Ca' Foscari University of Venice

Single Cycle Degree programme

in Economics and Finance
"EM50"

Final Thesis

INCOME INEQUALITY AND CREDIT

"Investigation of the causal relationship between
the Level of Credit to the households and Income Inequality"

Supervisor
Ch. Prof. Monica Billio

Assistant supervisor
Ch. Prof. Danilo Cavapozzi

Graduand
Esiona Mezuraj
Matriculation Number 827099

Academic Year
2016 / 2017

ABSTRACT

This study explores the income inequality – level of household indebtedness nexus for the case of the United States during the period 1980 – 2012. Co-integration techniques are used both in case of single equation and multivariate equation, by the means of Engle Granger Two Step Procedure and Vector Error Correction model. Empirical results from the estimation of the VECM and ECM indicate evidence of a unidirectional causality in the short – run flowing from Income Concentration to the Level of Domestic Credit for the Private Sector, weather in the long – run no evidence of causality was found for neither direction.

Key words: Income inequality, Credit Growth, causality and co-integration

Table of Contents

1. INTRODUCTION

In the three decades leading up to the financial crisis of 2007/2008, income inequality rose across most of the developed world (OECD, 2015). In the late 1990s, the level of income inequality in the United States rose to levels not seen since the Great Depression of 1929. As a result, the causes for the path of income distribution came into the attention of many popular books (i.e. Galbraith, 2012; Rajan, 2010), policy oriented papers (OECD, UNDP) and opinion editorials (i.e. Milanovic).

Only in the recent years there is an increasing attention on the topic from the academic literature, which mostly analyses the link between income inequality, credit boom (or household debt) and the probability of crises. The variables influencing the level of indebtedness and hence the probability of crises are complex and reflect both economic and social changes, but debates rose as to whether widening inequality was also to blame for the financial crises by driving private sector credit boom. Analyzing whether the concentration of incomes contributes to the excessive accumulation of debt, which in turn is recognized as being one of financial instability's drivers (Kumhof and Rancière, 2011) is a not a recent concern. This notion can be traced back to Fisher (1933) who argued that 'all great booms and depressions' are caused by two dominant factors, 'over indebtedness and deflation following soon after'. As Galbraith (2012, pg.3) suggests, between the rise of income inequality and financial crises, there might be a link, and this link is debt. Rajan (Fault Lines, 2010) argued in his book that the rising inequality in the last three decades caused a political pressure for income redistribution, but since the redistribution of income via taxes and social spending was not preferred, government chose instead to apply policies that would expand and ease the access to credit, such as deregulation of credit markets and encouraging lending to low-income households. This policy allowed the low-income household to have access to credit, especially to mortgage finance, which created a boom in house-pricing and later led to the banking crises of 2008. The tide of populism in the last two years, reflect the dissatisfaction of the society regarding globalization, offshore job, deregulation of the markets and the polarization of incomes.

Hence, based on the recent political and economic turbulence, the purpose of this paper is that of providing an empirical investigation of whether there exists a causality relationship between income concentration and private sector indebtedness, and possibly its direction.

Key words: Income inequality, Credit Growth, causality and co-integration

2. LITERATURE REVIEW

The global financial crises of 2008 generated a growing attention towards its causes, and in the recent years more attention is directed also to whether the distribution of incomes can be added to the list.
Branko Milanovic (2009) argues that apart all the other factors suggested from the literature[1], income inequality was the real cause of the crises. Growing incomes among the top households created a surplus of resources which were looking for investment opportunities. From the other hand, another group of households, middle and low incomes, in absence of real increase of incomes, where looking for more resources to keep up their consumptions demand. The problem was that middle and low-income households didn't keep up with what the growing economy was capable of producing, as larger share of incomes went to the people at the top (Reich, 2010). The Government meanwhile approached the problem of wealth concentration not through redistribution, but rather the solution came in the form of "easy credit". Because the redistribution of income via taxes and social spending was politically not preferred, government chose to follow policies that would expand the access to credit, such as deregulation of credit markets and encouraging lending to low-income households (Rajan, 2010). Hence the middle and lower class could consume and spend like the wealthiest, even though they didn't have the real resources, making these investments risker, which means higher rate of returns both for the intermediary and the investors. As the exchange of founds between the rich and the poor increase, so did the size of the financial sector as measured by total assets (or liabilities) to GDP. But Kumhof and Rancière (2010) showed that without the income's recovery of the middle and lower class over a reasonable period of time, the results would have been a major economic crisis. The higher the level of debts to income ratio for middle and low households, the higher the probability of default in case of unpredictable events such as the loss of the job or illness. As a result, a credit-fueled system was created, and once the middle class began defaulting, the system collapsed. Thus, based on this explanation, Milanovic argues that banks, hedge funds and other financial intermediaries are not the real cause of the financial crises, rather income inequality is.
Kumhof and Rancière (2010) explore the link between increase of income and wealth of the high-income households and contemporary a similar increase of debt-to-income ratio among poor and middle-income household. The key mechanism of their model, is that the additional income gained by the high-income household (the investors) is landed to the rest of the population (the workers), which as a result create an increase need for financial service and intermediation which are owned and controlled by 'the investors'. Kumhof and Rancière apply a Dynamic Stochastic General

[1] There are of course other factors for the financial crises of 2007/2008, like assets price bubbles and financial deregulation.

Equilibrium (DSGE) to a closed economy setting with only two groups of households: the investors (which own top 5% of incomes) and the workers (the other 95%) and then analysis the impact of an idiosyncratic shock in the income (like job lost or illness) of the workers through three scenarios: baseline, uncertainty and high leverage. By applying the model to the US data, both before and after the crises of Great Depression, the simulation shows that an increase in income inequality can lead to credit growth, higher leverage and increase the probability of crises. If the poor and middle class who borrow the money in order to sustain their consumption, will not recovery their level of income over a reasonable period of time, the loans will keep growing, therefore the leverage and the probability of a financial crises. When the debt-to-income ratios started to be perceived as unsustainable, it became a trigger for the crises.

However, their results (KR) are called into question by Bordo and Meissner (2012), who conclude that while financial crises are typically preceded by credit boom, inequality only occasionally increased during periods of credit expansions. Bordo and Meissner investigate whether there exists a relationship between income inequality, credit booms and financial crises. They found little evidence that a rise in top income shares leads to credit booms, while credit booms increase the probability of a banking crisis. They argue that whether changes in inequality generates credit growth is a matter of data. For instance, in Japan credit growth rises before the rise of the share of top incomes. In Australia, also credit growth was unrelated to the income concentration among the top 1%. Top incomes followed, rather than proceed the credit expansion in Australia. Instead they found that low interest rates and economic expansions are robust determinants of credit boom. The paper uses Bank Loans to the Price Index as a proxy for credit growth, while Real GDP, Index of Investment to the price level, M2 and Nominal Interest Rate are used as credit growth determinants, for a time period from 1920 up to 2008. Their proxy for income concentration and income inequality is the share of total income earned by the top 1% of individuals or households or tax units.

Similarly, Atkinson and Morelli (2011) conclude from their cross-country empirical research that outside the US, the history of systematic banking crises in different countries does not suggest that income inequality is a significant casual factor. In their paper, the authors address not only the impact of inequality in economic crises but they also reverse the question and analyze the impact of economic crises on the inequality of resources. They test whether the relationship between income inequality and greater risk of crisis is casual or co-incident, the latter referring to the possibility that both crises and the rise in inequality may have another third common cause like Credit Market Deregulation. Rather than level of indebtedness they focus in Systematic Banking Crises and Consumption and GDP collapses, for a period of 100 year over 25 countries, but because of the problem of missing data they divide the analysis in different sub-periods. Particularly the authors focus in the financial crises of the Nordic countries (Norway, Sweden and Finland) in the 1990s (to find out that these three countries differed in terms of their prior distributional experience) and that of Asian financial crises of 1997. The banking crisis in Sweden, followed a period of rising inequality; those in Norway and Finland were preceded by periods of relative stability in the distribution, hence in case of Nordic countries there is no general pattern. Out of 22 banking crises for which they had evidence of inequality, only 6 cases were clear evidence of inequality. They found few financial crises which were proceed by rising inequalities, but the predominant pattern isn't this. The classis pattern they found was that crises are not proceed by income inequality. The paper found that there is more evidence that financial crises are followed by rising inequality, but no causality was found between precede inequality and financial crises.

Acemoglu (2011) suggests the alternative hypothesis of Rajan; it was politics that drove both inequality and financial crises, hence between the recent paths of incomes distribution and the crises of 2007 there is a relationship of concomitance, not causation. For instance, a third variable might have served as link between this two, like the deregulation of the markets. Politicians implemented financial deregulation policies favoring high income households.

The key difference between Acemoglu's hypothesis and Rajan's hypothesis is the tail of the distribution from which the deregulation pressure comes (the middle low income in Rajan's hypothesis and top incomes as Acemoglu's hypothesis).

In his book "Fault Lines" (2010), Rajan argues that many low and middle-income households have reduced their saving and increased debt since income inequality in the US started to expand in the late 1970s. This helped in short-term to keep private consumption and employment high, but it also contributed to the creation of a credit bubble (Tree, 2014). With the downturn in the housing market and the sub-prime mortgage crises, the over indebtedness of the US householders came into light, and the economy experienced a crisis not seen since the Great Depression of the 1929. According to Rajan, the roots of the financial crisis lie in several fractures of the economy that existed before the crises itself, but those causes were ignored (intentionally or unintentionally) from the system. Among the unbalanced growth in the global economy, the wages' gap between under qualified and qualified employee, the reckless credit growth and the risk taker behavior of the financial system, Rajan put the rising income inequality in the US in the center of his analysis. Rajan argues that over time the gap between the earnings of educated and under educated individuals was rising, which pressured governments to enact policies aimed at improving the situation of the voters left behind. Because the redistribution of income via taxes and social spending was politically not preferred, government chose to follow policies that would expand the access to credit, such as deregulation of credit markets and encouraging lending to the so called "no income, no job, no assets" individuals. From the other side,

the monetary policy to keep interest rates low, added the incentives provided by the government for supporting low income mortgages which stimulated an extra ordinary credit boom. Such policies were an easy and fast fix of the problem; however, the real costs were experienced only in the future. To summarize, based on Rajan' hypothesis, the increase of income inequality has a casual effect on the level of domestic credit to the households, hence of the level of credit in general and the probability of crises.
Perugini, Holscher and Collie (2015), based in a panel of 18 OECD countries for the years 1970 – 2007 provide evidence of a positive relationship between income concentration and private sector indebtedness, one other traditional drivers are accounted for: the deregulation of financial system (Demirguc-Kunt and Detragiache, 1998, Ranciere 2006), accommodative monetary policy (Borio and White, 2003), rapid economic growth (Mendoza and Terrones, 2008), inflow of foreign capital (Elekdag and Wu, 2011). The authors use Domestic Credit to the Private Sector (in levels, as a percentage of GDP) as their dependent variable, while as a proxy of income concentration the share of total income going to the top 1% of earners. They include also the level of credit market (de) regulation as a key variable of their analysis since this variable is found to have a positive, statistically significant effect on private credit. Also, they find a positive and significant coefficient of the inequality variable (top 1%) suggesting that higher inequality directly drives credit. Therefore, they found clear cross -country evidence that inequality can directly impact on credit expansion. To check for the robustness of their analysis, Perugini, Holscher and Collie (thereafter PHC) carried additional estimation using also the share of income hold by top 5% and 10%, and confirm that higher inequality triggers higher level of indebtedness.

Empirical research of the last decade, have confirmed that episodes of financial instability are indeed precede by excessive levels of debt in some form or another. While there is a lot of attention in the link between economic growth and inequality of different forms, income inequality rarely plays a significant role in the large literature of financial instability and credit boom. Overall, as summarized in the previous pages the economic literature is inconclusive on the link between credit expansion and income distribution. The purpose of this paper is that of analyzing whether there exists a link between this two variables, once the other factors of credit growth are accounted for, and also test for the direction of the causality.

3. METHODOLOGY AND DATA

3.1. DATA DESCRIPTION

There are few empirical and econometric issues that need to be address before moving into the model. Data collection is a challenge, especially when we try to measure variables like inequality. In 2009, Milanovic wrote that 'to go to the origin of the crises, we need to go to rising income inequality within all countries in the world, but especially at the United States over the last thirty years. Hence, the US is the main country of analysis in the paper. It is known that unit root test and cointegration tests require a long-time span of data rather than merely a large number of observations (Luinetl and Khan, 1999). Hakkio and Rush (1991, p.573) point out that there is no universal answer to the question: how long is the long run? However, the length of the long run may vary between problems, that is, our data of 33 years would have been long enough to capture long-run relationship between private sector indebtedness and income concentration among the top 1% in case of panel data, but since I am working with time series, 33 observations is a limited sample size, which can come with some asymptotical limitations as we will see latter.

The Level of Domestic credit to the private sector (%GDP) is the dependent variable. This variable has some limitations which I have taken into account further in the paper. Explicative variables include income concentration top 1% as a proxy for income inequality (my main variable of interest), and other credit drivers in order to avoid endogeneity issues due to omitted variables bias: credit markets deregulation (Ranciere, 2006), Real Interest Rate and Board Money Supply to GDP ratio (M2) as proxies for the monetary policy (White 2003, Meissner 2012), and GDP per Capita (Bordo and Meissner, 2012) in order to count for level of credit growth due to the economy expansion.
Gross Fixed Capital Formation (%of GDP) and Portfolio Investment (%GDP) are also added in the regression in order to take into consideration the part of domestic credit which is not demanded by the households, but rather from the business.

$$Credit = f(top1, dereg, capform, ptfinv, M2, RIR, LGDP) \qquad \textbf{(1)}$$

DOMESTIC CREDIT TO THE PRIVATE SECTOR (%GDP)

A second major point is the choice of my dependent variable: it would have been preferable to use Household Debt to GDP or Income Ratio, but the data are not available for the time coverage of the study (OECD data starting frim 2005). Other proxies like credit card, consumer debt, mortgage loans etc., could have been used, but in order to have a higher comparability/homogeneity of the data I am, following the PHC study and use as dependent variable ***"The level of domestic credit to the private sector, % of GDP"*** from the World Development Indicators Database (World Bank, 2017)[2].

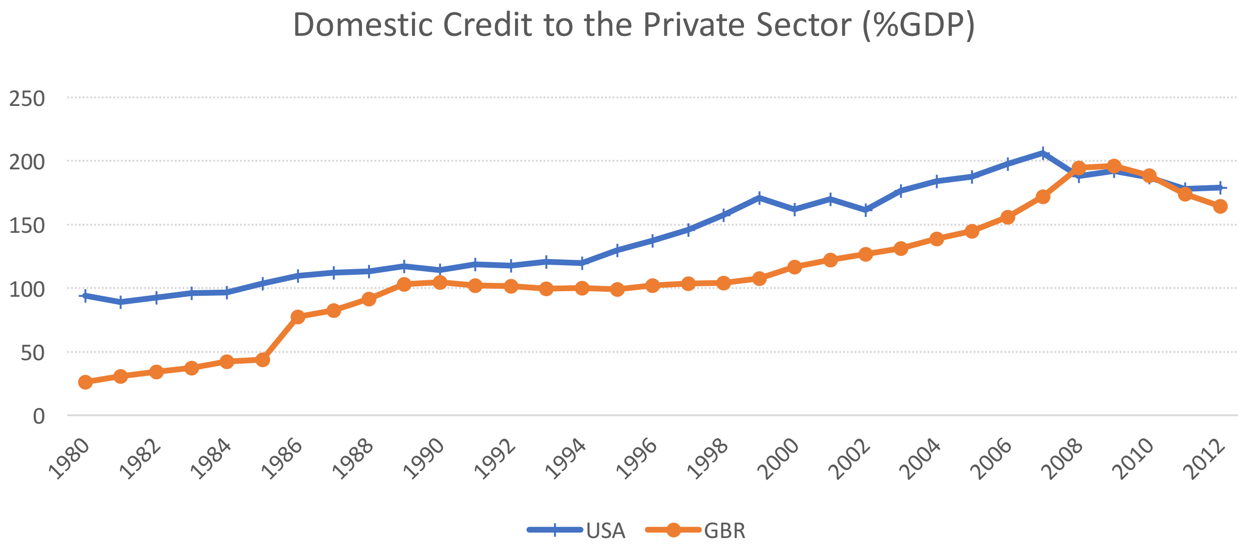


Fig. 1. *Evolution of the levels of Domestic Credit to the private sector in USA and UK, from 1980 to 2012. Both series show continuous increasing patterns until the 2007/2008 when the series start falling, which is a response to the financial crises.* *Sources: OECD data*

One important limitation of the dependent variable, is that it includes both household debt (which we are interested in this paper) but also debt to business and other private organizations.

To overcome this limitation, by following PHC paper, I will add the necessary proxies for the part of credit demanded by non-household private sector:

a) ***Gross Fixed Capital Formation (% of GDP)*** *as a proxy of credit demanded by firms for investment purposes*, and
b) ***Portfolio Investment (% of GDP***) *as a proxy of firms' credit demand driven by transactions in equity and debt securities.*

[2] http://data.worldbank.org/indicator/FS.AST.PRVT.GD.ZS
Domestic credit to the private sector refers to financial resources provided to the private sector by financial corporations such as through loans, non-equity securities, credits and other accounts that established a claim for repayment. The data are taken from the financial surveys of the IMF.

INCOME INEQAULITY PROXY: TOP 1% INCOME CONCENTRATION

My other variable of interest is the inequality of income, and as a proxy for it, the share of total income going to the top 1% of earners is used.

This is a ratio measurement (p99p100), meaning that it compares how much people at one level of income distribution (in our case people at the top 1% of the income) have compared to people at another.

Gini is another measure of income inequality but it will not be adopted in this paper not only for data availability reasons, but also because Gini cannot tell us where in the distribution the income is rising or falling (Atkinson, 2011). And we want to be able to distinguish between the top percentiles, 1% and the rest of the households. Furthermore, the share of top 1% of earners provides an excellent proxy since it follows the idea that the income concentrated among the richest one, is a fuel of credit expansion (Kumhof and Ranciere, 2010).

Data are taken from the World Top Incomes Database (http://wid.world)[3] which have been obtained from historical income tax records. Observations unit indicators also is market income (pre-tax and transfer), whereas it would be preferable to use disposable income[4], which bears more significantly on household consumption, investment and borrowing decisions.

The figure below (Fig.2), displays the evolution of Top 1% income concentration in the US, series of other countries are added for comparison reasons, during the period 1970 to 2012. As we can see from the figure, the income concentration in the US is the highest one compared to the other countries, the UK comes immediately after it, and is followed by China. Sweden, as it is expected, has the lowest Income Inequality 1% ratio during the period in consideration.

