Long-run inequality persistence in the U.S., 1870–2019

Carlos Gayán-Navarro*[a] and Marcos Sanso-Navarro[a,b]

[a]Departamento de Análisis Económico, Universidad de Zaragoza

[b]Instituto Universitario de Investigación en Empleo, Sociedad Digital y Sostenibilidad (IEDIS), Universidad de Zaragoza

carlosgayannavarro@gmail.com, marcossn@unizar.es

Abstract

This paper studies the long-run persistence of inequality in the U.S. With this aim, both income and wealth inequality measures covering the period from 1870 to 2019 have been analyzed. The persistent character of inequality has been assessed using unit root and structural break test statistics for time series. Furthermore, the determinants of inequality persistence have been explored by implementing Bayesian model averaging techniques in a generalized linear model framework. Our results suggest that the wealth-to-income ratio displays a non-stationary behavior throughout the whole sample period. On the contrary, the Gini index of disposable income and the top 10% income share alternate between I(0) and I(1) regimes. We also find that, while a higher level of globalization increases the persistence of income inequality, it is inversely related to the levels of educational attainment and trade union membership.

JEL classification: C11, C12, C22, D63.

Keywords: Inequality, unit roots, structural breaks, Bayesian model averaging.

*Corresponding author. Address: Facultad de Economía y Empresa. Departamento de Análisis Económico. Gran Vía 2. 50005 Zaragoza (Spain).

1 Introduction

The interest in the evolution of inequality dates back to the seminal contribution of Kuznets (1955), who established an inverted U-shaped relationship with economic growth. More recently, Milanovic (2016) has argued that inequality is cyclical – influenced by economic, demographic, and political factors – and, as such, it may not experience unlimited growth. Piketty (2014) contends that inequality has been increasing in developed countries since the 1980s. This author also expects that this trend will continue during the next century, driven by a greater rate of return to assets as compared to GDP growth. Corroborating this prediction, theories aligned with the Great Gatsby Curve (Durlauf, Kourtellos, and Tan, 2022) suggest that initial levels of inequality are associated with persistent intergenerational income.

Together with a wider availability of long-run inequality data, a growing body of literature investigating the persistence of its time series has appeared. This is the case of Islam and Madsen (2015), who study the stationary character of the Gini index and the top 10% income share for a panel of 21 OECD countries during the period 1870–2011. Their main conclusion is that income inequality can be considered to be a stationary – or I(0) – process. Quite the opposite is found by Christopoulos and McAdam (2017), who were not able to reject the null hypothesis that the Gini coefficient is a unit root – or I(1) – process in a panel of 47 countries during 1975–2013. Similarly, and for a panel covering 60 countries from 1989 to 2015, Ghoshray, Monfort, and Ordóñez (2020) show that the Gini indexes for net and market income exhibit a unit root. In a sample of 34 countries covering the period 1960–2020, Makhlouf (2023) concludes that inequality trends have been relatively stable in developed countries, while both increasing or decreasing trends can be found in developing countries. These results lead this author to claim that national factors have exerted a stronger influence on inequality than global ones.

In a comprehensive study covering the period from 1870 to 2020 and a sample of 21 OECD countries, Solarin et al. (2022) implement both fractional and non-fractional integration techniques to show that shocks to inequality generally exhibit a persistent nature. With a similar aim, but adopting a different empirical approach, Roine and Waldenström (2011) test for the presence of structural breaks in the top income shares of a group of developed countries applying both time series and panel data methods. These authors find common breaks after the Second World War (WWII) and during the 1970s, as well as differential patterns across country groups. Using a Gaussian mixture autoregressive models, Kalliovirta and Malinen (2020) show that the breaks detected by Roine and Waldenström (2011) are, in fact, regime changes.

Ghoshray, Malki, and Ordóñez (2021) apply time series methods that control for the presence of structural breaks both under the null and the alternative hypotheses when testing for a unit root. For time spans close to a century, they conclude that both top income and wealth shares present high levels of persistence in groups of Anglo-Saxon, continental European, and Asian countries. This framework is similar to that implemented by Sanso-Navarro and Vera-Cabello (2020), who, for a sample of both developed and emerging economies during the years from 1960 to 2017, show that the Gini coefficient for disposable income displayed a persistent behavior. These authors also assess the existence of changes in persistence finding that, in most of the countries included in their sample, inequality alternates between stationary and non-stationary regimes. The main exceptions are South Africa and the United States (U.S.).

The U.S. not only presents one of the highest levels of inequality among developed countries (Alvaredo et al., 2018), but it has also been shown to exert a non-negligible influence on the inequality of other advanced economies (Kalliovirta and Malinen, 2020). In the present paper, we are extending the sample back to 1870 in order to check whether the results obtained by Sanso-Navarro and Vera-Cabello (2020) are influenced by the time span covered, as it excludes significant periods – such as the pre-industrial era and the two World Wars – during which inequality is known to have behaved differently. As another contribution, we consider both income and wealth inequality measures. In particular, we will evaluate the persistence of the Gini index for disposable income, that provides an overall indication of income inequality; the top 10% income share, which reflects income concentration among the highest earners; and the wealth-to-income ratio, that represents wealth concentration.

Our empirical analysis follows the approach adopted by Sanso-Navarro and Vera-Cabello (2020), based on the implementation of unit root tests that allow for the presence of structural breaks both under the null and the alternative hypotheses. Proceeding this way, we deal with the circular problem that arises when both features are present in the data (Perron, 2006). Subsequently, we examine the presence of changes between I(0) and I(1) regimes in those inequality measures that have an overall non-stationary behavior. Although Sanso-Navarro and Vera-Cabello (2020) state that the persistence of income inequality appears to be influenced by tax progressivity, income for top earners, and working conditions, they do not establish any statistical relationships regarding the drivers of inequality persistence. To delve deeper into this aspect, we draw on the existing studies about the determinants of inequality in order to examine which socio-economic variables are associated with the probability of inequality being in a stationary or a non-stationary regime. This strand of the literature – reviewed by Förster and Tóth (2015) and Nolan, Richiardi, and Valenzuela (2019) – provides a wide range of theories and mechanisms through which factors such as globalization, technological change, macroeconomic and financial conditions, as well as labor institutions and demographics, affect income inequality. Therefore, and as another contribution, we examine the variables that are related to the persistence (rather than the level) of inequality. This has been done by implementing Bayesian model averaging (BMA) techniques (Raftery, 1995; Raftery, Madigan, and Hoeting, 1997) in a generalized linear model (logistic regression) framework.

The rest of the paper is organized as follows. Section 2 motivates the inequality measures that have been studied, describes the main data sources, and carries out the time series analysis. Section 3 deals with the study of those socio-economic factors that display a robust relationship with the persistence of income inequality. Section 4 interprets and discusses the main results obtained in the empirical analysis and, finally, Section 5 concludes.

2 A time series analysis of income and wealth inequality in the U.S.

2.1 Inequality measures

The standard measure for the study of income inequality is the Gini coefficient which takes a value of one when all income is concentrated in one unit of reference, zero when income is equally distributed across units. This information for the U.S. has been extracted from the Standardized Income Inequality Database (SWIID; Solt, 2020) for the period that covers the years from 1913 to 2019. The Gini index provided by this dataset refers to disposable income, and takes households weighted by their respective number of members as the unit of reference. This information has been extended back to 1870 with the data compiled by Madsen, Minniti, and Venturini (2018) implementing the procedure proposed by Prados De La Escosura (2008). Despite the widespread use of the Gini index as a measure of income inequality, it is not exempt of problems. On the one hand, and given the aggregate character of the coefficient, it is difficult to determine whether a given value reflects a high or a low level of inequality. On the other hand, and due to the mathematical properties of this index, it tends to disguise the changes experienced in both the lower and the upper tails of the distribution of income (Alvaredo et al., 2018). In order to overcome these limitations – especially the second one – we have also considered the share of income of the top 10% of the distribution as an alternative a measure of income concentration. This information has been collected from the World Inequality Database (WID) for the period 1913–2019 and, similarly to the Gini coefficient, completed until 1870 using the data set created by Madsen, Minniti, and Venturini (2018).

