The Role of Credit on the Evolution of Wealth Inequality in the USA

Rodolfo Oviedo Moguel
Banco de México

November 2020
The Role of Credit on the Evolution of Wealth Inequality in the USA*

Rodolfo Oviedo Moguel†
Banco de México

Abstract: In the USA, the share of household wealth held by the richest 1% increased from 23.5% in 1980 to 41.8% in 2012. This paper contributes to understanding the causes behind this increase. First, using an accounting decomposition, I show that more than half of the increase in the share of the top 1% can be attributed to a decrease in the saving rate of the bottom 99%. Second, using a heterogeneous agent model, I show that the decrease in the saving rate of the bottom groups cannot be rationalized by the reduction in the progressively of taxation or changes in the volatility and concentration of labor earnings. Lastly, I introduce a shock to the credit market into the model in the form of loosening the borrowing constraints of the economy. This shock can simultaneously match the increase in wealth concentration and the decrease of the saving rate of the economy.

Keywords: Credit, Debt, Saving Rate, Wealth Distribution

JEL Classification: D14, D31, D33, E21, E62, G51

Resumen: En Estados Unidos, el porcentaje de la riqueza de los hogares en manos del top 1% se incrementó de 23.5% en 1980 a 41.8% en 2012. Este artículo contribuye a entender las causas detrás de este incremento. Primero, utilizando un ejercicio de descomposición contable, demuestro que más de la mitad del incremento en la riqueza relativa del top 1% se puede atribuir a una caída en la tasa de ahorro de los individuos fuera del top 1%. Segundo, utilizando un modelo de agentes heterogéneos, demuestro que la caída en el ahorro de los individuos fuera del top 1% no puede ser racionalizada por la reducción en la progresividad de los impuestos o cambios en la volatilidad y concentración de los ingresos laborales. Por último, introduzco un choque al mercado de crédito en el modelo relajando las restricciones al endeudamiento. Este choque puede replicar de forma simultanea el incremento en la concentración de la riqueza y la caída en la tasa de ahorro de la economía.

Palabras Clave: Crédito, Deuda, Distribución de la Riqueza, Tasa de Ahorro

*I am indebted to my PhD advisor Michele Boldrin for his excellent guidance. Gaetano Antinolfi, Miguel Faria-e-Castro, Faisal Sohail, David Wiczer and Lijun Zhu provided valuable suggestions. I am grateful to Per Krusell and Joachim Hubmer for providing me with helpful references. I also had very useful conversations with Nicolás Amoroso, Santiago Bazdresch, Alfonso Cebreros, Juan Ramón Hernández, Raúl Ibarra, Gustavo Leyva, and Felipe Meza.

† Dirección General de Investigación Económica. Email: roviedom@banxico.org.mx.
1 Introduction

Wealth concentration in the USA has increased sharply during the last 40 years. According to Saez and Zucman (2016), the share of wealth held by the top 1% increased from 24% in 1980 to 42% in 2012. The Gini index of wealth inequality moved from 0.79 in 1983 to 0.87 in 2013, Wolff (2014). In recent years, there has been an intense debate about the causes of this increase, and several policies have been proposed to reverse this trend. In this paper, I contribute to our understanding of the possible causes of the increase in wealth concentration by evaluating the impact of changes in the credit conditions, defined as an increase in the ability to borrow, on the distribution of wealth after 1980.

In the first section of the paper, I decompose changes in the wealth holdings of the top 1%, the next 9% in the top decile (hereafter [90-99]%), and the bottom 90% into volume and capital gain effects. Changes in the stock of wealth due to savings are volume effects, while changes due to fluctuations in the market value of wealth (i.e. changes in the price of stocks or housing) are capital gains effects. I found that, between 1980 and 2012, most of the increase in the fraction of wealth held by the top 1% can be explained by volume effects: the rate of growth of wealth coming from savings for the top 1% was consistently higher relative to the other groups, and this fact explains most of the increase in concentration. The overall effect of differences in capital gains between groups was, in fact, slightly negative: it reduced wealth concentration. This is due to a well-known fact: the bottom 90% holds most of its wealth in housing which was the best performing asset class from 1980 to 2007. The increase in the price of housing during the period was a force toward equality.\(^1\)

To then understand the differences in the rate of wealth accumulation among these three groups, I compute the savings rate for each of them. Between 1950 and 1980, the savings rate of the bottom 90% and [90-99] % groups was stable, around 5 and 29%, respectively. In the next three decades, 1980-2010, the average saving rate of both groups decreased significantly to 2% and 15%, respectively. The pattern for the top 1% was different: its savings rate

\(^{1}\text{The subsequent crash reduced the equalizing effect of house pricing but did not eliminate it completely.}\)
increased from an average of 30% between 1950 and 1980 to 37% between 1980 and 2012. The divergence in saving rates between the top 1% and the remaining groups was the most important force driving wealth concentration.

To quantify the contribution of the changing saving patterns to the increase in wealth concentration, I construct hypothetical wealth holdings for the bottom 90% and the [90-99] % under the assumption that saving rates were kept at their 1950-1980 average during 1980-2012. Then, I estimate the wealth shares implied by these hypothetical series. If we assume that the bottom 90% kept saving at their historical average of 5%, then the share of the top 1% would have increased from 24% to 34% instead of 42%. This attributes around 40% of the top percentile’s increase to the change in the saving behavior of the bottom 90%. Assuming that both the bottom 90% and the top [90-99] % did not change their saving behavior during 1980 – 2012 implies that the share of the top 1% would have reached 30% in 2012 instead of the actual 42%.

In the second section of my paper, I build a general equilibrium model based on Bewley (1977), to quantify the relative effect of different plausible hypotheses on the evolution of the wealth distribution: i) changes in credit conditions, (ii) increase in the concentration and riskiness of labor income and, (iii) reforms to the tax code (reductions in corporate and personal income taxation). Changes in credit conditions are defined as increases in the ability of households to borrow and are modeled as a loosening of borrowing constraints. The initial stationary distribution of the model is calibrated to match the wealth distribution in 1980. The values for the taxation system and the labor income process after 1980 are taken from the data, and the path for the borrowing constraints is calibrated to match the evolution of the ratio of non-mortgage debt to disposable income. The model is successful at generating a transition path that matches the increase in wealth concentration observed in the data.

