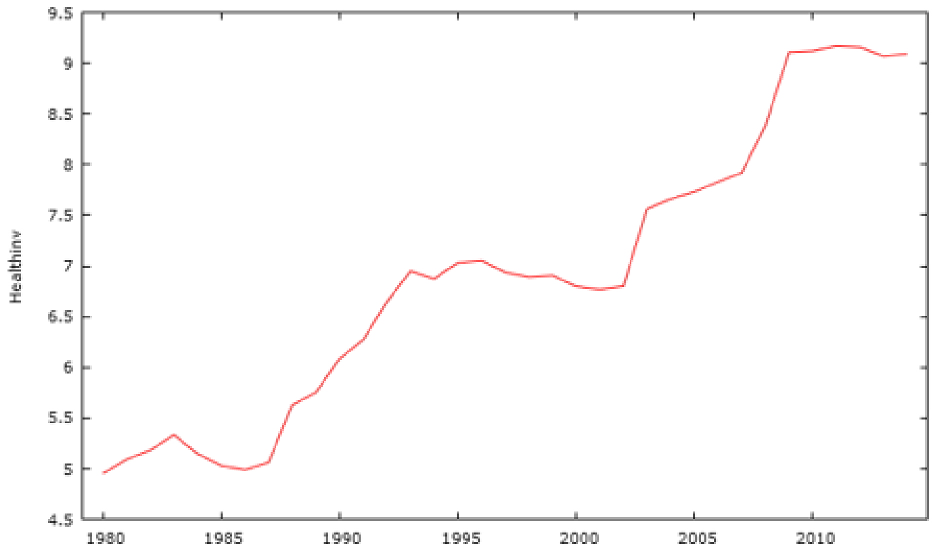
Figure 2: graph of the evolution of the spending in Education



Source: own elaboration based on data from Espuelas (2013) and Databank

There is a growing trend in public spending on education. It is worth noting the large increase at the beginning of the 1990s. At first sight it seems strange, because the effects of the crisis of the 1990s were evident until about the middle of that decade. If we take into account the source of the data used up to 1995, they probably do not accurately represent the value used at that time. However, both Calero and Gil (2013) and Espuelas (2013) highlight the expansion and consolidation of the welfare state in Spain in the 1990s. It is argued that the Spanish welfare state converged towards European levels, specifically from 1993 onwards. Therefore, the reason for such an abrupt increase, may also be due to the increase in public funds. In which part of these funds were allocated to education, increasing the numerator part of the ratio, thus compensating for the decrease in GDP due to the crisis. From 1996 onwards, the data are more reliable, as they come officially from Databank. During the expansionary phase between 1995-2007, we can see how this expenditure remains relatively stable. After the outbreak of the 2008 crisis, the effects on public spending on education were not immediate. On the one hand, from 2008 to 2010 the upward evolution of this variable stopped. It remained constant during those two years, due to the decrease in Spanish GDP after the crisis. It was not until 2010 when the Spanish government resorted to budget cuts. From then on, the percentage of public spending on education went downwards.

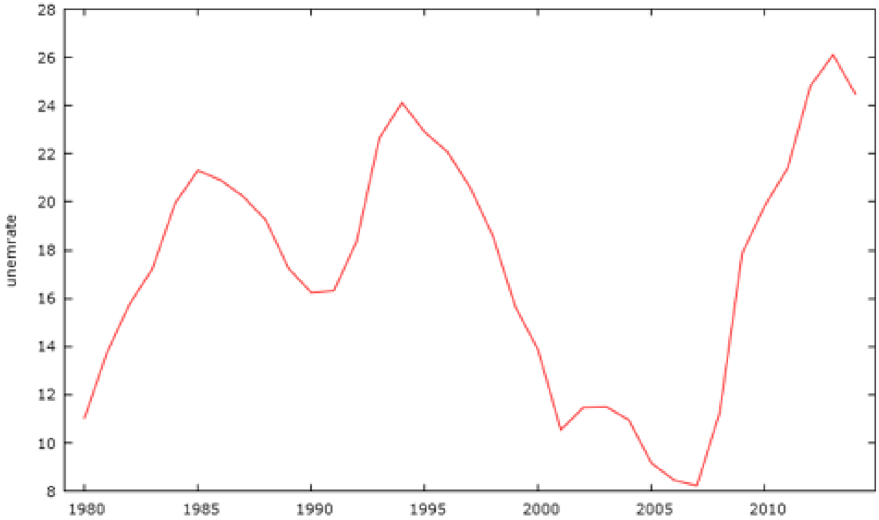
Figure 3: graph of the evolution of health investment



Source: own elaboration based on data from the OECD Databse

There is a growing trend in health investment. According to the analysis of Espuelas (2013), between 1981 and 1986, the universalisation of healthcare for the entire population was consolidated. The effects of universalisation became evident from 1987 onwards. This was the year in which Spain, after the brief recessionary period at the beginning of the 1980s, finally experienced an economic expansion after joining the European Economic Community in 1986. Both private and public investment in health care was encouraged until almost the middle of the 1990s. Despite expansionary phases from 1995 onwards, health investment remained constant until the early 2000s. The limitations imposed by the Maastricht Treaty and, subsequently, by joining the single European currency, affected public debt, the fiscal deficit and, therefore, social spending was affected. Thus, investment in health was also affected, especially public investment. However, as private health investment is also considered, the variable turned out to be more stable compared to the oscillations suffered by public spending on education. Once Spain adapted to the euro, the expansionary phase in GDP was also evident in health-related spending and investment until 2007. After the impact of the 2008 crisis, investment in health remained constant. In spite of budget cuts, some private investment compensated for this reduction.

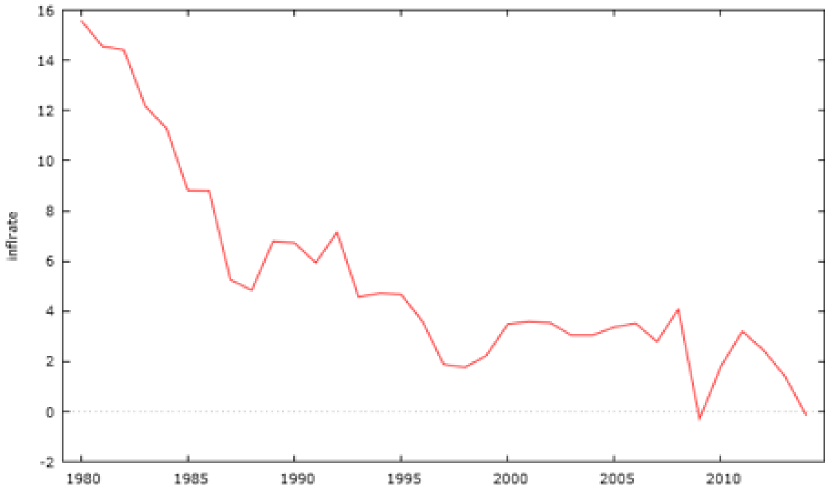
Figure 4: graph of the evolution of the unemployment rate



Source: own elaboration based on data from the World Economic Outlook

In this case, no trend seems to be observed. The time series seems to oscillate around a constant non-zero average. The Spanish economy, despite an improvement in income distribution in the 1980s, remained in a period of recession until 1985. This also affected the unemployment rate and, as we shall see below, inflation. After the accession of the socialist to government in 1982, from 1986 onwards, its policies became evident and led to an expansive economic phase. As a result, unemployment eased until the crisis emerged in 1990. After the economic expansion of 1995, unemployment went down partly due to the massive arrival of immigrants in Spain and the entry into the Monetary Union. In spite of reduced unemployment rates and economic growth, according to Ochando (2010), the labour market, characterised by a high level of temporary and precarious employment, and the occupation of a large number of immigrants in low-productive and low-paying jobs, led to an increase in inequality. After the bursting of the real estate bubble, and the strong influence of the housing sector on the Spanish economy (Daher 2013), unemployment began to increase drastically from around 2008 onwards.