[3] Other database for inequality measures are OECD, Luxemburg Income Studies (LIS), World Institute for Development Economics (WIDER) and University of Texas Inequality Project (UTIP).

[4] Income remaining after deduction of taxes and social security changes, available to be spent or saved (Wikipedia)

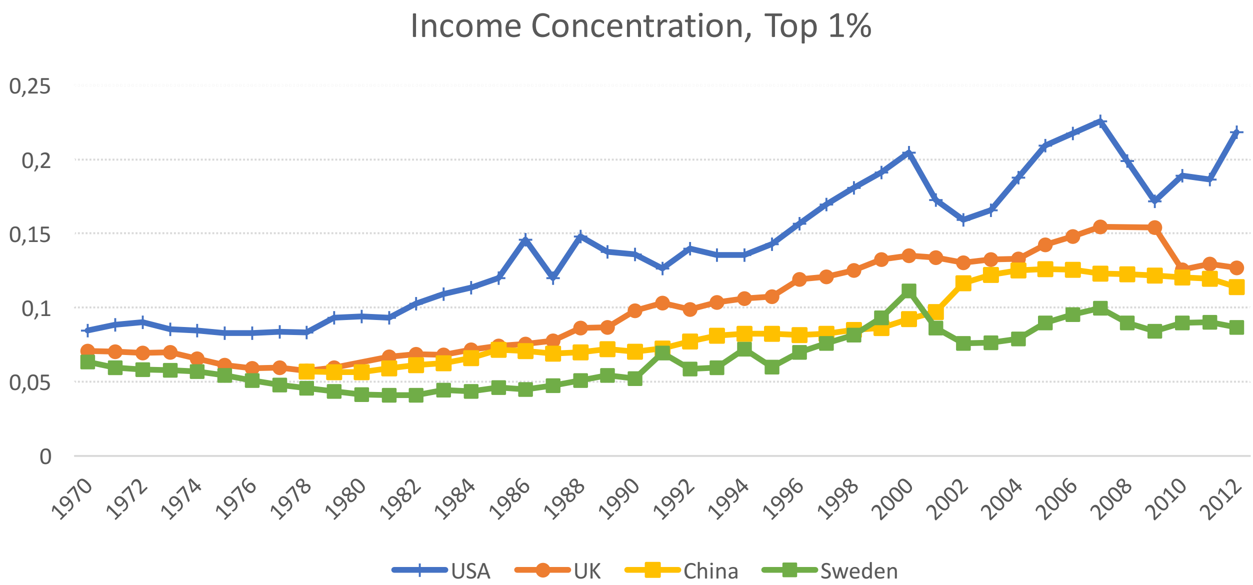


Fig. 2. *Evolution of the levels of Top 1 % Income Concentration in USA, UK, China and Sweden, from 1980 to 2012. All series overall have increasing patterns, they quite start form same levels of inequality in the 1970s but immediately after the Income Concentration in America increases much more than that on the other countries.* *Sources: WID.WORLD data*

In the figure below (fig.3) we can see with more precision the evolution of income distribution for the case of the UK and the US from the period which goes before the WWI until the recent times. Income and wealth inequality was very high a century ago, particularly in Europe, but dropped dramatically in the first half of the 20th century. For much of the 20th century, the gap in incomes between the well-off and less well-off is generally thought to have narrowed in much of the world.

In effect, the rich didn’t get much richer while the poor caught up a bit. According to research based on The World Top Incomes Database, this decline in inequality began in North America and much of Europe in around the 1920 and 1930s and a little later, perhaps the 1950s, in some developing countries. But then, in the 1970s and 1980s, the pattern began to reverse, and inequality began to rise again.

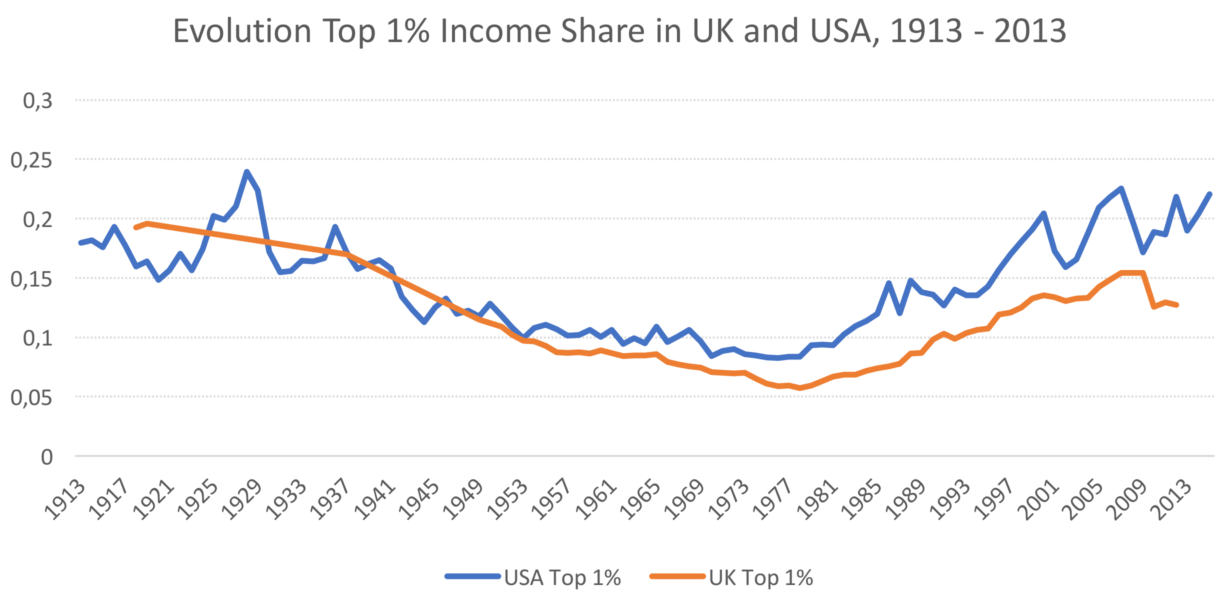


Fig.3. *Evolution of the levels of Top 1 % Income Concentration in the US and the UK before WWI. Income concentration follows the shape of a U, differently from Kuznets' theory, it is high at the beginning but it falls immediately after the two WWs and start increasing for both countries in the beginning of 1970s which coincides with the deregulation of the credit market.* *Sources: WID.WORLD database*

Regarding the recent political and economic upheavals, like Brexit vote and the rise of populism, Alesina and Perotti (1996) showed that income inequality, creates uncertainty in the political-economic environment by fueling social discontent. For example, perception growing gap between rich and poor was a big factor in Brexit vote, yet in 2015 Gini coefficient is no higher than it was in the late 1970s.

Analysis expectations: Based on Rajan's hypothesis, we expect a positive impact of the income inequality in the level of the domestic credit to the private sector.

KUZNETS CURVE THEORY

In the economics science, inequality is defined as the difference found in various measures of economic well-being among individuals in a group, among groups in a population, or among countries.[5] There are different forms of inequality: inequality of incomes, wealth, consumption and recently there are talks also about inequality in opportunities. In this paper are focus goes to income inequality, although they are often viewed as comparable indicators of the economic status of a family. Income inequality refers to the unequal distribution of household or individual income across the various participants in an economy.

In 1955, Simon Kuznets offered for the first time an argument relating inequality with the process of industrialization, hence a process of economic growth. According to Kuznets, inequality is low in pre-industrial societies, where most people live at subsistence levels. As industrialization begins, however, gaps start to widen thanks to the rising earnings of factory workers compared to those of farmers, and they continue to grow with the emergence of increasing specialization among industrial workers. But then, argued Kuznets, gaps start to narrow as the state begins collecting more taxes and distributing them as benefits.

Hence, based on Kuznets theory, the fundamental drivers behind inequality were the *changing structure* of the economy during a phase of development and changes in the payment that follows in the major sectors. The result is an inverted U-shape curve between inequality and per capita income. The main point in Kuznets' theory wasn't the discovery of some pattern between income and inequality, rather was the principle that change in inequality are driven by transitions in economic activity, and such transitions are a normal evolution of economic development, (Galbraith, 2012, pg. 48). *Hence, based on Kuznets theory, inequality level is a structural outcome of any economic life*.

Kuznets' hypothesis, as it became known, was influential in the 20th century, and the shape of inequality that it traced – an inverted-U – seemed to match the facts reasonably well. However, in recent years – rather than rising and then falling, the trajectory of inequality now appears to be more U-shaped: It was high at the start of the 20th century, fell in the middle of the century, but has been rising since the 1970s (Keeley, 2015).

In line with Kuznets theory, Thomas Pikkety in his book "Capital in the Twenty-First Century" argues that high levels of inequality are natural state of modern economies. Only unusual events, like the two World Wars and the Great Depression of the 1930s will disrupt that normal equilibrium.

A more recent theory is that of Branko Milanovic an economist at the Luxembourg Income Study Centre and the City University of New York. In his book (Global Inequality, 2016) Milanovic propose a theory of "Kuznets Waves" instead of Kuznets U-inverted curve. Across history, he argues, inequality has tended to move in cycles. In pre-industrial period, inequality would rise as countries enjoyed a good fortune and high income and latter it falls as war or famine dropped incomes back to their average. Then same waves are present during the industrialization also, but the forces are different: technology, openness and policy will affect income distribution. As workers moved from farms to factories, average income and inequality levels increased sharply, this time speeded up by

the global interconnection. Then a combination of factors (war, political upheaval and education) pulled inequality to the lows of 1970s. Starting from then, a new era of rising inequality has started. Milanovic explains how technological progress and trade combined, press workers, and make it easier for firms to substitute people with machines. Workers' decreasing economic power is then followed by a lost political power, since the very rich use their money to influence candidates and elections.
As technology is shaping the global market, so it has done with the distribution of incomes. Today there is a race between the speed of technological changes and education, and when technological advance leaves behind educational changes, level of inequalities will rise. As a result, people with lower level of education see their jobs taken away from technology, while people with high-level skills are well positioned in jobs mostly highly paid, by widening even more the differences in their income.

There are different ways to measure inequality, but it has always been a challenge, since for most of the countries measuring economics inequality never was part of official statistics routine, the results are spare and often inconsistent. In general, OECD countries prefer income survey, while expenditure survey are more often used in Asia (Galbraith, 2012, pg. 24). For instance, the inequality measurement can be expenditure based or income based, per capita or households, some are gross and some are after taxation. Also, measuring inequality within countries many time is different from inequality between countries.

Contributing factors to the income inequality: many factors explain the rise of income inequality. Some are economic, such as the role of technology and the globalization, others are social, such as shifts in who people marry, and some relates to the rising incomes of top earners. Education is known to affect equality. Competition for talent also creates a salary divide. In the recent years, because of the market competition an increase of salaries for people in executive roles has driven the concentration of incomes. Stagnant wages also play a big role.

CREDIT MARKET (DE)REGULATION

Based on Rajan's hypothesis (2010), Credit Market Deregulation since 1970s is one of the factors which shaped the distribution of incomes.

After the Great Depression that took place in 1929, the US government in order to protect the country from similar economical disasters, created a strict financial regulation that worked until the 1960s. The Glass-Seagall Act in 1933 separated commercial banks from investment one, and gave to Federal Reserve more centralized power. A period of stability followed, but it was criticized as it made American banks less innovative and competitive in the global market. As a result, three decades later, in 1980, the Congress passed the Depository Institution Deregulation Act which liberalized the financial sector, both within and across nations (Borio, White, 2004). The new area of market deregulation followed by rapid financial innovation is often referred as the New Financial Architecture. Galbraith (2012) in his book Inequality and Instability (2012) suggests that macro variables like New Financial Architecture including also globalization of finance, drive income inequality more than micro - country specific factors. Crotty, J. (2009) argue that the subprime mortgage crises in the US was generated from NFA because of its light institutions and financial practices. In 2010, Obama Administration passed the Dodd-Frank Wall Street Reform and Consumer Act protection as a response to the 2007 Financial Crisis.

Data for Credit Market Regulation[6] are taken from *Economic Freedom of the World, 2016 Annual Report*. This component reflects conditions in the domestic market, the extent to which (in a scale from 0 to 10, where 0 is a total regulated market like the case of China before the 1990s and 10 stands for total financial liberalization) the banking industry is privately owned, credit is supplied to the private sector and whether controls on interest rates interfere with the market for credit. Countries that use a private banking system to allocate credit to private parties and refrain controlling interest rates receive higher ratings.

Expectations: we expect a positive link between the level of deregulation (from 0 to 10) and the level of domestic credit to the private sector.

[6] https://www.fraserinstitute.org/resource-file?nid=10159&fid=4820

MONETORY POLICY VARIABLES

Real Interest Rate (lending interest rate adjusted for inflation) and Board Money Supply %GDP (M2)[7] are also added in the regression as variables of control for the monetary policies.

Fed and Bank of England conducts Monetary Policy by changing their official interest rate - known as Bank rate, as an attempt to influence the overall level of economy's activity. A reduction in interest rates make saving less attractive and borrowing more attractive, which stimulates spending. Lower interest rates can also affect consumers' and firms' cash-flow, a decrease of real interest rates reduces the income from saving and the interest payments due to loans. Lower interest rates make today's value of future profits higher, by giving corporations more incentive to invest. For instance, lower interest rates can boost the prices of assets such as shares and houses. [8]

Interest Rates in the US were constant for the late 1970s to the beginning of 2000. After NASDAQ Index crash in 2000-2001[9], FED tried to offset the decrease in investment by cutting short-term interest rates. From a level of 6.8 in 2000 interest rates were close to 1.5 percent in 2003. As a result, more people demand house's mortgages, as with lower interests rate they could better afford mortgage repayment. Increasing housing demand encouraged more home constructions, (Rajan, 2010, p. 106). Loose monetary policies accommodate the asset bubble, in particular the housing sector (Bernanke, 2005; D'Apice and Ferri, 2010). A low interest rate stimulated more and more household to buy homes, and the housing sector enjoyed high prices and high profits. Through securitization (the process of spreading the individual risk of subprime mortgages in many tranches) and innovative financial products, financial markets were avoiding risk (Tridicio, 2012).

Expectations: negative link with interest rates, since as the central banks decrease the interest rates, the higher will be the borrowing incentives for the households. Regarding the Money Supply M2, according to standard macroeconomics theory, an increase in the supply of money should lower the interest rates in the economy, which leads to more consumption and lending/borrowing, hence we expect a positive relationship between M2 and Level of domestic credit to the private sector.

Last exogenous variable is Log GDP per Capita. It is included as a variable of control for the pro-cyclicality of credit. Research studies find that the overall level of economic development, is the strongest predictor of financial progress and credit availability (Collins 2016, Adarov and Tchaidze

[7] Broad Money, M2, is a measure of the money supply that other physical money it also includes demand deposits at commercial banks and my money held in liquid accounts.

[8] http://www.bankofengland.co.uk/monetarypolicy/Pages/how.aspx

[9] Known as the Dot-Com Bust…

2011). Based on the economic theory, a positive effect is expected between GDP per capita and the level of credit to the private sector.

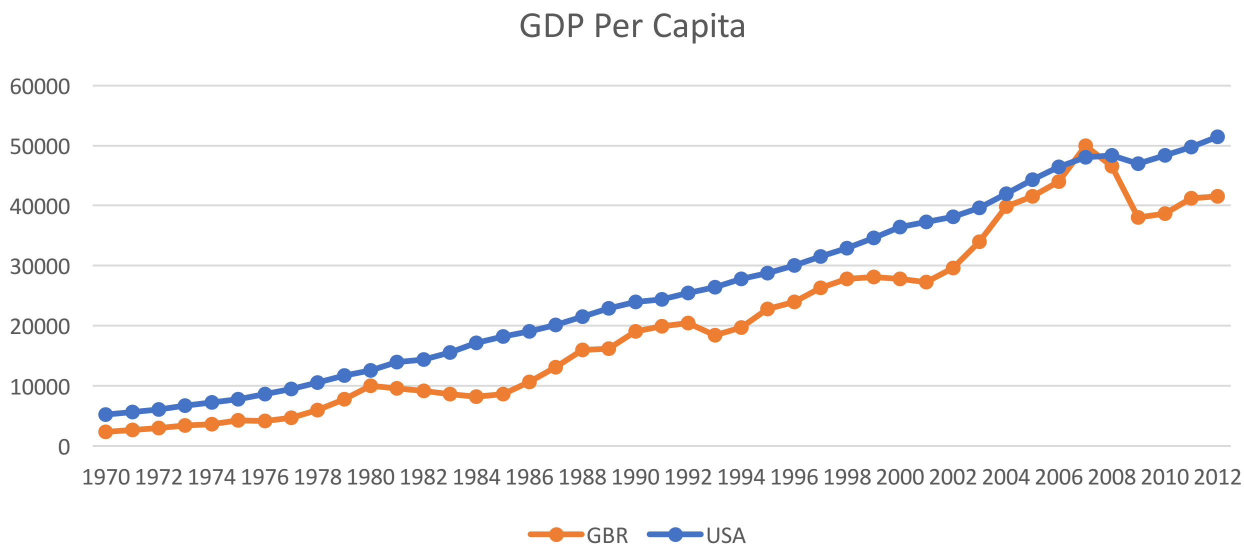


Fig.4. Data in US $
Source: World Development Indicators, World Bank

4. EMPIRICAL METHODOLOGY

4.1. Unit Root Test

In regression analysis involving time series data a critical assumption is that the time series under consideration is stationary. Working with non-stationary series, the regression analysis may lead to spurious or nonsense regression (i.e. obtaining a high R^2 or statistically significant coefficients). As a result, the first step before estimating the regression model is that of testing for stationary. In the literature, there are three ways to test for stationary: graphical analysis, correlogram and unit root analysis (Gujarati, 2011 pg. 208).

