A COMPARISON BETWEEN SPAIN AND ITALY BASED ON GRANULAR ECONOMY THEORY

Carlos Monlleó Durà

al226186@uji.es

Simone Alfarano

Bachelor's Degree in Economics

2016/2017

ABSTRACT

The main objective of this paper is to explore whether there exist statistic evidence to conclude that Spain and Italy are granular economies by comparing the results obtained between both countries. For In order to conduct the study, we have followed the Gabaix paper from 2011. For this purpose, we have used a database with 30.000 companies for each country.

One of the greatest indicators to know whether these countries present granular economies is the verification of an existing heterogeneity among the most powerful companies. In the present paper this is achieved in terms of sales and the rest of the companies. The second step has been to perform a simple regression analysis where we have the fictitious Gross Domestic Product (GDP) as the dependent variable and the sum of volume of sales of the 30.000 companies as the independent variable.

The results obtained may conclude that there are enough statistic evidences to confirm that Spain is a granular economy but not for Italy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This paper is based on Gabaix's article (2011) where it is proposed that the idiosyncratic crashes at the company level may explain a big proportion of the aggregate movements in the economy. For doing so, the researcher also suggests an analysis of the effects that individual shocks may have in the aggregate level with the main results supporting that the individual shocks of some big companies affect the economic cycle of the USA.

The researcher refers to each of these companies as “grains” and if the economy shows these features, it is said to be a granular economy. In the present study, we are going to do the same analysis by making a comparative between Spain and Italy. In the same vein, Gabaix claims that, as the companies are heterogeneous, their aggregate effect is heterogeneous too. These companies, which he considers a small sample, are the one hundred biggest companies in the USA. The analysis is conducted by performing a regression between the GDP and the granular companies. The results show how these one hundred companies may describe 30% of the GDP variation of the USA. This kind of study where the granularity of the economy is demonstrated has already been used in other papers but using other macroeconomic indicators such as the exports or the commercial balance, such as Di Giovanni & Levchenko (2012) and del Rosal (2013). Although the advanced economies with the same characteristics and relatively close such Spain and Italy should be granular economies, we will see throughout this project some disparities. For instance, in the case of Italy, the results are totally opposed to the expected outcomes due to the fact that the estimators in most cases fail to demonstrate the granularity of the Italian economy.

2. THEORETICAL FRAMEWORK

Most of the literature on economic cycles has been conducted from the aggregate perspective in which the fluctuations are popularly referred to and provoked by shocks. In other words, it is some type of variation due to any change in the macroeconomic framework what actually makes the economic system fluctuate, have imperfections and restructurings. With the analysis of these shocks, the causes of such fluctuations in economic cycles could be determined in order to predict the scope and consequences of economic fluctuations.

Regarding previous research, it is widely known that the pioneers in attempting to explain the regularities of economic cycles were Long and Poser (1983). The researchers created a useful and well-defined reference point for assessing the relative importance of economic changes or disturbances. In this way, they could demonstrate that the output growth rates in a multisectoral model may be utterly similar to the distribution obtained from a model in which the only system impulse is based on an aggregate shock.

In the same vein, Bak et al (1993) concluded that fluctuations in aggregate activity might be the result of many small and independent shocks of individual sectors. By examining these sectors, it was further addressed how the effects of these small independent shocks are not canceled in aggregate due to the presence of two non-standard assumptions: (a) the local introduction between productive units, and (b) convex technology. The association of both is crucial since none of the features alone would suffice.

There exist other studies as the one implemented by Di Giovanni & Levchenko (2012) by introducing a new mechanism in which the size of international trade affects the volatility of macroeconomics. In a study with many firms of different sizes and dimensions, when the size of multinationals follows a power law distribution with an exponent less than 1, there will be idiosyncratic shocks and the big companies will have a presumable influence on the aggregate volatility. In the case of small countries there are less numbers of companies so the volatility is also smaller.

Durlauf (1993) proposed an analysis of the evolution of an economy composed of a set of accounting firms, which have no connection among them and are only connected by technological complementarities. In this analysis it was observed that a long-term balance in the economic activity was produced when the connections among companies were stronger. Thus, this model seemed to explain that the growth of the leading sectors may cause a growth in the aggregate production of the whole economy.

Moving now to the study of the effect of links, Acemoglu et al. (2012) argue that the presence of intersectoral links between productive assets and idiosyncratic shocks might lead to aggregate fluctuations. The researchers claim that as the economy disintegrates, the rate at which the aggregate volatility falls is determined by the network structure that captures those links. In this vein, our main results provide a characterization of this relationship in terms of the importance of different sectors as suppliers of their immediate customers as well as their role as indirect suppliers. They also show that considerable aggregate volatility can be obtained from the idiosyncratic sectors.

The traditional argument claimed by Horvarth (1998) against the relevance of specific sectoral shocks to the aggregate phenomenon of economic cycles invokes the law of large numbers: positive shocks in some sectors are offset by negative shocks in other sectors. This article raises the hypothesis that the cancellation of specific sector shocks through the law of large numbers is affected by the interactions between the producing sectors. The analysis is performed within a multisectoral model framework similar to the one conducted by Long and Plosser (1983). It is shown that the rate at which the law of large numbers is applied, is further controlled by the rate of increase in the number of complete rows in the matrix of use of production goods and not by the rate of increases in the total number of sectors. As a matter of fact, the investigations on real arrays of use of goods of production of the American economy conclude that the number of complete rows increases much slower than the total number of rows when disaggregating and when these matrices of use of raw materials are used to parametrize the model. Consequently, the aggregate volatility of sector shocks decreases to less than half the rate implied by large numbers law. This fact leaves the conclusion open since it may be possible that a considerable part of aggregate volatility is caused by “smaller” shocks to individual sectors. Simple statistics performed on the model indicate that, approximately up to 80% of the

cases, the volatility in the USA gross domestic product growth rates could be the result of independent shocks for standard 2-digit industrial code sectors.