Wealth inequality has been proxied using the wealth-to-income ratio. This measure of wealth inequality is based on the 'second law of capitalism' formulated by Piketty (2014), which considers that a higher value of this ratio reflects an increase in the gains of capital obtained by its owners, hence increasing wealth inequality. In other words, the wealth-to-income ratio can be understood as a measure of wealth accumulation. The time series of this variable for the period 1870–2019 has been constructed by combining the information provided by the WID and by Piketty and Zucman (2014).

[Insert Figure 1 about here]

Figure 1 shows the evolution of the three above-mentioned inequality measures. It can be observed that income inequality followed a steady decreasing trend from the end of the 19th century until the First World War (WWI), experiencing a sudden increase after this historical event. In contrast, the wealth-to-income ratio followed an overall increase, with some episodes of high volatility. During the interwar period, and while the top 10% income share displayed an increasing trend, the Gini coefficient fluctuated around constant values with a slight drop before WWII. After that moment, the three time series present a significant drop, being more moderate in the case of the Gini index due to its lower starting values. The wealth-to-income ratio displayed a staggered upward trend afterwards, alternating sharp increases with stable periods. This measure of wealth inequality also experienced a short drop in the 2000s, that reversed after the Great Recession. Income inequality almost remained stable until the 1980s, when pronounced upward trends emerged. They moderated in the 1990s, especially for the Gini index.

2.2 Unit roots and trend shifts

As a first attempt to assess the persistence of inequality, we have performed the battery of unit root tests[1] discussed in Ng and Perron (2001). On the one hand, they propose to improve the low power of univariate unit root tests using the generalized least squares (GLS) detrending procedure of Elliott, Rothenberg, and Stock (1996). On the other hand, Ng and Perron (2001) develop a modified information criterion (MIC) to better select the truncation lag that mitigates size distortions. Therefore, these authors conclude that the use of the MIC with GLS-detrended data bring about unit root tests with desirable size and power properties. Resulting test statistics for the three inequality measures described in the previous subsection are presented in Table 1. Although it can be observed that the unit root null cannot be rejected, this conclusion can be considered as preliminary. The reason is that the unit root tests that have been implemented do not consider the possible presence of changes in the deterministic component of the time series under scrutiny. As demonstrated by Perron (1989), level and/or trend shifts reduce the power of unit root tests and, given the length of the time series that have been analyzed, this is an issue that should not be discarded in the present context.

[Insert Table 1 about here]

The potential limitations of the tests implemented in Table 1 are determined by the interrelation between testing for unit roots and for the presence of structural breaks exposed in Perron (2006). If we are analyzing the order of integration of a time series and it displays shifts in its determistic component that are not taken into account, the unit root test will

[1] Instead of considering stationarity as the null hypothesis, such as the Kwiatkowski et al. (KPSS, 1992) test.

tend to fail in rejecting the null hypothesis. Alternatively, if we are looking for shifts in the deterministic component and the time series is non-stationary, the structural change test will tend to over-reject the null hypothesis of an absence of breaks. To address these issues we first take advantage of the trend shift test proposed by Perron and Yabu (2009). This test is based on a quasi-feasible GLS detrending approach, and is valid for both I(0) and I(1) processes. We have performed the test considering a shift in the intercept and in the slope parameters, and allowing for several structural breaks by implementing the sequential procedure proposed by Kejriwal and Perron (2010), that consists of testing for the null hypothesis of l breaks against the alternative of $l+1$ breaks. Under the null hypothesis, break dates are the global minimizers of the sum of squared residuals from an ordinary least squares estimation, and are obtained using the algorithm developed by Bai and Perron (2003).

[Insert Table 2 about here]

Resulting trend shift test statistics and estimated break dates are reported in Table 2. These figures show that the two measures of income inequality and the wealth-to-income ratio experience a trend shift around WWII. The test statistic detects further breaks for the Gini coefficient and the top 10% income share, being some of them coincident. In particular, both time series present a trend shift at the beginning of the 20th century, and in the early 1980s. Moreover, the Gini coefficient displays two breaks in 1965 and in 1993, and the top 10% income share another one in 2001. At this point, it is worth noting that we are not yet interested in the particular interpretation of the shifts detected. The reason is that, as explained by Kurozumi (2005) and as it will be studied in the next subsection, they may be related to changes in the order of integration of the series. Nonetheless, it is necessary to detect these breaks in order to control for their presence when testing for unit roots.

[Insert Table 3 about here]

The study of the non-stationary character of the income and wealth inequality measures accounting for the detected trend shifts has been carried out using the procedure developed by Carrion-i-Silvestre, Kim, and Perron (2009). These authors propose to extend the quasi-GLS detrending method of Elliott, Rothenberg, and Stock (1996) to allow the unit root tests implemented before to consider structural breaks both under the null and the alternative hypotheses. Resulting test statistics are presented in Table 3, only providing some weak evidence against the unit root null in the case of the Gini coefficient for disposable income. Therefore, it can be claimed that the inequality measures that have been analyzed are non-stationary and, as a consequence, have a persistent character. However, the Gini index and the top 10% income share are bounded variables. Although this is not the case of the wealth-to-income ratio in theory, it is implausible that this variable may display an explosive behavior in practice. For these reasons, it is difficult to consider that these inequality measures are I(1). In fact, and following the arguments put forward by Christopoulos and McAdam (2017), the unit root behavior of inequality measures might be reflecting that they are not adjusting to a long-run mean or, at most, doing it very slowly, hence showing high persistence.

2.3 Persistence changes

Similarly to Sanso-Navarro and Vera-Cabello (2020), and given the long sample period covered by the time series being studied, we can go a step further in the analysis of the measures of inequality. It is reasonable to check whether these variables behave as a unit root process from 1870 to 2019 or if they alternate between I(0) and I(1) regimes. This has been made using the persistence change test statistic developed by Leybourne, Kim, and Taylor (2007), which contrasts the null hypothesis that the time series is I(1) against the alternative hypothesis that it experiences, at least, one change to an I(0) regime. This testing procedure is based on doubly-recursive sequences of augmented Dickey and Fuller (ADF, 1979) unit root test statistics, and has the advantage of being valid when the time series have multiple changes in persistence. That is to say, the persistence change test can be implemented in a sequential manner in order to detect multiple breaks. In this way, when an I(0) regime is detected, the test is applied again to the remaining I(1) subperiods of the time series. The minimum time span allowed for these I(1) regimes is 20 years, as determined by the critical values calculated in Leybourne, Kim, and Taylor (2007).

[Insert Table 4 about here]

Table 4 reports the resuls obtained when the persistence change test is applied sequentially. These figures show that the Gini coefficient for disposable income is the indicator that has experienced a higher number of regime changes. This may be because this index is a measure of overall inequality, rather than an indicator of income concentration. The Gini index displays a persistent character from 1870 to 1907, when it turned into a stationary process. This first I(0) regime covers WWI, the Great Depression, and the beginning of WWII. A second short non-stationary regime began in 1944 and lasted five years. The upward change of income inequality experienced at the end of the 1970s as a consequence of policy shifts towards free market and tax cuts triggered a new I(1) regime that lasted until the 1990s. After that moment, the Gini coefficient for disposable income behaved again as a stationary process until the aftermath of the financial crisis that began in 2007.

The persistence change test detects a more limited number of regime shifts in the time series of the top 10% income share. Similarly to the Gini coefficient, a first stationary regime started in 1909, lasting until the beginning of WWII. However, the top 10% income share behaves as a non-stationary process until the late 1990s. Also in line with the other measure of income inequality, the top 10% is in an I(1) regime at the end of our sample period. Finally, the results reported in Table 4 shows that the wealth-to-income ratio does not suffer any regime change, what can be interpreted as an indication that wealth inequality is more persistent and difficult to mitigate than income inequality. We will discuss further these results in Section 4.