Consider the following AR (1) process:

$$y_t = \rho y_{t-1} + X_t\varphi + \varepsilon_t$$

for $t \in (1, \ldots, T)$. The X_t represent the exogenous variables, including a constant or a constant and a trend, ρ are the autoregressive coefficients, and the errors ε_t are assumed to be independent and identically distributed. Now, if $\rho < 1$ the series is said to be weakly stationary, while if $\rho = 1$ then the process y_t contains a unit root. However, as was shown in the seminal paper of DF (1979), under the null that $\rho = 1$ the standard t-ratio does not have a t-distribution, not even asymptotically. The reason for this is the nonstationary of the process invalidating standard result on the distribution of the OLS estimators. Hence, critical values have to be taken from the appropriate distribution, which under the null hypothesis of nonstationary is nonstandard. The distribution is skewed to the left, with a long left-hand tail so that critical values are smaller than those for the normal approximation of the t distribution. As a result, if we use the standard t tables we may reject the unit root too often. The table below present 1% and 5% critical values for Dickey-Fuller test:

Table 1. *ADF C.V for Unit Root Tests*

	Without trend		With trend	
Sample size	1 %	5 %	1 %	5 %
$T = 25$	−3.75	−3.00	−4.38	−3.60
$T = 50$	−3.58	−2.93	−4.15	−3.50
$T = 100$	−3.51	−2.89	−4.04	−3.45
$T = 250$	−3.46	−2.88	−3.99	−3.43
$T = 500$	−3.44	−2.87	−3.98	−3.42
$T = \infty$	−3.43	−2.86	−3.96	−3.41

Source: Fuller, W. A., (1976), Introduction to Statistical Time-series, p.373, Verbeek,2012, p.352

4.2. Co-integration Test Analysis

Once the variables of the time-series are characterized by the presence of unit root in the levels, the next step consist in testing for existence of long-run relationship. In general, a linear combination of two or more time-series will be nonstationary if one or more of them is nonstationary, and the degree of integration of the combination will be equal to that of the most highly integrated individual series. For example, a linear combination of an I (1) series and an I (0) series will be I (1), that of two I (1) series will be also I (1), and that of an I (1) series and an I (2) series will be I (2). However, if there is a long-run relationship between the series, the outcome may be different, and that is the case when the series are co-integrated.

Co-integration was first introduced by Granger in 1981 and then extended by Granger and Engel in 1987. Consider two variables Y_t and X_t that are I (1). Then Y_t and X_t are said to be co-integrated if there exist a β such that:

$$Y_t - \beta X_t \sim I(0)$$

What the previous concept means is that the regression in case of bivariate equation:

$$Y_t = \beta X_t + \mu_t$$

makes sense since Y_t and X_t do not drift too far apart over time. Thus, we say that there is a long run equilibrium relationship between them. If Y_t and X_t are not cointegrated, that is:

$$Y_t - \beta X_t \sim I(1)$$

then Y_t and X_t would drift apart from each other over time, and the relationship between them would be spurious.

Although there are several ways to test for co-integration, in this paper the Engle and Granger and Johansen Test are been used to investigate for co-integration among the I(1) variables. In case of Engle and Granger test, the DF and ADF unit root tests is performed in the residuals of the static equation, but by using adjusted critical values (table below). Tests for unit roots are performed on single time series, whereas co-integration deals with the relationship among a group of variables, each having a unit root.

Critical Values for the Dickey-Fuller test for no co-integration are given by:

Table 2. *C.V of ADF for test of co-integration*

No. of Variables N + 1	Sample Size	Critical Values 10%	5%	1%
2	50	3.28	3.67	4.32
	100	3.03	3.37	4.07
	200	3.02	3.37	4.00
3	50	3.73	4.11	4.84
	100	3.59	3.93	4.45
	200	3.47	3.78	4.35
4	50	4.02	4.35	4.94
	100	3.89	4.22	4.75
	200	3.89	4.18	4.70
5	50	4.42	4.76	5.41
	100	4.26	4.58	5.18
	200	4.18	4.48	5.02
6	500	4.43	4.71	5.28

Source: "Forecasting and testing in co-integrated systems", Journal of Econometrics, Vol. 35, 1987, pg 157.

4.2.1 Engel Granger Two –Step Method of Co-integration

If a trended variable is regressed in another trending variable, we often find significant t and F statistics and a high R^2 even though there might not be a true relationship between the two. In this case our regression is spurious, and very often it is characterized by a low Durbin – Watson d statistics. For example, the figure 5 below shows that both level of Credit and Income Top 1% in the US have a similar path, both grow on time.

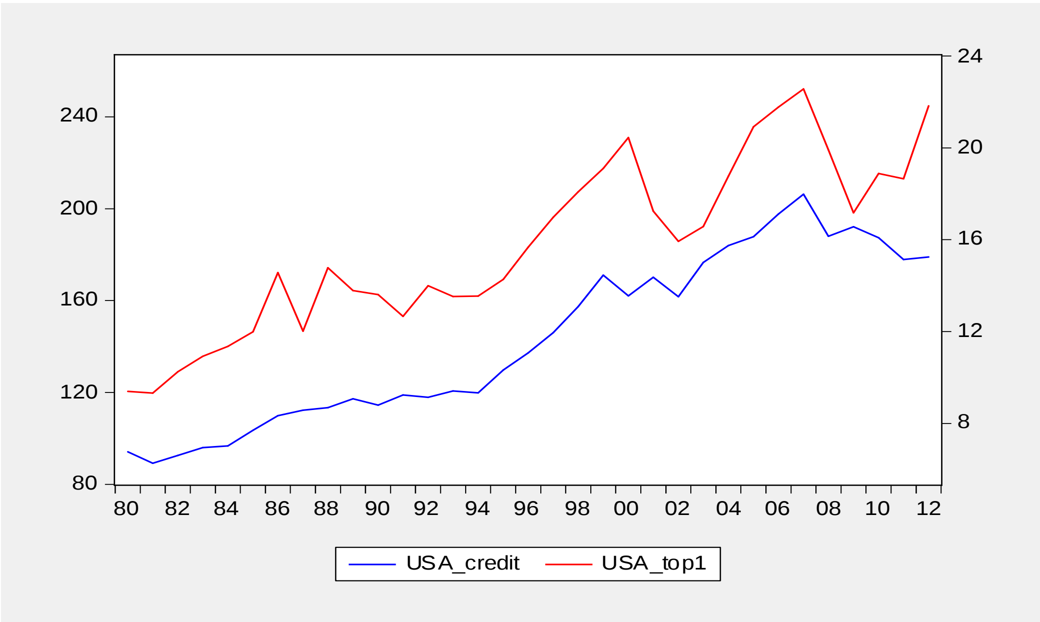


Fig.5. *Level of Credit and Income Top 1% series, US 1980 -2012. We can see from the graph that the series follow a common trend, which is an indication for possible co-integration.*

The series seem to have a trend in common, hence we deduct that they are integrated of order one, which is further confirmed by the ADF tests of unit root for each variable (Appendix). One way of resolving the problem of spurious regression is to transform each series into stationary and then use the stationary series for the regression analysis, but it is not ideal because we lose information about long-run equilibrium. Instead if two or more variables are nonstationary, but the error term of their combined regression is stationary, then we say that the variables are co-integrated. The concept of co-integration was introduced by Granger in 1981, which also developed a test for it that we can be applied in case of single equation.

The Engle-Granger two step Method is run as followed:

First step, we test the variables for their order of integration, applying the DF and ADF test, but by paying attention to use not the normal critical values, rather the C.V for Co-integration test (table 2). There are three cases:

a) all variables are stationary, it is not necessary to proceed,
b) the variables are integrated of different orders, we cannot apply EG Two Step procedure, rather an ARDL model is used,
c) all variables are integrated of same order, and then we proceed with the second step

Second step, the long run, static equation is estimated as specified by the equations:

$$Credit_t = \propto + \beta_1 Top1_t + \beta_2 CapForm_t + \beta_3 PtfInv_t + \varepsilon_t \quad \textbf{(2)}$$

$$Income\ Top1_t = \propto + \beta_4 Credit_t + \beta_5 CapForm_t + \beta_6 PtfInv_t + u_t \quad \textbf{(3)}$$

and save the residuals for each of them. Then, the ADF test with modified CV, is performed in the residuals of the static equation, ε_t and u_t. If the residuals are stationary, I (0), then the variables in the equation are co-integrated and there is no risk of obtaining a spurious regression.

After confirming that the residuals of the static equation are stationary, the residuals are used to analyze the long run and short run effects and to compute the adjustment coefficient, which is the coefficient of the lagged error correction term. First lag of residuals is used as the error correction terms, $\mathrm{ECT}_{\mathrm{i},t-1}$, for the dynamic error correction model (ECM).

The Error Correction Term, ECT is obtained as:

$$ECT_{t-1} = (Credit_{\ t-1} - \hat{\beta}_1 Top1_{t-1} - \hat{\beta}_2\ CapForm_{\ t-1} - \hat{\beta}_3\ PtfInv_{t-1}) \quad \textbf{(4)}$$

and,

$$ECT_{t-1} = (Top1_{\ t-1} - \hat{\beta}_4 Credit_{t-1} - \hat{\beta}_5\ CapForm_{\ t-1} - \hat{\beta}_6\ PtfInv_{t-1}) \quad \textbf{(5)}$$

where, $\hat{\beta}_i$ for $i = 1\ \dots 6$, are the estimated coefficients of long-run relationship.

For the long-run equilibrium to be accepted, the ECT needs to satisfy the STABILITY CONDITION, the coefficient has to be a number in the interval $[-2, 0\,]$ and statistically significant.

Yet, there are some drawbacks of the EG approach. If we have more than two variables (as in this paper), there might be more than one co-integration relationship (the rule is that the number of co-integration relationship is at most (n-1), where n is the number of variables we are testing). The EG two step procedure does not allow for estimation of more than one co-integration regression. As a result, the Johansen test will be used in order to assess the number of co-integration relationship and test for causality as a robust check of the Engle Granger two step procedure. Since in Johansen test it is necessary to specify the number of lags among the variable, a VAR model will be first estimated in order to select the proper number of lags. Another problem with the EG methodology in dealing with multiple time series is that we not only have to consider finding more than one co-integration relationship, but that we will also have to deal with the Error Correction Term (ECM) for each co-integrating relationship. As a result, the Vector Error Correction Model need to be considered (VECM). A last drawback of the two step EG test for co-integration, is that since the model is a two-step procedure, any error generated in the first step is transmitted in the second step, by risking in having non-consistent estimation.

4.2.2 Johansen Test of Co-integration and VECM

In the economic science, it is quite common that some variables are not only explanatory for the dependent variable, but that they also might be explained by the variable they are used to explain. For example, in our case, it might be that Income Inequality can play a role in explaining the Level of Credit to the private sector, but there is also the possibility that it itself can be explained by the level of credit. In such cases, we are dealing with simultaneous equations, all variables that enter into the equation can be treated as endogenous, and this is done by the Vector Autoregressive Model (an autoregressive model extended to more than one dependent variable). Suppose we have a bivariate VAR, $X_t$and Y_t. In VAR model, X_t is affected not only by its past (lagged value of X), but also by current and lagged value of $Y_{t,}$ and vice-versa. Hence, one important factor of the VAR model is the estimation of the lagged values, and for the number of lag equal of p, we write VAR(p). Of course, this bivariate VAR can be extended to more variables.

When all the variables used in the VAR are non-stationary but co-integrated of the same order than the Error Correction Term can be included in the system, and the model is called: Vector Error

Correction Model, VECM, which is a form of restricted VAR. As in case of single equation, VECM is preferred over VAR in case the variables are non-stationary at level but co-integrated because by including the error correction term we have information about the long-run relationship and the adjustment terms.

Our VECM model can be expressed as follow:

$$\Delta Credit_t = \sum_{k=1}^{q} \phi_k \, \Delta Credit_{\,t-k} + \sum_{k=0}^{q} \phi_{1k} \, \Delta Top1_{\,t-k} + \sum_{k=0}^{q} \phi_{2k} \, \Delta CapForm_{\,t-k} + \sum_{k=0}^{q} \phi_{3k} \, \Delta PtfInv_{\,t-k} + \alpha_1 ECT_{t-1} + \nu_t \tag{6a}$$

$$\Delta Top1_t = \sum_{k=1}^{q} \phi_{5k} \, \Delta Top1_{\,t-k} + \sum_{k=0}^{q} \phi_{6k} \, \Delta Credit_{\,t-k} + \sum_{k=0}^{q} \phi_{7k} \, \Delta CapForm_{\,t-k} + \sum_{k=0}^{q} \phi_{8k} \, \Delta PtfInv_{\,t-k} + \alpha_2 ECT_{t-1} + \vartheta_t \tag{6b}$$

$$\Delta CapForm_t = \sum_{k=1}^{q} \phi_{9k} \, \Delta CapForm_{\,t-k} + \sum_{k=0}^{q} \phi_{10k} \, \Delta Credit_{\,t-k} + \sum_{k=0}^{q} \phi_{11k} \, \Delta Top1_{\,t-k} + \sum_{k=0}^{q} \phi_{12k} \, \Delta PtfInv_{\,t-k} + \alpha_3 ECT_{t-1} + \xi_t \tag{6c}$$

$$\Delta PtfInv_t = \sum_{k=1}^{q} \phi_{13k} \, \Delta PtfInv_{\,t-k} + \sum_{k=0}^{q} \phi_{14k} \, \Delta Credit_{\,t-k} + \sum_{k=0}^{q} \phi_{15k} \, \Delta CapForm_{\,t-k} + \sum_{k=0}^{q} \phi_{16k} \, \Delta Top1_{\,t-k} + \alpha_4 ECT_{t-1} + \vartheta_t \tag{6d}$$

Where, $k = 1 \dots q$ is the lag length which is determined by the Lag length criteria in VAR equation. ECT_{t-1} is the first lag of the estimated error correction term, which is obtained after estimating the long-run relationship for each equation. The coefficients α_i represent the adjustment coefficient toward the long run equilibrium, known as adjustment factor. In order for the model to be stable, this coefficient, must be in the interval $[-2, 0]$. ν_t, ϑ_t, ξ_t, and ϑ_t are the disturbance term assumed to be uncorrelated with zero mean. The short – run causality is assessed by examining the statistical significance of the lagged variables in the equations above, using WALD Statistic tests. The long-run causality instead, is assessed by the statistical significance of the coefficient of error correction term α_i, using t-test.

The adjustment factor is the inverse of the absolute value of the coefficient associated to the ECT, and it can be interpreted as the number of years that it takes any deviation from the equilibrium to return back to equilibrium.

As anticipated above, when we have more than two variables in the model, multivariate equation, there is the possibility of having more than one co-integration vector. In general, in a system of p number of variables, we can have at most (p-1) co-integration vector, in our case since we have four variables, at most there can be 3 co-integration relationships. To find out, Johansen methodology for co-integration is used.

There are two types of Johansen test, and Eviews reports both of them: Trace and Eigenvalue. The null and alternative hypothesis for the Trace is that the number of co-integration vectors are that:

$$H_0: r = r^* < p$$
$$H_1: r = p$$

When:

r=0, there is no co-integration, hence no long run equilibrium between the variables, and the VECM cannot be estimated

$0 < r < p$, there are r co-integration vectors, VECM can be estimated

r = p, all variables are already stationary, there is no need in estimating VECM, since VAR in the levels is good.

Johansen test procedure:

Step 1: test for the order of integration for all the variables,
Step 2: Lag selection. The procedure usually consists in estimating a VAR model of the variables in level and starting from a large number of lags, weather it is possible based on the size of our sample size. Also, we have to be careful, that a big number of lags can reduce the degrees of freedom of the model. The appropriate number of lags is than choose based on the Lag Selection Criteria, the VAR with the lowest AIC or SC is preferred.
Step 3: For the lag length selected above we perform the Johansen Test for Co-Integration. For the Trace test, the null hypothesis is rejected (r<p) is the test statistics is greater than the critical values.
Step 4: VECM estimation. VECM automatically converts the variables into the first difference (table below), while with EG two step procedure, we have to put manually the Dynamic Equation in the differences of the variables.

As the last step, we save the systems that VECM has estimated, and by the means of the OLS, we estimate each Error Correction Model.
Model diagnostics conclude the analysis, in order to understand the goodness of our estimation, usually by analyzing R-squared, F-test, and residuals analysis: test for correlation, test for heteroscedasticity and Jarque Bera test for the Normality.

5. RESULTS AND DISCUSSION

5.1. Unit Root Tests Results

As previously discussed, a necessary condition before testing for possible existence of long-run relationship between the Level of Domestic Credit to the private Sector and Income Concentration of Top 1%, all the variable should be co-integrated of order one. To examine this condition, I performed the ADF Unit Root Test for the --- countries. The results are shown in the table below.

***Table 3** Results of Unit Root Tests. ADF TEST for each country*

	Credit	D(Credit)	Top1%	D(Top1%)	CapForm	D(CapForm)	PtfInv	D(PtfInv)
US	-1,49 (0,81)	-6,07 (0,00)	-4,73 (0,03)	-5,23 (0,00)	-3,50 (0,06)	-3,81 (0,00)	-3,96 (0,02)	-10,27 (0,00)

***t-statistic values, in parenthesis the respective p-values*

It can be seen from the table that the null hypothesis of the unit root cannot be rejected, hence we conclude that the variables of our series are integrated of first order. All the variables are integrated of the same order, and none was find of order two, so we are dealing with a balanced regression equation.

5.2. Co-integration Tests Results

After the unit root tests confirm that all variables are I (1) in level, the next step is to test for the existence of a long-run relationship, hence co-integration.

The first step of the EG procedure suggests estimating with OLS the coefficients of the static regression, and proceed with testing the stationarity of the residuals of the static equation, by applying ADF test (null hypothesis: there is unit root).

Table 4. ADF test on Static Equation residuals (ECM)

	Credit Dependent Variable		Top 1% Dependent Variable	
ADF test	t-statistic	Probability	t-statistic	Probability
	-5,106716	0,0000	-3,838860	0,0004
C.V* 1% level	4,94		4,94	
5% level	4,35		4,35	
10%level	4,02		4,02	

**modified c.v for co-integration test, for n+1=4 variables, and sample size less than 50 observations*

Also, the correlogram of the residuals (Appendix) shows that there is no indication of Unit Rot, furthermore this is confirmed from the ADF test above. The absolute value t-statistics of 5,10 has to be compared with Critical Values in Table 2 (ADF C.V for test of no co-integration), which in our case (n=4 variables, sample < 50 observation) is equal of 4,35 at 5% significance level, hence we reject the null of unit root in our ECM.

Table 5. Johansen Test for Co-integration

Hypothesized No. of CE(s)	Trace Statistics	0.05 Critical Value	Probability
None	61,10723	47,85613	0,0018
At most 1	31,08944	29,79707	0,0353
At most 2*	10,61321	15,49471	0,2364
At most 3	1,328408	3,841466	0,2491

Trace Statistics of Johansen Co-Integration test, indicates at most 2 co-integration equations, since for 2 co-integration equations the Trace Statistics 10,61 is lower than the 5% Critical Values 15,49, hence we fail to reject the null hypothesis that "at most two integration equations are present in our VECM".

Different from the case of test for co-integration in the EG two steps, when using Johansen test there is no need to estimate the reverse equation where Income is the dependent variable and to test again for co-integration presence. The Johansen test is based on a VAR equation, and as a result it already estimates the maximum number of co-integration relationship for each equation.