3. THE DATABASE

The database employed to obtained the total volume of sales in the Spanish case has been gathered from the “Sistema de Balances Ibéricos” (SABI) by Bureau Van Dijk. On the other hand, we also include the nominal GDP and the real GDP, which were obtained from the Statistics National Institute and from macro dates respectively. Finally, the variation of the IPC was also retrieved from the Statistics National Institute. The present project was implemented between 2002 and 2015, so in order to access the total volume of sales I downloaded the total amount of 30.000 companies with the higher volume of sales of the database. However, in order to perform the analysis and to prove the above mentioned hypothesis, I have tried to conduct the study with several sample amounts for 100, 200, 500, 1.000, 20.000 and 30.000.

To study the Italian environment, the database from where companies have been obtained was told to us by Marko Petrovic. The nominal GDP from Italy’s ISTAT; and the real GDP from macro data such as in the Spanish analysis. The IPC variation has been also retrieved from macro data. The proper study was conducted between 2001 and 2013. The corresponding data base given consisted of more than 40.000 samples, but as in the Spanish case/example we have just obtained 30.000 companies. We have also chosen for the Italian case the 30.000 companies with the highest volume of sales. Thus, we can equate the dates and do a more accurate and precise comparison. Finally, as I did in the Spanish example, I have performed the analysis several times for different sectors of samples: 100, 200, 500, 1.000, 20.000, and 30.000.

4. HYPOTHESIS WE WANT TO DEMONSTRATE

The main goal of this project is to perform a similar study for the Spanish and Italian contexts in order to make a further comparison between both countries. In this work, we will try to see and demonstrate the aspects that change in the highest companies in terms of sales, that is to say, if the companies with the highest volume of sales in the country can be correlated with the real GDP variation. Once these hypotheses are tested and the existent correlation between both variables is verified (real GDP and total volume of sales) for a very big sample of companies, that is to say, trying to demonstrate the granular hypothesis for both countries, we will go beyond and we will test if there is enough statistic evidence to affirm the hypothesis postulated by Gabaix in which he claimed that shocks of few companies (which present a higher volume of sales), may explain the greater proportion of the variability of economic cycle or in this case the real GDP.

Once we have the main results and conclusions, we will make a comparison between the Italian and Spanish situations. To be able to understand the explanation along with the comparison and to know how we manage the database, we have to identify the purpose, the variables and indicators that we will use. Even though Gabaix in his study regress the GDP per capita, in this work we will do it by regressing the real GDP of both countries, the variable in which we will try to see whether there is certain correlation in the volume of sales of all the companies chosen of every country by doing a summation with every amount to finally perform the regression analysis. This regression will be tested for small and large samples. In this way, we will be able to see how the results evolve according to the sample size and, the most important aspect, we will observe from which sample evidence is shown to test the hypothesis. For this, we are going to suppose a model like the one proposed by Gabaix, where the production is exogenous and there are no links among the companies, that is to say, no one is supplier of anyone and everything that they produce are the final products. We will thus refer to the summation of all the amounts of volume of sales that we will use to test the correlation with the real GDP as Fictitious GDP and there will be several quantities of this variable, that is to say, a fictitious GDP for each sample.

All in all, the approach would be the following: to perform this analysis we will make a simple regression so as to have the real GDP as the dependent variable and the summation of all the amounts of volume of sales of the companies that have this amount higher the independent variable. The regression would be like this:

$$Real\ GDP = \beta_0 + \beta_1 + PIB\ ficticious + \mu.$$

The fictitious GDP or summation of sales would be like this:

$$\gamma_t = \sum_{t=1}^{T} sales, i.t.$$

With this regression we will try to demonstrate what sort of changes in the sales of big companies may affect the variability of the real GDP and, in our specific case, to what extent they affect to real GDP and to observe whether these changes have a significant linear dependence to be able to explain the variation of real GDP departing from changes in the summation of the amount of sales volume.

5. HETEROGENEITY

One of the most important factors for granular economies to take place in the great existent heterogeneity among the companies of the same countries would be the great difference in the sales variables between big and small companies. However, there are a lot of existing differences among the biggest companies. For instance, we can imagine the difference that may exist in terms of sales among the highest companies and the middle and small companies. It would be reasonable to suppose that this condition is given in Spain and Italy, since they are two European developed countries with unquestionably conditions similar to the ones given in the USA, which was reflected on Gabaix's conclusions regarding the heterogeneity among the companies. In order to demonstrate it, we are going to perform a small study in which, on the one hand the average of every company we have in our database is going to be calculated, and on the other, we will repeat this process with only the highest 100 companies. In this way we will be able to see the existing differences among the companies with higher size and the rest of the companies. Also, we will see whether these great differences arise among the best companies in terms of sales, and to this goal we will make a comparison among the companies with higher volume of sales. Due to the fact that it was not possible to find a reliable source from where to obtain the real number of companies in the country, we assume that as the countries share a high number of similarities it would have more or less the same quantity of companies.