3 Searching for robust determinants of income inequality persistence

3.1 Data and variables

Apart from the larger time span analyzed in the present paper, as well as to the consideration of wealth inequality, we further extend the study carried out by Sanso-Navarro and Vera-Cabello (2020) by trying to disentangle the variables that determine the probability of being in a stationary or a non-stationary regime of income inequality. Despite we are interested in finding the factors that affect the persistence of inequality (instead of its level), a reasonable starting point would involve examining whether the factors that contribute to a higher level of income inequality are also associated with a greater persistence, and vice versa.

In this line of research we come across studies that consider a wide set of determinants such as Hailemariam, Sakutukwa, and Dzhumashev (2021), who apply panel vector autoregressions to a dataset comprising OECD countries. These authors conclude that real interest rates, government spending shares, and financial development indicators are inversely related to inequality. Another remarkable work is that of Furceri and Ostry (2019), who implement model averaging techniques and find that unemployment, technological change, and globalization are robust determinants of income inequality.

To provide a comprehensive overview of both the theoretical backgorund and empirical evidence, we draw upon the reviews of Förster and Tóth (2015) and, more specifically, Nolan, Richiardi, and Valenzuela (2019), as they focus on the determinants of income high-income countries. These authors divide the main driving forces behind income inequality, as well as their transmission channels, into several groups. Taking these references into account, we have encompassed them taking into account data availability without favoring any particular one. As a result, we have considered six groups of covariates, which are listed in Table 5, along their definitions, and data sources.

[Insert Table 5 about here]

The first group includes variables related to globalization, which is one of the most commonly studied driver of inequality. The way that this phenomenon may have affected inequality depends on the theoretical framework adopted. For example, in models *à la* Heckscher-Ohlin (Wood, 1995), trade is supposed to increase inequality in developed countries because they are more abundant in skilled labor and capital, hence increasing the wage gap. Predictions about capital flows go in the same line, as they should move towards the more productive sectors. This would increase the demand for skilled labour and capital gains both in developed and developing countries (Feenstra and Hanson, 1996). Trade volume has been proxied using the openness ratio (imports plus exports as a percentage of GDP), and capital flows have been measured by non-residential investment, which is also an indicator of financial globalization. As another dimension of globalization, the number of foreign patent applications tries to capture technological transfers.

The wage gap – a factor driving income inequality – is supposed to depend on technological change, our second set of determinants. It includes the stock of patents and the expenditure on research and development (R&D) as a percentage of GDP. Claessens and Perotti (2007) claim that the link between financial development and inequality is not clear. The development of the financial sector may facilitate the access to credit of the poorer share of population, enabling their ability to get their productive projects off the ground and, as a consequence, reducing inequality. However, if credit facilities increase in a context of asymmetric information, only the richer population will take advantage. Whatever the case may be, financial development has been proxied by the percentages over GDP of gross private savings and of the bank credit to the private sector. The role of financial institutions through this channel has been captured using the interest rate of long-term government bonds.

Without question, economic policy can also influence income inequality. On the one hand, monetary policy effectiveness has been captured using the inflation rate. On the other, we have included tax revenues and government expenditures as variables related to fiscal policy. Demographics and social structure have been proxied with life expectancy – reflecting the income gap between age groups and the economic weigth of pensions – and the schooling enrollment rate, which is expected to reduce the skill wage gap. In addition, it is worth highlighting that wages are related to the difference between the returns of labor and capital. Given that labor markets are not perfectly competitive, and that wages are determined through negotiations between firms and workers, labor institutions and regulations may play a crucial role in determining inequality. Therefore, the bargaining power of workers has been proxied using the unemployment rate and the share of trade union membership. Moreover, the latter might be a mitigating factor of the disparities between skilled and unskilled workers and, as shown by Farber et al. (2021), has been closely linked with inequality in the U.S.

3.2 Methodology and results

Given that we are giving the same plausibility to all the potential drivers of income inequality described above as well as their related theories, and due to the large number of covariates involved, we are applying Bayesian model averaging (BMA) techniques

(Raftery, 1995; Raftery, Madigan, and Hoeting, 1997) to assess which of them display a more robust relationship with the probability of being in an I(0) or an I(1) regime of the income inequality measures. BMA deals with variable selection, estimation, and inference simultaneously by enumerating the full model space and assigning a posterior inclusion probability (PIP) to each regressor. In fact, these inclusion probabilities are considered as one of the main advantages of these methods, see Steel (2020). BMA also permits to obtain the mean coefficients and their standard deviations for each covariate and, more importantly for our context, has been extended to be applied in a generalized linear model framework; see Raftery (1996) and Hoeting et al. (1999). This paradigm includes a logit estimation when the link function between the dependent variable and the regressors is binomial. Therefore, and according to the results in Table 5, we have created indicator variables that take a value of one if the Gini index for disposable income or the top 10% income share are in an I(1) regime, zero otherwise. Proceeding this way, we will be able to disentangle those factors that display a more robust relationship with income inequality persistence.

In order to be implemented, BMA methods require to control for the level of uncertainity about the importance of the variables and the value of their corresponding coefficients by selecting prior probability distributions for the model space and their related parameters (Olmo and Sanso-Navarro, 2021). In our case, we have established an uniform prior for the model space which assigns equal probability to all combinations of covariates and, following the recommendations in Ley and Steel (2012) and Li and Clyde (2018), we have elicited the hyper-g\n and benchmark priors for model-specific parameters. BMA has been specified in a logit regression framework using the Bayesian Adaptative Sampling (BAS) R package (Clyde, Ghosh, and Littman, 2011) which, as is our case, enumerates all models when there are less than nineteen regressors included.

Results for the Gini index and for the top 10% income share are shown, respectively, in Tables 6 and 7. It can be observed that the PIPs received by the covariates are high, especially when the Gini coefficient for disposable income is considered as the measure of inequality. In fact, these inclusion probabilities are higher than 0.47 in all cases. This can be interpreted as evidence of the suitability of the regressors considered as potential determinats of income inequality. Nonetheless, and as it has been pointed out by Li and Clyde (2018), it is worth noting that elicited priors tend to favor large models. For this reason, we include in the Appendix the PIPs obtained using alternative priors for model-specific parameters. Although the inclusion probabilities displayed in Tables A1 and A2 tend to be lower, the main conclusions drawn about the relative importance of the regressors do ont change significantly.

[Insert Tables 6 and 7 about here]

Our results suggest that globalization – in terms of investment and trade – favors both overall and cumulative income inequality to be more persistent. Technology transfers, captured by foreing patent applications, make the Gini coefficient (top 10% income share) to be less (more) persistent. We also find that the higher the stock of patents and R&D expenditures, the higher the probability of being in a non-stationary regime. This is also the case of gross private savings as a percentage of GDP, what might be an indication that people with higher incomes have a higher propensity to save, hence perpetuating income inequality. While there is an inverse relationship between bank credit to the private sector and the persistence of the concentration of income, the converse is true when inequality is measured using the Gini coefficient. This reflects that credit rationing mainly affects people with lower levels of income. In addition, a higher long-term interest rate of government bonds increases the probability of income inequality being in an I(0) regime.