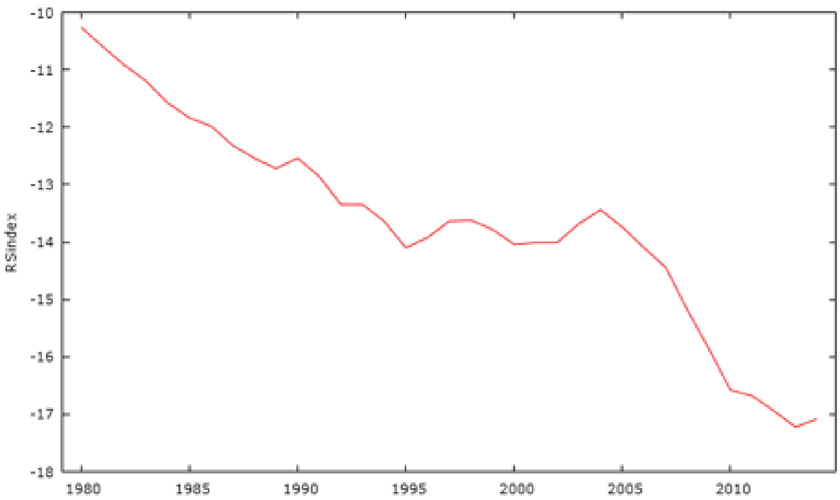
Figure 5: graph of the evolution of the inflation rate



Source: own elaboration based on data from the World Economic Outlook Database

With respect to the inflation rate, a negative time trend can be observed. Since the beginning of the 1980s, the Spanish government aimed to maintain stable inflation levels. The reason was that, until 1980, Spain maintained high levels of inflation, derived from a previous economic recessions. Spain gradually achieved its objective after devaluing the peseta and controlling wages. However, after the crisis of the 1990s, this downward trend continued, but with a more gentle slope. Following the second expansionary phase, and after having adapted to the euro, Spanish inflation slightly gradually increased from 2000 onwards, until it reached a relative peak in 2008. This was largely due to the expansion of the real estate sector (Daher, 2013). However, the 2008 financial crisis subsequently caused both house price indices and consumer price indices to fall.

Figure 6: graph of the evolution of the Reynolds-Smolensky index



Source: own elaboration based on data from the SWIID

The Reynolds-Smolensky index also has a negative trend. This index is calculated as the difference between the gini index of net income and the gini index of gross income. Thus, the greater the difference between the two variables, the greater the redistributive capacity of the state to tax. Graphically, it can be seen that, in general, this index maintains a decreasing trend. That is, as time goes by, the difference between the two variables becomes larger and larger, and even more negative. Apparently, the redistributive ability of the state appears to be increasingly effective over time. It is worth noting the interruption of the downward trend in the 10 years between 1995-2005. In the research work by Fernandez and Sanchez (2012), they describe in detail the income tax reforms during this period. Basically, they conclude that all these reforms, in sum, did not bring about any substantial improvement in the redistributive process. In fact, they explain that both the progressivity of the tax and its redistributive properties were indirectly affected by all the reforms. In the face of the 2008 crisis, the slope was slightly softened, as tax revenues were lower.

Despite the two previous interruptions in the trend, when using the Reynolds-Smolensky index, in general, the redistributive capacity of Spanish personal income tax is undeniable.

4. Source of data

As can be seen in table 1, the source of the data we will use is indicated. In addition, we cannot ignore the following observation:

Given the difficulty of obtaining data on Public Expenditure on Education, the data which will be used to analyse the period 1980-1995, come from a report from the Bank of Spain, carried out by Espuelas (2013). The data were obtained by linear interpolation techniques, from data available from alternative official sources, and from data obtained by simulation techniques. Whereas, from 1996 onwards, the data are obtained from the General Government Accounts. These have been obtained from official Databank data from the World Bank.

5. Methodology

So as to test the previous hypotheses, the simple OLS method will be used using time series. In this way, we will be able to estimate the influence of the explanatory factors on the gini index over the selected period.

However, although all variables (except for the unemployment rate) seem to have an apparent deterministic trend, it may not be enough to include it in the model. Nevertheless, our variables could be influenced by a stochastic trend. In order to check its existence, we will resort to the analysis of stationarity for each of the variables. Depending on whether they are stationary or not, we would adequately specify our econometric model of study. The fulfilment of the required assumptions for the proper use of OLS, will also be briefly analysed. In addition, the possibility of including some lag of the endogenous variable in the model will also be considered. In this way, we will be able to properly specify the model and extract the linear relationships with the highest possible reliability. Therefore, in this section we are going to explain how it works each stationarity test, and to summarize the methodoly employed related to OLS with its required transformations, in order to carry out our research. The results of them will be explained in detail in section 6.

5.1 Stationarity and stationary tests

Stationarity is a fundamental requirement for the consistency of OLS estimators. In view of the existence of a clear trend in almost all the variables, these could indeed satisfy the criterion of stationarity over a trend. Thus, these variables would oscillate about the trend, with a constant variance over time. However, in this study we will try to ensure that each variable fulfills the condition of stationarity in covariance. In which, both the mean and the variance of the respective series remain constant over time.

The importance of compliance with stationarity remains in the stability of our variables over time. Thus, we will be able to extract linear relationships with the greatest possible reliability and precision, using OLS. Otherwise, the predictions would be disturbed by the existence of a pattern that is not predicted by our model, caused by the presence of a "shock". This may be caused either by the presence of a unit root, or due to the existence of a deterministic trend. In the former, it happens when stochastic shocks have permanent effects on the variable, so that the process is not reversible to the mean. In the second, it is when the shocks have only transitory effects on the variable. Where the variable will tend over time to return to a stationary mean, which evolves deterministically along a trend.

In view of the evolution of our variables, it would make sense to consider the possible existence of stationarity, around a deterministic trend (apart from the unemployment rate). However, our variables might suffer from some kind of stochastic trend, caused by the existence of persistent unit roots over time.

We will proceed to analyse whether the variables show any kind of trend. To do so, first, we will use the Augmented Dickey-Fuller test (ADF). Secondly, an additional test will be used to complete and improve the first one used. For this purpose, the ADF test will also be used, using Generalised Least Squares (GLS). Finally, the KPSS test will be employed to complement the previous tests, with the aim of analysing with greater precision whether the series itself has a unit root, or is a stationary series which oscillates around a trend.