Counter-factual simulations indicate that the credit market channel explains approximately half of the increase in wealth concentration. When all the other shocks, except the credit

---

2 Even if saving rates had remained constant, there would have been an increase of 6 points in the share of the top 1%. This is because the share of total labor income going to the richest 1% increased from 4.7% in 1980 to 8.9% in 2012.
channel, are fed into the model, wealth concentration measured as the share of the top 1% increases only half as much compared to the baseline scenario in which the credit market channel is included. The relaxation of the borrowing constraints causes the originally constrained households to accumulate more debt and decreases the precautionary savings of those “close” to the constraint since the likelihood of hitting it decreases. In addition to these two forces, the increase in the real interest rate caused by the contraction in the overall supply of savings increased the saving rate of the top groups, which further increased wealth concentration. The credit channel is fundamental to match the pattern of savings observed in the data: a decreased in the overall saving rate fueled by the change of behavior of the bottom groups and a slight increase in the saving rate of the top 1%.

The increase in concentration of labor income and the tax reforms that took place during this period also contributed to the concentration of wealth. The top wage earners tend to be the wealthiest agents in the economy and a higher share of total income going to this group mechanically increases their savings flow relative to other groups, thereby increasing concentration. The reduction in the progressivity of personal and corporate taxation increased the incentive to save for individuals in the top groups and especially in the top 1%, since they faced significantly lower tax rates on the return to capital. A higher savings rate for top groups combined with a stable saving rate for other groups generated wealth concentration.

Still, when the higher concentration of labor income and the changes in taxation are not included but the relaxation of the credit constraints is, the share of wealth owned by the top 1% increases by 45% compared with the baseline case. This suggests that the credit market channel plays a major role by itself, and that only the interplay between this channel and the other two exogenous changes allows us to fully understand the observed evolution in wealth concentration.

Finally, I should note that the increased riskiness of labor income was a force toward equality. A riskier income process increases overall precautionary savings and this effect was particularly strong for the bottom groups, which were closer to the borrowing constraint. When only this shock is considered, the share of the top 1% in total wealth decreases.
1.1 Related Literature

The main contribution of this paper is to study the role of changes in credit conditions on the evolution of the distribution of wealth in the context of a heterogeneous-agent macroeconomic model calibrated for the USA. As mentioned before, the main findings are that a loosening in credit constraints is key to simultaneously match the increase in wealth concentration and the decrease of the saving rate of the overall economy. Furthermore, it is shown that fiscal and labor earning shocks are not able to match the behavior of the saving rate. This paper is directly related to three strands of literature. First, to a literature that aims to replicate the upper tail of the distribution of wealth in heterogeneous agent models. Second, to a number of studies that build on these models to quantify the causes of changes in the distribution of wealth in the USA over time. Lastly, to an empirical and theoretical literature that explores the relationship between credit conditions and the decrease in the saving rate. I now proceed to describe the relationship of this paper with this vast literature.

The Bewley model is the standard workhorse used in quantitative macroeconomics to model the distribution of wealth. Aiyagari (1994), and Huggett (1996) solved general equilibrium versions of the Bewley model and documented that adding labor earning heterogeneity using data from the PSID is not enough to generate the degree of wealth concentration observed in the data. De Nardi, Fella, and Paz-Pardo (2016) showed that the same result holds even if administrative data from the W2 forms are used to calibrate different versions of the Bewley model. Based on this observation, several studies explored the mechanisms that are able to generate concentration that is quantitatively similar to the one observed in the data. Krusell, Smith, and Jr. (1998) proposed the introduction of stochastic discount factors as a reduced form way of generating saving behavior that correlates positively with wealth in line with the data; Castaneda, Díaz-Giménez, and Ríos-Rull (2003) showed that an earning process that displays enough negative skewness at the top of the earning distribution is able to generate

3Aiyagari (1994) considers the infinite horizon case while Huggett (1996) explicitly models the life-cycle. Hansen and Imrohoroglu (1992) also proposed a general equilibrium version of the Bewley model with infinite horizon, but their focus was to study the effects of unemployment insurance rather than the distribution of wealth.
high precautionary savings at the top and hence a long right tail on the wealth distribution; De Nardi (2004) explicitly modeled the effect of non-homothetic bequest motives and inheritance of ability across generations to rationalize higher accumulation at the top of the wealth distribution; Quadrini (2000) and Cagetti and De Nardi (2006) introduced entrepreneurship to a basic Aiyagari model generating higher saving at the top by entrepreneurial income that is riskier than labor earnings (Benhabib, Bisin, and Zhu (2011) generalized their result to any type of capital income risk). Recently, Hubmer, Krusell, and Smith (2020), introduced the assumption that the risk and expected return of assets depends on the level of wealth and showed that this mechanism is able to generate a concentration of wealth that matches the data.\(^4\) De Nardi and Fella (2017) provides an excellent survey of all the mechanisms explored in the literature to generate a realistic concentration of wealth. In this paper, I follow Krusell, Smith, and Jr. (1998) and employ stochastic discount factors as the reduced form mechanism that allows the model to properly capture the initial stationary distribution of wealth. This decision was based on the simplicity of this approach relative to the others used in the literature, as well as on the fact that the logic of the main result of the paper does not crucially depend on stochastic discount factors.

The studies described above focus on generating a stationary distribution of wealth that is similar to the one observed in the data. This paper directly contributes to a literature that uses different versions of the Bewley model to quantify the drivers behind the increase in wealth concentration in the USA. The general structure of these studies is to first use some of the mechanisms previously described to match the initial concentration of wealth observed in the data and then proceed to feed the model with the policy and technological changes faced by the USA economy. The goal is to quantify the effect of each one of these changes into the evolution of the distribution of wealth which is endogenously generated by the model. The studies more closely related to this paper are the following:

1. Hubmer, Krusell, and Smith (2020) (from now on HKS) use stochastic discount factors and are a risk-return profile of assets that depend on the level of wealth to match the

\(^4\)This assumption is justified by the empirical findings of Fagereng et al. (2020) and Calvet, Bach, and Sodini (2015) using administrative data for Norway and Sweden respectively.
initial stationary distribution of wealth in an Aiyagari model. Then, the model is fed
with the observed changes in the volatility and concentration of labor earnings (as
computed by Heathcote, Storesletten, and Violante (2010) and Piketty and Saez (2003)
respectively), changes in the tax system, and the observed returns by asset class over the
period of study. The model is successful at replicating the low-frequency movements
due to asset fluctuations and also does a good job at matching the overall long term
increase in wealth concentration. The main quantitative finding of the paper is that the
reduction in progressivity of taxation is the main driver behind the increase in wealth
concentration.