The variables related to fiscal and monetary policy receive the lower PIPs. Inflation is expected to be harmful to lower incomes and/or higher labor income shares. For this reason, the inflation rate has a direct relationship with the persistence of the Gini coefficient of disposable income. Contrarily, higher inflation lessens the probability of being in an I(1) regime of the top 10% income share, reflecting the erosion of incomes in the upper tail of the distribution. Overall tax revenues as a percentage of GDP increases (decreases) the persistence of overall income inequality (income concentration). In this regard, it is worth noting that an alternative measure of tax progressivity may have lead to different results. However, data availability has forced us to use this general indicator. Similar results are found for government expenditure, which seems to increase the persistence of overall inequality of income and reduce that of its concentration. Again, it would have been more appropriate to use a covariate reflecting social spending rather than the total amount of government expenditures. As expected, a higher school enrollment rate reduces income inequality, reflecting the favorable effects of education on labor market prospects and wages. Nonetheless, differences across age groups, captured by life expectancy, are associated to more persistent income inequality. Furthermore, union membership reduces the probability of being in a persistent regime of income inequality. This influence is more pronounced in overall inequality than in the share of the top 10%. Finally, we find that the unemployment rate is inversely related to the persistence of income inequality. This striking result may be determined by the fact that unemployment reduces the wages of less skilled workers and, as a consequence, increases the wage gap.

4 Discussion

This section tries to make a depiction of long-run inequality in the U.S. according to the results obtained so far. The wealth-to-income ratio has behaved as an I(1) process during the whole sample period analyzed. This implies that the time series of our measure of wealth inequality is a stochastic process whose shocks have permanent effects. Similarly to related studies about wealth inequality, such as Piketty and Zucman (2014), the most prominent shifts of this variable are the sharp drop in the aftermath of WWII, and its subsequent change of tendency. These shocks can be considered to have affected the evolution of the wealth-to-income ratio in a permanent way, suggesting that redistribution policies are expected to have had long lasting effects.

Grounded on the results displayed in Table 4, and in contrast to the wealth-to-income ratio, Figures 2 and 3 show, respectively, that both the Gini coefficient for disposable income and the top 10% income share have experienced persistence changes. Overall income inequality, as measured by the Gini index, behaved as an I(1) process between 1870 and 1907. This is also the case of income concentration, captured by the top 10% share, whose non-stationary regime lasted until 1909. This is in line with BMA results showing that globalization has been found to make income inequality more persistent, as this non-stationary period matches the so-called 'First Globalization' (Daudin, Morys, and O'Rourke, 2010). Moreover, the following structural breaks take place in a decade when union membership – a factor inversely related to persistence – experienced a rapid growth (Taft, 1976). After that moment, both time series follow deterministic trends;

slightly decreasing in the case of the Gini coefficient, and increasing in that of the top 10% share. During this stationary period, policies intended to reduce inequality only had transitory effects. According to this, drops of income inequality due to the WWI or the Great Recession and the 'New Deal' were just temporary changes that did not alter its trend. Nonetheless, we can consider that these shocks could have induced reductions on overall inequality time serie by altering the mean of the time series that have been analyzed.

[Insert Figures 2 and 3 about here]

The Gini coefficient and, especially, the top 10% share exhibit their largest drop during WWII. This evolution is in parallel to that of the wealth-to-income ratio. This is in line with Kopczuk and Saez (2004) who show that, to a great extent, top incomes are derived from capital. Furthermore, this reduction in the top 10% income share takes place during an I(1) regime that lasted until 1997. The second non-stationary period of the Gini coefficient only lasted five years, starting in 1944 and ending in 1949. These I(1) regimes corroborate the results obtained in the BMA exercise because their beginning coincide with the 'Second Globalization' marked by the Bretton Woods Agreements and the increase in variables with a direct relationship with income inequality persistence, such as R&D expenditures or the amount of credit to the private sector as a percentage of GDP. The short duration of the second non-stationary regime of the Gini coefficient may be linked to the endorsement of the 'Fair Deal' in 1949. Among others, this package of structural reforms improved the ways of action of trade unions, and increased the minimun wage. Therefore, it can be claimed that this latter influence compensates that of technological change on the persistence of income inequality.

The decades after WWII were characterized by a stable and low income inequality until the early 1980s. While the top 10% share was in an I(1) regime during this period, the Gini coefficient changed to a persistent regime in 1978. This regime shift may have been driven by a period of globalization in terms of trade, a recovery of R&D expenditures, and a decline in unionization as a consequence of the 1973 energy crisis. In addition, tax cuts during the Reagan mandate (1981-1989) provoked a change in redistribution than induced a permanent shock in income inequality. During the 1990s, the increasing trend of our measures of income inequality became again a deterministic trend. The corresponding stationary period has a shorter duration in the Gini coefficient (until 2006) that in the top 10% income share (until 2018). This result may be related to the different influence of credit on the two measures of income inequality. Although the last change in the top 10% income share may need to be checked in a wider temporal perspective, it is worth noting that all measures of inequality behave as non-stationary time series at the end of our sample period. Therefore, redistributive policies intended to reduce both income and wealth inequality are nowadays expected to exert permanent effects.

5 Concluding remarks

The increasing social concern about inequality has created a necessity for a proper knowledge of its behavior and determinants. This paper deals with the long-run evolution of income and wealth inequality in the U.S. To do so, we have carried out a time series analysis of the Gini coefficient for disposable income, the top 10% income share, and the wealth-to-income ratio covering the period 1870–2019. The study has been structured in three stages. First, we have assessed the unit root character of the variables under scrutiny allowing for the presence of multiple trend shifts both under the null and the alternative hypotheses. Proceeding this way, we have circumvented the circular problem between these two data features posed by Perron (2006). Second, we have analyzed whether the time series have experienced structural breaks alternating between stationary and non-stationary regimes throughout the sample period. Finally, we have implemented Bayesian model averaging techniques in a logistic regression framework to study those socio-economic variables that display a more robust relationship with the persistence of inequality.

We find that overall income inequality and income accumulation have experienced several shifts between I(0) and I(1) regimes during the sample period covered in our study. Contrarily, our proxy for wealth inequality behaves as a unit root process, even after accounting for shifts in its deterministic component. In any case, the three time series are in a non-stationary regime at the end of the sample period. Therefore, contemporary redistributive efforts are expected to have persistent effects. The great majority of the potential determinants of income inequality persistence affect the two measures analyzed in the same direction. On the one hand, globalization, technological change, and life expectancy are directly related to the probability of being in a persistent regime of income inequality. On the other hand, a higher school enrollment, unemployment and union membership rates display an inverse relationship. The regressors that try to capture the effects of fiscal and monetary policies are those that have a less robust link with inequality persistence.

The empirical analysis conducted in the present paper and its main limitations open up avenues for future research. The most simple is to take advantage of the wider availability of long-term inequality data and extend the study to other countries. Another alternative is to go beyond trend breaks and implement unit root tests in the presence of other types of non-linear structures such as threshold (Enders and Granger, 1998) or smooth transition (Leybourne, Newbold, and Vougas, 1998; Kapetanios, Shin, and Snell, 2003) models. In addition, the presence of persistence changes can be considered in the more flexible framework provided by fractionally integrated models, see Hassler and Meller (2014) and Martins and Rodrigues (2014).

References

Alvaredo, Facundo et al. (2018). *World Inequality Report 2018*. Cambridge, MA: Harvard University Press. ISBN: 9780674984769. DOI: 10.4159/9780674984769.

Bai, Jushan and Pierre Perron (2003). "Computation and analysis of multiple structural change models". *Journal of Applied Econometrics* 18.1, pp. 1–22. DOI: 10.1002/jae.659.

Carrion-i-Silvestre, Josep Lluís, Dukpa Kim, and Pierre Perron (2009). "GLS-based unit root tests with multiple structural breaks under both the null and the alternative hypotheses". *Econometric Theory* 25.6, pp. 1754–1792. DOI: 10.1017/S0266466609990326.

Christopoulos, Dimitris and Peter McAdam (2017). "On the persistence of cross-country inequality measures". *Journal of Money, Credit and Banking* 49.1, pp. 255–266. DOI: 10.1111/jmcb.12374.