ADF test

First, it is worth noting the logarithmic transformation of each of the variables, to ensure the stationary condition of constant variance is satisfied. With the exception of the inflation variable and the Reynolds-Smolensky index, because their representative sample shows negative observations values. Therefore, it is not possible to apply logarithms to them.

The ADF test has the advantage of directly exposing the existence of a unit root in the null hypothesis. It can also use a version which is capable of performing the test by incorporating a deterministic trend.

The regression for the test will consist of regressing the variable of interest on the same lagged variable. In addition, it will also be regressed on the deterministic trend (except for the unemployment rate variable):

$$Y_t = \alpha + \delta time + \rho Y_{t-1} + \varepsilon_t \text{ (1)}$$

The purpose of applying this test is to study the relevance of the correlation between a given variable at the present time and its first-order lag. Using equation (1), we can check if persistence really exists. Thus, the contrast to execute the Dickey-Fuller test would look like this:

$$H_0: \rho = 1$$

$$H_1: \rho < 1$$

In case of H_0 being true, the series under study would have a strong dependence on its lag. That is, it would show a strong persistence over time, breaking one of the key assumptions for the fulfilment of stationarity. In case of rejecting it, the variable under study would have a stationary AR (1) process.

However, in the case of the unemployment rate, in the absence of an apparent trend, the H_1 would indicate the fulfilment of stationarity over a constant mean.

For simplicity, when carrying out the contrast, it is usual to subtract Y_{t-1} from both sides of equation 1. Thus, the equation for the contrast would result:

$$\Delta Y_t = \alpha + \delta time + \theta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \cdots + \gamma_p \Delta Y_{t-p} + \varepsilon_t \text{ (2)}$$

The possibility of increased inclusion of time delays of ΔY_t has been taken into account, to delete any type of serial correlation. As we are working with annual data, it should be enough to include only a couple of lags for each of the selected variables. However, we will trust the AIC criterion when choosing the optimal number of lags.

DFA-GLS test

For the series used in our analysis, it would be asymptotically better to apply this test. Because it would efficiently estimate the deterministic component, and the parameters of equation 2, having applied GLS to the data set.

GLS are used to eliminate any kind of trend in any series. In this way, for those variables where the presence of a unit root can be suspected, this trend would be eliminated. Therefore, we will be able to work with these transformed data, to efficiently test for the existence of unit roots. That is, on equation 2, the data is transformed to eliminate any type of tendency.

Given the above, and knowing this test controls by default for a linear trend in time with higher precision, the hypotheses for almost all series will be:

$$H_0: the\ variable\ is\ affected\ by\ a\ random\ walk$$

$$H_1: the\ variable\ is\ stationay\ around\ the\ trend$$

We will apply this test in section 6.

KPSS test

In recognition of Kwiatkowski et al. (1992), is used to complement those performed previously. Given the low empirical power of acceptance of the alternative hypothesis of stationarity of the other two tests, we use this test which establishes, as the null hypothesis, the presence of stationarity around a trend.

Moreover, it is more precise when analysing the possible existence of stationary processes with trend, as the methodology of the other two tests may lead to the confusion of these with stochastic processes.

Following this new approach, we will be able to consider both empirical methodologies simultaneously, when determining whether our variables have unit roots. In section 6, we will show the results obtained through this test.

5.2 Differentiation

As we will explain later in section 6, we are going to need to apply first differences. So, we have to differentiate the variable at time (t), with respect to the same lagged variable. That is, $\Delta Y_t = Y_t - Y_{t-1}$ (3) for each variable, in order to delete any kind of trend.

Once the differentiation would be applied, we will re-apply the three previous tests to the already differenced variables.

5.3 Inclusion of lags of the dependent variable

In order to complement the model, and not to leave out any important explanatory factors, we will consider the possibility of incorporating some lag of the proposed dependent variable:

$$\Delta log\_Gini_t = \beta_0 + \beta_j x_j + \beta_2 \Delta log\_Gini_{t-1} + \cdots + \beta_k \Delta log\_Gini_{t-k} + v \quad (4)$$

Depending on the significance of the lagged variables, we will include a required number of lags. We will carry out a restricted test in the next section, so as to include the optimal number of them.

5.4 Asymptotic Normality

By applying the differencing method, we will assure the compliance with stationarity in covariance. This turns out to be a very important requirement for the consistency, of the estimators to be estimated in the OLS regression, even for those of stochastic inference. Besides, we will check later in section 6 wheter our variables fulfill the property of perfect non-collinearity and zero conditional mean. In this way, consistency in our regressors would be fully guarantee:

$$p\hat{\beta}_j = \beta_j \quad (5)$$

That is, the regression estimators would be consistent in the sense that, as the size of the observations increases, they would converge towards the true population value.

On the other hand, we also need to demonstrate the compliance with stochastic inference requirements. We need to check if the autocorrelation and heterocedasticity assumptions are given.

Autocorrelation in the errors is a very important assumption to take into account, together with compliance with homoscedasticity. The fact of incurring autocorrelation and heterocedasticity problems, would have a negative impact on the assumption of maintaining the minimum possible variance. Therefore, the inference procedures of our model would be affected.

We will start by applying the Durbin-Watson test, applicable for those regressors that are strictly exogenous with the error term. We will perform this test in section 6. Given that the objective is to test for the possible existence of some kind of correlation in the error, between the current and the past, we will work with the following equation:

$$u_t = \rho u_{t-1} + \varepsilon_t \quad (6)$$

However, it is very difficult to assume that none of the explanatory variables will be correlated with the error term. Therefore, we will use the Breusch-Godfrey test. Thus, we would be working with the following equation:

$$u_t = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \rho u_{t-1} + \varepsilon_t \quad (7)$$

To use the LM and F statistics in this test, we shall ensure that there is no problem of heteroscedasticity.

Once all these assumptions are fulfilled, the estimators used for our investigation will have asymptotically normal distributions. Both the standard errors and all the statistics obtained will be asymptotically valid. That is, as the sample increased, these statistics would approach the true population value.

5.5 Econometric Model

After we have proved the compliance with all the previous assumptions, we will be able to estimate this proposed econometric model, by employing time series. So that, both our reserch and started hypothesis can be checked.

$$\Delta log\_Gini_t = \beta_0 + \beta_1 \Delta log\_educ_t + \beta_2 \Delta log\_inv\_health_t + \beta_3 \Delta log\_unem\_rate_t + \beta_4 \Delta inf\_rate_t + \beta_5 \Delta RS\_ind_t + \beta_6 \Delta log\_Gini_{t-1} + u_t \quad (8)$$

Given the logarithmic and differentiation transformations that we will perform, our model used will be one of growth. We proceed to present in next section the statistical and empirical results introduced in this section.

6. Results

In the first place, all the tests and assumptions from the previous section are going to be checked. After that, in subsection 6.2, we will show the estimation results of the research, explaining its economic implications.