First of all, we are going to do it in the Spanish case. We can see that the average of the companies with greater volume of sales is of 3.673.253 million €, while the average of the companies of the sample is of 3.724.578 million €. Thus, it is evident that the difference is amazing because the highest 100 companies with major volume of sales is 71 times higher than the other 29.000 that we have in the sample. Specifically, only the 5 companies with major volume of sales already present more than the other 29.000 companies together. Mercadona already has almost half the amount of sales figures that the rest of the sample. We can also see how there are big differences between the first companies of the sample, for example, Mercadona doubles in sales to the second one included in the table. In order to have a clearer and wider view of the comparison, I have obtained from INE, the Spanish business demography evolution, in other words the total quantity of companies that there are in Spain. In the following figure it can be observe:

Fig. 1: Spanish Business Demography. *Source:* INE

If we perform the same process but in the Italian case we can see how the 100 companies with higher volume of sales present a sales average of 282.204.930 million €. On the other hand, the average of the 30.000 companies chosen is of 951.798.552 million€. It is possible to predict that for the Italian case there is no heterogeneity as in the Spanish case since the small companies sample represents 30% of the total amount of companies and it continues presenting great heterogeneity because only 100 companies represent 30% of the total. It is necessary X companies for them to add the same amount of the other 29.000. In the Italian case the same happens as in Spanish and the second largest company of the sample already represents twice as much as the 4.5 or 6, which have a similar amount.

With the analysis made, data reflect two economies where their companies present a clear heterogeneity, although in the Spanish case to a much greater extent than in the Italian. We will see as we analyse in our analysis if this may influence the final conclusions. The heterogeneity is a determining factor for our hypothesis and we can say that they are two granular economies, which is an indispensable requirement for the hypothesis of granularity.

6. EXPLANATION OF THE ANALYSIS

To begin with, we have obtained the following tables by analysing and explaining the experiment, and we will explain how we have calculated each column in order to perform the study. In the previous section, I introduced the aspects included in each column and now we will go a little bit further and explain the objectives or functions obtained and how we have used all these data in order to perform the regression analyses.

Table 1: Data used for the Analysis of the Independent Variable *Source:* Own elaboration from SABI Database

	NOMINAL GDP	VARIATION OF REAL GDP	VARIATION OF NOMINAL GDP
2005	930.566 €	0,90%	X
2006	1.007.974 €	2,00%	8,3%
2007	1.080.807 €	1,50%	7,2%
2008	1.116.207 €	-1,10%	3,3%
2009	1.079.034 €	-5,50%	-3,3%
2010	1.080.913 €	1,70%	0,2%
2011	1.070.413 €	0,60%	-1,0%
2012	1.039.758 €	-2,80%	-2,9%
2013	1.025.634 €	-1,70%	-1,4%
2014	1.034.951 €	0,10%	0,9%

6.1. EXPLANATION OF TABLE 1

At first sight, we observe that the nominal GDP is the one that takes into account the goods and the services at the price of the current year. As the nominal GDP is not accurate to make comparisons with previous years, for this study we downloaded the GDP simply because it is considered important when performing the analysis and to do check and compare ir with the real GDP. The real GDP is the important one for the present analysis that we are going to carry out and it is in fact the one that acts like the independent variable in the model. With the real GDP we can value goods and services at the price of a given year and, consequently, their variations or growth will allow us to thee if the economy as a whole has grown. This will also allow us to make annual comparisons since it does not consider inflation because it compares the amounts with a base year. On the other hand, we have the variations of the real GDP, which was obtained by performing the following operation:

$$\Delta Real\ GDP = \frac{Real\ GDP_{t-}\ _{Real\ GDP_{t-1}}}{Real\ GDP_{t-1}}\ .$$

As such, this variable is the one that we are going to use to introduce it into the dependent variable. The Table that we can see corresponds to the Italian example, however, for Spain it would have the same structure but with the corresponding macroeconomic indicators. (Quintanilla, 2014)

Table 2. Data used for the analysis of the Dependent Variable. *Source:* Own elaboration from SABI Database

Year	GDP/CPI	VARIA-TION	GDP (Summation 100 companies)	FICTITIOUS CPI	VARIATION OF CPI
2002	1736645,89	0,53%	185.487.674	106,808	4,00%
2003	1811437,373	4,31%	198.506.379	109,585008	2,60%
2004	1944438,83	7,34%	219.899.948	113,0917283	3,20%

2005	2202534,289	13,27%	258.304.680	117,2761222	3,70%
2006	2416615,267	9,72%	291.063.372	120,4425775	2,70%
2007	2296495,86	-4,97%	288.212.908	125,5011658	4,20%
2008	2360437,751	2,78%	300.385.017	127,2581821	1,40%
2009	2113719,72	-10,45%	271.140.034	128,2762475	0,80%
2010	2306878,764	9,14%	304.795.284	132,124535	3,00%
2011	2530050,524	9,67%	342.304.511	135,2955238	2,40%
2012	2602577,784	2,87%	362.328.521	139,219094	2,90%
2013	2730923,672	4,93%	381.337.310	139,6367513	0,30%
2014	2995787,114	9,70%	414.138.760	138,2403838	-1,00%
2015	3032557,428	1,23%	420.186.113	138,5583366	0,23%

6.2. EXPLANATION OF TABLE 2

Before starting to explain each of the columns and variables that we have in the tables, it should be mentioned that for each sample type there will be a different table, since the sum of the sales figure amount, will not be the same for a sample of 100 or 200 companies that for one of 30,000. The quantities and variations of the CPI will coincide in all cases, but the sum of sales and variation of these sales will not be the same since there will be different amounts for each sample. Here, we are going to explain it in a standard way although the table we present is a table for a sample of 100 companies and it refers to the Spanish case, which is the same but with different data for the Italian case.