Claessens, Stijn and Enrico Perotti (2007). "Finance and inequality: Channels and evidence". *Journal of Comparative Economics* 35.4, pp. 748–773. DOI: 10.1016/j.jce.2007.07.002.

Clyde, Merlise A., Joyee Ghosh, and Michael L. Littman (2011). "Bayesian adaptive sampling for variable selection and model averaging". *Journal of Computational and Graphical Statistics* 20.1, pp. 80–101. DOI: 10.1198/jcgs.2010.09049.

Daudin, Guillaume, Matthias Morys, and Kevin H. O'Rourke (2010). "Globalization, 1870-1914". In: *The Cambridge Economic History of Modern Europe*. Ed. by Stephen Broadberry and Kevin H. O'Rourke. Vol. 2. Cambridge, UK: Cambridge University Press, pp. 5–29. DOI: 10.1017/CBO9780511794841.003.

Dickey, David A. and Wayne A. Fuller (1979). "Distribution of the estimators for autoregressive time series with a unit root". *Journal of the American Statistical Association* 74.366, pp. 427–431. DOI: 10.2307/2286348.

Durlauf, Steven N., Andros Kourtellos, and Chih Ming Tan (2022). "The Great Gatsby Curve". *Annual Review of Economics* 14.1, pp. 571–605. DOI: 10.1146/annurev-economics-082321-122703.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock (1996). "Efficient tests for an autoregressive unit root". *Econometrica* 64.4, pp. 813–836. DOI: 10.2307/2171846.

Enders, Walter and Clive W. J. Granger (1998). "Unit-toot tests and asymmetric adjustment with an example using the term structure of interest rates". *Journal of Business & Economic Statistics* 16.3, pp. 304–311. DOI: 10.2307/1392506.

Farber, Henry S. et al. (2021). "Unions and inequality over the twentieth century: New evidence from survey data". *Quarterly Journal of Economics* 136.3, pp. 1325–1385. DOI: 10.1093/qje/qjab012.

Feenstra, Robert C. and Gordon H. Hanson (1996). "Globalization, outsourcing, and wage inequality". *American Economic Review* 86.2, pp. 240–245. ISSN: 00028282. DOI: 10.3386/w5424.

Förster, Michael F. and István György Tóth (2015). "Cross-country evidence of the multiple causes of inequality changes in the OECD area". In: *Handbook of Income Distribution*. Ed. by Anthony B. Atkinson and François Bourguignon. Vol. 2. Handbook of Income Distribution. Amsterdam: Elsevier, pp. 1729–1843. DOI: 10.1016/B978-0-444-59429-7.00020-0.

Friedman, Gerald (1999). "U.S. Historical statistics: New estimates of union membership the United States, 1880–1914". *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 32.2, pp. 75–86. DOI: 10.1080/01615449909598928.

Furceri, Davide and Jonathan D. Ostry (2019). "Robust determinants of income inequality". *Oxford Review of Economic Policy* 35.3, pp. 490–517. ISSN: 0266-903X. DOI: 10.1093/oxrep/grz014.

Ghoshray, Atanu, Issam Malki, and Javier Ordóñez (2021). "On the long-run dynamics of income and wealth inequality". *Empirical Economics*, pp. 1–34. DOI: 10.1007/s00181-021-02043-1.

Ghoshray, Atanu, Mercedes Monfort, and Javier Ordóñez (2020). "Re-examining inequality persistence". *Economics* 14.1, p. 20200001. DOI: 10.5018/economics-ejournal.ja.2020-1.

Hailemariam, Abebe, Tutsirai Sakutukwa, and Ratbek Dzhumashev (2021). "Long-term determinants of income inequality: Evidence from panel data over 1870–2016". *Empirical Economics* 61.4, pp. 1935–1958. DOI: 10.1007/s00181-020-01956-7.

Hassler, Uwe and Barbara Meller (2014). "Detecting multiple breaks in long memory the case of U.S. inflation". *Empirical Economics* 46, pp. 653–680. DOI: 10.1007/s00181-013-0691-8.

Hoeting, Jennifer A. et al. (1999). "Bayesian model averaging: A tutorial (with comments by M. Clyde, David Draper and E. I. George, and a rejoinder by the authors)". *Statistical Science* 14.4, pp. 382–417. DOI: 10.1214/ss/1009212519.

Islam, Md. Rabiul and Jakob B. Madsen (2015). "Is income inequality persistent? Evidence using panel stationarity tests, 1870–2011". *Economics Letters* 127, pp. 17–19. DOI: 10.1016/j.econlet.2014.12.024.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor (2017). "Macrofinancial history and the new business cycle facts". *NBER Macroeconomics Annual* 31.1, pp. 213–263. DOI: 10.1086/690241.

Kalliovirta, Leena and Tuomas Malinen (2020). "Non-linearity and cross-country dependence of income inequality". *Review of Income and Wealth* 66.1, pp. 227–249. DOI: 10.1111/roiw.12377.

Kapetanios, George, Yongcheol Shin, and Andy Snell (2003). "Testing for a unit root in the nonlinear STAR framework". *Journal of Econometrics* 112.2, pp. 359–379. DOI: 10.1016/S0304-4076(02)00202-6.

Kejriwal, Mohitosh and Pierre Perron (2010). “A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component”. *Journal of Time Series Analysis* 31.5, pp. 305–328. DOI: 10.1111/j.1467-9892.2010.00666.x.

Kopczuk, Wojciech and Emmanuel Saez (2004). “Top wealth shares in the United States, 1916–2000: Evidence from estate tax returns”. *National Tax Journal* 57.2, pp. 445–487. DOI: 10.17310/ntj.2004.2S.05.

Kurozumi, Eiji (2005). “Detection of structural change in the long-run persistence in a univariate time series”. *Oxford Bulletin of Economics and Statistics* 67.2, pp. 181–206. DOI: 10.1111/j.1468-0084.2004.00116.x.

Kuznets, Simon (1955). “Economic growth and income inequality”. *American Economic Review* 45.1, pp. 1–28. URL: http://www.jstor.org/stable/1811581.

Kwiatkowski, Denis et al. (1992). “Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?” *Journal of Econometrics* 54.1–3, pp. 159–178. DOI: 10.1016/0304-4076(92)90104-Y.

Lebergott, Stanley (1957). “Annual estimates of unemployment in the United States, 1900-1954”. In: *The measurement and behavior of unemployment.* National Bureau of Economic Research project report, pp. 211–242. ISBN: 0-691-04144-X.

Ley, Eduardo and Mark F. J. Steel (2012). “Mixtures of g-priors for Bayesian model averaging with economic applications”. *Journal of Econometrics* 171.2, pp. 251–266. DOI: 10.1016/j.jeconom.2012.06.009.

Leybourne, Stephen, Tae-Hwan Kim, and A. M. Robert Taylor (2007). “Detecting multiple changes in persistence”. *Studies in Nonlinear Dynamics and Econometrics* 11.3. DOI: 10.2202/1558-3708.1370.

Leybourne, Stephen, Paul Newbold, and Dimitrios Vougas (1998). “Unit roots and smooth transitions”. *Journal of Time Series Analysis* 19.1, pp. 83–97. DOI: 10.1111/1467-9892.00078.

Li, Yingbo and Merlise A. Clyde (2018). “Mixtures of g-priors in generalized linear models”. *Journal of the American Statistical Association* 113.524, pp. 1828–1845. DOI: 10.1080/01621459.2018.1469992.

Madsen, Jakob B., Md. Rabiul Islam, and Hristos Doucouliagos (2018). “Inequality, financial development and economic growth in the OECD, 1870–2011”. *European Economic Review* 101, pp. 605–624. DOI: 10.1016/j.euroecorev.2017.11.004.

Madsen, Jakob B., Antonio Minniti, and Francesco Venturini (2018). “Assessing Piketty’s second law of capitalism”. *Oxford Economic Papers* 70.1, pp. 1–21. DOI: 10.1093/oep/gpx040.