2. Kaymak and Poschke (2016) (from now on KP) follows Castaneda, Díaz-Giménez, and
Ríos-Rull (2003) and calibrate an extraordinary state of labor earnings, not observed in
the data, to match the initial stationary distribution of wealth in a framework that explic-
itly models retirement and a non-homothetic bequest motive (as in De Nardi (2004)).
Changes to the taxation scheme for corporations, individuals, and estates are introduced
into the model along with changes to Social Security transfers and labor earnings. The
model is successful at replicating the overall increasing pattern for wealth concentra-
tion and, according to the counterfactual exercises, higher wage dispersion explains
between 50-60% of the rise in wealth concentration.

3. Benhabib, Bisin, and Luo (2019) (from now on BBL) introduce capital income risk
(idiosyncratic returns), a non-homothetic bequest motive and stochastic labor earnings
in an OLG model, and estimate the contribution of each one of these forces using
the method of simulated moments (MSM) targeting certain parts of the distribution of
wealth in 2007 as well as social mobility as calculated by Charles and Hurst (2003)
using data from the PSID. While the main goal of this paper is to match both social
mobility and the wealth distribution in 2007, they also include a transition dynamics
exercise in which they simulate the economy starting at a distribution of wealth that
is less concentrated compared to the estimated stationary distribution and find that a
stronger bequest motive, as well as a return of capital that have a higher mean and stan-
standard deviation, are able to generate an increase in wealth concentration that is broadly consistent with the data.

In addition to these studies, this paper is also related to Capital in the Twenty-First Century by Thomas Piketty where he famously proposed that the concentration of wealth is a positive function of \([r - g]\), where \(r\) is the average real return on capital after taxes and \(g\) is the real growth rate of net production.\(^5\) While his argument is general, and not quantified specifically for the USA, our finding that an increase in the after-tax return on capital associated with a decrease in the progressivity of taxation is broadly in line with his claim.\(^6\)

This paper shares the methodology of HKS and KP in the sense of first calibrating the stationary wealth distribution to the second half of the twentieth century, and then studying the dynamics generated by changes in the labor earning process and the tax system. There are, however, important features that differentiate this paper from HKS and KP. The goal of this paper is to construct a story that is consistent with both the increase in wealth concentration and the decrease in the overall saving rate that is observed in the data. Consistent with the results of HKS and KP, I found that changes in labor earnings and the tax system are able to generate more concentration, but they do so through an increase in the overall saving rate of the economy, a result that is counterfactual as the aggregate saving rate decreased from 12% in 1980 to 2% in 2007. It is then shown, that adding a credit shock that is meant to replicate the loosening of credit conditions observed in the data, is able to simultaneously replicate a decrease in the overall saving rate of the economy, that is mainly concentrated in the bottom 90% of the wealth distribution, while at the same time generating an increase in the concentration of wealth. Comparing our results with the transition dynamic exercise presented in BBL, we find that even when a stronger bequest motive and a higher expected return on capital are both important forces to generate more wealth concentration they both increase the overall saving rate of the economy and hence are not able to replicate the observed behavior of the overall saving rate. Summarizing, the main contribution of this paper to the quantita-

\(^5\)A formal version of the \([r-g]\) argument can be found in Piketty and Zucman (2015).
\(^6\)Aoki and Nirei (2017) also studies how progressive taxation affects the distribution of income and wealth in an heterogeneous agent model with investment risk.
tive literature that studies the drivers behind the increase of wealth concentration in the USA, is to add the credit channel as a shock that is both quantitatively relevant, and consistent with the increase in wealth concentration and the decrease in the saving rate of the economy.

2 Facts About Wealth Inequality in the USA

Saez and Zucman (2016) estimated the top wealth shares for USA using the capitalization method for the period between 1913 and 2012. Wealth is defined as the current market value of all assets owned by households, net of all debts.\footnote{Assets include all the non-financial and financial assets over which ownership rights can be enforced and that provide economic benefits to their owners.} Figure 1 presents the share of total net wealth held by the wealthiest 1% households in the USA:

\begin{figure}
\centering
\includegraphics[width=\textwidth]{wealth_share.png}
\caption{Wealth Share of the Top 1%.
Source: Saez and Zucman (2016).}
\end{figure}

The share of wealth owned by the top 1% reached a peak of 51.4% in 1928 followed by a dramatic decrease during the Great Depression and World War II. It stabilized at around 29% between 1950 and 1968 and then suffered another significant decrease that started in 1968, reaching its lowest level at 22.9% during 1978. Starting in the early 1980s, and accelerating
around 1986, the wealth share of the top 1% experienced a dramatic recovery reaching 41.8% in 2012. In this paper, I focus on studying the causes of the increase in wealth inequality between 1980 and 2012.

There was heterogeneity in the changes experimented by bottom 99% wealth holders. For the sake of simplicity, we examine the population according to their wealth holdings in three categories: families in the bottom 90%, the [90-99%] (the top decile excluding the top 1%) and the top 1%. The key events of the period between 1950 and 2012 were the following. The share of the group [90-99%] was stable at approximately 43% between 1950 and 1978; it then monotonically decreased between 1978 and 1988, stabilizing at around 38%. The Great Recession did not affect the share of this group significantly. The share of the bottom 90% went from 26% in 1962 to 36% in 1986. In 1987, a decreasing trend began that was intensified by the Great Recession, reaching 23% in 2012.

Thus far, we have talked about the share of wealth going to each group in the distribution. In Figure 2, we can see the evolution of the stock of wealth of the three groups, in real terms. The wealth of the three groups increased almost monotonically between 1980 and 2000. It then decreased as a result of the bust of the dot-com bubble — mainly for the top 1% — and increased between 2002 and 2007. The Great Recession affected the three groups significantly. However, the top 1% had already surpassed its 2007 level of wealth by 2012 while the bottom 90% and the [90-99%] were still considerably below their pre-crisis levels.

2.1 A Framework for Understanding Changes in Wealth Inequality

In this subsection, I introduce an accounting framework that makes possible to decompose the changes in the level of wealth holdings of the top 1%, the [90-99] %, and the bottom 90% into volume and capital-gain effects. Changes in the stock of wealth due to savings are volume effects, while changes due to fluctuation in the market value of wealth (i.e. changes in the price of stocks or housing) are capital gains effects.
Figure 2: Wealth by group in 2010 USD [1980 =1].
Source: Saez and Zucman (2016).