6.1 Statistical results

Considering the equation 2, the hypotheses would remain:

$$H_0: \theta = 0$$

$$H_1: \theta < 0$$

The relationship between Y_{t-1} and ΔY_t has to be considered as before in equation 1, through the coefficient θ. We will consider the asymptotic p-value approach associated to Y_{t-1},to see whether this relationship turns out to be strong or weak. We will accept or reject the H_0 respectively. In addition, we will see if the trend is significant by applying the test.

If we apply the ADF test for each of the variables:

Table 3: ADF results

Variable	Optimal number of lags	Asymptotic p-value	Trend
Gini index	1	0.1906	0.000589308 **
Education spending	0	0.8784	0.00117921
Health investment	1	0.3049	0.00554192 **
Unemployment rate	1	0.1150	...
Inflation rate	0	0.2266	−0.0724072 *
Reynolds index	1	0.4237	−0.0272266 **

Source: own elaboration (Gretl)

According to the asymptotic p-value approach, we would not be able to reject the H_0 which indicates the existence of a unit root, as long as we use a significance level of 1%. That is, there would be a strong dependence of the variable at time (t) on its lag of order 1. If we analyse the significance of the trend, all are significant at 5% (**), with the exception of the inflation rate, which is significant only at 10% (*). In the case of the trend in public expenditure on education, it has not been statistically significant. It is probably due to the abrupt increase in the 1990s, from the data used up to 1995, the regression contrast has not been able to detect a clear linear trend. Still, since there is a clear upward trend, we cannot ignore the trend when applying the test. In the case of unemployment, no trend has been included, as discussed above. The sign of the estimated trend coefficient for each variable (see the trend column in table 3) coincides with its respective evolution over time (section 3.3)

Given we are working with a relatively small sample, it has been difficult to reject the null hypothesis by performing this test. If the series had a relevant stochastic trend, and we applied the test with equation 2, it would only affect the efficiency of the estimation of the trend parameter. Therefore, the statistics used to carry out the test contrasts would also have efficiency and reliability problems.

In the following, we will use the ADF test using GLS to solve this problem. In this way, we will obtain more reliable conclusions on the study of compliance with stationarity.

So, we are making again the unit root search, by using this second test:

Table 4: test ADF-GLS results

Variable	Number of lags	t statistic
Gini index	K= 1	-2.6026
Education spending	K= 0	-1.43611
Health investment	K= 0	-1.93742
Inflation rate	K= 0	-2.20704
Reynolds-Smolensky index	K= 1	-2.27178
Variable	**Number of lags**	**Asymptotic p-value**
Unemployment rate	K= 1	0.01752

Source: own elaboration (Gretl)

Table 5: tabulated critical values by EPS (1996)

Critical values	Significance level
-2.89	∝= 10%
-3.19	∝= 5%
-3.77	∝= 1%

Source: own elaboration

The optimal number of lags was selected for each of the variables, based on the AIC criterion. We obtained a t-statistic (for all variables with a trend) lower than all the tabulated critical values. That is to say, this test seems to corroborate with the previous assumption, that all these variables, effectively have a random walk, using a ∝= 1%.

When analysing the unemployment variable, the modification of the alternative hypothesis must be considered:

$$H_0: \rho = 1; random\ walk$$

$$H_1: \rho < 1; stationarity\ over\ a\ nonzero\ mean$$

Applying the same test for the unemployment rate would result in an asymptotic p-value of 0.01752. Using this approach, with ∝= 1%, the null hypothesis of the presence of a unit root could not be rejected. For this series without trend, we would use the critical values of the ADF without constant or trend. For simplicity, the p-value approach has been used. As there is not apparently any trend to be considered, the parameters efficiency of the test should not be affected. So the critical values of the ADF can be used.

However, the evidence for rejecting the H_0 with ∝= 5% and ∝=10% cannot be ignored. In my opinion, it seems to me more convenient to apply the smaller nivel of significance. Thus, we would reduce the probability of incurring the type I error. This test would then support the above evidence on the apparent existence of a unit root.

Following this, we are going to apply the another complementary test:

Table 6: test KPSS results

Variable	p-value
Gini index	0.089
Education spending	0.027
Health investment	>0.10
Unemployment rate	>0.10
Inflation rate	0.019
Reynolds-Smolensky index	>0.10

Source: own elaboration (Gretl)

In view of the results obtained, we could reject the null hypothesis, at ∝= 5%, in the case of public expenditure and inflation. In the case of the gini index, we can reject it, but with ∝= 10%. However, we prefer working with ∝=1%, as we explained before, where the null hypothesis of stationarity on a trend cannot be rejected. In the case of the unemployment rate, as there does not seem to be any kind of trend visually, the null hypothesis established is that of stationarity over a constant, where this hypothesis neither could be rejected.

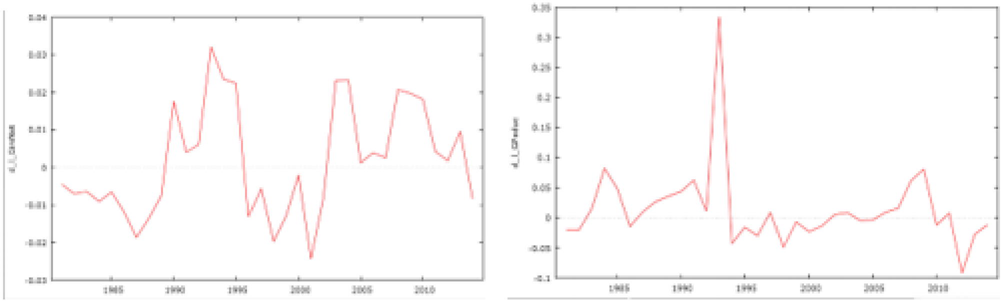
Considering the results obtained by each test, there seems to be a contradiction in them. On the one hand, in case of following the criteria of the first two tests, anyone would conclude that the variables would be characterised by the presence of unit roots. On the other hand, using the KPSS approach, the variables would be considered as stationary, around a trend (except for unemployment). As can be seen, in both cases their respective null hypothesis could not be rejected. By using the KPSS test, one would expect it to support the results obtained by the other tests. Unfortunately, this is not the case. It is probably due to the small number of observations obtained for our sample, that both typologies cannot reach the same conclusions.

In general, there are indications to apply differentiation in order to eliminate any existing trends. Either for the deterministic trend, or for the stochastic trend. In this way, we would be consider the two possible results analysed in the previous tests. So, we will carry out the method of differentiation.