We see that we have another table where we try to obtain the independent variable. In the first column it appears each of the years or periods in which we are going to make the study, and in the third column we have the sum of all the amount of sales of the n largest companies depending on their sales and the type of sample. On the other

hand, we have the fifth and sixth column whose main objective is to simulate a fictitious CPI in order to deduct the fictitious GDP that I have just calculated with the sum of all sales. In this way, we only have the second column left, which shows the fictitious GDP deducted from the CPI. It has to be mentioned that all these operations that we have had to do to deduct the IPC to the fictitious GDP have as main function to get a standardized GDP and, only if we can standardize it, we can buy it with real GDP since none of the two variables would take into account the variation of prices. This variation of the fictitious GDP or variation of the sum of the sales is the variable that we will use to introduce it in the regression. It is worth mentioning that I have calculated the CPI because it has been difficult for me to find a good indicator of the evolution of the GDP which was reliable.

Once all the information and necessary data are shown in order to obtain and carry out the regression proposed above, we perform a quick compilation to clarify concepts: we will use the variation of the real GDP as a dependent variable and the variation of the fictitious GDP as an independent variable. This process will be performed for each one of the selected n, that is to say, we will do it as many times as samples we have, and we will do everything for Italy and Spain. In total, there will be 9 different sample types being the smallest of 100 and the largest of 30.000. The main reason why I have performed a regression analysis for different sample types is to track the evolution, check for which type of sample the estimators give more favorable results depending on our hypothesis, and especially to see from which samples companies start explaining most of the variability of the real GDP, so that they can respond to the question of whether shocks in large companies may affect the economic cycle of a country and see if we have enough evidence to affirm that Spain and Italy are two granular economies.

7. COLUMNS

This section is similar to the previous one but with the difference that here we explain how we have made the calculations to know the way we obtained each one of the variables defined in the previous section.

In the first column that we see in Table 1, we observe the years of study and the period of time that we are going to analyse while in the second column we have the nominal GDP obtained from INE in Spain's case and from Macro Data referring to Italy, which shows the production of goods and services. In the third column we can see the variation of the real GDP obtained from Macro Data for Spain and for Italy.

In the third column we can see the real GDP variation obtained from INE as regards as Spain's case and macro data for the Italian one. To finish, in the fourth column nominal GDP variation appears, it has been obtained by ourselves; the formula used to obtain it is the one that can be observed below, mathematically expressed.

$$\Delta\, Nominal\ GDP = \frac{Nominal\ GDP_t - Nominal\ GDP_{t-1}}{Nominal\ GDP_{t-1}}$$

This first table is the one used to calculate the independent variable that will be used later in the real GDP variation to use it in the regression.

In this second table that we will use to calculate the dependent variable, we will use the variation of the sum of sales. As in the previous table, in the first column we have the years of study and the period of time that we are going to analyze. Now, the fourth column contains the sum of all the amounts of sales figures that we have obtained from SABI. For doing so, we have simply added up all the sales figure amounts on the n larger companies depending on the size of the sample, mathematically express would be as follows:

$$\gamma_t = \sum_{t=1}^{T} sales, i.t.$$

We would be thus talking about the dependent variable which we will use in the simple regression, in each one of the regressions that we do for each sample. As previously explained, we have used several samples of different size, therefore, we will repeat such formula for each of the samples. Regarding column 6, I have used it to obtain the variation of the CPI of such site for Spain and Such a site for Italy. Regarding the variation of the CPI, we have tried to simulate an IPC, operation that is reflected in the 5th column. To calculate it, we have selected the number 100 and from the results obtained, I made a geometric progression from all the variations of the CPI obtained in column number 6 so we can get a variable that simulates the CPI. These last data obtained in column number 5 have been used to reduce to the sum of sales of all the companies of the sample the IPC or inflation. In this way, we have a standardized GDP in which we do not take into account the evolution of the prices and they may be compared with other variables from different years, mathematically it would result like this:

$$GDP\ Ficticious\ standard = \frac{Ficticious\ GDP}{IPC}$$

It is necessary to perform this standardization operation to be able to compare the real GDP with the sum of all sales, because in this way none of the variables will include inflation. All this process will be explained in more detail later. Finally, in the second column, we have the variation of the standardized or non-inflationary GDP obtained:

$$\Delta\ Ficticious\ GDP = \frac{Ficticious\ GDP_t\ \_\ Ficticious\ GDP_{t-1}}{Ficticious\ GDP_{t-1}}$$

And this last column that I have described is the one to be used as a dependent variable in the previously proposed simple regression.