Let C denote a group of individuals with wealth in period t equal to $W^C_t$, which is equal to the sum of the market value of all asset classes held by this group (housing + equities + sole proprietorship and partnerships + fixed income + pension funds - debt):

$$W^C_t = \sum_{i \in I} W^C_t(i)$$

The wealth of group C in $t + 1$ is a function of changes in price of wealth of group C, $q^C_t$, and the flow of savings, $S^C_t$:

$$W^C_{t+1} = [1 + q^C_t][W^C_t + S^C_t] = W^C_t[1 + q^C_t]\left[1 + \frac{S^C_t}{W^C_t}\right]$$  \hspace{1cm} (1)

In the previous expression, it is assumed that savings are made before the realization of price changes. Also, $q^C_t$ is defined to be the price change of wealth between period t and period t+1 and is a weighted average of the price change of every asset class, $q_t(i)$:

$$q^C_t = \sum_{i \in I} \left[\frac{W^C_t(i) + S^C_t(i)}{W^C_t + S^C_t}\right]q_t(i)$$  \hspace{1cm} (2)
In equation 2, the implicit assumption is that the return of group C, within each asset class, i, is equal to the aggregate return of that asset (i.e the average change in the prices of houses or equities). This assumption reflects the lack of individual level data on the composition of the portfolio for different wealth groups. I will discuss the potential effects of this assumption on the decomposition exercise at the end of the section.

Using equation 1, we can obtain an expression for the share of wealth going to group C as follows:

\[
\frac{W_{C,t+1}}{W_{t+1}} = \frac{W_t^C}{W_t} \left[ 1 + \frac{s_t^C Y_t^C}{W_t^C} \right] \frac{1 + q_t^C}{1 + q_t} \frac{1 + s_t Y_t}{1 + s_t W_t} \]

(3)

Notice that, in equation 3, the flow of savings, \(S_t^C\), is expressed as \(s_t^C Y_t^C\), where \(s_t^C\) and \(Y_t^C\) denote the net saving rate and disposable income of group C respectively. Changes in the share of wealth owned by group C are a function of the ratios \(s_t^C/s_t\), \(q_t^C/q_t\) and \(Y_t^C/Y_t\.

Intuitively, if the saving rate, the capital gains or the income of group C are above the average then the share of wealth going to this group will increase. If, instead of considering the change between t and t+1, we are interested in the change between t and t+j, it is possible to iterate on equation 3 and obtain:

\[
\frac{W_{C,t+j}}{W_{t+j}} = \frac{W_t^C}{W_t} \prod_{i=t}^{t+j-1} \left[ 1 + \frac{s_i^C Y_i^C}{W_i^C} \right] \frac{1 + q_i^C}{1 + q_i} \frac{1 + s_i Y_i}{1 + s_i W_i} \]

(4)

Let \(a_t^{t+j}\) and \(b_t^{t+j}\) be defined as:

\[
a_t^{t+j} = \prod_{i=t}^{t+j-1} \left[ 1 + \frac{s_i^C Y_i^C}{W_i^C} \right] \frac{1 + q_i^C}{1 + q_i} \frac{1 + s_i Y_i}{1 + s_i W_i} \\
b_t^{t+j} = \prod_{i=t}^{t+j-1} \left[ 1 + \frac{s_i^C Y_i^C}{W_i^C} \right] \frac{1 + q_i^C}{1 + q_i} \frac{1 + s_i Y_i}{1 + s_i W_i} 
\]

Then,

\[
\frac{W_{C,t+j}}{W_{t+j}} = \frac{W_t^C}{W_t} \ a_t^{t+j} \ b_t^{t+j} 
\]

(4)

Thus, \(a_t^{t+j}\) and \(b_t^{t+j}\) summarize the change in the wealth share of group C between t and t+j due to capital gains and savings respectively. To bring this framework to the data requires
series for \( \{W_t, q_t\} \) and \( \{W_t^C, q_t^C\} \). Given that my interest is in the increase in wealth concentration and that the groups inside of the top 1% played an important role in it, I use the data provided by Saez and Zucman (2016) to compute equation 4. Their estimations of the top wealth shares are based on administrative tax data and have the advantage of capturing the movements at the very top of the wealth distribution. Alternative estimates based on the Survey of Consumer Finance (Wolff (2014)) sub-estimate families at the very top and hence are not ideal for my purposes.

\( W_t \), the value of total net household wealth, is taken from the US Financial Accounts. For each type of asset, \( i \), \( q_t(i) \) is estimated as a residual using \( W_t(i) \) from the Financial Accounts and the Investment flows for each period which are obtained from additional sources. The movements that are not explained by investment flows are attributed to price changes. The aggregate capital gains, \( q_t \), are then defined as the weighted average of each asset capital gains, \( q_t(i) \):

\[
q_t = \sum_{i \in I} \left[ \frac{W_t(i) + S_t(i)}{W_t + S_t} \right] q_t(i)
\]

The value of \( W_t^C \) for the bottom 90% and different groups inside the top 10% of the wealth distribution comes from the estimations of Saez and Zucman (2016). \( W_t^C \) is obtained using the capitalization method. For each asset class, \( i \), we observe both its aggregate stock, \( W_t(i) \), (from the US financial Accounts) and its total capital income reported by taxpayers to the IRS, \( I_t(i) \). The capitalization factor for this asset is then defined as \( F_t(i) = \frac{W_t(i)}{I_t(i)} \), which is used to estimate \( W_t^C(i) \) as \( I_t^C(i)F_t(i) \). Furthermore, the series of \( q_t^C \) makes it possible to construct \( q_t^C \) using equation 2.

Figure 3 presents the values of \( a_{t+j}^t \) and \( b_{t+j}^t \) for the top 1% for \( t = 1980 \) and \( j = 1, ..., 32 \). The bulk of the increase in the wealth share of the top 1% was due to volume effects (changes in relative savings and disposable income). The role of changes in relative capital gains was marginal and negative in the years after the bust of the dot com bubble. The reason for this finding is the following: the bottom groups held most of their wealth in housing, and this
was the best-performing asset in terms of capital gains. The increase in the price of housing between 1980 and 2007 was a force toward equality. The subsequent crash partially reverted this situation.

Figure 3: \( a_t^{t+j}, b_t^{t+j} \) and \( a_t^{t+j} b_t^{t+j} \) for the Top 1%.
Source: Saez and Zucman (2016).