Once we have applied the differentiation method, by using the equation 3, we can see every series now have a more stable evolution:

Figure 7: (1) graph of the evolution of the gini index, (2) graph of the evolution of the spending in education

(1) (2)

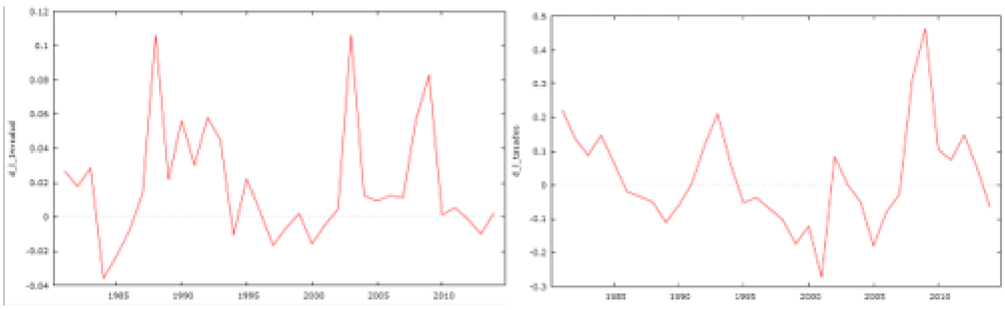


Note. The series have been transformed by applying logarithms and the differentiation method

Source: own elaboration, (1) based on data from the SWIID, (2) data from the research by Espuelas (2013) and Databank

Figure 8: (3) graph of the evolution of the investment in health, (4) graph of the evolution of the unemployment rate

(3) (4)

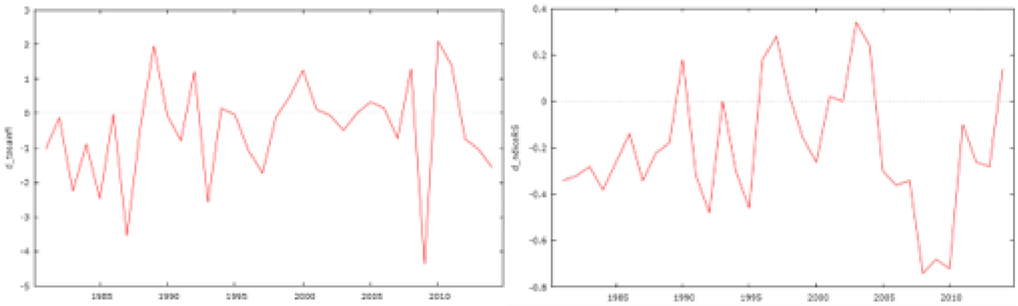


Note. Both the differentiation method and logarithms have been applied

Source: own elaboration, (3) based on data from the OCDE Database, (4) data from the World Economic Outlook Database

Figure 9: (5) graph of the evolution of the inflation rate, (6) graph of the evolution of the Reynolds-Smolensky index

(5) | (6)



Note. Only the differentiation method has been applied to both series

Source: own elaboration, (5) based on data from the World Economic Outlook Database, (6) data from the SWIID

At first sight, we could assume they already fulfill the necessary stationarity requirements for OLS. To make sure, we are going to re-apply the three previous tests to the already differenced variables.

Starting with the ADF, we would work with the following equation due to differentiation, keeping the same previous assumptions:

$$\Delta\Delta Y_t = \alpha + \theta\Delta Y_{t-1} + \varepsilon \text{ (9)}$$

As there does not seem to be any trend, we are not taking it into account. In case of the alternative hypothesis being correct, the series would become stationary processes, with constant mean and variance.

Table 7: ADF results on differentiated series

Variable	Number of lags	Asymptotic p-value
Gini index	K= 0	0.04324
Education spending	K= 0	6.19e-005
Health investment	K= 0	0.001337
Unemployment rate	K= 0	0.03868
Inflation rate	K= 0	1.92e-006
Reynolds index	K= 0	0.02232

Source: own elaboration (Gretl)

According to the AIC criterion, any time delay of $\Delta\Delta Y_t$ has been used neither. In view of the results, we can reject the null hypothesis at $\propto$=5%. That is, according to this test the variables would be stationary after differencing. However, this assumption is being made allowing for a larger margin of type I error. Maybe some of the variables, as in the case of the Gini index, require the application of second differences for greater certainty. For the sake of simplicity, the null hypothesis will be rejected by assuming that ∝.

Table 8: ADF-GLS results on differentiated series

Variable	Number of lags	Asymptotic p-value
Gini index	K= 0	0.002602
Education spending	K= 1	0.001068
Health investment	K= 0	8.473e-006
Unemployment rate	K= 0	0.0077
Inflation rate	K= 2	0.00378
Reynolds-Smolensky index	K= 0	0.0008409

Source: own elaboration (Gretl)

If we re-applied the ADF-GLS test for all variables, after differencing, we would obtain asymptotic p-values of less than ∝= 1%. That is, we would now have more statistical evidence that each series would maintain stationarity in covariance, as they no longer appear to maintain any kind of trend. Given that this test is more accurate than the previous one, we can now reject the null hypothesis with greater certainty, at ∝=1%, assuming a smaller type I error than in the previous test.

On the other hand, by applying the KPSS we have obtained greater statistical evidence in Gretl of not rejecting the null hypothesis, after differentiation. The p-values obtained are well below the 1% significance level.

We can conclude that, after having applied the method of differencing by eliminating any kind of trend, the three tests now establish the variables are stationary about a constant mean. There is no doubt that our variables already fulfill the requirements of stationarity.

Number of lags

Table 9: regression of lags

	$\Delta\log\_Gini_t$	t statistic	p-value
constant	0.00144802 (0.00318406)	0.4548	0.6542
$\Delta\log\_educ_t$	0.0677309 (0.0356926)	1.898	0.0723 *
$\Delta\log\_inv\_health_t$	0.0870807 (0.0723616)	1.203	0.2429
$\Delta\log\_unem\_rate_t$	0.0452044 (0.0208316)	2.170	0.0422 **
Δin_rate	0.00300969 (0.00154912)	1.943	0.0662 *
ΔRS_ind_t	0.00992055 (0.00908975)	1.091	0.2881
$\Delta\log\_Gini_{t-1}$	0.473302 (0.161272)	2.935	0.0082 ***
$\Delta\log\_Gini_{t-2}$	0.0459830 (0.199171)	0.2309	0.8198
$\Delta\log\_Gini_{t-3}$	−0.155303 (0.197390)	−0.7868	0.4406
$\Delta\log\_Gini_{t-4}$	0.133040 (0.174700)	0.7615	0.4552
N= 30 $R^2 = 0.655423$ F (9, 20) = 4.226902 ^ P-value (F) = 0.003492			

Source: own elaboration (Gretl)

In order to decide how many lags to incorporate in the model, the equation 4 has been estimated, with the maximum possible number of lags of the Gini index. Apparently, only the first-order lag is significant. To not overspecify the model, we proceed to make an exclusion restriction for the other three lags. For this purpose, a restricted model estimation has been performed, without the three variables, to test the following hypotheses:

$$H_0: \beta_{t-2} = \beta_{t-3} = \beta_{t-4} = 0$$

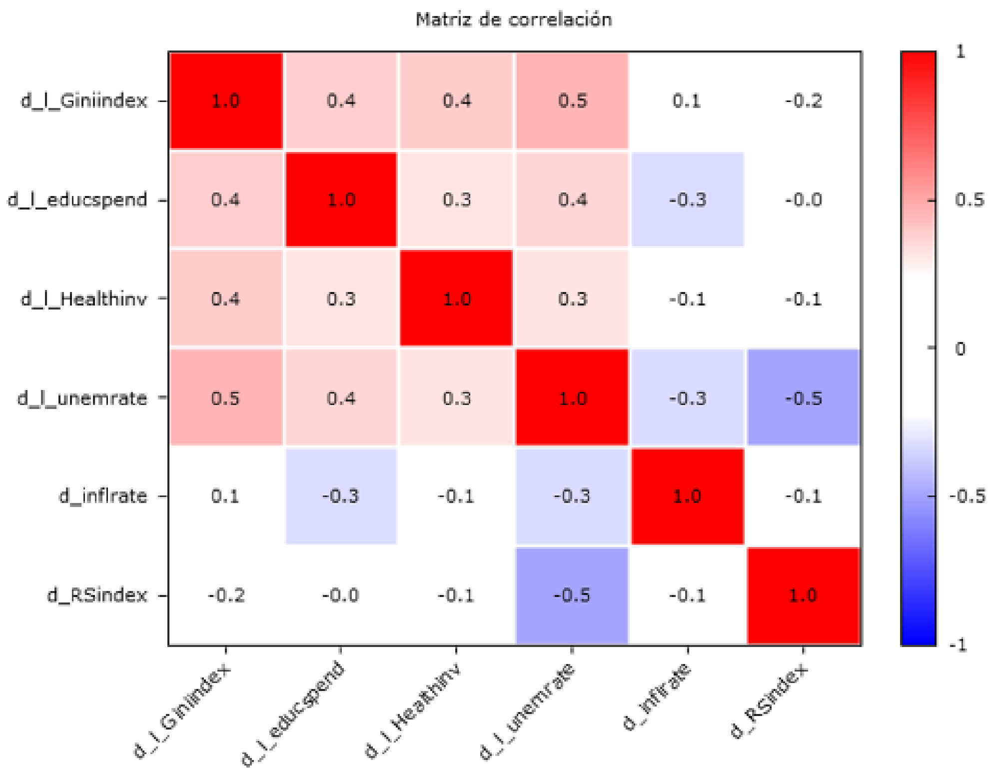
$$H_1: H_0 \; is \; false$$

Considering the unrestricted model, which only includes the first-order delay, and the restricted model, an F-statistic with an associated p-value of 0.841976 has been obtained in Gretl. That is, at ∝=1%, there is statistical evidence not to reject the null hypothesis of joint non-significance. So only a single lag of the Gini index will be incorporated, to explain the evolution of the Gini index.

Assumptions of Asymptotic Normality

After our variables have been diferenciated, we need to verify wether the other assumptions of consistency are given or not.

Figure 10: Correlation matrix



Source: own elaboration (Gretl)

When obtaining the correlation matrix, in this figure it can be seen there is some correlation between the different variables, without reaching any perfect correlation equivalent to unity.

After have assured the stationarity, and have specified the model by approximating the economic literature as far as possible, our explanatory variables should be (at least) contemporaneously exogenous with the error term. Therefore, although the property of unbiasedness is probably not comply due to the contemporaneity of our regressors, everything seems to indicate that the various consistency assumptions would be fulfilled. So the equation 5 can be applied to our variables.

In our case, given the limited data availability required for the analysis, the series have few observations. Therefore, the linear relationships cannot be taken at face value, since the estimated coefficients are probably still far from the true population value. It is logical that, if we wanted to consider a larger time range using more observations, the coefficients would gradually converge to their true population value. Nevertheless, we can still extract interesting linear relationships, because the series satisfy the assumptions which guarantee the consistency of OLS.

On the other hand, we have to check the stochastic inference assumptions of autocorrelation and homocedasticity. We will start first by estimating equation 6 related to autocorrelation, assuming that our regressors are strictly exogenous. Where only a single lag of the error term u_t has been selected, since we are working with annual data, and because we have chosen to incorporate only a single lag of the dependent variable. Assuming that ε_t fulfills the assumption of zero conditional mean and constant variance, we apply the Durbin-Watson test in Gretl, to contrast these following hypothesis:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Bearing in mind that the statistic has to be $DW \approx 2$ in order not to reject the null hypothesis, this statistic has been estimated with a value of DW= 2.42631, with a p-value= 0.838049. The statistic is close to 2, as well as showing a p-value of well over ∝= 1%. Therefore, we have statistical evidence of not rejecting the null hypothesis of no autocorrelation, if we assume strict exogeneity of the regressors.

However, if we assume that our explanatory variables are correlated with the error term, we have to estimate the equation 7. To do that, we have estimated the model (equation 8) with our differenced variables and the lag of the dependent variable, including robust HAC errors, in case there is any heteroscedasticity. We have been able to verify that the test results are the same as those shown in the model with standard errors without applying the HAC. Therefore, there does not seem to be any problem of heteroscedasticity.

After estimating the equation 7 in Gretl, we have obtained a p-value associated with the LM and F-statistics with a value of 0.0901, maintaining the previous hypotheses of autocorrelation. We cannot reject the null hypothesis with a $\propto$=1%. That is, we have statistical evidence that the model would not be autocorrelated, with contemporaneously exogenous explanatory variables.

Normally, when models are not autocorrelated, they will also not suffer from heteroscedasticity in (at least) their weakest version. In our case, this is no exception. We have executed several tests, such as the Breusch-Pagan test and the White test, to make sure that there is no heteroskedasticity in the weak version.

After all this, we have been able to verify that our model would now be accurate to extract asymptotically reliable linear relationships. That is, our estimators would be ready to have an asymptotically normal distribution. Where, as the sample size increases, the distribution of the estimators would arbitrarily approximate a normal distribution. Therefore, we can proceed to estimate our econometric model.

6.2 Empirical results

Using the Gretl programme, the econometric growth model of equation 8 has been estimated:

Table 10: model estimation

	$\Delta\log\_Gini_t$	t statistic	p-value
constant	0.000551567 (0.00265495)	0.2078	0.8370
$\Delta\log\_educ_t$	0.0596613 (0.0301099)	1.981	0.0582 *
$\Delta\log\_inv\_health_t$	0.103690 (0.0561529)	1.847	0.0762 *
$\Delta\log\_unem\_rate_t$	0.0392816 (0.0172883)	2.272	0.0316 **
Δinf_rate	0.00298756 (0.00137604)	2.171	0.0392 **
ΔRS_ind_t	0.00929701 (0.00793303)	1.172	0.2518
$\Delta\log\_Gini_{t-1}$	0.504777 (0.124204)	4.064	0.0004 ***
N= 33 R^2 = 0.633975 F (6, 26) = 7.505559 ^ P-value (F) = 0.000097			

Source: own elaboration (Gretl)

We can see that all the variables, with the exception of the Reynolds-Smolensky index (RS_ind), have been found to be statistically significant. Next, we will execute another estimation, but without including the Reynolds-Smolensky index variable. In this way, the model will not be over-specified, and we can obtain more accurate estimates:

Table 11: final estimate

	$\Delta\log\_Gini_t$	t statistic	p-value
constant	−0.00142016 (0.00206796)	−0.6867	0.4981
$\Delta\log\_educ_t$	0.0632647 (0.0301589)	2.098	0.0454 **
$\Delta\log\_inv\_health_t$	0.110004 (0.0562790)	1.955	0.0611 *
$\Delta\log\_unem\_rate_t$	0.0281967 (0.0145711)	1.935	0.0635 *
Δinf_rate	0.00254649 (0.00133270)	1.911	0.0667 *
$\Delta\log\_Gini_{t-1}$	0.491963 (0.124575)	3.949	0.0005 ***
N= 33 R^2 = 0.614640 F (5, 27) = 8.612860 ^ P-value (F) = 0.000056			