The results for VAR Lag Selection are reported in the table below, lag 3 is selected.

__Table 6__. VAR Lag Order Selection, Endogenous variables: Credit, Top1, Cap.Form, Ptf.Inv

Lag	LR	FPE	AIC	SC	HQ
0	-	6000	20,0	20,2	21,1
1	133,4	70	15,5	16,5	15,8
2	43,5	25	14,5	16,2	15,0
3	31,5	13,1*	13,2*	16,1*	14,4
4	17,8	13,4	13,6	16,4	14,2*

5.3. Long Run Dynamics (Static Equation)

The acceptance of co-integration hypothesis, EG procedure and Johansen Test, allows us to consider the coefficients of the static model as long-run coefficients and verify if their signs reflect the expected relationships between the economic variables in the model. Results of the OLS coefficients estimation for the US case are reported in table 7.

The estimation in the OLS method was made by applying the correction to the variance – covariance matrix estimates of the estimators proposed by Newey – West (1987), called HAC. This because usually in the static equation we have presence of heteroscedasticity and autocorrelation in the residuals. HAC correction though has no effect on the estimate of parameters but only on standard errors.

Table 7. Parameter Estimation using OLS, 1980 – 2012, US case

Variables	Credit		Income Top 1%	
	Coefficients	Probabilities	Coefficients	Probabilities
Credit	-	-	0,094408	0,000
Income Top1	7,104210	0,000	-	-
Capital Formation	-4,123724	0,004	0,157437	0,4013
Ptf. Investment	-4,52268	0,086	-0,093167	0,7511
R^2	0,90		0,88	
Adj. R^2	0,89		0,86	
F-stat.	94,51		71,68	
Prob. F-stat.	0,00		0,00	
DWstat.	1,78		1,26	

Analysis of the long run coefficient estimates: For the regression estimation where Level of Credit is the Dependent Variable, all the coefficients are statistically significant at 5% level, except Portfolio Investment which is significant only at 10% level. The coefficient sign for Capital Formation is not in line with economics' theory as we expect capital formation to be positively related to credit since capital investment need to be financed with credit. The size of Portfolio Investment is non-significant at 5% level, but its sign is in line with our expectations: credit expansion is lower in the presence of large outflows of portfolio investments. When we reverse the regression, and Income Top1% becomes the dependent variable, only Credit is significant. Regarding the coefficient signs, for an increment of 1% of the level of domestic credit to the private sector, the ratio of income concentration among the top 1% increase by 9%. Whether it seem like Capital Formation and Portfolio Investment have no contribution in explaining the income inequality top 1% ratio.

5.4. Short Run Dynamics (Dynamic Equation)

The acceptance of the hypothesis of no unit root in the residuals of the Static Model, allow us to estimate the dynamic equation in the ECM form, second step EG methodology. Once the long-run equilibrium is established, we estimate the Error Correction Model (ECM) and examining the causality direction between the variables. For the dynamic equation, we do not apply HAC Correction, since for the dynamic equation we assume (and will be checked latter) that disturbances are white noise: lack of autocorrelation and heteroscedasticity between the residuals, and normal distribution. In case of EG two step procedure, the Short Run Dynamic can be estimated only for equation 3. (below) since for equation 4. (reverse of equation 3, with Income Top 1% as the dependent variable) the condition of co-integration in the first step was not satisfied.

$$Credit_t = \propto + \beta_1 Top1_t + \beta_2 CapForm_t + \beta_3 PtfInv_t + \varepsilon_t$$

Table 8. *Dynamic OLS Model, Engel Granger Two Steps Methodology*

	Source of causation (independent variables) Short – Run									Long – Run
	$\Delta \boldsymbol{Top1}$			$\Delta \boldsymbol{CapForm}$			$\Delta \boldsymbol{PtfInv}$			$\boldsymbol{ECT}$
	Lag1	Lag 2	Lag 3	Lag 1	Lag 2	Lag 3	Lag1	Lag 2	Lag 3	
$\Delta Credit(eq.4a)$	0,53 (0,69)	-2,26 (0,11)	-1,20 (0,35)	6,95 (0,06)	1,64 (0,65)	4,52 (0,14)	2,06 (0,20)	-1,13 (0,39)	-3,33 (0,01)	***-0,222753 (0,1589)***

The ECT_{credit} equal to -0,22 satisfies the stability condition since it is included in the interval $[-2, 0]$. It can be described as the adjustment factor, about 22,27% of the discrepancy between long-term and short-term level of credit, is corrected within a year. But, being statistically not significant we say that the variable is not adjusting to the long run equilibrium, no causality run *from* top 1%, capital formation and portfolio investment *to* the level of credit.
The goodness of our results is confirmed by the residuals analysis of the model, and it is confirmed the hypothesis that they are a realization of white noise: zero autocorrelation, homoscedastic, and normally distributed (Appendix), overall the model is very satisfactory.

On the other hand, Johansen Co-Integration test, estimated 2 co-integrations equations when all the four variables enter the equation as exogenous. The VECM automatically converts the variables into the first difference (table below), while with EG two step procedure, we have to put manually the Dynamic Equation in the differences of the variables. Results are summarized in Table 9, OLS Estimation in Appendix.

Table 9. VECM equation estimation from Johansen Procedure

	Dependent Variable	Equation Summary
1.	$\Delta Credit_t$	D(CREDIT) = C(1)*(CREDIT(-1) + 8.12739882001*CAPFORM(-1) + 22.9486909619*PTFINV(-1) - 274.189255317) + C(2)*(TOP1(-1) + 0.85604605912*CAPFORM(-1) + 1.95482529247*PTFINV(-1) - 30.5241479246) + C(3)*D(CREDIT(-1)) + C(4)*D(CREDIT(-2)) + C(5)*D(CREDIT(-3)) + C(6)*D(TOP1(-1)) + C(7)*D(TOP1(-2)) + C(8)*D(TOP1(-3)) + C(9)*D(CAPFORM(-1)) + C(10)*D(CAPFORM(-2)) + C(11)*D(CAPFORM(-3)) + C(12)*D(PTFINV(-1)) + C(13)*D(PTFINV(-2)) + C(14)*D(PTFINV(-3)) + C(15)
2.	$\Delta Top1_t$	D(TOP1) = C(16)*(CREDIT(-1) + 8.12739882001*CAPFORM(-1) + 22.9486909619*PTFINV(-1) - 274.189255317) + C(17)*(TOP1(-1) + 0.85604605912*CAPFORM(-1) + 1.95482529247*PTFINV(-1) - 30.5241479246) + C(18)*D(CREDIT(-1)) + C(19)*D(CREDIT(-2)) + C(20)*D(CREDIT(-3)) + C(21)*D(TOP1(-1)) + C(22)*D(TOP1(-2)) + C(23)*D(TOP1(-3)) + C(24)*D(CAPFORM(-1)) + C(25)*D(CAPFORM(-2)) + C(26)*D(CAPFORM(-3)) + C(27)*D(PTFINV(-1)) + C(28)*D(PTFINV(-2)) + C(29)*D(PTFINV(-3)) + C(30)
3.	$\Delta CapForm_t$	D(CAPFORM) = C(31)*(CREDIT(-1) + 8.12739882001*CAPFORM(-1) + 22.9486909619*PTFINV(-1) - 274.1892553) + C(32)*(TOP1(-1) + 0.85604605912*CAPFORM(-1) + 1.95482529247*PTFINV(-1) - 30.5241479246) + C(33)*D(CREDIT(-1)) + C(34)*D(CREDIT(-2)) + C(35)*D(CREDIT(-3)) + C(36)*D(TOP1(-1)) + C(37)*D(TOP1(-2)) + C(38)*D(TOP1(-3)) + C(39)*D(CAPFORM(-1)) + C(40)*D(CAPFORM(-2)) + C(41)*D(CAPFORM(-3)) + C(42)*D(PTFINV(-1)) + C(43)*D(PTFINV(-2)) + C(44)*D(PTFINV(-3)) + C(45)
4.	$\Delta PtfInv_t$	D(PTFINV) = C(46)*(CREDIT(-1) + 8.12739882001*CAPFORM(-1) + 22.9486909619*PTFINV(-1) - 274.189255) + C(47)*(TOP1(-1) + 0.85604605912*CAPFORM(-1) + 1.95482529247*PTFINV(-1) - 30.5241479246) + C(48)*D(CREDIT(-1)) + C(49)*D(CREDIT(-2)) + C(50)*D(CREDIT(-3)) + C(51)*D(TOP1(-1)) + C(52)*D(TOP1(-2)) + C(53)*D(TOP1(-3)) + C(54)*D(CAPFORM(-1)) + C(55)*D(CAPFORM(-2)) + C(56)*D(CAPFORM(-3)) + C(57)*D(PTFINV(-1)) + C(58)*D(PTFINV(-2)) + C(59)*D(PTFINV(-3)) + C(60)

VECM estimates 15 coefficients for each equation, having selected VAR lag 3. c(1), c(16), c(31) and c(46) are the Error Correction Terms for each system, the speed of adjustment towards equilibrium. Since my interest are the first two equation, the table below summarizes the information regarding the Error Correction Term for each equation, while for more information please see Appendix.

Table 10. *Estimated ECT, Long Run Causality from Income Top1, Capital Formation and Portfolio Investment to Credit*

$causality\ vs \Rightarrow Credit$	ECT Coefficient	ECT P-Value
EG Two Step Procedure	-0,22	0,15
VECM, Johansen Test	-0,26	0,50

And,

Table 11. *Estimated ECT, Long Run Causality from Level of Credit, Capital Formation and Portfolio Investment to Income Inequality*

$\boldsymbol{causality\ vs \Rightarrow Top1}$	ECT Coefficient	ECT P-Value
EG Two Step Procedure*	-	-
VECM, Johansen Test	0,0332	0,7479

**Since no co-integration relationship was found in the two steps procedure EG, the long run causality test that goes from Credit to Income will be investigated by estimating a VAR of the stationary variables, I (0), and Granger test for causality will be applied.*

5.5. Causality Analysis

The causality in the short – run is assessed by examining the statistical significance of the lagged variables in the dynamic equation, using the Wald statistical test; Since I am interested in the causality nexus between Top1% of Income and the Level of credit, the Wald Test is performed for the following null hypothesis:

a) Jointly Income Concentration coefficients are zero, c(6)=c(7)=c(8)=0). If the null hypothesis cannot be rejected, we say that past values of income concentration does not cause present value of credit.

H_0: Jointly Coefficient (Top1) = 0

H_1: Null Hypothesis is not true

Table 12. Wald test for short –run causality running from Income to Credit, case 1 Table 9

WALD test, Null Hypothesis Jointly coeff. (top1) = 0		
Test Statistics	Value	Probability
t-statistic	2.472908	0.1044
Chi-square	7.418724	0.0597

We reject the null hypothesis quite at 5 percent level when Chi-square statistic is used, and at 10 percent significance level for the t –statistic, meaning that in the short – run, there exist a causality link running from Income Concentration Top 1% to the Level of Domestic Credit for the private sector and it is statistically significant. While there is no causality effect in short – run from credit to income concentration, which means that in short-run, the level of domestic credit to the private sector has no causality effect in the level of income concentration, measured by the ratio of top 1% over the rest 99% of the population.

In the long – run, the causality is assessed by the statistical significance of the error correction terms, for each equation, using a t-test. For the causality running from Income Inequality to the level of credit (table 10), the Error Correction Term in EG two step procedure equal of -0,22 is not significant at none of our levels of significance, 1, 5 or 10%. Same result is confirmed from the VECM results, the error correction term equal of -0,26 is not significant. In both cases, the ECTs satisfy the Stability Condition, because they are both negative and between -2 and 0, but the p-value is higher that our normal statistical significance. Hence, there is no long run causality from Income inequality to the Level of Domestic Credit to the Private Sector.

Regarding the reverse equation, the causality running from Level of Credit to the Income Top1% ratio (table 11) can be interpreted as follow: the ECT in VECM model is equal 0,032 with p-value of 0,74. The error correction term does not satisfy neither of the Stability condition, it's positive and statistically non-significant, hence based on VECM there is no long - run causality from level of credit to income inequality.

Since no co-integration relationship was found in the two steps procedure EG when Income is the dependent variable, the long run causality test that goes from Credit to Income will be investigated by estimating a VAR of the stationary variables, I (0), and Granger test for causality. This is one of the good features of VAR model, they allow to test the direction of the causality. The concept of causality is not the same as the concept of regression equations. Usually the regression tells us whether there is some sort of relationship between two variables, suppose X_t and Y_t, and does not tell the nature of the relationship, such as whether X_t causes Y_t or Y_t causes X_t. The concept of causality is as follow: if X Causes Y, the changes of X happened first then they were followed by changes of Y. The graph below plots changes of Top 1% Income Concentration and Level of Domestic Credit to the Private sector in America, for the period from 1980 - 2012.

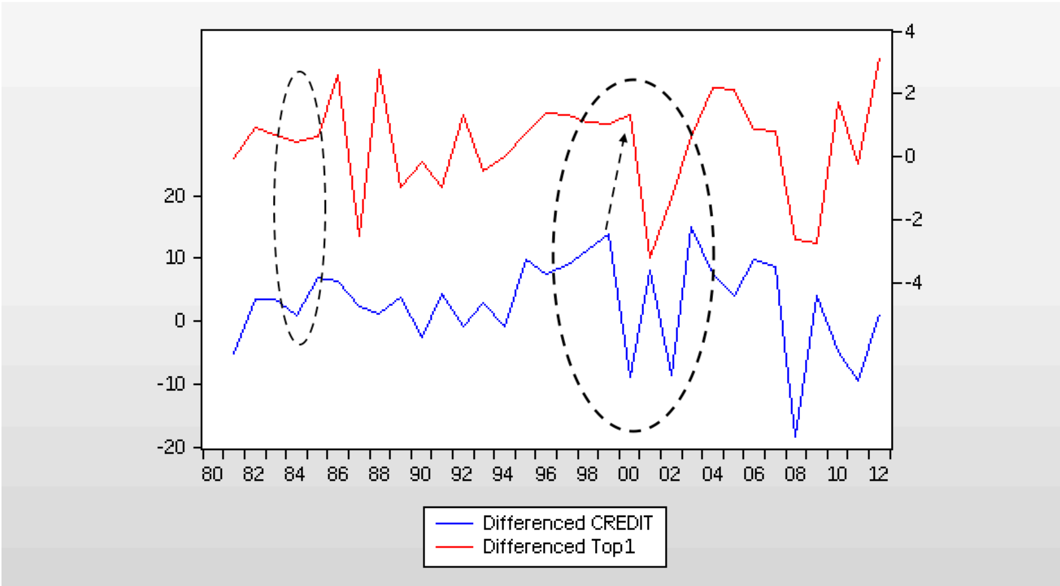


Fig.6. *Level of Credit and Income Concentration Top 1% in their first difference. As we can see, it seems like a change in the level of credit is followed by changes in the level of income inequality (credit changes happen first), hence we would expect the level of Credit to Granger Cause the level of Income Inequality (as argued from Rajan).*

Table below reports the results of the Granger Causality test for Level of Credit and Income Concentration Top 1%, both series must be transformed first into stationary, d(top1) and d(credit).

Table 13*. Granger Causality Test, Lag Order 3*

Null Hypothesis: d(X) does not Granger Cause d(Y)	**Obs.**	**F-stat.**	**Prob.**
D(Credit) does not Granger Cause D(Top1)	29	1,69305	0,1976
D(Top1) does not Granger Cause D(Credit)	29	0,45750	0,7147

As the p-value for both tests are higher than 5% and 10% significance level, the null hypothesis is rejected and based on the Granger Test, there is a long run causality for both direction, bilateral causality. But we have to be careful with this results, since in the Granger Causality test, the Error Correction Term is not included as it happens in case of Johansen test, as a result outputs and conclusions might be contradictory.

6. CONCLUSIONS

Based on the recent argument of Rajan in his book “Faul Line” for the existence of a nexus between income inequality and the level of indebtedness of households, in this paper I empirically investigated this hypothesis by the use of co-integration techniques both for single equation applying EG Two Step Methodology and multivariate technique by the means of Vector Error Correction Models.

Even though four variables enter the equation, I focused only on the causality between Income Inequality measured as the ratio of income going to the top 1% of the household to the income going the rest 99%, and the Level of Domestic Credit to the Private Sector which is my proxy for Household Indebtedness.

EG Co-integration test provided evidence for long run equilibrium when Credit is the dependent variable, since we found the residuals of the static equation stationary. This wasn’t the case when the equation was reversed and Income Inequality Top 1% was used as dependent variable. The residuals of the static equation didn’t satisfy the stationary condition, hence based on EG test there was found no co-integration relationship. This contradictory result may be due to one of the main drawbacks of the EG procedure which is the order of the variables. When estimating the long –run relationship we have to place one variable on the left side and use the others as regressors. The test does not say anything about which of the variables can be used as regressors and why. For example, in case of only two variables, Income Top1 and Level of Credit, we can either regress Credit on Income or choose the reverse and regress Income on Credit. Asymptotically, when the sample size goes to infinity, the test for co-integration on the residuals of these two variables is the same, but since in this paper the sample size was limited to only 33 variables, the test for co-integration based on EG two step procedure test didn’t produce equivalent results. On the other hand, Johansen test for co-integration confirmed the existence of co-integration among our four variables, at most 2-co-integration relationship were found. In Johansen test all variables enter as Endogenous, hence once the co-integration is confirmed there is no need to reverse the equation. A VECM lag 3 was estimated, and the respective dynamic equations for all the four variables were produced. The causality between Income Inequality and Level of Credit was tested both in short-run and long – run. In the first case, the short – run causality was assessed by examining the statistical significance of the lagged variables in the dynamic equations. A causality link was found from Income Inequality to the level of credit, by the means of Wald test on the dynamic coefficients estimated following EG two step procedure. But no causality effect was found for the reverse equation, the level of credit in short-run does not cause the level of income inequality. Regarding the long-run equilibrium, Error Correction Terms were analyzed, weather only in case the ECT were going to be significant and included in the interval (-2, 0) we would have say that there exists a long run equilibrium among the variables and compute the speed of adjustment. For the long-run causality going from Income Inequality to Credit level, the ECT coefficient was found to be negative, but not significant. Same result was confirmed from the dynamic equation estimated through the Vector Error Correction Model. Hence, in the long run we found that Income Inequality does not cause the Level of Credit, while as argued above, in the short run, Income Inequality was found to have a causal effect on the level of credit. Lastly, neither in case of causality going from Credit to Income, the ECM term was positive and negative, hence we reject the hypothesis of a long run nexus and equilibrium between our two variables of interest.