Figure 4 presents the compounded capital gains by each asset class, \( C_t^{d+j} (i) \), for \( t = 1980 \) and \( j = 1, 2, ..., 32 \).

\[
C_t^{d+j} (i) = \prod_{k=t}^{t+j-1} \left[ 1 + q_k^C (i) \right] = [1 + q_t^C (i)][1 + q_{t+1}^C (i)] \cdots [1 + q_{t+j-1}^C (i)]
\]

From Figure 3, it can be concluded that most of the increase in wealth concentration was due to volume effects. In other words, the growth rate of wealth of the top 1% coming from savings was consistently higher than that of the average between 1980 and 2012:

\[
\frac{s_t^{Top1\%} Y_t^{Top1\%}}{W_t^{Top1\%}} \geq \frac{s_t Y_t}{W_t} \quad \text{for} \quad t = 1980, ..., 2011
\]

To understand the mechanics behind this fact, it is necessary to observe the behavior of \( s_t^e/s_t \) and \( Y_t^e/Y_t \) for different groups. From equation 1 one obtains \( s_t^e \) as a residual:
Figure 4: Compounded capital gains by asset class.
Note: author calculation using data from Saez and Zucman (2016).

\[ s_t^C = \left[ \frac{W_{t+1}^C}{q_t^C} - W_t^C \right] \frac{1}{Y_t^C} \]  

The saving rate of the bottom 90%, the [90-99]% and the top 1% are presented in Figure 5:

Figure 5: Saving rate by wealth class
Source: Saez and Zucman (2016).
The saving rate of the bottom 90% decreased continuously from a stable average of 5% between 1950 and 1980 to -8% in 2006. It then returned to 0% after the Great Recession. The saving rate of the group [90-99]% followed a similar pattern. It was stable at approximately 29% between 1950 and 1980 and then started falling until reaching 0% in 2000. It then partially recovered and reached 14% in 2011. The saving rate of the top 1% showed much more variability than that of the other groups. However, overall, it went from an average of 30% between 1950 and 1980 to 37% between 1980 and 2012. It is important to mention that these saving rates are “synthetic” in the sense that they abstract from mobility between groups. This limitation comes from the fact that it is not possible to follow a particular group of individuals over time using administrative tax data. The same situation emerges when using the Survey of Consumer Finances.8

The fact that the bottom 90% and the [90-99]% decreased their saving rates while the top 1% increased slightly explains an important fraction of the increase in wealth concentration. In addition to this fact, there was also movement in the share of total income going to each wealth group, \( Y^C_t / Y_t \). Figure 6 summarize these changes. The income share of the bottom 90% was very stable at approximately 69% between 1950 and 1985. It then decreased continuously until reaching 60% in 2012. The income share of the group [90-99]% was stable at approximately 22% between 1950 and 2012. There was not a major trend during this period.

The share going to the richest 1% went from 9% in 1980 to 17.9% in 2012. It almost doubled in a period of 32 years.

2.2 Counter-Factual Exercise: The Effect of Changes in Saving Patterns on Wealth Inequality

To quantify the contribution of the changing saving patterns to the increase in wealth concentration, I construct hypothetical wealth holdings for the bottom 90% and the [90-99] % under the assumption that their saving rates are kept at their 1950-1980 average between 1980 and

8Bosworth and Anders (2008) estimates the saving rates by wealth groups using supplementary questions of the PSDI and concludes that measurement errors are a particularly serious problem with this panel.
2012. Then, I estimate the wealth shares implied by these hypothetical series. The original accumulation equation is given by:

\[ W_{t+1}^C = [1 + q_t^C][W_t^C + s_t^C Y_t^C] \]

To construct the hypothetical wealth holdings for the bottom 90% and the [90-99]% groups between 1980 and 2012, I substitute \( s_t^C \) by \( \bar{s}_t^C \) which is defined as the average saving rate of group C between 1950 and 1980. The hypothetical wealth holdings are given by:

\[ \widetilde{W}_{t+1}^C = [1 + q_t^C][W_t^C + \bar{s}_t^C Y_t^C] \]

I then use the series of \( \widetilde{W}_t^C \) to compute hypothetical wealth shares under three different assumptions: i) the bottom 90% is assumed to save its historical average of 5% between 1980 and 2012, ii) the group [90-99]% is assumed to save its historical average of 29% between 1980 and 2012, iii) both the bottom 90% and the [90-99]% are assumed to save their 1950-1980 average between 1980 and 2012. The results of these simulations are presented in Figure 7.

Figure 6: Share of Disposable Income by wealth class
Source: Saez and Zucman (2016).
If we assume that the bottom 90% kept saving at their historical average of 5%, then the share of the top 1% would have increased from 24% to 34% instead of 42%. This attributes around 40% of the top percentage’s increase to the change in the saving behavior of the bottom 90%. Assuming that both the bottom 90% and the [90-99]% did not change their saving behavior between 1980 and 2012, the share of the top 1% would have reached 30% in 2012 instead of the actual 42%. Notice that even if the saving rates of the bottom 90% and the [90-99] had remained constant, there would have been an increase of 6 points in the share of the top 1%. This is mainly because the share of total labor income going to the richest 1% increased from 4.7% in 1980 to 8.9% in 2012. It is also because of the slight increase in the saving rate of this group. It is important to mention that these estimations ignore general equilibrium effects, and they are an accounting exercise that provide us with a first-imperfect-estimation of the importance of the change in the pattern of savings.
3 Model

In Section 2, we established that approximately 40% to 60% of the increase in the wealth share of the top 1% can be attributed to the change in the saving patterns of the bottom 90% and the [90-99]%. Given this fact, it logically follows that studying the causes of the decrease in the saving rates of these groups is fundamental to understanding the increase in wealth concentration. The empirical literature has found that changes in credit conditions, defined as shocks that increase the household’s ability to borrow, contributed significantly to the decrease in the saving rate of the bottom groups. Mian and Sufi (2011) show that the increase in the price of housing, which is typically used as collateral for borrowing, explains a good portion of the increase in non-mortgage debt and the decline in the saving rates between 2002 and 2006. Aladangady (2017) confirmed this result using a richer data set. This evidence suggests that an important proportion of households were credit constrained. Carroll (1997) argues that various financial innovations allowed households to transform future income into current purchasing power, thereby reducing the saving rates of previously credit-constrained households. Along the same lines, Parker (2000) concludes that the increase in debt can explain one-third of the observed decline in the overall saving rate.