Source: own elaboration (Gretl)

When re-estimating the model, the estimates have had some modification. First, the joint significance of the explanatory variables has increased, with a higher F-statistic and an increase in their respective p-values. In addition, the individual significance of some variables has also been modified. For example, public expenditure is now significant at 5%, while unemployment and inflation are significant at 10%. In the first case, individual significance has increased; in the second case, it has decreased. On the other hand, the value of the estimated coefficients have been modified slightly. In the case of expenditure on education and investment in health, the coefficient increases moderately. The remaining explanatory variables show a decrease in their respective coefficients, compared to the first estimation. Lastly, although the R^2 has decreased slightly, it still remains a decent goodness-of-fit. Where approximately 60% of the variation in the gini index is explained by our explanatory variables. However, we are interested in interpreting the linear ceteris paribus relationships, which we will discuss below.

First, we will comment on the associated results of the unemployment and inflation indicators in the gini index. Second, we will explain the results obtained from the distributional measures. Finally, we will analyse the effect of the lag of the endogenous variable:

- In the case of the unemployment rate (log_unem_rate), our model has estimated it has a direct relationship with the gini index. According to the estimation, a 1% increase in the unemployment rate would imply a 0.028% increase in the gini index, ceteris paribus. In other words, the higher the unemployment rate in the Spanish economy, the greater the inequality of disposable income in Spain.

- In the case of the inflation rate (inf_rate), according to our model this macroeconomic indicator would also keep a direct relationship with the dependent variable. In other words, a 1% increase in the inflation rate would imply (ceteris paribus) an increase of 0.254649%. This relationship would imply that the higher the Spanish inflation rate, the greater the inequality in the distribution of income in Spain.

The two relationships extracted from both macroeconomic variables coincide with a large part of the economic literature. Where both relationships are also defended for the case of Spain, for example, in (Doncel et al.); or, for other OECD countries, such as Jantti and Jenkins (2009) in the UK, and Mocan (1999) for the USA. In our case, both rates are only justified at only ∝=10%. Perhaps, with a larger number of observations in the sample, we would have obtained more evidence. Nevertheless, though not having the desired evidence, both results make economic sense:

On the one hand, when the unemployment rate rises, it means a lost of wages and income for those people who have been affected. Those with lower incomes would be affected exclusively by being unemployed. Especially after the impact of an economic crisis. For example, as we saw in section 3, after the crisis of the 1990s and 2008, unemployment increased considerably, negatively affecting inequality. Therefore, the inequality gap between rich and poor would increase even more, assuming that higher income individuals were not as badly affected by rising unemployment.

On the other hand, when the inflation rate rises, it implies a general increase in consumer prices. For those people with lower incomes, inflation is a problem. The reason is it creates a greater obstacle for them in the acquisition of goods and services, given their tighter budgets. We saw during the recessionary phase of the 1990s, Spain had problems controlling inflation. Moreover, during the second expansionary phase, prices in the Spanish economy gradually increased until 2008. In part, this was due to the climate of confidence that emerged largely due to the housing boom. Inflation was rising, along with house prices. Both developments coincided with an increase in inequality. We have therefore found some evidence that when prices in the Spanish economy rise, lower income groups are the most affected.

Therefore, in the case of the crisis of the 1990s, unemployment and inflation increased. This affected exclusively those individuals with the lowest purchasing power. Then, during the second expansionary phase, these groups were also affected by the inflationary effect. Although having a low unemployment rate in Spain, they were mostly subjected to low-productive and low-paid jobs, characterised by a high degree of temporariness and precariousness. After the impact of the 2008 financial crisis, the effects on unemployment were devastating. As a result, inequality increased considerably.

We have some evidence that the inequality gap increases in Spain when unemployment and inflation rise. Therefore, the Spanish Government should be careful with the design of its structural employment policies, with the aim of reducing inequality, and not making its past mistakes. With regard to inflation, since Spain is part of the common European monetary policy, they should control it as much as possible through better control of wages and prices. Because the higher the wages, the higher the inflation rate. That is, as people have greater purchasing power, prices in the economy will tend to increase.

If we look at the effects of redistributive measures:

- In the case of public spending on education (log_GP_educ), an increase of 1% leads to an increase in the gini index of 0.0632647%, ceteris paribus. That is, according to the estimation, there would be a direct effect between such expenditure and inequality in Spain, increasing the latter one.
- For health investment (log_inv_health), the model has estimated that (ceteris paribus), an increase of 1% would imply an increase in the dependent variable of approximately 0.11%. According to our estimation, there would also be a positive relationship between health investment and inequality in Spain.

In view of our results, both variables would against to their economic nature. In fact, the significance of both variables cannot be ignored. That of education expenditure is 5%, and that of health investment is 10%.

It is weird to observe both variables increase inequality in Spain, when they should reduce it. This is probably because we are working with a relatively small sample, and the OLS model has not been able to accurately capture the distributional effects of both measures. Both health investment and public spending on education may take some time to consolidate their distributional effects. That is, as they are normally effective investments in the medium/long term, the methodology used may not have accurately captured the effects of both redistributive measures. Even more so when we are using such a basic model as OLS. For example, in the two papers by Calero and Gil (2013) and (2014), among others, they demonstrate the redistributive effectiveness of both measures in the years 2005 and 2008, and for the year 2010 (respectively), using a more accurate, and advanced methodology.

However, it could be the case that both measures are not truly efficient. In fact, the report by García-Altés and Ortún (2014) comments on the redistributive inability of such measures, and on the need to reform the policy in Spain.

We strongly believe that the lack of data is the reason why the model could not accurately estimate the effect of both redistributive measures. In fact, as we are working with a small sample, our estimates are not 100% reliables, considering what we explained before about the consistency of our parameters. Moreover, given the period considered, the different cycles may have influenced the delayed effect of the measures. This interfers with the capturing of the results.

Finally, if we consider the effect of the lagged endogenous variable:

- An increase (ceteris paribus) of 1% in the gini index corresponding to year (t-1) would increase the gini index at time (t) by 0.50%. That is, according to this estimate, there would be a transitory effect on income inequality with its first-order lag.

It is curious to have obtained this relationship when the economic literature talks about the persistence of inequality in Spain and in other OECD countries. Probably, the reason why we have obtained a transitory effect on inequality is due to the stationary transformations mentioned above. However, even though having applied the differencing method, the estimated relationship between the two variables is exactly half of unity. Actually, when the coefficient is exactly equal to unity, there is persistence. Moreover, we must consider the convention in the literature of accepting the persistence hypothesis when the coefficient is close to 0.8. In our definitive estimation, our coefficient is not very far from this value.

Therefore, though having obtained a transitory effect on income inequality, we cannot ignore the existence of persistence in inequality in Spain. Probably, if we performed a more extensive study on persistence in Spain, looking at a longer period, we would agree with the convention established in the literature.