Summarizing, the only causality link that is find is in short-run and moving from Income Inequality to the Level of Domestic Credit to the Private Sector. This result is in line and confirms Rajan's hypothesis that in short term, people crushed by unemployment, job market disruption, and other factors, in order to maintain their consumption made use of the credit offered by the financial institutions even in cases where there was no eligibility for such credits, as a result in long term, together with other economic factor, it triggers the financial crises of 2007/2008.

BIBLIOGRAPHY

Atkinson, A. and Morelli, S. 2011. *'Economic Crises and Inequality'*, Human Development Research Paper no. HDRP-2011-06, Human Development Report Office, United Nations Development Program

Alesia. Perotti., 1996. Income distribution, political instability and investment. European Economic review, Vol. 40, 1203-1228

Acemoglu, D., 2011. Thoughts on Inequality and the financial crisis, MIT https://economics.mit.edu/files/6348

Barlas. Y., 2012. Book review: Falut Lines, Central bank Review, Vol. 12, pp.37-44, http://www.tcmb.gov.tr/wps/wcm/connect/27e11fa8-af8c-454f-85ea-c3c3074ab711/july12-4.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE27e11fa8-af8c-454f-85ea-c3c3074ab711

Borio, C. and White, W. R. 2003. Whither monetary and financial stability: the implications of evolving policy regimes, Proceedings - Economic Policy Symposiom - Jackson Hole, Federal Reserve Bank of Kansas City, 131–211

Bordo, M. D. and Meissner, C. M. 2012. *Does inequality lead to a financial crisis*? Journal of International Money and Finance, vol. 31, no. 8, 2147–61

Costantini, M., 2010. Panel unit root and cointegration methods, University of Vienna, http://homepage.univie.ac.at/mauro.costantini/master_class_2010.pdf

Crotty, J. 2009. Structural causes of the global financial crisis: a critical assessment of the 'new financial architecture' Cambridge Journal of Economics, vol. 33, 563-580

Charfeddine, Kahia, Ben Aissa, 2016. Impact of renewable and non-renewable energy consumption on economic growth: new evidence from the MENA Net Oil Exporting Countries, Researchgate

Dimitrov, V., Palia.D., and Lang.Leo., 2015. *Impact of the Dodd-Frank act on credit ratings*, Journal of Financial Economics, vol.115, 505-520.

Economic Freedom of the World, 2016 Annual Report, Fraser Institute

Financial Liberalization: what went right, what went wrong? Chp.7, World Bank

Finance and Inclusive Growth; Boris, Denk, Hoeller; OECD Policy Paper, June 2015, No.14

Finkel, E. S., *Casual Analysis with Panel Data*, Department of Government and Foreign Affairs, University of Virginia

Fisher, I. 1933 "The Debt Deflation Theory of Great Depressions." Econometrica 1 (4). Pp. 337-357.

Galbrith, J.K., 2012. Inequality and Instability: *A study of the world economy just before the great crisis*, Oxford University Press

Granger.W.J., Engle.F.R., 1987. *Co-integration and error correction: representation, estimation and testing*, Econometrica, Vol. 55, No. 2, 251-276

Griffiths, W.E., Hill.R.C., Lim.G.C., 2012. Using Eviews for principals of econometrics, John Wiley &Sons, fourth edition

Gujarati, D., 2011. Econometrics by examples, Iacoviello, M., 2008. *Household Debt and Income Inequality, 1963-2003*; Journal of Money, Credit and Banking, Vol. 40, No. 5

Keeley, B. (2015), Income Inequality: The Gap between Rich and Poor, OECD Insights, OECD Publishing, Paris.

Krugman, P. Inequality and crises: coincidence or causation? Princeton University, https://www.princeton.edu/~pkrugman/inequality_crises.pdf

Maddala GS, Wu S. 1999. A comparative study of unit root tests with panel data and a new simple test. Oxf Bull Econ Stat. Vol. 61:631-652.

Milanovic, B. 2016. Global Inequality: *A new approach for the age of globalization*, Harvard University Press

Milanovic, B. 2009. Two views on the cause of the global crisis—part I, Yale Global Online, http://yaleglobal.yale.edu/content/two-views-global-crisis

Tridico, P. 2012. Financial crisis and global imbalances: its labor market origins and the aftermath, Cambridge Journal of Economics, vol. 36, no. 1, 17–42

Hakkio, C.S., Rush, M., 1991. Cointegration: how short is the long-run? Journal of International Money and Finance 10, 571–581.

Maddala GS, Wu S.1999. A comparative study of unit root tests with panel data and a new simple test. Oxf Bull Econ Stat.61:631-652.

Ranci.re, R., Tornell, A. and Westermann, F. 2006. 'Decomposing the Effects of Financial Liberalization: Crises vs. Growth', Working Paper no. 74, Institut für Empirische Wirtschaftsforschung

Phillips PCB, Hansen BE.1990. Statistical inference in instrumental variables regression with I (1) processes. Rev Econ Stud Vol. 57:99e125

Phillips PCB, Moon HR. 1999. Linear regression limit theory for nonstationary panel data. Econometrica Vol. 67:105 - 112

Piketty, Th., 2013. *Capital in the twenty-first century,* Belknap Press

Kao C. 1999. Spurious regression and residual-based tests for cointegration in panel data. J Econ 90. 1-44.

Reich, R. (2010), Aftershock: The Next Economy and America's Future, New York: Random House.

Scenario, Stress Test and Strategies – The rise of populism. MSCI second quarter 2016 Report

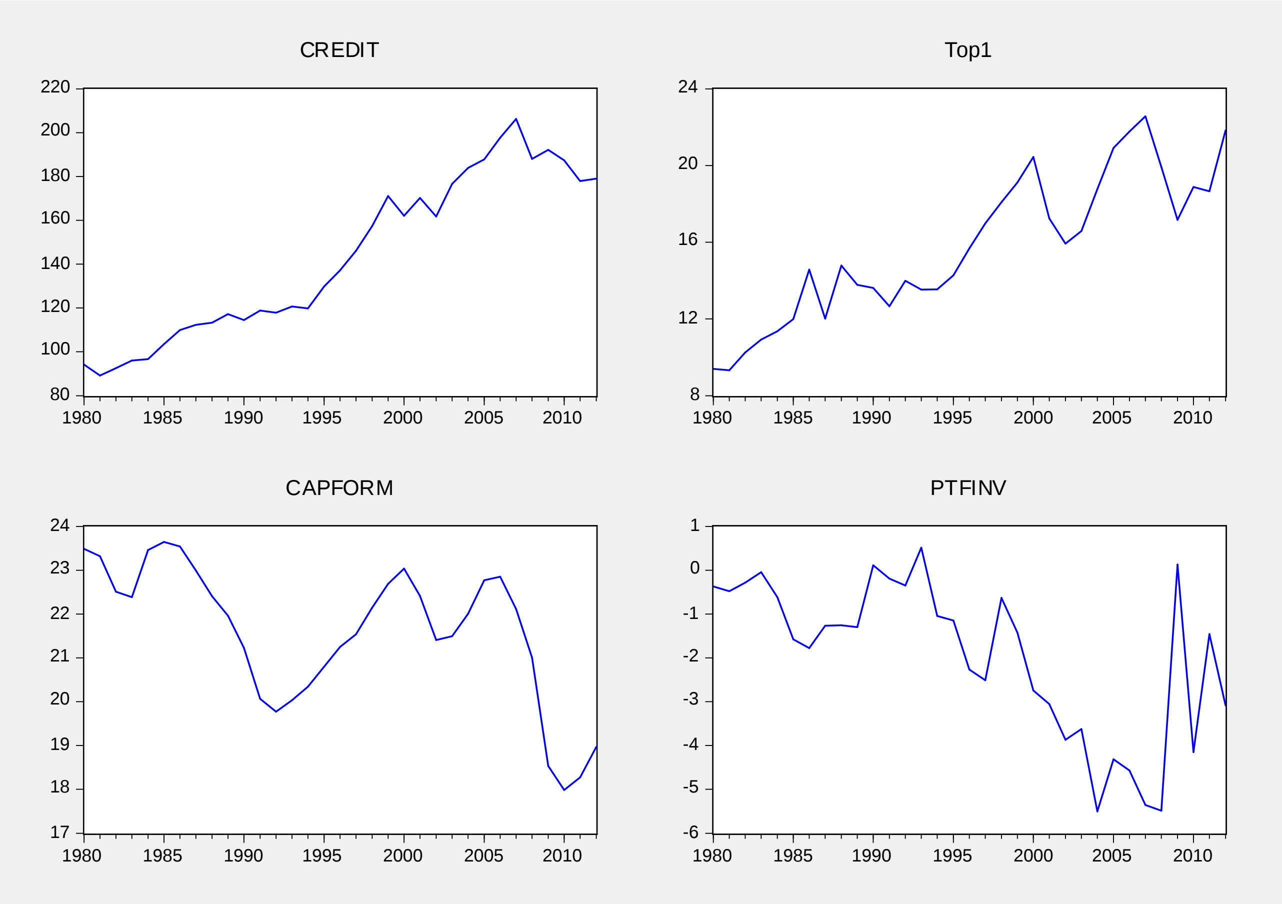
Stiglitz, J. 2009. The global crisis, social protection and jobs, International Labour Review, vol. 148

http://www.economist.com/news/books-and-arts/21695853-surprisingly-little-known-about-causes-inequality-serbian-american-economist

K.B. Luintel, M. Khan., 1999. A quantitative reassessment of the finance-growth nexus: evidence from a multivariate VAR, Journal of Development Economics 60, 381–405

7. APPENDIX

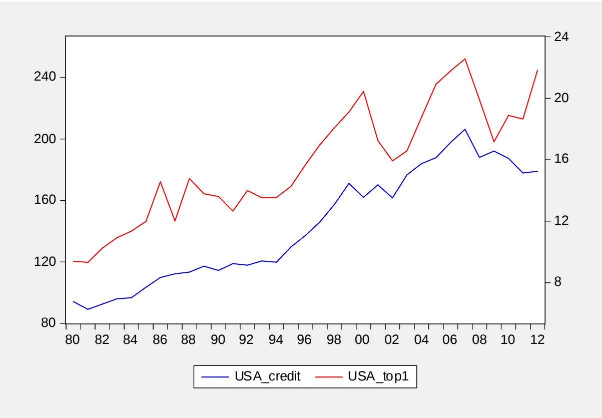
Descriptive statistics

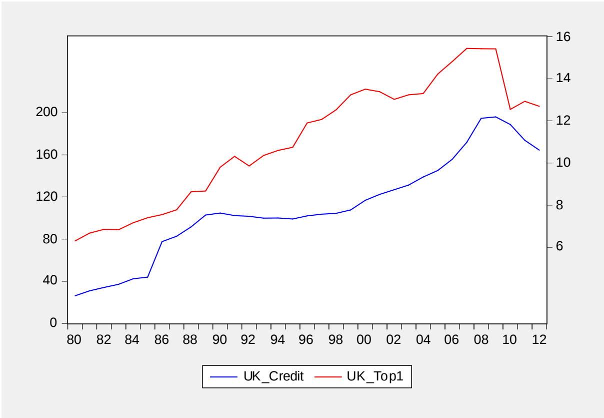


	CREDIT	LCREDIT	TOP1	LTOP1	CAPFORM	LCAPFORM	PTFINV
Mean	143.2955	4.931518	15.77394	2.729335	21.58838	3.069399	-1.969798
Median	137.2055	4.921480	15.67300	2.751939	22.00648	3.091337	-1.425270
Maximum	206.3028	5.329345	22.56200	3.116267	23.64299	3.163067	0.512407
Minimum	89.12920	4.490087	9.323000	2.232484	17.98343	2.889451	-5.502142
Std. Dev.	36.99042	0.264691	3.767982	0.248321	1.592399	0.076249	1.789341
Skewness	0.104779	-0.106525	0.067176	-0.313113	-0.764832	-0.887165	-0.619042
Kurtosis	1.533589	1.580010	1.998850	2.186741	2.663793	2.879866	2.195021
Jarque-Bera	3.017128	2.834922	1.402984	1.448629	3.372747	4.348683	2.998662
Probability	0.221227	0.242328	0.495845	0.484657	0.185190	0.113683	0.223279
Sum	4728.752	162.7401	520.5400	90.06806	712.4165	101.2902	-65.00334
Sum Sq. Dev.	43785.31	2.241956	454.3260	1.973229	81.14355	0.186043	102.4558
Observations	33	33	33	33	33	33	33

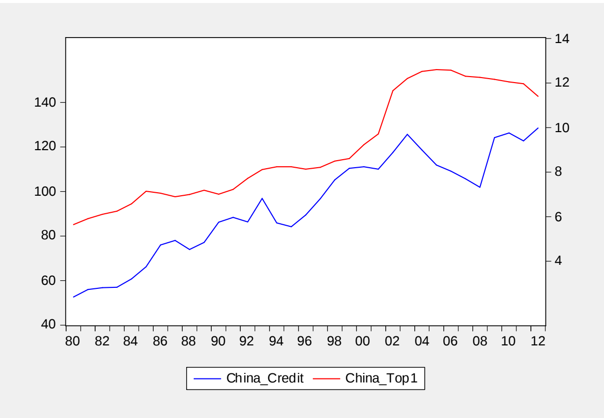
Level of Credit and Income Top1, comparative group graph

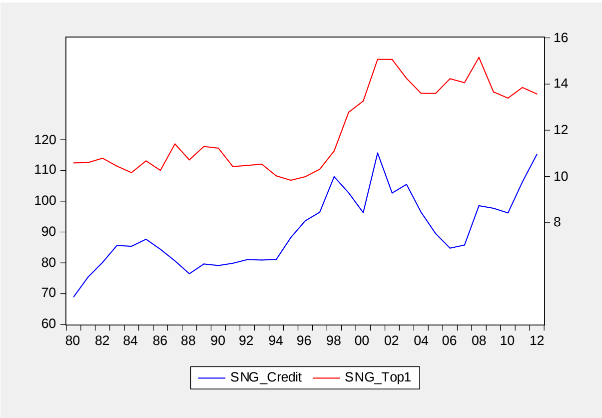
a) Anglo – Saxon Countries



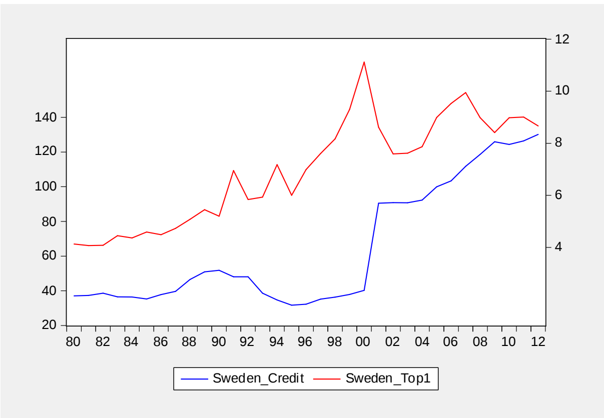


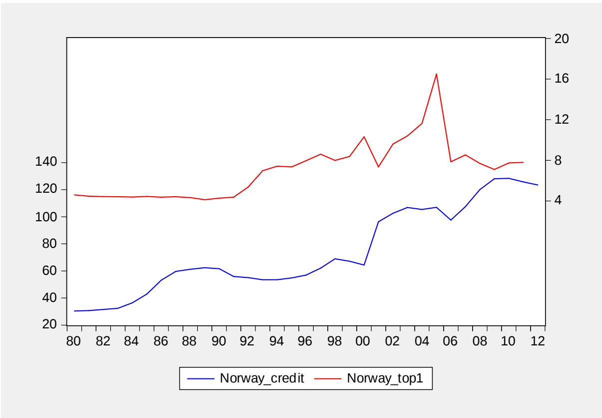
b) Emerging Market Countries





c) Nordic Countries





UNIT ROOT TESTS *(t-statistics are compared with Critical Values of ADF for Unit Root Test)*

Unit Root Test for **Credit – USA**, in levels and 1st difference

Null Hypothesis: CREDIT has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.493407	0.8109
Test critical values:	1% level	-4.273277	
	5% level	-3.557759	
	10% level	-3.212361	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: D(CREDIT) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob
Augmented Dickey-Fuller test statistic		-6.139427	0.00
Test critical values:	1% level	-4.284580	
	5% level	-3.562882	
	10% level	-3.215267	

*MacKinnon (1996) one-sided p-values.

Unit Root Test for **Income Top 1% – USA**, levels and 1st difference

ull Hypothesis: TOP1 has a unit root
xogenous: Constant, Linear Trend
ag Length: 2 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.*
ugmented Dickey-Fuller test statistic		-4.733582	0.0035
est critical values:	1% level	-4.296729	
	5% level	-3.568379	
	10% level	-3.218382	

MacKinnon (1996) one-sided p-values.

Null Hypothesis: D(TOP1) has a unit root
Exogenous: Constant
Lag Length: 4 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.993878	0.0004
Test critical values:	1% level	-3.699871	
	5% level	-2.976263	
	10% level	-2.627420	

*MacKinnon (1996) one-sided p-values.

Unit Root Test for **Capital Formation** in levels and 1st difference

ıll Hypothesis: CAPFORM has a unit root
ogenous: Constant, Linear Trend
g Length: 1 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.*
ıgmented Dickey-Fuller test statistic		-3.501860	0.0567
st critical values:	1% level	-4.284580	
	5% level	-3.562882	
	10% level	-3.215267	

lacKinnon (1996) one-sided p-values.

Null Hypothesis: D(CAPFORM) has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-3.785575	0.007
Test critical values:	1% level	-3.670170	
	5% level	-2.963972	
	10% level	-2.621007	

*MacKinnon (1996) one-sided p-values.

Unit Root Test for **Portfolio Investment** in levels and 1st difference

ull Hypothesis: PTFINV has a unit root
ogenous: Constant, Linear Trend
ag Length: 0 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.*
ıgmented Dickey-Fuller test statistic		-3.969062	0.0203
st critical values:	1% level	-4.273277	
	5% level	-3.557759	
	10% level	-3.212361	

lacKinnon (1996) one-sided p-values.

Null Hypothesis: D(PTFINV) has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-10.13992	0.0000
Test critical values:	1% level	-3.661661	
	5% level	-2.960411	
	10% level	-2.619160	

*MacKinnon (1996) one-sided p-values.