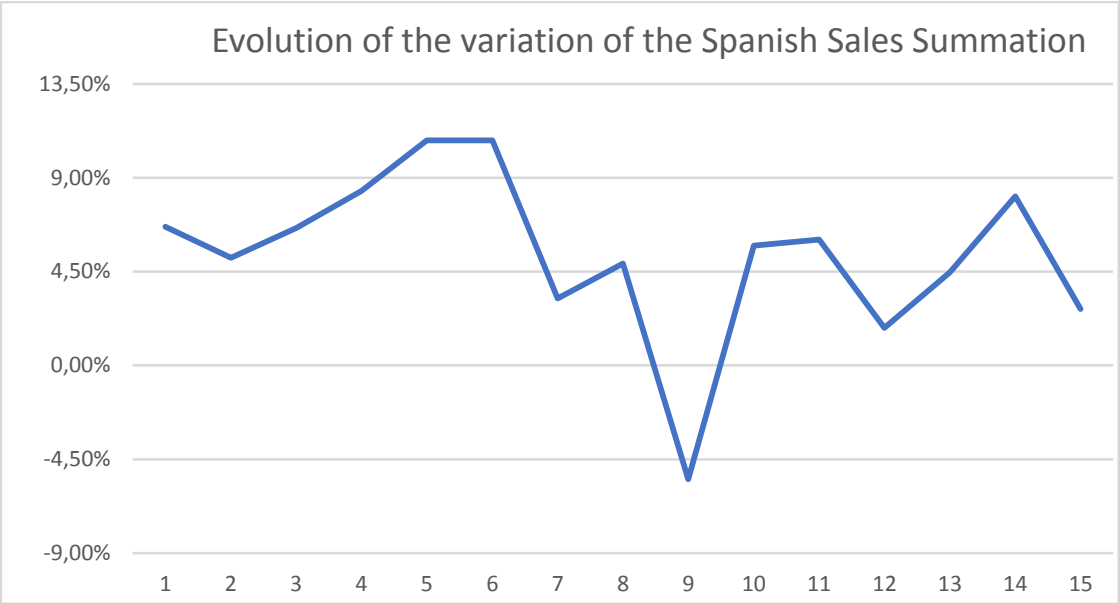

Fig 2. Evolution of the variation of the Spanish Sales Summation *Source:* Own elaboration from SABI Database

Even though it has previously been explained how to calculate the variation of the sum of sales, this table shows the evolution of these calculations for a sample of 100 companies. The decision to select a sample of 100 companies lies in the fact that the rest of samples would not serve, at first sight, to test the hypothesis. This table represents the case of Spain, and it can be observed that it follows a similar trend to the real GDP. Between 2002 and 2007, there exists a rise in the level of sales, from which we can deduce that it was caused because of the increase in consumption and investment since we were coming from a period of economic growth where the level of sales was actually chaining a process of prolonged and sustained growth. This is also reflected in the table since this strong growth of the level of sales is followed by a great drop in sales. The main reason for this is the fall in consumption and investment suffered a few years ago due to the financial crisis that occurred in 2007 with the fall of Lehman Brothers, which was quickly spread around the world. In the case of Spain, the country was experiencing a strong economic expansion where the proper increase in the level of the Spanish citizens' income due to supply and demand pushed very strongly the real estate business producing a real estate bubble. This triggered a crisis that provoked a drop in sales that,

as we can see even in 2015, has not been able to recover the levels preceding the year 2017 (Alonso & Furió, 2017).

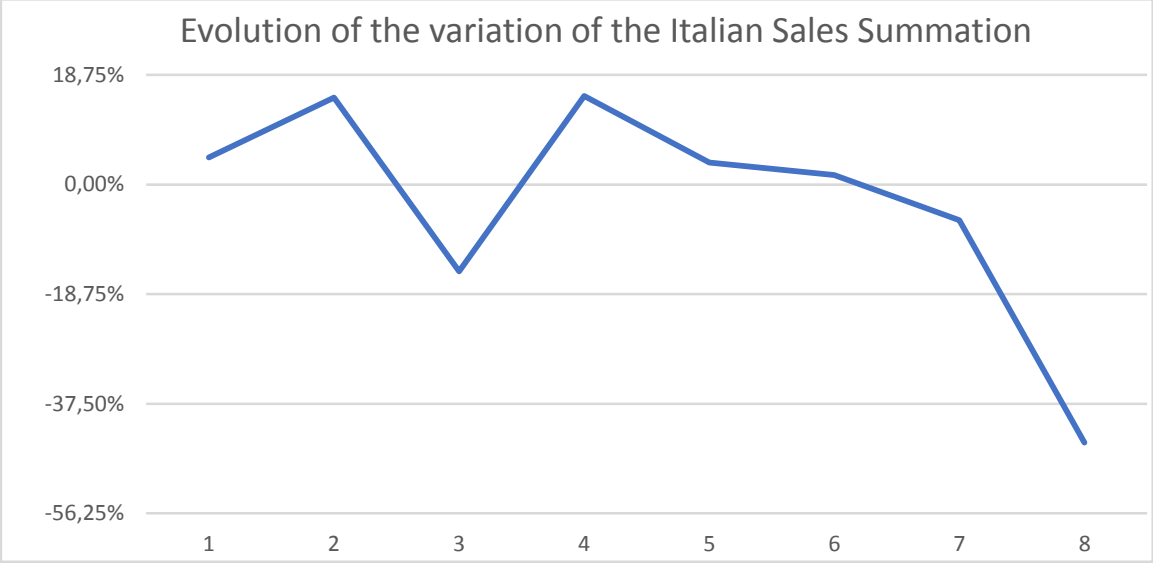

Fig 3. Evolution of the variation of the Italian Sales Summation *Source:* own elaboration from data provided by Marko Petrovic