In the next section, I construct a general equilibrium version of the Bewley model, to quantify the importance of different plausible hypotheses on the evolution of the wealth distribution: i) changes in credit conditions, ii) increase in the concentration and riskiness of labor income, and iii) reforms to the tax code (reduction in corporate and personal income taxation). Changes in credit conditions are defined as increases in the ability of households to borrow and are parsimoniously modeled as a loosening of borrowing constraints.9 The initial stationary distribution of the model is calibrated to match the wealth distribution in 1980. The values for the taxation scheme and the labor income process after 1980 are taken from the data, and the path for the borrowing constraints is calibrated to match the evolution of the ratio of non-mortgage debt to disposable income.

---

9Carroll, Slacalek, and Sommer (2019) followed a similar strategy to explain the decrease in the savings rate. Their goal was not to match the distribution of wealth.
As mentioned in Section 1.1, the two papers most similar to mine are those of Kaymak and Poschke (2016) and Hubmer, Krusell, and Smith (2020). Both papers use different variations of the Bewley model to study the effects of i) changes in the labor income process (higher concentration and riskiness) and ii) reforms in the tax code (reductions in progressivity and corporate taxation) on the distribution of wealth in the previous 50 years. While there is no doubt that these shocks are important for the saving decisions of individuals, the overall effect of them in the Bewley model is an increase in the total saving rate of the economy. Given that the leading papers in the field abstract from the forces that decreased the saving rate of the bottom groups, my paper contributes to the literature by explicitly modeling one of the key forces behind this decrease: changes in credit conditions. I find that this channel is crucial to matching the observed patterns of the saving rates and hence the causes responsible for the increase in wealth concentration.

3.1 Description of the Model

The economy is populated with a continuum of infinitely lived, ex-ante identical agents that choose streams of consumption to maximize their expected lifetime utility:

$$E_0 \left\{ \sum_{t=0}^{\infty} \left[ \prod_{s=0}^{t} \beta_s \right] u(c_s) \right\}$$  \hspace{1cm} (6)

The instantaneous utility function takes the CRRA form with parameter $\sigma$:

$$u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma}$$

Each period, the agent faces the following budget constraint:

$$c_t + [a_{t+1} - a_t] = [w_t l_t(p_t, v_t) + r_t a_t][1 - \tau_t(I_t)] + T_t$$  \hspace{1cm} (7)

$$I_t = w_t l_t(p_t, v_t) + r_t a_t$$  \hspace{1cm} (8)
In equation 7, \( c_t \) denotes the level of the consumption good and \( a_{t+1} \) is an asset that provides \((1 + r_{t+1})a_{t+1}\) units of the consumption good the next period. \( w_t \) and \( r_t \) are the market wage and real interest rate respectively. The agent does not value leisure and always supplies the full amount of effective units of labor, \( l_t(p_t, v_t) \). \( I_t \) denotes the total pretax income from labor and capital which is taxed at a rate \( \tau_t(I_t) \). The tax rate \( \tau_t(I_t) \) is an increasing function of total income \( I_t \). Lastly, \( T_t \) is a lump-sum transfer financed with the proceeds from taxation. In addition to the budget constraint, the agent also faces a borrowing constraint:

\[
a_t \geq a_t
\]  
(9)

There are two sources of uncertainty for the agent: the discount factor, \( \beta_t \) and the effective units of labor, \( l_t(p_t, v_t) \). At every period \( t \), the factor at which the agent discounts the next period consumption, \( \beta_{t+1} \), is known. However, the discount factor \( \beta_{t+2} \) is stochastic and assumed to follow a Markov process:

\[
\beta_t = \rho^\beta \beta_{t-1} + (1 - \rho^\beta)\mu^\beta + \epsilon_t^\beta \quad \epsilon_t^\beta \sim N(0, \sigma^\beta)
\]  
(10)

Adding heterogeneity in the discount factor was introduced by Krusell, Smith, and Jr. (1998) and is a parsimonious way of obtaining a realistic level of wealth concentration. As mentioned in Section 1.1, adding labor income heterogeneity to an Aiyagari model is not enough to replicate the right tail of the distribution of wealth observed in the data. This indicates that there are other factors affecting the saving decisions of households and hence the distribution of wealth. De Nardi and Fella (2017) summarized all of the mechanisms explored in the literature to generate a realistic concentration of wealth. The decision to use heterogeneous discount factors rather than bequest motives, capital income risk, non-observable extraordinary states of labor earnings, or returns that depending of the level of wealth, was based on the simplicity of this approach and on the fact that the logic of the main result of this paper does not crucially depend on stochastic discount factors. Heterogeneous discount factors can be seen as a reduced form to introduce bequest motives, or entrepreneurial income risk.
These motives generate a certain group of individuals that save more than others and keep accumulating wealth in spite of being at the top. The introduction of heterogeneous discount factors generates the same result (a percentage of agents that are more prone to saving) and it comes with the benefit of simplifying the computation of transitions.

The effective units of labor, \( l_t(p_t, v_t) \), are a function of \( p_t \) and \( v_t \), the persistent and transitory components of labor income. They are assumed to behave according to:

\[
v_t \sim N(0, \sigma^v_t) \tag{11}
\]

\[
p_t = \rho^p p_{t-1} + \epsilon_t^p \quad \epsilon_t^p \sim N(0, \sigma^p_t) \tag{12}
\]

Piketty and Saez (2003) estimated top wage shares for the USA and found that the share of the total wage bill going to the top wage earners has increased sharply in the previous 35 years. To be able to match this fact, I follow Hubmer, Krusell, and Smith (2020) and use the following functional form for \( l_t(p_t, v_t) \):

\[
l_t(p_t, v_t) = \psi_t(p_t) \exp(v_t) \tag{13}
\]

Let \( F_{p_t} \) denote the unconditional cumulative distributive function of \( p_t \). The function \( \psi_t(p_t) \) is defined as follows: if the value of \( p_t \) is less than or equal to the 90 percentile of \( F_{p_t} \) then \( \psi_t(p_t) = \exp(p_t) \) (the standard case); if the realization of \( p_t \) is in the top 10% then \( \psi_t(p_t) \) takes the following form:

\[
F_{\kappa_t}^{-1}\left(\frac{F_{p_t}(p_t) - 0.9}{1 - 0.9}\right) \tag{14}
\]

\( F_{\kappa_t} \) is the cumulative distributive function of a Pareto distribution with coefficient \( \kappa_t \) and lower bound \( \exp(F_{p_t}(0.9)) \). Since the right tail of the labor income distribution is very well approximated by a Pareto distribution, changes in the tail parameter \( \kappa_t \) can be used to match the fraction of labor income going to the top 10%, top 1% and the top .1%. Adding a Pareto tail to the, otherwise standard, labor income estimation process using data from the PSID allow us to replicate the deviations from log-normality (high levels of kurtosis and negative
skewness) documented by Guvenen et al. (2015) using administrative data.