SERIES CORRELOGRAM (at levels)

Level of Credit

Sample: 1980 2012
Included observations: 33

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.936	0.936	31.642	0.000
		2	0.866	-0.085	59.596	0.000
		3	0.779	-0.178	82.934	0.000
		4	0.682	-0.112	101.47	0.000
		5	0.588	-0.014	115.76	0.000
		6	0.492	-0.068	126.11	0.000
		7	0.397	-0.054	133.12	0.000
		8	0.315	0.041	137.71	0.000
		9	0.222	-0.169	140.07	0.000
		10	0.136	-0.030	141.01	0.000
		11	0.054	-0.042	141.16	0.000
		12	-0.035	-0.127	141.22	0.000
		13	-0.119	-0.080	142.05	0.000
		14	-0.206	-0.101	144.63	0.000
		15	-0.285	-0.025	149.83	0.000
		16	-0.341	0.066	157.71	0.000

Income Top 1%

Sample: 1980 2012
Included observations: 33

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prol
		1	0.820	0.820	24.291	0.0
		2	0.685	0.037	41.785	0.0
		3	0.543	-0.088	53.132	0.0
		4	0.488	0.176	62.628	0.0
		5	0.442	0.037	70.702	0.0
		6	0.400	-0.022	77.543	0.0
		7	0.377	0.086	83.845	0.0
		8	0.275	-0.231	87.329	0.0
		9	0.195	-0.031	89.168	0.0
		10	0.148	0.101	90.263	0.0
		11	0.103	-0.122	90.820	0.0
		12	0.047	-0.083	90.944	0.0
		13	-0.054	-0.155	91.111	0.0
		14	-0.131	-0.088	92.161	0.0
		15	-0.222	-0.078	95.337	0.0
		16	-0.297	-0.125	101.35	0.0

Capital Formation

Sample: 1980 2012
Included observations: 33

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.830	0.830	24.850	0.000
		2	0.526	-0.522	35.156	0.000
		3	0.225	-0.027	37.105	0.000
		4	0.002	0.000	37.105	0.000
		5	-0.109	0.057	37.591	0.000
		6	-0.162	-0.156	38.708	0.000
		7	-0.188	-0.038	40.273	0.000
		8	-0.218	-0.119	42.468	0.000
		9	-0.265	-0.109	45.842	0.000
		10	-0.301	-0.046	50.394	0.000
		11	-0.299	0.003	55.089	0.000
		12	-0.222	0.115	57.801	0.000
		13	-0.107	-0.055	58.463	0.000
		14	0.013	0.043	58.473	0.000
		15	0.127	0.100	59.505	0.000
		16	0.221	0.092	62.815	0.000

Portfolio Investment

Sample: 1980 2012
Included observations: 33

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.589	0.589	12.533	0.000
		2	0.633	0.439	27.485	0.000
		3	0.478	0.019	36.278	0.000
		4	0.410	-0.056	42.958	0.000
		5	0.262	-0.151	45.786	0.000
		6	0.241	0.024	48.276	0.000
		7	0.139	-0.012	49.136	0.000
		8	0.134	0.042	49.970	0.000
		9	0.006	-0.140	49.971	0.000
		10	0.022	-0.011	49.996	0.000
		11	-0.030	0.031	50.045	0.000
		12	-0.169	-0.257	51.618	0.000
		13	-0.214	-0.144	54.255	0.000
		14	-0.264	-0.034	58.477	0.000
		15	-0.246	0.139	62.375	0.000
		16	-0.178	0.238	64.539	0.000

Comment: for all the four series, the p-values are zero, which means that we reject the null hypothesis, and conclude that variables are not stationary at levels.

ENGLE GRANGER TWO STEPS PROCEDURE - OLS REGRESSION

a) Dependent variable “Level of Domestic Credit to the private sector”

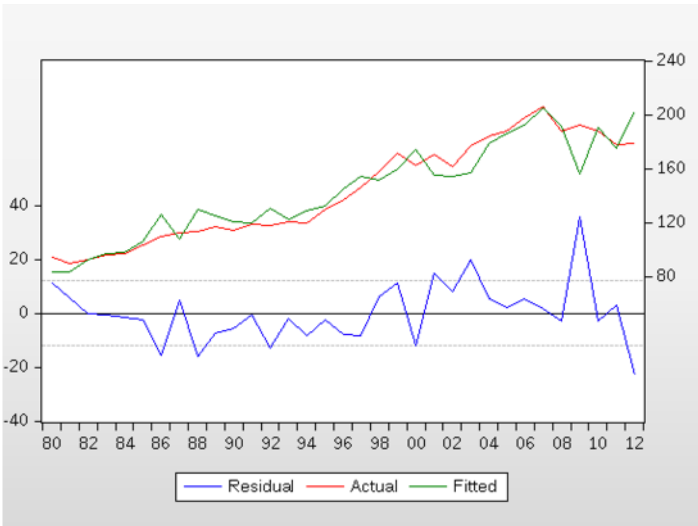
Static Equation, case USA

Dependent Variable: CREDIT
Method: Least Squares
Date: 06/17/17 Time: 15:29
Sample: 1980 2012
Included observations: 33
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	111.3499	30.02193	3.708950	0.0009
TOP1	7.104210	0.982203	7.232933	0.0000
CAPFORM	-4.123724	1.320872	-3.121971	0.0040
PTFINV	-4.522685	2.552004	-1.772209	0.0869

R-squared	0.907216	Mean dependent var	143.2955
Adjusted R-squared	0.897617	S.D. dependent var	36.99042
S.E. of regression	11.83594	Akaike info criterion	7.893372
Sum squared resid	4062.598	Schwarz criterion	8.074767
Log likelihood	-126.2406	Hannan-Quinn criter.	7.954406
F-statistic	94.51740	Durbin-Watson stat	1.781759
Prob(F-statistic)	0.000000	Wald F-statistic	140.3609
Prob(Wald F-statistic)	0.000000		

Residuals of the static equation (ECM) and ADF test on the residuals*



Null Hypothesis: ECM has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.106716	0.0000
Test critical values:	1% level	-2.639210	
	5% level	-1.951687	
	10% level	-1.610579	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(ECM)
Method: Least Squares
Date: 06/18/17 Time: 21:52
Sample (adjusted): 1981 2012
Included observations: 32 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ECM(-1)	-0.968493	0.189651	-5.106716	0.0000

- T-test is compared not with the one reported by the Eviews above, but with the ADF C.V for co-integration test (4,94 at 1%, 4,35 at 5% and 4,02 at 10%, for n=4 variables, sample<50 obs.)

Dynamic Equation,

Dependent Variable: D(CREDIT)
Method: Least Squares
Date: 06/19/17 Time: 01:57
Sample (adjusted): 1984 2012
Included observations: 29 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.362411	2.796789	2.632451	0.0219
D(CREDIT(-1))	-0.733631	0.261962	-2.800526	0.0160
D(CREDIT(-2))	-0.120183	0.293517	-0.409460	0.6894
D(CREDIT(-3))	0.105245	0.256446	0.410397	0.6887
D(TOP1)	2.258350	1.061731	2.127044	0.0548
D(TOP1(-1))	0.530598	1.332104	0.398315	0.6974
D(TOP1(-2))	-2.264372	1.314819	-1.722193	0.1107
D(TOP1(-3))	-1.203117	1.239719	-0.970476	0.3510
D(CAPFORM)	0.272273	2.876467	0.094655	0.9262
D(CAPFORM(-1))	6.957199	3.423324	2.032293	0.0649
D(CAPFORM(-2))	1.646062	3.586669	0.458939	0.6545
D(CAPFORM(-3))	4.529182	2.890334	1.567010	0.1431
D(PTFINV)	0.507460	1.180720	0.429789	0.6750
D(PTFINV(-1))	2.066341	1.541712	1.340290	0.2050
D(PTFINV(-2))	-1.131395	1.282840	-0.881946	0.3951
D(PTFINV(-3))	-3.330365	1.239434	-2.687005	0.0198
ECM(-1)	-0.222753	0.148277	-1.502269	0.1589

R-squared	0.842436	Mean dependent var	2.862928
Adjusted R-squared	0.632350	S.D. dependent var	7.566130
S.E. of regression	4.587660	Akaike info criterion	6.174642
Sum squared resid	252.5595	Schwarz criterion	6.976160
Log likelihood	-72.53231	Hannan-Quinn criter.	6.425668
F-statistic	4.009958	Durbin-Watson stat	2.179475
Prob(F-statistic)	0.009617		

Dependent Variable: D(CREDIT)
Method: Least Squares
Date: 06/17/17 Time: 19:01
Sample (adjusted): 1984 2012
Included observations: 29 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob
C	4.416944	1.387927	3.182404	0.004
D(CREDIT(-1))	-0.504395	0.195819	-2.575828	0.01
D(TOP1)	1.851745	0.840449	2.203280	0.03
D(CAPFORM(-1))	4.746940	2.020043	2.349920	0.02
D(PTFINV(-3))	-2.637055	0.958924	-2.750015	0.01
ECM(-1)	-0.131196	0.141987	-0.923996	0.36

R-squared	0.468044	Mean dependent var	2.8629
Adjusted R-squared	0.352402	S.D. dependent var	7.5661
S.E. of regression	6.088729	Akaike info criterion	6.6327
Sum squared resid	852.6704	Schwarz criterion	6.9156
Log likelihood	-90.17484	Hannan-Quinn criter.	6.7213
F-statistic	4.047334	Durbin-Watson stat	2.6602
Prob(F-statistic)	0.008812		

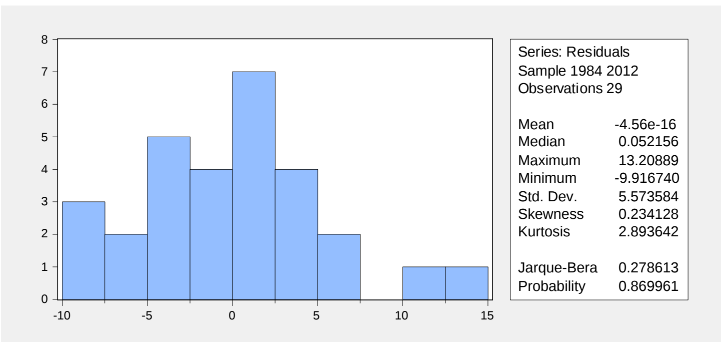
Residual diagnostics of the dynamic model (from the left to the right: test of autocorrelation, test of heteroscedasticity, and normality test)

Q-statistic probabilities adjusted for 5 dynamic regressors

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.278	-0.278	2.4867	0.115
		2	0.162	0.091	3.3571	0.187
		3	0.209	0.301	4.8652	0.182
		4	-0.313	-0.238	8.3955	0.078
		5	0.108	-0.133	8.8361	0.116
		6	0.141	0.251	9.6171	0.142
		7	-0.268	-0.084	12.555	0.084
		8	0.230	-0.039	14.828	0.063
		9	-0.186	-0.151	16.382	0.059
		10	-0.124	-0.079	17.115	0.072
		11	0.024	-0.121	17.143	0.104
		12	-0.079	0.040	17.476	0.133

*Probabilities may not be valid for this equation specification.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.047	0.047	0.0715	0.789
		2	-0.024	-0.026	0.0900	0.956
		3	-0.041	-0.038	0.1472	0.986
		4	0.299	0.303	3.3568	0.500
		5	0.149	0.127	4.1891	0.523
		6	-0.185	-0.209	5.5232	0.479
		7	-0.090	-0.054	5.8533	0.557
		8	-0.097	-0.188	6.2557	0.619
		9	-0.125	-0.261	6.9538	0.642
		10	-0.164	-0.090	8.2236	0.607
		11	-0.011	0.104	8.2297	0.693
		12	0.021	0.115	8.2521	0.765



b) Dependent variable “Income Top 1%”

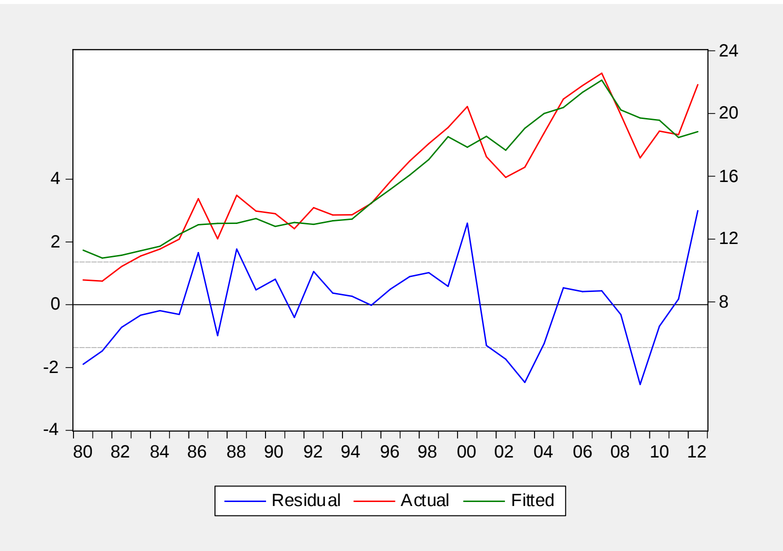
Static Equation

Dependent Variable: TOP1
Method: Least Squares
Date: 06/19/17 Time: 18:29
Sample: 1980 2012
Included observations: 33
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.336558	4.735423	-0.282247	0.7798
CREDIT	0.094408	0.014221	6.638589	0.0000
CAPFORM	0.157437	0.184826	0.851808	0.4013
PTFINV	-0.093167	0.290925	-0.320246	0.7511

R-squared	0.881170	Mean dependent var	15.77394
Adjusted R-squared	0.868877	S.D. dependent var	3.767982
S.E. of regression	1.364421	Akaike info criterion	3.572550
Sum squared resid	53.98768	Schwarz criterion	3.753944
Log likelihood	-54.94707	Hannan-Quinn criter.	3.633583
F-statistic	71.68185	Durbin-Watson stat	1.264871
Prob(F-statistic)	0.000000	Wald F-statistic	71.48428
Prob(Wald F-statistic)	0.000000		

Residuals of the static equation (ECM) and ADF test on the residuals



Null Hypothesis: RESID_TOP1 has a unit root
Exogenous: None
Lag Length: 2 (Automatic - based on SIC, maxlag=8)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.838860	0.0004
Test critical values:	1% level	-2.644302	
	5% level	-1.952473	
	10% level	-1.610211	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESID_TOP1)
Method: Least Squares
Date: 06/20/17 Time: 15:42
Sample (adjusted): 1983 2012
Included observations: 30 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID_TOP1(-1)	-0.989957	0.257878	-3.838860	0.0007
D(RESID_TOP1(-1))	0.279860	0.230030	1.216627	0.2343
D(RESID_TOP1(-2))	0.452953	0.185658	2.439718	0.0215

t-statistic equal of -3,83, we fail to reject the null of unit root test, hence the reverse co-integration relationship by applying EG two step methodology is not confirmed. We stop here, Dynamic Equation cannot be estimated, because there was found no cointegration. Since, a cointegration relationship was found when the same regression was made with “Credit” as the dependent variable, we would expect to confirm the cointegration also when another variable is the dependent variable.

But, according to Asteriou (2007) one of the main drawbacks of the EG two steps procedure, is the order of the variables. When estimating the long –run relationship we have to place one variable on the left side and use the others as regressors. The test does not say anything about which of the variables can be used as regressors and why. For example, in case of only two variables, Income Top1 and Level of Credit, we can either regress Credit on Income or choose the reverse and regress Income on Credit. Asymptotically, when the sample size goes to infinity, the test for co-integration on the residuals of these two variables is the same. But this is not our case, we found that the series

$$Credit_t = \propto\ + \beta_1 Top1_t + \beta_2 CapForm_t + \beta_3 PtfInv_t + \ \varepsilon_t$$

exhibit co-integration, while series

$$Top1_t = \propto\ + \beta_1 Credit_t + \beta_2 CapForm_t + \beta_3 PtfInv_t + \ \varepsilon_t$$

does not. As a result, the Johansen test for Co-Integration is proposed.

JOHANSEN CO-INTEGRATION TEST

Step 1: test the order of co-integration of all the variables, which is already performed above, all variables are I (1).

Step 2: Set the appropriate lag length for the model. This is done by estimating a VAR model and using the Lag Length Criteria option. Lag 3 is selected.

VAR Lag Order Selection Criteria
Endogenous variables: CREDIT TOP1 CAPFORM PTFINV
Exogenous variables: C
Date: 06/20/17 Time: 20:05
Sample: 1980 2012
Included observations: 29

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-286.7403	NA	6000.849	20.05106	20.23965	20.11012
1	-206.1392	133.4086	70.68391	15.59581	16.53877	15.89113
2	-174.5816	43.52778	25.94601	14.52287	16.22020	15.05445
3	-145.9615	31.58079*	13.13297*	13.65252	16.10422*	14.42036
4	-124.3390	17.89448	13.44403	13.26476*	16.47083	14.26886*

Step 3: Determining the number of co-integration equations. The rule is that for an equation of n-variables, at most (n-1) co-integration relation can exists. In our case, 2 co-integration equations are found.

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.644808	61.10723	47.85613	0.0018
At most 1 *	0.506423	31.08944	29.79707	0.0353
At most 2	0.273971	10.61321	15.49471	0.2364
At most 3	0.044774	1.328408	3.841466	0.2491

Trace test indicates 2 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Step 4: Vector Error Correction Model, VECM

Included observations: 29 after adjustments
Standard errors in ( ) & t-statistics in [ ]

Cointegrating Eq:	CointEq1	CointEq2
CREDIT(-1)	1.000000	0.000000
TOP1(-1)	0.000000	1.000000
CAPFORM(-1)	8.127399	0.856046
	(2.13627)	(0.26583)
	[3.80448]	[3.22034]
PTFINV(-1)	22.94869	1.954825
	(1.39107)	(0.17310)
	[16.4971]	[11.2932]
C	-274.1893	-30.52415

Since Johansen Co-Integration test, estimated 2 cointegrations equations, the table above shows both of them. The VECM automatically converts the variables into the first difference (table below), while with EG two step procedure, we have to put manually the Dynamic Equation in the differences of the variables.

The values inside the red box are the Error Correction Terms, while since in this paper we are interested about the causality between Income and Level of Credit, the equations regarding the dependent variables D(credit) and D(income) will be analyzed further (green box).