On the other hand, this table introduces the case of Italy between 2001 and 2015. It has also been selected a sample of 100 companies for the same reasons than in the analysis of Spain. At first, all European countries, some to a greater extent than others, experienced a similar sales trend. This is indicated in the table since we can see that the sales in the first years had a rise basically due to the same factors that in Spain, which are the increased consumption and investment in a time of economic boom that increased sales. It is true that we can later appreciate a sharp drop in sales between 2002 and 2003 and then it returns to rise to the original levels of the 21st century. After this, we observe a fall due to the financial crisis caused by the collapse of Lehman Brothers, and which, as stated before, expanded worldwide causing havoc in both investment and consumption as a consequence of sales by the fall in per capita income and the uncertainty that triggered the financial crisis, which made us tend to save more than actually consuming. Even though it has to be stated that in the case of Italy we can observe certain points or nuances in the evolution of sales that do not correspond to the economic situation of the time, we will later on be able to explain and deepen these nuances to see if we can give

some explanation, such as the downturn in 2002 and the subsequent shred drop in sales volume in the year 2014.

8. INTERPRETATION OF RESULTS

Having analyzed the graphs, we can state that in the first box we have the samples used to perform the regression analyses and in the second column we have the slope or parameter that reflects the effect of the explanatory variable on the variable explained when the rest of factors remain fixed (in our case it would include the effect of the sum of the sales figure amount on the real GDP). The third column shows the constant parameter that contains the real GDP value when the sum of the sales is zero while in the fourth we find the error term, an estimator that indicates those factors other than the independent variable or the sum of all the sales that can affect the dependent variable or real GDP. Finally, the last column incorporates the coefficient of determination that determines the proportion of the variation explained in comparison to the total variation. The fraction of the sample variation is shown in the real GDP, which is explained by the sum of the amounts of sales figures of the selected companies. The regression that we have done to get all these estimators and get to test the initial hypothesis would thus be:

$$Real\ GDP = \beta_0 + \beta_1 + PIB\ ficticious + \mu.$$

However, as we have said, our hypothesis goes further and our main goal is not only to show that there is correlation between the variation of sales and the real GDP but to analyze whether shocks in very few companies explain most of the variation of the economic cycle or real GDP. In order to do so, we will analyze in detail all the estimators of any country and examine the evolution of the estimates obtained from the regression according to the different samples sizes that we have selected to make the proposed analysis.

Table 3. Estimators of the regression for Spanish Case *Source*: own elaboration from SABI Database

SPAIN			
N	**X**	**Y**	**Chi Square**
100	0,7289(0,66)	0,0338	0,0912
200	0,7457(0,57)	0,0366	0,1233
500	0,9627(0,51)	0,0374	0,2282
1000	0,9784(0,44)	0,0381	0,2858
3000	1,0332(0,36)	0,0387	0,403
5000	1,0406(0,33)	0,0393	0,4494
10000	0,8704(0,48)	0,0333	0,213
20000	0,9295(0,44)	0,0337	0,2658
30000	0,9814(0,42)	0,0337	0,3098
Between 2002 and 2015			

In the first place, we will analyze the estimators obtained for the case of Spain. There are different sample sizes, the greater of 30.000 and the smaller of 100, and initially it would be logical to think that if there really exists a correlation between both variables, real and summation of sales.

The larger the number of companies in the sample (in this case 30.000 would be the case in which the sample would have the largest number of companies) the more

precise and consistent are our estimates according to our interests although in our case it can be seen that the slope, the constant term, the error and the coefficient of determination show a model with more precise estimates when we have a sample that has around 3000 or 5000 companies. However, it is also true that our database has some limitations such as the lack of data in some small companies, which may hinder the analysis. Having said all this, with the results obtained we can affirm that the larger the number of companies considered in the sample, the better the estimators obtained up to a certain sample size, specifically from 3.000 to 5.000, because once it reaches this point estimators worsen a bit and they still remain interesting. If we try to follow the line marked to demonstrate the initial hypotheses, what I just explained is not the most relevant aspect, because what we want to show is that the samples that are really important are the small ones, and with the data we have, we see that for very few companies with the case of 100 or 200 already obtain some really good estimators that would come closer to what we are looking for and in this way we would be able to affirm our initial hypothesis which claims that very few companies can explain most of the variation of the economic cycle.

We see how for 100 companies we have a slope of 0.8285, close to one, which means that increasing the percentage of the sales figure by a percentage of the real GDP would increase by 0.82%. Although it is not a datum, which indicates a perfect correlation as for other cases, it is true that we have enough evidence to be able to say that for 100 companies the two variables are quite correlated. Since we do not look for a perfect correlation, with few companies explaining most of the variation of the economic cycle would be sufficient and it is precisely what we find in the estimations obtained from the regression: the 100 companies explain most of the variation of the real GDP.