So far, we described the problem that each agent faces taking the sequence of \( w_t, r_t \) and the taxation scheme as given. The production function is Cobb-Douglas, firms are perfectly competitive and the overall supply of effective units of labor is equal to 1. The market wage and real interest rate are then:

\[
\begin{align*}
    r_t &= F^K(K_t, L_t) - \delta \\
    w_t &= F^L(K_t, L_t)
\end{align*}
\]  

The government budget constraint is always balanced: the homogeneous transfer, \( T_t \), is financed with resources coming from taxation. A stationary equilibrium in this economy is a level of capital \( K^* \) and the prices associated to it \( r^* \) and \( w^* \) such that: given \( r^*, w^* \) and a taxes and transfer system \( (\tau^*, T^*) \), the policy function of the agents induce an invariant distribution over the unique asset in the economy \( a_t \). In equilibrium, the assets held by the agents are equal to the aggregate capital \( K^* \).

### 3.2 Description of the Simulations and Calibration

The objective of the model is to quantify the relative importance of (i) changes in credit conditions, (ii) increase in the concentration and riskiness of labor income and, (iii) reforms to the tax code (reduction in corporate and personal income taxation) on the evolution of wealth concentration after 1980. To achieve this goal, I use the following strategy:

1. Calibrate the parameters \( \rho^\beta, \mu^\beta, \sigma^\beta, \) and \( \delta \) to match the share of the top 10%, top 1% and top .1% as well as the ratio K/Y observed in the data in 1980.

2. I then fed the observed paths for the taxation scheme \( (\tau_t) \), the concentration of labor income \( (\kappa_t) \) and the riskiness of labor income \( (\sigma^p_t \) and \( \sigma^v_t) \) between 1980 and 2012 into the model.

3. The path of \( a_t \) between 1980 and 2012 is set to match the change in the ratio of non-mortgage debt to disposable income for the bottom 90% (Figure 8).
The values for $\tau_t$, $\kappa_t$, $\sigma^p_t$, $\sigma^v_t$ and $\omega_t$ are assumed to remain fixed at their last observed value. The model eventually converges to a new stationary distribution associated to these values. During the transition from the original stationary distribution to the new one, it is assumed that the agents have full knowledge of the path of the shocks. This is an assumption that will be relaxed as a robustness check of my results. Once the transition is computed, then it is possible to compare the prediction of the model with the data and perform counterfactual exercises to estimate the relative importance of each one of the shocks in the evolution of the wealth distribution.

The key assumption behind calibrating $\rho^\beta$, $\mu^\beta$, $\sigma^\beta$ to 1980 is that the distribution of wealth was stationary in the years close to 1980. Based on this assumption, we can measure the role of taxes, labor income on the distribution of wealth and then we compute the role of the heterogeneous $\beta$s factor as a residual. This residual is meant to capture the forces driving concentration in addition to the ones modeled explicitly. The identifying assumption is then that these fundamental forces are stable over time and then we could measure the role of taxes, labor income and the borrowing constraint on the distribution of wealth. As it was mentioned before, heterogeneous discount factors are a reduce form way to generate a saving behavior that is consistent with the one produced by non-homothetic bequest motives and heterogeneous capital income risk. In other words, their role is to generate more savings at the top of the wealth distribution. Given that these mechanisms are not able to generate less saving at the bottom, as observed in the data, it is reasonable to assume that changes in the parameters describing the stochastic behavior of the heterogeneous discount factor are not a plausible candidate to explain the saving patterns of the data.\footnote{The alternative to calibrating the parameters $\beta$, $\mu^\beta$ and $\sigma^\beta$ using 1980 as reference would be to calibrate a path for them over the transition. The issue with this approach is that it would generate indeterminacy given than both the borrowing constraints and these parameters could be used to match the observed ratio of net debt over disposable income. In other words, there are infinite combinations of paths for both the borrowing constraint and the parameters of the $\beta$ process that are able to generate the path of net debt over disposable income.}

The path of $(\sigma^p_t, \sigma^v_t)$ is taken from Heathcote, Storesletten, and Violante (2010) who estimated the values for the standard deviations of the permanent and transitory shocks to labor income from 1967 to 2000 using data from the PSID (Figure 9). The path for $\kappa_t$ is chosen to
match the share of total wage income going to the top wage earners estimated by Piketty and Saez (2003). Since the right tail of the distribution of labor income is approximately Pareto, changing the parameter $\kappa_t$ in equation 16 allows me to closely match the top wage shares every period. Figure 10 presents the top wage shares in the last 50 years.

Figure 9: Cross-sectional Standard Deviations
Source: Heathcote, Storesletten, and Violante (2010).
Piketty and Saez (2007) estimated the effective tax rate for eleven income brackets for the years 1967 to 2000. The function $\tau(I_t)$ is a step function that is calibrated to match the changes in corporate and progressive taxation after 1980. After 2000, it is assumed that taxes are fixed at this level. The average of the ratio capital-output during 1960 - 1980 -using Piketty and Zucman (2014) data- was 3.9 which implies that $\mu^\beta = .917$ and $\delta = .05$. The parameters $\rho^\beta$ and $\sigma^\beta$ are calibrated to match as closely as possible the shares held by the top 10% and top 1% of the wealth distribution in 1980 ($\rho^\beta = .986$, $\sigma^\beta = .0041$).

The value of $a_t$ in the initial stationary distribution is the one that match the average ratio debt to income for the bottom 90% between 1960 - 1980. ($a_{ss} = -.41$). The shares of the top groups come from Saez and Zucman (2016) and the share of the Bottom 50% comes from the Survey of Consumer Finances. Given that the top of the wealth distribution is very well approximated by a Pareto distribution and that the type of uncertainty in the model yields a stationary distribution that is Pareto at the top it is possible to match very closely the distribution in 1980:
### Table 1: Distribution of Wealth in 1980: data and stationary distribution

Data: Saez and Zucman (2016).