Makhlouf, Yasmine (2023). “Trends in income inequality: Evidence from developing and developed countries”. *Social Indicators Research* 165, pp. 213–243. DOI: 10.1007/s11205-022-03010-8.

Martins, Luis F. and Paulo M.M. Rodrigues (2014). “Testing for persistence change in fractionally integrated models: An application to world inflation rates”. *Computational Statistics & Data Analysis* 76, pp. 502–522. DOI: 10.1016/j.csda.2012.07.021.

Mayer, Gerald (Jan. 2004). “Union membership trends in the United States”. *Federal Publications*.

Milanovic, Branko (2016). “Income inequality is cyclical”. *Nature* 537.7618, pp. 479–482. DOI: 10.1038/537479a.

Mitchell, Brian R (2007). *International historical statistics 1750-2005: Americas*. Germany: Springer. ISBN: 9781137305688.

Ng, Serena and Pierre Perron (1995). “Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag”. *Journal of the American Statistical Association* 90.429, pp. 268–281. DOI: 10.1080/01621459.1995.10476510.

— (2001). “Lag length selection and the construction of unit root tests with good size and power”. *Econometrica* 69.6, pp. 1519–1554. DOI: 10.1111/1468-0262.00256.

Nolan, Brian, Matteo G. Richiardi, and Luis Valenzuela (2019). “The drivers of income inequality in rich countries”. *Journal of Economic Surveys* 33.4, pp. 1285–1324. DOI: https://doi.org/10.1111/joes.12328.

Olmo, Jose and Marcos Sanso-Navarro (2021). “Modeling the spread of COVID-19 in New York City”. *Papers in Regional Science* 100.5, pp. 1209–1229. DOI: 10.1111/pirs.12615.

Perron, Pierre (1989). “The great crash, the oil price shock, and the unit root hypothesis”. *Econometrica* 57, pp. 1361–1401. DOI: 10.2307/1913712.

Perron, Pierre (2006). "Dealing with structural breaks". In: *Palgrave Handbook of Econometrics*. Ed. by K. Patterson (Eds.) H. Hassani T. C. Mills. Vol. 1. Econometric Theory. Basingstoke: Palgrave Macmillan, pp. 278–352. ISBN: 978-1-4039-4155-8.

Perron, Pierre and Tomoyoshi Yabu (2009). "Testing for shifts in trend with an integrated or stationary noise component". *Journal of Business & Economic Statistics* 27.3, pp. 369–396. DOI: 10.1198/jbes.2009.07268.

Piketty, Thomas (2014). *Capital in the twenty-first century*. Cambridge, MA: Harvard University Press. ISBN: 9780674430006.

Piketty, Thomas and Gabriel Zucman (2014). "Capital is back: Wealth-income ratios in rich countries 1700–2010". *Quarterly Journal of Economics* 129.3, pp. 1255–1310. DOI: 10.1093/qje/qju018.

Prados De La Escosura, Leandro (2008). "Inequality, poverty and the Kuznets curve in Spain, 1850–2000". *European Review of Economic History* 12.3, pp. 287–324. DOI: 10.1017/S1361491608002311.

Raftery, Adrian E. (1995). "Bayesian model selection in social research". *Sociological Methodology* 25, pp. 111–163. DOI: 10.2307/271063.

— (1996). "Approximate Bayes factors and accounting for model uncertainty in generalised linear models". *Biometrika* 83.2, pp. 251–266. DOI: 10.1093/biomet/83.2.251.

Raftery, Adrian E., David Madigan, and Jennifer A. Hoeting (1997). "Bayesian model averaging for linear regression models". *Journal of the American Statistical Association* 92.437, pp. 179–191. DOI: 10.2307/2291462.

Roine, Jesper and Daniel Waldenström (2011). "Common trends and shocks to top incomes: A structural breaks approach". *Review of Economics and Statistics* 93.3, pp. 832–846. DOI: 10.1162/REST_a_00112.

Sanso-Navarro, Marcos and María Vera-Cabello (2020). "Income inequality and persistence changes". *Social Indicators Research* 152.2, pp. 495–511. DOI: 10.1007/s11205-020-02444-2.

Solarin, Sakiru Adebola et al. (2022). "Inequality persistence of 21 OECD countries from 1870 to 2020: Linear and non-linear fractional integration approaches". *Social Indicators Research* 164.2, pp. 711–725. DOI: 10.1007/s11205-022-02982-x.

Solt, Frederick (2020). "Measuring income inequality across countries and over time: The standardized world income inequality database". *Social Science Quarterly* 101.3, pp. 1183–1199. DOI: 10.1111/ssqu.12795.

Steel, Mark F. J. (2020). "Model averaging and its use in economics". *Journal of Economic Literature* 58.3, pp. 644–719. DOI: 10.1257/jel.20191385.

Taft, Philip (1976). "Expansion of unionization in the early 20th century". *Monthly Labor Review* 99.9, pp. 32–35. DOI: 10.2307/41840306.

Troy, Leo (1965). "Introduction to "Trade union membership, 1897–1962"". In: *Trade Union Membership, 1897–1962*. National Bureau of Economic Research, pp. 1–10. ISBN: 0-87014-406-5.

Vernon, James R. (1994). "Unemployment rates in postbellum America: 1869–1899". *Journal of Macroeconomics* 16.4, pp. 701–714. DOI: 10.1016/0164-0704(94)90008-6.

Wood, Adrian (July 1995). *North-south trade, employment and inequality: Changing fortunes in a skill-driven world*. Oxford University Press. DOI: 10.1093/0198290152.001.0001.

Tables and figures

Table 1: Unit root testing

	P_t	MP_t	ADF	Z_α	MZ_α	MSB	MZ_t
Gini coefficient	42.5610	38.7540	-0.5390	-1.5320	-1.4190	0.4150	-0.5390
Top 10% income share	25.1630	23.7990	-1.1360	-3.4380	-3.3390	0.3300	-1.1030
Wealth-to-income ratio	13.6223	13.7191	-1.5111	-6.8241	-6.6782	0.2600	-1.7361

Note: P_t is the Point Optimal test proposed by Elliott, Rothenberg, and Stock (1996). ADF is the augmented Dickey-Fuller test (Dickey and Fuller, 1979). Z and M-class tests are discussed in Ng and Perron (2001).The deterministic component contains a constant and a linear trend, and the number of augmentation lags has been selected by the modified AIC.

Table 2: Trend shift test statistic (Perron and Yabu, 2009) and estimated break dates.

	Exp - W_{FS}	F_T (2|1)	F_T (3|2)	F_T (4|3)	F_T(5|4)	TB_1	TB_2	TB_3	TB_4	TB_5
Gini coefficient	5.9028***	4.5744**	10.3449***	30.7594***	9.0741***	1980	1905	1993	1946	1965
Top 10% income share	4.3999**	17.0284***	99.3356***	6.0225***		1942	1900	1983	2001	
Wealth-to-income ratio	2.9010*					1940				

Note: The trimming parameter has been set to 0.15. Break dates are reported in the order of appearance in the sequential implementation (Kejriwal and Perron, 2010). * $p<0.10$, **$p<0.05$, *** $p<0.01$.

Table 3: Unit root testing in the presence of trend shifts.

	P_t	MP_t	ADF	Z_α	MZ_α	MSB	MZ_t
Gini coefficient	13.3416	11.3442	-4.9168**	-41.3771*	-35.4032	0.1188	-4.2070
Top 10% income share	15.6839	13.2450	-4.0565	-29.6265	-26.6507	0.1369	-3.6490
Wealth-to-income ratio	12.7447	11.3418	-2.8039	-15.2779	-14.4416	0.1835	-2.6504

Note: P_t is the Point Optimal test proposed by Elliott, Rothenberg, and Stock (1996). ADF is the augmented Dickey-Fuller test (Dickey and Fuller, 1979). Z and M-class tests are discussed in Ng and Perron (2001). The deterministic component contains a constant and a linear trend, and the number of augmentation lags has been selected by the modified AIC. Estimated break dates reported in Table 2 have been controlled for (Carrion-i-Silvestre, Kim, and Perron, 2009). * p<0.10, **p<0.05,***p<0.01.