7. Conclusions

After having reviewed the literature, analysed our variables and the historical context, and obtained the results, we have been able to answer some of our hypotheses set out at the beginning of the study.

First, we have some evidence that both inflation and unemployment have had a negative influence on income inequality in Spain. Taking into account the period analysed, and in view of recent economic cycles, it is not surprising that both macroeconomic indicators increase inequality. Especially in times of crisis. For example, in section 3 we saw that both the 1990 and 2008 crises caused unemployment to rise considerably. In addition, we showed the change in the downward trend of inflation due to the crisis of the 1990s. We could see during the second expansionary phase up to 2008, there was also a certain increase in inflation. We observe that, in all these cases, income inequality in Spain worsened as a consequence of the influence of both variables.

Secondly, we have not obtained the expected effects of the redistributive measures, neither in the health nor in the education sector. This is probably due to the presence of several cycles in such a short period that we have not been able to collect the redistributive effects with complete precision. In addition, this would also explain why we have not obtained any evidence of the redistributive effectiveness of Spanish personal income tax. We strongly believe that these three measures have fulfilled their redistributive role. Taking into account, there are more studies in the literature which demonstrate the effectiveness of these measures. Moreover, when looking at the evolution of the indicator we have used to measure the effectiveness of the tax in section 3, we have seen how it has become increasingly redistributive over time.

Finally, although it was not the main objective of the paper, we have found that there is a transitory and temporary effect on income inequality in Spain. However, taking into account the methodology we have followed to achieve stationarity, and the different persistence studies discussed in section 2, we consider there is indeed persistence in Spain.

Perhaps future research, either using more data or more complex methodologies, could analyse the persistence of inequality in Spain, better highlighting the effects of inflation and unemployment on inequality. Indeed, with a longer time range, they could even accurately capture the redistributive effects of investment in health and education, and of Spanish personal income tax. Our study could serve as an inspiration for more elaborate work.

8. References

Adiego, M., & Ayala, L. (2013). The structure of inequality in the long run. Journal of Applied Economics, 66, 5-35.

Anghel, B., Basso, H. S., Bover Hidiroglu, O., Casado García, J. M., Hospido Quintana, L., Kataryniuk, I., ... & Vozmediano Peraita, E. (2018). Income, consumption and wealth inequality in Spain. *Occasional Papers/Banco de España, 1806*.

Argimon, I., González-Páramo, J. M., & Salas, R. (1986). *Progresividad y capacidad redistributiva del impuesto sobre la renta en España 1979-1984* (No. 86-14). Complutense University of Madrid, Faculty of Economics and Business.

Ayala, L., Martínez, R. and Ruiz-Huerta, J. (1993): "La distribución de la renta en España en los años ochenta: una perspectiva comparada", in AA.VVV: I Simposio sobre Igualdad y Distribución de la Renta y la Riqueza, vol. II. Madrid: Fundación Argentaria.

Blinder, A. and Howard, E. (1978): "Macroeconomic activity and income distribution in the post-war United States". Review of Economics and Statistics, 60, pp 604-9.

Bourguignon, F. (2015). The globalization of inequality. In The Amazon book review. pp. 224

Buman, S., Lensink, R. (2015). Capital Account Liberalization and Income Inequality. Journal of International Money and Finance, forthcoming.

Calero Martínez, J., & Gil Izquierdo, M. (2013). The distributional incidence of public spending on education and health in Spain. *Welfare state: sustainability and reforms. XX Encuentro Economía Pública (2013), p 1-40*.

Card, D., & DiNardo, J. E. (2002). Skill-biased technological change and rising wage inequality: Some problems and puzzles. Journal of Labor Economics, 20(4), 733-783.

Christopoulos, D., & McAdam, P. (2017). On the persistence of cross-country inequality measures. Journal of Money, Credit, and Banking, 49(1), 255-266.

Daher, A. (2013). The real estate sector and economic crises. *Eure (Santiago)*, *39*(118), 47-76.

Dollar, D. (2005). Globalization, poverty, and inequality since 1980. The World Bank Research Observer, 20(2), 145-175.

Dollar, D., & Kraay, A. (2002). Growth is good for the poor. Journal of Economic Growth, 7(3), 195-225.

DONCEL PEDRERA, L. M., GÓMEZ GÓMEZ, N. E., GRAU CARLES, P., RODRÍGUEZ HERNÁNDEZ, J. G., & SANTERO SÁNCHEZ, R. INFLUENCIA DEL CICLO ECONÓMICO SOBRE LA DISTRIBUCIÓN DE LA RENTA EN ESPAÑA.

Fernández, J. O., & Sánchez, F. P. (2012). Progressivity, redistribution and welfare through Spanish personal income tax in the period 1999-2007. In *XIX Encuentro de Economía Pública: Políticas Públicas para la salida de la crisis (*p. 69).

Galbraith, J. K. (2011). Inequality and economic and political change: A comparative perspective. Cambridge Journal of Regions Economy and Society, 4(1), 13-27.

García-Altés, A., & Ortún, V. (2014). Functioning of the social lift in Spain and possible improvements. SESPAS Report 2014. *Gaceta Sanitaria*, *28*, 31-36.

Gil-Alana, L. A., Škare, M., & Pržiklas-Družeta, R. (2019). Measuring inequality persistence in OECD 1963-2008 using fractional integration and cointegration. *The Quarterly Review of Economics and Finance*, *72*, 65-72.

Griliches, Z. (1969). Capital-skill complementarity. The Review of Economics and Statistics, 51(4), 465. Available at:. http://www.jstor.org/stable/ 1926439?origin=crossref

Jäntti, M., & Jenkins, S. P. (2009). The impact of macroeconomic conditions on income inequality. Journal of Economic Inequality, 8(2), 221-240.

Milanovic, B. (2016). Global inequality. A new approach for the age of globalization. pp. 311. Harvard University Press.

Mocan, H. N. (1999). Structural unemployment, cyclical unemployment, and income inequality. The Review of Economics and Statistics, 81(1), 122-134.

Ochando Claramunt, C. (2010). La distribución de la renta en España en el periodo de crecimiento económico: 1998-2005/The Distribution of the Income in Spain in the Period of Economic Growth: 1998-2005. *Estudios de economía aplicada*, *28*, 691-22á.

Onrubia, J. and Picos, F. (2013), "Desigualdad de la renta y redistribución a través del IRPF, 1999-2007", Revista de Economía Aplicada, 63: 75-115.

Organisation for Economic Co-operation and Development (OECD). (2014). *All on board: Making inclusive growth happen*. Paris: OECD.

Pena Trapero, B.(director) (1996): "La distribución personal de la renta en España". Ediciones Pirámide S.A.

Pijoan-Mas, J., and Sánchez-Marcos, V. (2010): "Spain is Different: Falling Trends of Inequality", Review of Economic Dynamics, vol. 13, pp. 154-178.

Sarmiento, S. (2017). Evolution of income inequality in Ecuador, period 2007-2015, Revista Analítika, 13(1), 49-79.

Sturm, J. E., & De Haan, J. (2016). Finance and income inequality revisited.