Turning now to the constant term, in case the sum of sales were zero, we see that for all samples is around 0.03, which means that it would be a good result for a small sample of 100. This would mean that if the sum of the sales were 0 or the real GDP would be 0.03, this is a good data that gives us enough statistical evidence to affirm the hypothesis raised due to the fact that it denotes the great importance that the amount of sales figure has on the real GDP. If this variable did not exist, the real GDP would be zero. On the other hand, it can also be observed how the error term for all cases is kept around 0.05%, which represents another good datum that indicates that Spain is a granular economic, because there are few variables other than the sum of sales that can affect the

real GDP. Finally, in order to finish with the Spanish analysis, the coefficient of determination that appears in the table as we see, although the rest of the estimators are favorable to affirm that the initial hypothesis is fulfilled and Spain is a granular economic, we have a very low coefficient of determination since the highest case is for a sample of 30.000 companies that has a R^2 of 0.30. These data means that 30% of the variation of the sum of the amount of sales figure depends on the actual GDP. It is a low value, which does not coincide with the other estimators, but as we have explained there are many limitations and points to improve in the database, which will be explained in more detail late. However, a coefficient of low determination does not mean that the model is bad or inaccurate.

Table 4. Estimators of the regression for Italy's case. *Source:* Own elaboration from data provided by Marko Petrovic

ITALY			
N	**X**	**Y**	**Chi Square**
100	2,57(0,20)	0,0213	0,1024
200	1,73(0.19)	0,0017	0,0441
500	1,42(0,20)	-0,0037	0,0288
1000	1,397(0,20)	-0,0054	0,0282
3000	1,417(0,20)	-0,0068	0,0292
5000	1,433(0,20)	-0,0071	0,0299
10000	1,478(0.20)	-0,0067	0,0318
20000	1,489(0,20)	-0,0068	0,0327
30000	1,479(0,19)	-0,0073	0,0326
Between 2006 and 2013			

Moving now to the case study of Italy, we are going to conduct a deep and broad analysis as we have done in the previous case of Spain. Without specifying or drawing conclusions, we see that the slope, which reflects the effect of the sum of sales on the real GDP, in all samples it is positive, which implies that there is a positive correlation between the two mentioned variables, that is to say, that an increase in the sum of the sales is always paired with an increase in the real GDP. The two variables go in the same direction: if one increases the other does too, if the other variable falls, the other does too. This is a good indicator for our hypothesis to be fulfilled because we have that the slope has a positive connotation, but also has its negative part. I say this because it has a value really over 1 for all sample sizes. This reveal a sign of a poor correlation between the variables studied, because the increase in sales per unit percentage, the real GDP would be greatly affected. In some cases, we see how it increases even more than double, meaning that by increasing sales by one percentage point, the real GDP increases by almost 2.5 percentage points, which is practically unsustainable and denotes a null correlation between the dependent and the independent variable. Summarizing, we see on the one hand that both variables have a positive correlation, but this correlation is unsustainable because when it increases in one of the variables it also carry very large deviations from the real GDP. Focusing now on the constant parameter is obvious that in some cases the parameter is negative, which theoretically indicates that if the sum of the sales was null, the real GDP would be negative, this is totally without economic sense, since we cannot give a negative real GDP, therefore, it is a clear indicator that for the Italian case we start to see estimators that are not in the ideal way for our hypothesis to be fulfilled. From what we can see from the error term, we see that it is quite low, there would be little influence of other variables other than the sum of the sales. In order to conclude with the estimators, the coefficient of determination can be observed as it is quite low for most of the samples, which allows to glimpse the little explanatory capacity that the model has for the Italian case, there is little proportion of the real GDP that is explained by the sum of the sales.

Here we can see the table where it can be observed all the estimators obtained in the case of Italy for the regression:

$$Real\ GDP = \beta_0 + \beta_1 + PIB\ ficticious + \mu.$$

We see as for all sizes shows from the smallest to the largest, the pending parameter is above 1, this estimator reflects the effect of the sales of the companies on the real GDP when the other factors remain fixed. To demonstrate our hypothesis, we would be interested in that it was close enough to 1, but in the table we can see as in all the cases are well over one. If it were close to one, it would mean that there is a correlation between both variables, because when increasing sales by a percentage unit, it would increase the real GDP also by the same amount. In this case, for Italy, as we see is not so, in fact, is above enough one, especially for when the sample is small, which is 2.57, practically unviable because as there were variations, deviations would occur of more than double between a real GDP and sales of these companies. Therefore, far from demonstrating the hypothesis initially raised. We see that when the smallest is the most distant sample of one and therefore as the sample grows, it approaches one, which indicates that as the sample increases there is a greater correlation between the dependent and independent variables and when they are small the sum of the amount of sales figures explains practically nothing the variation of the real GDP. For example, for 30000 companies we have the constant parameter of 1.4 which would mean that for each percentage point that increases the sum of sales the real GDP grows by 1.4 percent, an estimator rather better than for 100 companies that we remember was 2.57, but even so, we are far from being able to affirm that there is a correlation between the sales of the country and the variation of the real GDP and much further from our initial hypothesis, since 2.57 is very high to say that few companies explain the variations of the real GDP, that is why the data obtained for the case of Italy there is very little correlation between the dependent variable and independent. By focusing now the constant parameter we can see as in small samples of 100 and 200 is very low and positive and from sample 1000 starts to be negative, and as, we have said before, this indicates a negative correlation between the dependent and independent variables. A negative constant parameter means that in case the sum of the sales were zero, the real GDP would be negative, which is absolutely lack of economic sense. Therefore, this estimator also shows nothing positive to demonstrate our initial hypothesis. Following the same negative trend to prove that shocks of few large companies affect the economic cycle the coefficient of determination is very low for all samples, both large and small is close to zero, therefore the fraction of the variation of the sample in the real GDP that is explained by the sum of sales is very low, but as in the case of the slope it we can see that the bigger the sample the sales explain the greater variation, specifically for 30.000 companies, to approximately 60% of the