Error Correction:	D(CREDIT)	D(TOP1)	D(CAPFORM)	D(PTFINV)
CointEq1	-0.267374	0.092019	0.095146	-0.104391
	(0.39476)	(0.07654)	(0.02709)	(0.07206)
	[-0.67732]	[1.20230]	[3.51175]	[-1.44866]
CointEq2	1.016545	-1.406497	-0.790701	0.513226
	(3.43119)	(0.66524)	(0.23550)	(0.62634)
	[0.29627]	[-2.11426]	[-3.35759]	[0.81940]
D(CREDIT(-1))	-0.548408	0.030488	-0.029153	-0.011168
	(0.36485)	(0.07074)	(0.02504)	(0.06660)
	[-1.50311]	[0.43100]	[-1.16420]	[-0.16769]
D(CREDIT(-2))	0.000889	0.095861	0.007299	0.090291
	(0.30759)	(0.05964)	(0.02111)	(0.05615)
	[0.00289]	[1.60743]	[0.34572]	[1.60806]
D(CREDIT(-3))	-0.065050	0.035322	-0.014075	0.044352
	(0.30166)	(0.05849)	(0.02070)	(0.05507)
	[-0.21564]	[0.60393]	[-0.67981]	[0.80542]
D(TOP1(-1))	-1.007160	0.176326	0.563218	-0.427281
	(2.94204)	(0.57041)	(0.20192)	(0.53705)
	[-0.34233]	[0.30912]	[2.78925]	[-0.79560]
D(TOP1(-2))	-3.489850	-0.092644	0.219372	-0.042610
	(2.13053)	(0.41307)	(0.14623)	(0.38892)
	[-1.63802]	[-0.22428]	[1.50021]	[-0.10956]
D(TOP1(-3))	-2.275069	-0.251482	0.065654	-0.200846
	(1.54275)	(0.29911)	(0.10589)	(0.28162)
	[-1.47469]	[-0.84077]	[0.62005]	[-0.71318]
D(CAPFORM(-1))	11.06561	0.703132	0.155301	0.189039
	(4.48519)	(0.86960)	(0.30784)	(0.81875)
	[2.46714]	[0.80857]	[0.50449]	[0.23089]
D(CAPFORM(-2))	5.624269	1.876402	-0.145196	-0.481678
	(3.67146)	(0.71183)	(0.25199)	(0.67020)
	[1.53189]	[2.63603]	[-0.57620]	[-0.71870]
D(CAPFORM(-3))	3.624661	-0.154213	0.161117	0.138686
	(3.46165)	(0.67115)	(0.23759)	(0.63190)
	[1.04709]	[-0.22977]	[0.67814]	[0.21947]
D(PTFINV(-1))	5.700582	0.931882	-0.373072	0.446464
	(3.43157)	(0.66532)	(0.23552)	(0.62641)
	[1.66122]	[1.40066]	[-1.58402]	[0.71273]
D(PTFINV(-2))	1.432886	0.778211	-0.294711	0.463938
	(2.65220)	(0.51421)	(0.18203)	(0.48414)
	[0.54026]	[1.51340]	[-1.61901]	[0.95826]
D(PTFINV(-3))	-1.979804	0.476937	-0.072243	0.097777
	(1.98480)	(0.38482)	(0.13623)	(0.36231)
	[-0.99748]	[1.23939]	[-0.53032]	[0.26987]
C	10.45998	0.438970	-0.289980	-0.298099
	(3.57358)	(0.69285)	(0.24527)	(0.65234)
	[2.92703]	[0.63357]	[-1.18229]	[-0.45697]

R-squared	0.748319	0.811250	0.881568	0.831832
Adj. R-squared	0.496638	0.622499	0.763136	0.663664
Sum sq. resids	403.4191	15.16454	1.900366	13.44293
S.E. equation	5.368021	1.040760	0.368430	0.979903
F-statistic	2.973280	4.298000	7.443661	4.946443
Log likelihood	-79.32308	-31.74834	-1.633101	-30.00100
Akaike AIC	6.505040	3.224024	1.147110	3.103517
Schwarz SC	7.212262	3.931246	1.854332	3.810739
Mean dependent	2.862928	0.376207	-0.117859	-0.104696
S.D. dependent	7.566130	1.693915	0.757016	1.689649

Determinant resid covariance (dof adj.)	1.652886
Determinant resid covariance	0.089777
Log likelihood	-129.6456
Akaike information criterion	13.63073
Schwarz criterion	16.83681

The test does not report the p-values for the coefficients, only the t-statistics in () and the standard deviations in [ ].

CASE 1: Level of domestic credit as the Dependent Variable (testing for the causality that goes form Income Top1, Capital formation and Portfolio Investment to Credit)

$$\Delta Credit_t = \sum_{k=1}^{q} \phi_k \,\Delta Credit_{\,t-k} + \sum_{k=0}^{q} \phi_{1k} \,\Delta Top1_{\,t-k} + \sum_{k=0}^{q} \phi_{2k} \,\Delta CapForm_{\,t-k} + \sum_{k=0}^{q} \phi_{3k} \,\Delta PtfInv_{\,t-k} + \alpha_1 ECT_{t-1} + \nu_t$$

Included observations: 29 after adjustments
D(CREDIT)=C(1)*(CREDIT(-1) + 8.1273*CAPFORM(-1) + 22.9486*PTFINV(
-1) - 274.1892) + C(2)*(TOP1(-1) + 0.8560*CAPFORM(-1) + 1.9548
*PTFINV(-1) - 30.5241) + C(3)*D(CREDIT(-1)) + C(4)*D(CREDIT(-2)) +
C(5)*D(CREDIT(-3)) + C(6)*D(TOP1(-1)) + C(7)*D(TOP1(-2)) + C(8)
*D(TOP1(-3)) + C(9)*D(CAPFORM(-1)) + C(10)*D(CAPFORM(-2)) +
C(11)*D(CAPFORM(-3)) + C(12)*D(PTFINV(-1)) + C(13)*D(PTFINV(-2))
+ C(14)*D(PTFINV(-3)) + C(15)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.267359	0.394758	-0.677273	0.5093
C(2)	1.016389	3.431238	0.296217	0.7714
C(3)	-0.548417	0.364850	-1.503129	0.1550
C(4)	0.000890	0.307592	0.002892	0.9977
C(5)	-0.065041	0.301667	-0.215606	0.8324
C(6)	-1.007055	2.942056	-0.342296	0.7372
C(7)	-3.489780	2.130511	-1.638001	0.1237
C(8)	-2.275028	1.542725	-1.474682	0.1624
C(9)	11.06554	4.485233	2.467104	0.0271
C(10)	5.624258	3.671454	1.531889	0.1478
C(11)	3.624727	3.461637	1.047113	0.3128
C(12)	5.700529	3.431560	1.661206	0.1189
C(13)	1.432866	2.652191	0.540258	0.5975
C(14)	-1.979803	1.984797	-0.997484	0.3355
C(15)	10.46031	3.574730	2.926181	0.0111

R-squared	0.748319	Mean dependent var	2.862928
Adjusted R-squared	0.496638	S.D. dependent var	7.566130
S.E. of regression	5.368020	Akaike info criterion	6.505040
Sum squared resid	403.4189	Schwarz criterion	7.212262
Log likelihood	-79.32307	Hannan-Quinn criter.	6.726533
F-statistic	2.973281	Durbin-Watson stat	2.130279
Prob(F-statistic)	0.025179		

Where C(1) is the Error Correction Term, ECT. In this case it does satisfies the Stability Condition, because is negative and between -2 and 0, but the p-value is higher that both 5 and 10 percent. Hence, there is no long run causality.

Estimated ECT, Long Run Causality from Income Top1, CapForm and PtfInv to Credit

	Coefficient	P-Value
EG Two Step Procedure	-0,22	0,15
VECM, Johansen Test	-0,26	0,50

Short Run causality, from Income Top1 to Credit, we run the WALD Test for the join significance of the Top1 coefficients (c(6)=c(7)=c(8)=0), if the null hypothesis will be accepted, then we say that they are jointly zero, so there is no short-run causality running from Income Concentration among the richest one and Level of Credit.

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	2.472908	(3, 14)	0.1044
Chi-square	7.418724	3	0.0597

Null Hypothesis: C(6)=C(7)=C(8)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(6)	-1.007055	2.942056
C(7)	-3.489780	2.130511
C(8)	-2.275028	1.542725

Chi- square P-value equal of 0,0597, we reject null hypothesis at 10% critical value, meaning that there is a short run causality going from Income Top1, Capital Formation and Portfolio Investment.

Step 5: VECM model diagnostic checking

R-squared	0.748319	Mean dependent var	2.862928
Adjusted R-squared	0.496638	S.D. dependent var	7.566130
S.E. of regression	5.368020	Akaike info criterion	6.505040
Sum squared resid	403.4189	Schwarz criterion	7.212262
Log likelihood	-79.32307	Hannan-Quinn criter.	6.726533
F-statistic	2.973281	Durbin-Watson stat	2.130279
Prob(F-statistic)	0.025179		

R squared is 74% which is a good level of significance, same for F-statistics that have a p-value less than 5%, meaning that our data are fitting well. Durbin Watson test is very close to two, which means that there is no serial correlation among the residuals.

Then the test on the Residuals is performed,

a) Serial Correlation test, Null Hypothesis: there is no serial correlation, p-value is greater than both 5 and 10 percent significance value, which means that we fail to reject the null of NO SERIAL CORRELATION of the residuals. This is a very satisfactory result.

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.177340	Prob. F(2,12)	0.8397
Obs*R-squared	0.832534	Prob. Chi-Square(2)	0.6595

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 06/20/17 Time: 21:38
Sample: 1984 2012
Included observations: 29
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.082435	0.445412	-0.185075	0.8563
C(2)	1.211879	4.181165	0.289842	0.7769
C(3)	0.296083	0.630983	0.469242	0.6473
C(4)	-0.021869	0.376898	-0.058025	0.9547
C(5)	-0.086556	0.357010	-0.242448	0.8125
C(6)	-0.571328	3.279547	-0.174210	0.8646
C(7)	-0.250031	2.306478	-0.108404	0.9155
C(8)	0.552280	1.888194	0.292491	0.7749
C(9)	0.200694	4.810241	0.041722	0.9674
C(10)	-2.683645	5.965254	-0.449879	0.6608
C(11)	-1.059732	4.403130	-0.240677	0.8139
C(12)	-0.589382	3.910861	-0.150704	0.8827
C(13)	-1.031556	3.375726	-0.305581	0.7652
C(14)	-0.112670	2.236418	-0.050380	0.9606
C(15)	-1.219202	4.468306	-0.272856	0.7896
RESID(-1)	-0.359856	0.619633	-0.580756	0.5722
RESID(-2)	0.183784	0.506996	0.362495	0.7233

b) Heteroscedasticity test, null hypothesis of no heteroscedasticity is accepted.

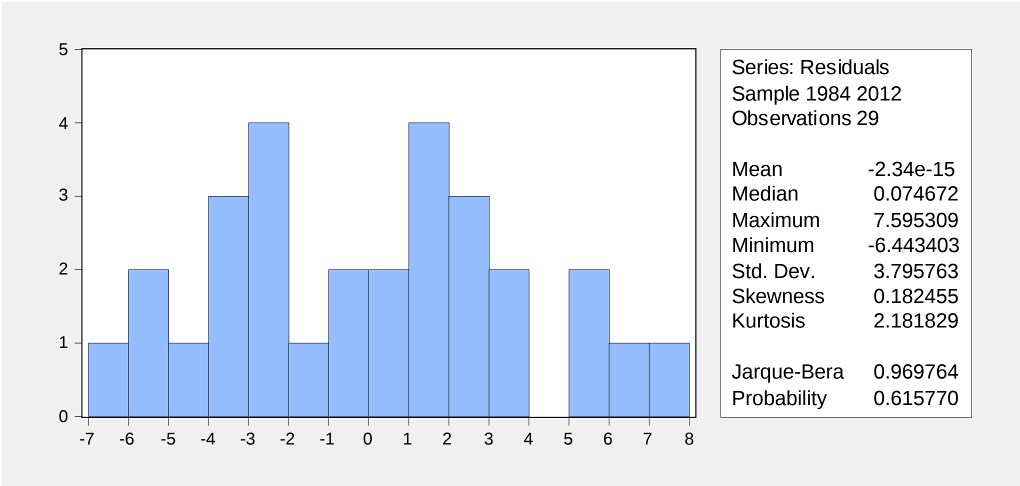
Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.681071	Prob. F(16,12)	0.7669
Obs*R-squared	13.80159	Prob. Chi-Square(16)	0.6135
Scaled explained SS	1.900702	Prob. Chi-Square(16)	1.0000

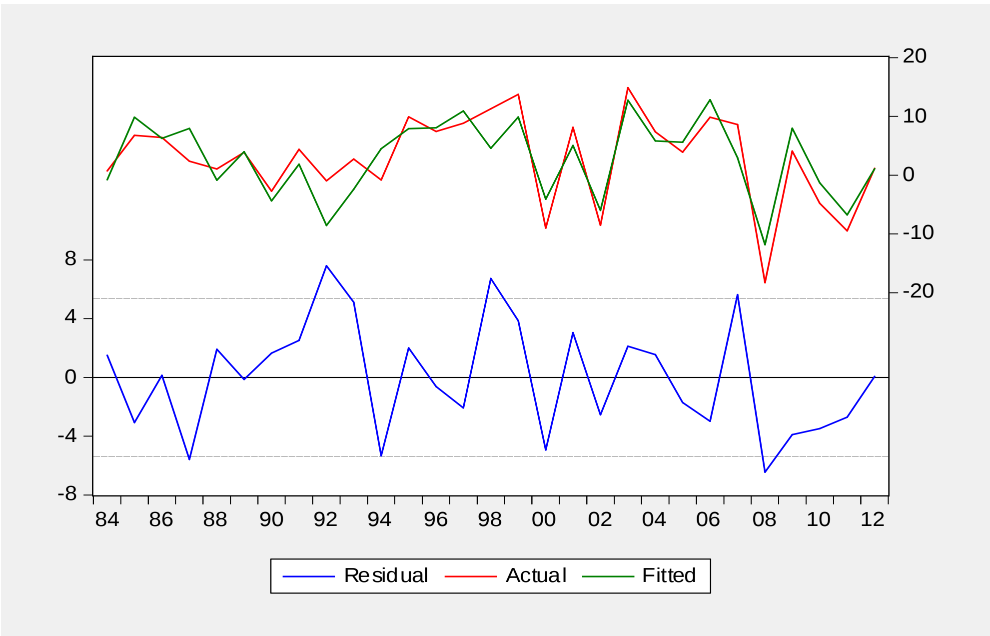
Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 06/20/17 Time: 21:43
Sample: 1984 2012
Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	145.3295	150.9303	0.962892	0.3546
CREDIT(-1)	1.021063	0.891477	1.145361	0.2744
CAPFORM(-1)	-24.47933	13.29885	-1.840710	0.0905
PTFINV(-1)	-2.966741	4.593373	-0.645874	0.5305
TOP1(-1)	2.192724	5.557088	0.394581	0.7001
CREDIT(-2)	0.218142	1.253679	0.174001	0.8648
CREDIT(-3)	0.501083	1.118301	0.448076	0.6621
CREDIT(-4)	-1.263852	1.022213	-1.236389	0.2400
TOP1(-2)	0.032076	4.932711	0.006503	0.9949
TOP1(-3)	-3.423958	5.176721	-0.661414	0.5208
TOP1(-4)	-3.437673	5.010524	-0.686090	0.5057
CAPFORM(-2)	14.27570	18.09849	0.788778	0.4455
CAPFORM(-3)	5.577783	15.19073	0.367183	0.7199
CAPFORM(-4)	-1.671594	11.00587	-0.151882	0.8818
PTFINV(-2)	4.819344	4.771405	1.010047	0.3324
PTFINV(-3)	1.408134	5.103888	0.275894	0.7873
PTFINV(-4)	0.664006	6.385951	0.103979	0.9189

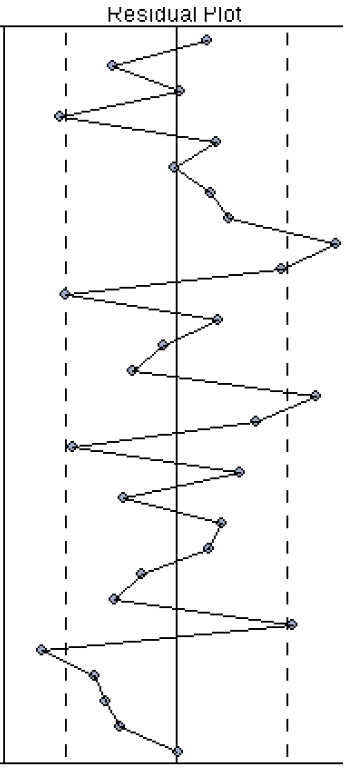
c) Check for Normality of the Residuals, p-value of Jarque-Bera test is 61% which is more than 5%.



Short - Run Curve



Actual	Fitted	Residual
0.72130	-0.77958	1.50088
6.77734	9.85317	-3.07583
6.40951	6.26117	0.14834
2.36301	7.94092	-5.57791
1.06827	-0.84517	1.91344
3.86479	3.99902	-0.13423
-2.72947	-4.38643	1.65695
4.39476	1.87394	2.52082
-0.97037	-8.56568	7.59531
2.73983	-2.37761	5.11743
-0.81076	4.51790	-5.32866
9.92651	7.91447	2.01204
7.44980	8.07215	-0.62235
8.84715	10.9182	-2.07101
11.2945	4.56084	6.73367
13.7439	9.88495	3.85897
-9.00569	-4.07629	-4.92940
8.12236	5.06981	3.05254
-8.51952	-5.98479	-2.53473
14.8763	12.7552	2.12111
7.37183	5.81332	1.55850
3.91119	5.61426	-1.70307
9.85961	12.8466	-2.98701
8.59566	2.97075	5.62491
-18.2787	-11.8353	-6.44340
4.10920	8.00096	-3.89176
-4.78302	-1.30028	-3.48274
-9.47901	-6.77152	-2.70749
1.15463	1.07996	0.07467



CONCLUSION: overall our VECM model is correct hence we accept the conclusion about the causality among the variables that:

- There is no long run causality that goes from Income Concentration to the Level of Credit
- But, there exist a short – run causality, we say that in short-run the level of income inequality does cause the level of domestic credit to the private sector.