Table 4: Persistence change test (Leybourne, Kim, and Taylor, 2007).

	Period	M test	I(0) start	I(0) end
Gini coefficient	1870 - 2019	-5.5960***	1907	1944
	1870 - 1906	-1.9290	1879	1901
	1945 - 2019	-4.7730*	1990	2006
	1945 - 1989	-4.6380**	1949	1978
Top 10% income share	1870 - 2019	-4.9640**	1909	1940
	1870 - 1908	-1.9290	1879	1901
	1941 - 2019	-5.3810**	1997	2018
Wealth-to-income ratio	1870 - 2019	-4.1420	1944	2018

Note: The deterministic component contains a constant and a linear trend, and the lag length has been selected using the procedure proposed by Ng and Perron (1995). * p<0.10, **p<0.05, ***p<0.01.

Table 5: Potential determinants of income inequality persistence: Definition and data sources.

Groups and variables	Definition	Sources
Globalization		
Foreign inverstment	Non-residential investment ratio (% of GDP)	Madsen, Islam, and Doucouliagos (2018) / FRED[1]
Trade volume	Ratio of imports plus exports to GDP	J-S-T Macrohistory Database[2] / WDI[3]
Foreign patents	Patent applications by foreign residents	USPTO[4]
Technological change		
Patent stock	Patent stock (15% of yearly depreciation)	Madsen, Islam, and Doucouliagos (2018) / USPTO[4]
R&D expenditure	Expenditure on research and development (% of GDP)	Madsen, Islam, and Doucouliagos (2018) / WDI[3]
Financial development		
Savings	Gross private savings ratio (% of GDP)	Madsen, Islam, and Doucouliagos (2018) / FRED[1]
Credit	Bank credit to the non-bank private sector (% of GDP)	Madsen, Islam, and Doucouliagos (2018) / FRED[1]
Interest rate	Real interest rate, long-term goverment bonds	Madsen, Islam, and Doucouliagos (2018) / FRED[1]
Fiscal and monetary policy		
Inflation	Inflation rate, consumer price index	J-S-T Macrohistory Database[2] / WDI[3]
Tax revenue	Ratio of overall tax revenues to GDP	Mitchell (2007)/WDI[3]
Gov. expenditure	Ratio of goverment expenditure to GDP	J-S-T Macrohistory Database[2] / WDI[3]
Demographics and societal structure		
Enrollment rate	Gross school enrollment rate	Madsen, Islam, and Doucouliagos (2018) / WDI[3]
Life expectancy	Life expectancy at age 10	Madsen, Islam, and Doucouliagos (2018) / SSA[5]
Labor institutions and regulations		
Unemployment	Unemployment share of the civilian labor force	Lebergott (1957) / Vernon (1994) / BLS[6]
Union membership	Union membership share of the civilian labor force	Troy (1965) /Friedman (1999) / Mayer (2004) / BLS[6]

[1]Federal Reserve Economic Data, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/ [2]Jordà-Schularick-Taylor Macro-history Database; Jordà, Schularick, and Taylor (2017) [3]World Development Indicators, The World Bank Group; https://databank.worldbank.org/ [4]United States Patents and Trademark Office, Open Data Portal; https://www.uspto.gov/ [5] Social Security Agency, Data Page; https://www.ssa.gov/data/ [6]U.S. Bureau of Labour Statistics, data tools; https://www.bls.gov/data/.

Table 6: Bayesian model averaging: Persistence of the Gini coefficient.

	Hyper-g/n prior			Benchmark prior		
	PIP	Mean Coef.	Std. Dev.	PIP	Mean Coef.	Std. Dev.
Foreign investment	0.9466	7.8807	6.9365	0.9447	8.6896	7.3623
Trade volume	0.9187	0.3544	0.2983	0.9240	0.3910	0.3111
Foreign patents	0.9998	-5.0943	2.8316	0.9996	-5.5847	2.8957
Patent stock	0.9036	-4.8449	4.7259	0.8990	-5.3654	5.0518
R&D expenditure	0.9988	10.3137	6.3778	0.9988	11.3199	6.6373
Savings	0.9955	0.4540	0.2735	0.9956	0.4989	0.2881
Credit	0.6390	0.0164	0.0327	0.6287	0.0180	0.0349
Interest rate	0.8846	-0.3884	0.3803	0.8824	-0.4298	0.4054
Inflation	0.5276	0.0681	0.2090	0.5174	0.0743	0.2226
Tax revenue	0.4912	0.0651	0.2910	0.4762	0.0699	0.3073
Gov. expenditure	0.5875	0.0684	0.1835	0.5824	0.0761	0.1951
Enrollment rate	1	-0.5234	0.2754	1	-0.5759	0.2880
Life expectancy	0.9940	1.1566	0.7698	0.9941	1.2741	0.8096
Unemployment	0.9996	-0.6319	0.3673	0.9993	-0.6947	0.3818
Union membership	1	-0.9774	0.5062	1	-1.0720	0.5273

Note: A uniform prior has been established for all models.

Table 7: Bayesian model averaging: Persistence of the top 10% income share.

	Hyper-g/n prior			Benchmark prior		
	PIP	Mean Coef.	Std. Dev.	PIP	Mean Coef.	Std. Dev.
Foreign inverstment	0.9197	16.0662	23.0919	0.9148	21.6182	27.0119
Trade volume	0.4775	0.0095	0.3223	0.4676	0.0120	0.3784
Foreign patents	0.5352	0.7064	3.5332	0.5164	0.9334	4.1038
Patent stock	0.9770	-10.6351	15.6420	0.9739	-14.2584	18.1841
R&D expenditure	0.5934	2.0500	7.4528	0.5793	2.7433	8.6937
Savings	0.9997	0.8693	1.2000	0.9993	1.1660	1.3985
Credit	0.9867	-0.1554	0.1942	0.9857	-0.2087	0.2262
Interest rate	0.9189	-1.2215	2.0067	0.9078	-1.6354	2.3480
Inflation	0.5202	-0.0389	0.3365	0.5082	-0.0542	0.3896
Tax revenue	0.4919	-0.0764	0.6696	0.4871	-0.1031	0.7879
Gov. expenditure	0.7890	-0.3561	0.7970	0.7709	-0.4761	0.9374
Enrollment rate	0.6463	-0.1307	0.4221	0.6309	-0.1760	0.4971
Life expectancy	0.9995	2.2578	2.9767	0.9994	3.0340	3.4614
Unemployment	1	-1.2278	1.5216	1	-1.6491	1.7703
Union membership	0.5165	-0.0549	0.6771	0.5077	-0.0747	0.7955

Note: A uniform prior has been established for all models.

Figure 1: Gini coefficient for disposable income, top 10% income share, and wealth-to-income ratio in the U.S., 1870-2019.