variation in the real GDP is determined by the sales variable, a variable that, within the poor results of the estimators, is quite good, but remember that the initial hypothesis wanted to show that few companies already explained the greater part of the variation of the real GDP and in this case the estimators begin to be good when the sample is very large, for few companies as 100 or 200 only 10 and 4% respectively the variation of the real GDP is determined by sales. A low coefficient of determination does not have to mean that the model is bad, but seeing the results of these estimators, we cannot affirm in any case that Italy is a granular economy. In conclusion, we will talk about the error, it is low for all samples, both large and small, it remains constant around 0.20, but they are still very poor estimators to have enough evidence.

Having conducted the analysis of both countries, we are going to compare them together, for the case of Spain, there is enough empirical evidence to state that few companies have the variations of the real GDP, if the hypothesis is fulfilled because only 100 companies have a 75 percent correlation and for the 500 largest companies there is a 97 correlation, for the case of Italy is totally different we have seen as none of the estimators gave positive results to affirm that there is evidence and can confirm our hypothesis. From the results obtained, we see that as we add companies to the sample there is more correlation between both variables, the real GDP and sales of companies with higher sales amount, that is why with the data of the estimators we could say that in the case of Italy the larger companies do not have enough strength in terms of sales to describe the variability of the real GDP, therefore we reject the hypothesis, shocks of a small portion of the most powerful firms do not affect the economic cycle, therefore it is not a granular economy. If this is not the case, we can say that it is an economy where small firms have more importance in the variation of the real GDP, because in the meanwhile there are more companies there is a better correlation, therefore, it is an economy where many more companies are needed to describe the variability of the business cycle. The larger the sample, the better all the estimators analyzed in the previous section.

9. CONCLUSION

In the case of Spain we have enough empirical evidence to claim that it is a granular economy because we see clearly how few companies, in particular 100, are already beginning to describe most of the economic cycle in Spain. Despite the fact that in this work we have done the analysis where the economic cycle indicator is the real GDP, it can also be analyzed and performed with other indicators such as imports.

In the case of Italy we do not have enough empirical evidence to claim that it is a granular economy. In all cases we have seen how few companies were not strong enough to describe the variations of the real GDP of Italy. It was also observed that when adding more companies to the sample, the estimators had worse interpretations and therefore less evidence to affirm that Italy is a granular economy.

Although there are many factors that we have not taken into account, which would be interesting to consider for future research, such as the period of time we are analyzing because it is very brief, and that many times we do not find data in the database as companies with less important relative to the rest, other types of analyses such as taking into account different indicators to the real GDP can also help us to go deeper into the analysis and have more evidence to affirm that shocks in the most representative companies of the country can affect the economic cycle. If we can show that with more than one indicator, we can test the hypothesis claimed by Gabaix.

By performing this type of work, we can acknowledge the great importance that this small percentage of companies has on the economy of a country. They could include very fragile economies since it would be interesting to consider to what extent they can influence the management teams or the interests of this type of companies on the society or even on government decisions. Thus, it is of great interest to raise future works to analyse or make a civic reflection on the influence that this type of signatures may have on us, either directly or indirectly through government decisions.

10. BIBLIOGRAPHY

Acemoglu, D., Carvalho, V. M., Ozdaglar, A. & Tahbaz-Salehi, A. (2012), `The network origins of aggregate fluctuations', Econometrica **80**(5), 1977–2016.

Bak, P., Chen, K., Scheinkman, J. & Woodford, M. (1993), `Aggregate fluctuations from independent sectoral shocks: self-organized criticality in a model of production and inventory dynamics', Ricerche Economiche **47**(1), 3–30.

Del Rosal, I. (2013), `The granular hypothesis in eu country exports', Economics Letters 120(3), 433–436.

Di Giovanni, J. & Levchenko, A. A. (2012), `Country size, international trade, and aggregate fluctuations in granular economies', Journal of Political Economy **120**(6), 1083–1132.

Durlauf, S. N. (1993), `Nonergodic economic growth', The Review of Economic Studies **60**(2), 349–366.

Gabaix, X. (2009), The granular origins of aggregate fluctuations, Technical report, National Bureau of Economic Research.

Gabaix, X. (2011), `The granular origins of aggregate fluctuations', Econometrica **79**, 733–772.

Horvath, M. (1998), `Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks', Review of Economic Dynamics **1**(4), 781–808.

Long Jr, J. B. & Plosser, C. I. (1983), `Real business cycles', The Journal of Political Economy pp. 39–69.

Matilde Alonso Pérez et Elíes Furió Blasco, « La economía española », *Cahiers de civilisation espagnole contemporaine* [En linea], 6 | 2010, publicado en linea el 22 junio 2011, consultado el 3 febrero 2017. URL: http://ccec.revues.org/3212 ; DOI : 10.4000/ccec.3212

Quintanilla, W. (2014), Diferencia entre el pib nominal y PIB real concepto de inflación y sus clases concepto de I.P.C y P.I.B y sus diferencias, Unidade Tecnologias de Santander.

Wooldridge, J. (2015), Introductory econometrics: A modern approach, Nelson Education.