CASE 2: Income Concentration Top 1% as the dependent variable (testing for the causality that goes form Credit, Capital formation and Portfolio Investment to Income Concentration 1%)

$$\Delta Top1_t = \sum_{k=1}^{q} \phi_{5k}\, \Delta Top1\,_{t-k} + \sum_{k=0}^{q} \phi_{6k}\, \Delta Credit\,_{t-k} + \sum_{k=0}^{q} \phi_{7k}\, \Delta CapForm\,_{t-k} + \sum_{k=0}^{q} \phi_{8k}\, \Delta PtfInv\,_{t-k} + \alpha_2 ECT_{t-1} + \vartheta_t$$

Included observations: 29 after adjustments
D(TOP1)=C(16)*(CREDIT(-1) + 8.1273*CAPFORM(-1) + 22.9486*PTFINV(-1) - 274.1892) + C(17)*(TOP1(-1) + 0.8560*CAPFORM(-1) + 1.9548*PTFINV(-1) - 30.5241) + C(18)*D(CREDIT(-1)) + C(19)*D(CREDIT(-2)) + C(20)*D(CREDIT(-3)) + C(21)*D(TOP1(-1)) + C(22)*D(TOP1(-2)) + C(23)*D(TOP1(-3)) + C(24)*D(CAPFORM(-1)) + C(25)*D(CAPFORM(-2)) + C(26)*D(CAPFORM(-3)) + C(27)*D(PTFINV(-1)) + C(28)*D(PTFINV(-2)) + C(29)*D(PTFINV(-3)) + C(30)

	Coefficient	Std. Error	t-Statistic	Prob.
C(16)	0.092024	0.076536	1.202369	0.2492
C(17)	-1.406555	0.665247	-2.114333	0.0529
C(18)	0.030483	0.070737	0.430940	0.6731
C(19)	0.095865	0.059636	1.607510	0.1303
C(20)	0.035327	0.058487	0.604007	0.5555
C(21)	0.176359	0.570405	0.309182	0.7617
C(22)	-0.092641	0.413063	-0.224277	0.8258
C(23)	-0.251495	0.299103	-0.840832	0.4146
C(24)	0.703065	0.869596	0.808496	0.4323
C(25)	1.876413	0.711820	2.636077	0.0196
C(26)	-0.154216	0.671141	-0.229782	0.8216
C(27)	0.931871	0.665310	1.400657	0.1831
C(28)	0.778219	0.514206	1.513438	0.1524
C(29)	0.476954	0.384812	1.239448	0.2356
C(30)	0.437858	0.693068	0.631769	0.5377

R-squared	0.811253	Mean dependent var	0.376207
Adjusted R-squared	0.622506	S.D. dependent var	1.693915
S.E. of regression	1.040750	Akaike info criterion	3.224005
Sum squared resid	15.16425	Schwarz criterion	3.931227
Log likelihood	-31.74807	Hannan-Quinn criter.	3.445498
F-statistic	4.298100	Durbin-Watson stat	1.727180
Prob(F-statistic)	0.005007		

Where C(16) is the Error Correction Term, ECT. In this case it does not satisfies neither the Stability Condition, nor it is significant. Hence, there is no long run causality going from Level of Credit to Income Concentration.

Then, for short –run causality is tested using WALD Test, where in this case the null hypothesis is:

$$H_0: c(18)=c(19)=c(20)=0$$

Wald Test:
Equation: VECM_TOP1

Test Statistic	Value	df	Probability
F-statistic	0.869002	(3, 14)	0.4803
Chi-square	2.607005	3	0.4563

Null Hypothesis: C(18)=C(19)=C(20)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(18)	0.030483	0.070737
C(19)	0.095865	0.059636
C(20)	0.035327	0.058487

Restrictions are linear in coefficients.

P- value is quite high, which means that we cannot reject the null hypothesis of jointly all the past values of credit are non-significant.

Conclusion: no causality running from the level of domestic credit to the private sector to the level of income inequality was found, neither in long run nor in short run.

As in the first case, the last step is that of checking the goodness of the VCEM model when the dependent variable is the Income Top 1% Concentration, results are reported below:

R-squared	0.811253	Mean dependent var	0.376207
Adjusted R-squared	0.622506	S.D. dependent var	1.693915
S.E. of regression	1.040750	Akaike info criterion	3.224005
Sum squared resid	15.16425	Schwarz criterion	3.931227
Log likelihood	-31.74807	Hannan-Quinn criter.	3.445498
F-statistic	4.298100	Durbin-Watson stat	1.727180
Prob(F-statistic)	0.005007		

And test of residuals:

a) Serial Correlation test, Null Hypothesis: there is no serial correlation

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.548223	Prob. F(2,12)	0.2523
Obs*R-squared	5.948218	Prob. Chi-Square(2)	0.0511

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 06/20/17 Time: 22:39
Sample: 1984 2012
Included observations: 29
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(16)	0.033244	0.101067	0.328935	0.7479
C(17)	-0.409823	1.018085	-0.402543	0.6944
C(18)	-0.058411	0.116371	-0.501940	0.6248
C(19)	0.036524	0.069028	0.529127	0.6064
C(20)	0.032893	0.061504	0.534809	0.6026
C(21)	0.007491	0.712500	0.010514	0.9918
C(22)	0.161075	0.478539	0.336596	0.7422
C(23)	-0.020398	0.294014	-0.069378	0.9458
C(24)	-0.061672	0.843743	-0.073093	0.9429
C(25)	0.501663	0.878077	0.571320	0.5783
C(26)	-0.079857	0.909317	-0.087821	0.9315
C(27)	-0.008672	0.640775	-0.013534	0.9894
C(28)	0.175267	0.536969	0.326402	0.7497
C(29)	-0.065472	0.382927	-0.170978	0.8671
C(30)	-0.024797	0.738144	-0.033594	0.9738
RESID(-1)	0.657571	0.622983	1.055521	0.3120
RESID(-2)	-0.649173	0.439820	-1.475995	0.1657

b) Heteroscedasticity test, null hypothesis of no heteroscedasticity

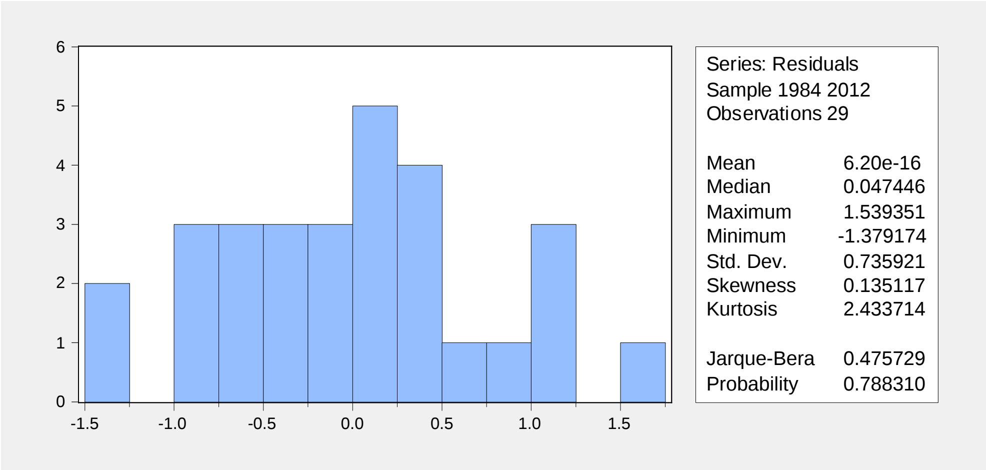
Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.395036	Prob. F(16,12)	0.9575
Obs*R-squared	10.00496	Prob. Chi-Square(16)	0.8664
Scaled explained SS	1.671505	Prob. Chi-Square(16)	1.0000

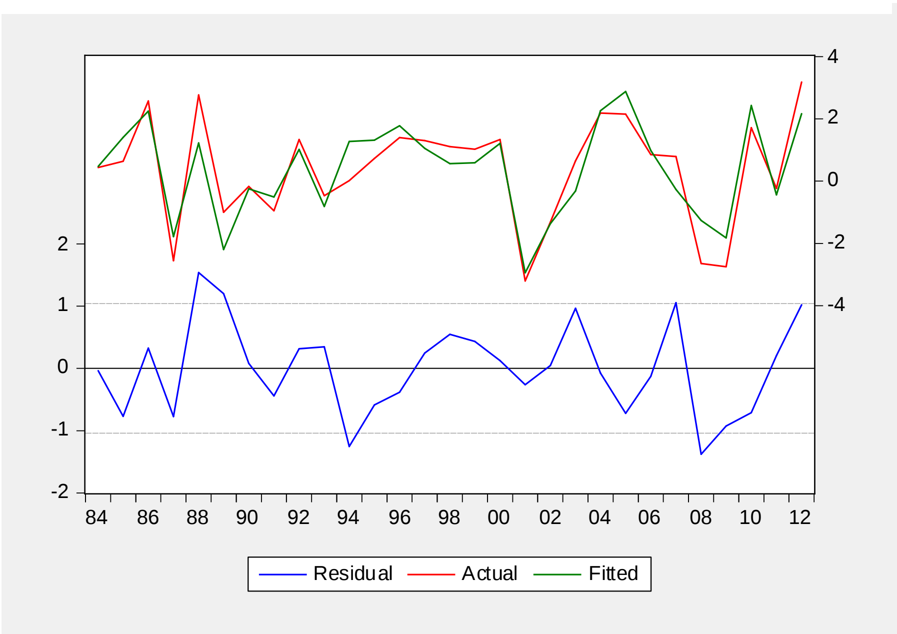
Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 06/20/17 Time: 22:41
Sample: 1984 2012
Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.037821	6.985797	0.578004	0.5739
CREDIT(-1)	0.009068	0.041262	0.219769	0.8297
CAPFORM(-1)	-0.055994	0.615536	-0.090968	0.9290
PTFINV(-1)	-0.167111	0.212604	-0.786020	0.4471
TOP1(-1)	-0.069103	0.257209	-0.268665	0.7928
CREDIT(-2)	-0.020257	0.058026	-0.349097	0.7331
CREDIT(-3)	-0.038727	0.051760	-0.748198	0.4688
CREDIT(-4)	0.012561	0.047313	0.265482	0.7951
TOP1(-2)	0.263405	0.228310	1.153717	0.2711
TOP1(-3)	0.208006	0.239604	0.868124	0.4024
TOP1(-4)	-0.195966	0.231912	-0.845002	0.4146
CAPFORM(-2)	-0.206623	0.837687	-0.246658	0.8093
CAPFORM(-3)	-0.169005	0.703102	-0.240370	0.8141
CAPFORM(-4)	0.312383	0.509406	0.613231	0.5512
PTFINV(-2)	-0.188580	0.220844	-0.853903	0.4099
PTFINV(-3)	0.020454	0.236233	0.086582	0.9324
PTFINV(-4)	-0.150877	0.295573	-0.510456	0.6190

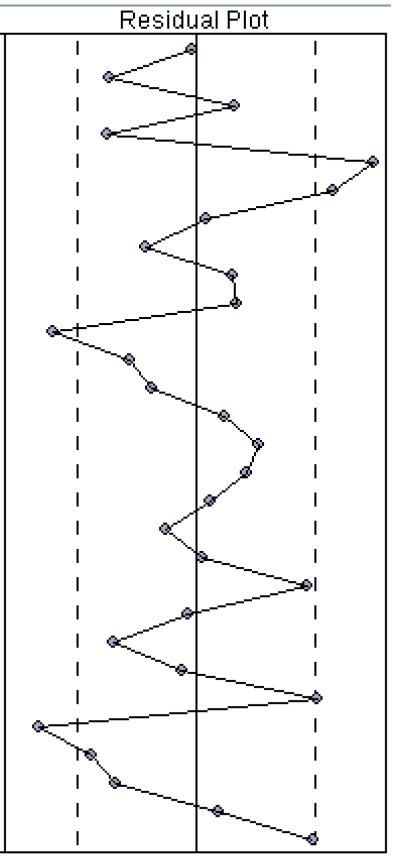
c) check of normality, Jarque Bera Test



Short - Run Curve



Actual	Fitted	Residual
0.43500	0.47376	-0.03876
0.63600	1.40691	-0.77091
2.57500	2.25031	0.32469
-2.55500	-1.78064	-0.77436
2.76700	1.22765	1.53935
-0.99900	-2.19893	1.19993
-0.17000	-0.25012	0.08012
-0.95200	-0.50910	-0.44290
1.33400	1.01852	0.31548
-0.46600	-0.81149	0.34549
0.01700	1.27222	-1.25522
0.73100	1.31647	-0.58547
1.39600	1.77897	-0.38298
1.30300	1.05500	0.24800
1.10900	0.56132	0.54768
1.02500	0.59238	0.43262
1.33700	1.21291	0.12409
-3.20300	-2.94076	-0.26224
-1.31700	-1.36445	0.04745
0.65100	-0.31603	0.96703
2.18800	2.26340	-0.07540
2.15500	2.87597	-0.72097
0.85300	0.98036	-0.12736
0.78800	-0.26725	1.05525
-2.64400	-1.26483	-1.37917
-2.74700	-1.82336	-0.92364
1.71400	2.42539	-0.71139
-0.23400	-0.43827	0.20427
3.18300	2.16369	1.01931



CONCLUSION: overall our VECM model is correct hence we accept the conclusion about the causality among the variables that:

- There is no long run causality that goes from the Level of Credit to Income Concentration
- And, there is neither short – run causality, we say that in short-run the level of credit has no causality effect on the level of income inequality.

Estimated ECT, **Long Run Causality** from Level of Credit, CapForm and PtfInv to Income Inequality

	Coefficient	P-Value
EG Two Step Procedure*	-	-
VECM, Johansen Test	0,0332	0,7479

- Since in the first step of the EG procedure, the residuals of the Static Equation were not stationary, we rejected the hypothesis of co-integration and estimate a VAR model with stationary variables, and test for Granger Causality.

GRANGER CAUSALITY TEST

Sample: 1980 2012
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Prob.
D(CREDIT) does not Granger Cause D(TOP1)	29	1.69305	0.1976
D(TOP1) does not Granger Cause D(CREDIT)		0.45750	0.7147
D(CAPFORM) does not Granger Cause D(TOP1)	29	3.76824	0.0253
D(TOP1) does not Granger Cause D(CAPFORM)		2.23556	0.1125
D(PTFINV) does not Granger Cause D(TOP1)	29	1.15620	0.3488
D(TOP1) does not Granger Cause D(PTFINV)		0.38961	0.7616
D(CAPFORM) does not Granger Cause D(CREDIT)	29	2.43349	0.0920
D(CREDIT) does not Granger Cause D(CAPFORM)		4.86605	0.0096
D(PTFINV) does not Granger Cause D(CREDIT)	29	3.03329	0.0508
D(CREDIT) does not Granger Cause D(PTFINV)		4.87198	0.0095
D(PTFINV) does not Granger Cause D(CAPFORM)	29	0.57933	0.6348
D(CAPFORM) does not Granger Cause D(PTFINV)		1.81143	0.1745

STATIC GENERAL EQUATION ESTIMATE (REGRESSORS SELECTION)

Dependent Variable: CREDIT
Method: Least Squares
Date: 06/18/17 Time: 11:21
Sample: 1980 2012
Included observations: 33
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-885.7215	198.9529	-4.451915	0.0002
TOP1	0.279551	1.398433	0.199903	0.8432
CAPFORM	4.556765	1.902052	2.395710	0.0244
PTFINV	-0.618691	1.419165	-0.435954	0.6666
RIR	-0.427394	1.295956	-0.329790	0.7443
LGDP	84.92912	19.21978	4.418839	0.0002
REG	1.733423	2.993450	0.579072	0.5677
M2	0.522308	0.231133	2.259770	0.0328

R-squared	0.961173	Mean dependent var	143.2955
Adjusted R-squared	0.950301	S.D. dependent var	36.99042
S.E. of regression	8.246388	Akaike info criterion	7.264644
Sum squared resid	1700.073	Schwarz criterion	7.627434
Log likelihood	-111.8666	Hannan-Quinn criter.	7.386712
F-statistic	88.41058	Durbin-Watson stat	1.134349
Prob(F-statistic)	0.000000	Wald F-statistic	72.98114
Prob(Wald F-statistic)	0.000000		

Dependent Variable: CREDIT
Method: Least Squares
Date: 06/18/17 Time: 11:21
Sample: 1980 2012
Included observations: 33
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-911.8645	165.2394	-5.518445	0.0000
TOP1	0.048315	1.125239	0.042937	0.9661
CAPFORM	4.833985	1.556351	3.105974	0.0044
PTFINV	-0.535655	1.390161	-0.385319	0.7030
LGDP	88.70637	14.86338	5.968117	0.0000
M2	0.514147	0.212758	2.416583	0.0227

R-squared	0.960878	Mean dependent var	143.2955
Adjusted R-squared	0.953634	S.D. dependent var	36.99042
S.E. of regression	7.965094	Akaike info criterion	7.150980
Sum squared resid	1712.954	Schwarz criterion	7.423073
Log likelihood	-111.9912	Hannan-Quinn criter.	7.242531
F-statistic	132.6310	Durbin-Watson stat	1.160188
Prob(F-statistic)	0.000000	Wald F-statistic	91.12213
Prob(Wald F-statistic)	0.000000		

High p-values:
-might be due to the small sample size, 33 observations.
-removing the monetary explanatory variables, the remaining one do make a significant contribution to the model.
-Multicollinearity check: Variance Inflator Factor: it measures how much the variance (squared of estimated standard errors) of an estimated coefficient is increased because of collinearity.
Interpretation: the square root of VIF indicates how much larger the standard error is, if compared with what it would have been if it was uncorrelated with the other explanatory variables)
Variance Inflation Factors:

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	39582.26	28426.59	NA
TOP1	1.955614	406.3916	10.03385
REG	8.960741	574.1245	2.072010
CAPFORM	3.617803	1180.476	4.073893
PTFINV	2.014029	14.67255	4.243047
M2	0.053423	191.3203	1.867312
RIR	1.679503	30.07347	6.688554
LGDP	369.4001	28952.35	20.68908

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	901.3164	187.7849	NA
TOP1	0.964723	56.32554	5.187115
CAPFORM	1.744702	179.9678	1.032953
PTFINV	6.512727	11.06125	5.228967

VIF Interpretation: VIF=1, the variables are not correlated, 1<VIF<5, moderately correlated and VIF>5, highly correlated.

Dependent Variable: CREDIT
Method: Least Squares
Date: 06/18/17 Time: 11:23
Sample: 1980 2012
Included observations: 33
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	111.3499	30.02193	3.708950	0.0009
TOP1	7.104210	0.982203	7.232933	0.0000
CAPFORM	-4.123724	1.320872	-3.121971	0.0040
PTFINV	-4.522685	2.552004	-1.772209	0.0869

R-squared	0.907216	Mean dependent var	143.2955
Adjusted R-squared	0.897617	S.D. dependent var	36.99042
S.E. of regression	11.83594	Akaike info criterion	7.893372
Sum squared resid	4062.598	Schwarz criterion	8.074767
Log likelihood	-126.2406	Hannan-Quinn criter.	7.954406
F-statistic	94.51740	Durbin-Watson stat	1.781759
Prob(F-statistic)	0.000000	Wald F-statistic	140.3609
Prob(Wald F-statistic)	0.000000		