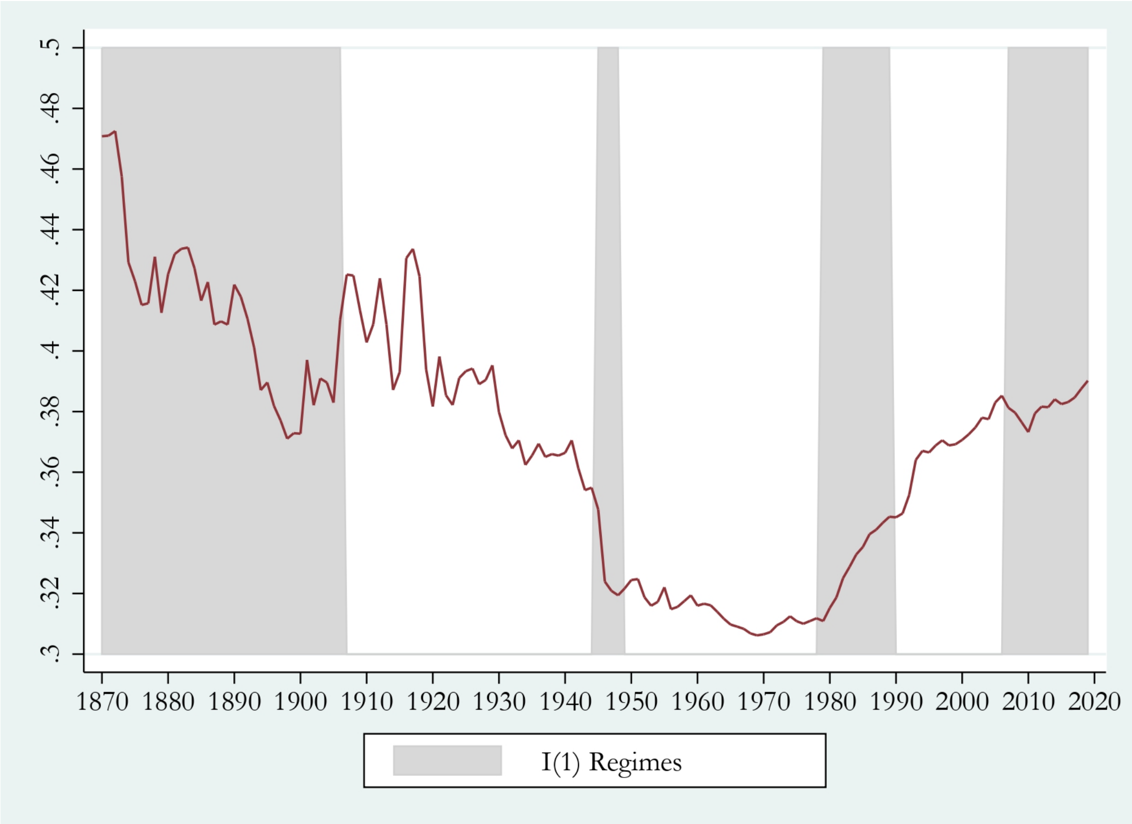

Figure 2: U.S. Gini coefficient for disposable income, 1870-2019.

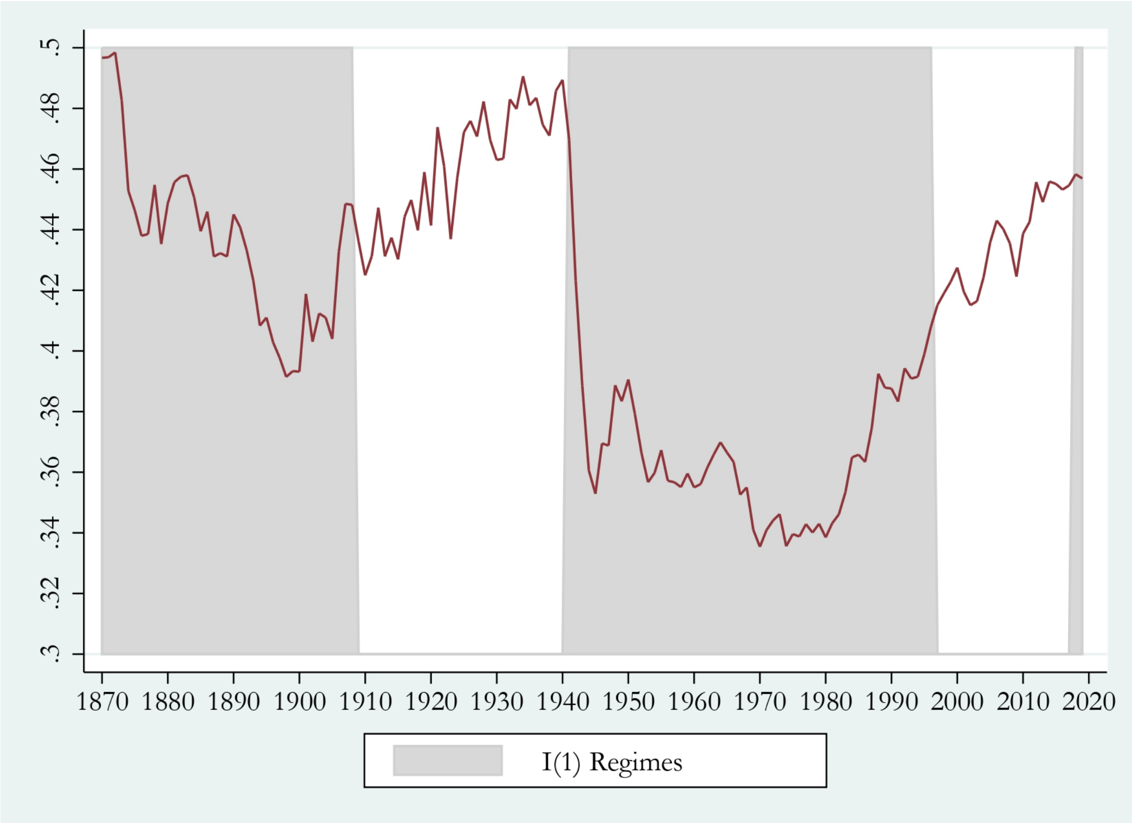

Figure 3: U.S. top 10% income share, 1870-2019.

Appendix

Table A1: BMA results for the Gini coefficient: PIPs using alternative priors for model-specific parameters.

	Beta prime	ZS adapted	Robust	Intrinsic
Foreign investment	0.8706	0.8691	0.9111	0.9129
Trade volume	0.9243	0.9252	0.9351	0.9307
Foreign patents	0.9998	0.9997	0.9995	0.9996
Patent stock	0.5098	0.4853	0.6649	0.6525
R&D expenditure	0.9975	0.9978	0.9982	0.9978
Savings	0.9786	0.9727	0.9837	0.9821
Credit	0.3053	0.2837	0.3987	0.3988
Interest rate	0.5285	0.5051	0.6679	0.6595
Inflation	0.2473	0.2333	0.3374	0.3279
Tax revenue	0.2450	0.2238	0.3243	0.3207
Gov. expenditure	0.3393	0.3374	0.4319	0.4260
Enrollment rate	1	0.9998	0.9999	1
Life expectancy	0.9500	0.9419	0.9733	0.9743
Unemployment	0.9949	0.9922	0.9961	0.9958
Union membership	1	1	0.9998	1

Note: A uniform prior has been established for all models.

Table A2: BMA results for the top 10% income share: PIPs using alternative priors for model-specific parameters.

	Beta prime	ZS Adapted	Robust	Intrinsic
Foreign inverstment	0.9013	0.8965	0.8926	0.8910
Trade volume	0.2349	0.2342	0.3031	0.3078
Foreign patents	0.2387	0.2398	0.3226	0.3276
Patent stock	0.8960	0.8932	0.9160	0.9113
R&D expenditure	0.3680	0.3672	0.4444	0.4470
Savings	0.9945	0.9931	0.9957	0.9958
Credit	0.9722	0.9703	0.9738	0.9736
Interest rate	0.7610	0.7487	0.7886	0.7849
Inflation	0.4187	0.4190	0.4628	0.4618
Tax revenue	0.2599	0.2590	0.3352	0.3378
Gov. expenditure	0.4296	0.4251	0.5286	0.5234
Enrollment rate	0.4246	0.4285	0.4819	0.4934
Life expectancy	0.9975	0.9965	0.9971	0.9971
Unemployment	0.9999	0.9998	0.9999	0.9999
Union membership	0.3263	0.3263	0.3962	0.4055

Note: A uniform prior has been established for all models.