<table>
<thead>
<tr>
<th></th>
<th>Top 10%</th>
<th>Top 1%</th>
<th>Bottom 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>65.1 %</td>
<td>23.5 %</td>
<td>1.1 %</td>
</tr>
<tr>
<td>Model</td>
<td>64.8 %</td>
<td>25.79 %</td>
<td>1.5 %</td>
</tr>
</tbody>
</table>

### 4 Results

Figure 11 presents the share of the top 1% estimated by Saez and Zucman (2016) against the share of the top 1% implied by the model. The fit of the model is good, except for the Dot-Com bubble and the Great Recession in which the predictions of the model and the data diverge. The change in the distribution of wealth in a Bewley model depends on the saving behavior of different groups, and it abstracts from changes in the relative prices of assets. For this reason, any change in the shares that comes from large movements on the price of assets (such as the Dot-Com bubble and the aftermath of the Great Recession) is not captured in this class of models. For example, the increase in the share of the top 1% between 2007 and 2012 was mostly due to the fact that the price of housing, the main asset held by the bottom 90%, decreased significantly relative to other assets.

As we can see in Table 2, the model does a better job at capturing the overall direction of the shares of the Top 1% and the Top 10% relative to the Top .1% and .01%. This is explained due to the wealth of individuals inside of the Top .1% being highly reactive to the price of assets (a component not captured in the model).

### Table 2: Top shares in 2012: data and model with all shocks

Data: Saez and Zucman (2016).

<table>
<thead>
<tr>
<th></th>
<th>Top 10%</th>
<th>Top 1%</th>
<th>Top.1%</th>
<th>Top.01%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>77.2 %</td>
<td>41.8 %</td>
<td>22.0%</td>
<td>11.2%</td>
</tr>
<tr>
<td>Model</td>
<td>79.3 %</td>
<td>37.8 %</td>
<td>16.7%</td>
<td>6.9%</td>
</tr>
</tbody>
</table>

In Figure 12, I evaluate the importance of the credit market channel (loosening of borrowing constraints). When all the other shocks, except for the changes in borrowing constraints, are fed into the model, the share of the top 1% increases only half as much compared to
the baseline scenario in which the credit market channel is included. The relaxation of the borrowing constraints causes the originally constrained households to accumulate more debt and decreases the precautionary savings of those “close” to the constraint since the likelihood of hitting it decreases. In addition to these two forces, the increase in the real interest rate caused by the contraction in the overall supply of savings increased the saving rate of the top groups, which further increased wealth concentration. The credit channel is fundamental to matching the pattern of savings observed in the data: a decrease in the overall saving rate fueled by the change of behavior of the bottom groups and a slight increase in the saving rate of the top 1%.

<table>
<thead>
<tr>
<th>Year</th>
<th>Data</th>
<th>All shocks</th>
<th>All shocks - $\alpha_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>24.3 %</td>
<td>24.6 %</td>
<td>24.6 %</td>
</tr>
<tr>
<td>2012</td>
<td>41.8 %</td>
<td>37.8 %</td>
<td>32.4 %</td>
</tr>
<tr>
<td>Change</td>
<td>17.4 %</td>
<td>13.2 %</td>
<td>7.8 %</td>
</tr>
</tbody>
</table>

Table 3: Wealth share of the top 1%: data and simulations
Data: Saez and Zucman (2016).
The increase in the concentration of labor income and the tax reforms during this period also contributed to the concentration of wealth. The top wage earners tend to be the wealthiest agents in the economy, and a higher share of total income going to this group mechanically increases their savings flow relative to other groups, thereby increasing concentration. The reduction in the progressivity of personal and corporate taxation increased the incentive to save for individuals from the top groups, especially from the top 1%, since they faced significantly lower tax rates on the return to capital. A higher saving rate for top groups combined with a stable saving rate for other groups generated wealth concentration. When all the shocks except for the higher concentration of labor income and the changes in taxation are included, the share of wealth owned by the top 1% increased by only 45% compared with the baseline case. The last of the shocks, the increase in the riskiness of labor income, was a force toward equality. A riskier income process increases overall precautionary savings, particularly for the bottom groups, which are motivated to avoid the borrowing constraint. If we consider this shock only, the share of the top 1% would decrease from 24% to 21%.
In Figure 13, I consider the effect of three types of shocks in isolation: i) changes in the credit conditions or loosening in the borrowing constraints (dotted line), ii) increase in concentration of labor income (κt) and changes in taxation (dashed line), and iii) increase in the riskiness of labor income (solid line).

Figure 13: Wealth share of the top 1%: each shock in isolation
Data: Saez and Zucman (2016).

<table>
<thead>
<tr>
<th>Year</th>
<th>Data</th>
<th>All shocks</th>
<th>dt</th>
<th>τt(·) &amp; κt</th>
<th>σv^p &amp; σv^p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>24.3%</td>
<td>24.6%</td>
<td>24.6%</td>
<td>24.6%</td>
<td>24.6%</td>
</tr>
<tr>
<td>2012</td>
<td>41.8%</td>
<td>37.8%</td>
<td>29.39%</td>
<td>32.4%</td>
<td>21.7%</td>
</tr>
<tr>
<td>Change</td>
<td>17.4 %</td>
<td>13.2 %</td>
<td>4.79%</td>
<td>7.8%</td>
<td>-2.9%</td>
</tr>
</tbody>
</table>

Table 4: Wealth share of the top 1%: data and simulations
Data: Saez and Zucman (2016).

The overall effect of changes in the tax system, τt(·), and changes in the labor income process, (κ, σ^p_t, σ^v_t) is an increase in wealth inequality. The increase in concentration driven by τt(·) and κt is larger than the decrease induced by σ^p_t and σ^v_t. The total effect of these shocks is an increase in the overall saving rate of the economy driven by higher precautionary sav-
ings and higher savings for the top groups induced by lower taxation. This is a counterfactual result since, in the data, the saving rate went from 12% in 1980 to 2% in 2007. This suggests that the credit market channel plays a major role, and that only the interplay between this channel and the other two shocks allows us to match both the observed level of concentration and the behavior of the overall saving rate.

5 Conclusions

In this paper, I first examine the data and conclude that between 40% and 60% of the increase in wealth concentration, between 1980 and 2012, can be attributed to the dramatic decrease in the saving rate of the households at the bottom 99% of the wealth distribution. I then construct a general equilibrium model to evaluate the importance of three forces on the evolution of the wealth distribution: i) changes in credit conditions (increase in the ability to borrow), ii) reforms to the tax code, and iii) changes in the labor income process. The model captures the increase in concentration. According to counterfactual exercises, the loosening of borrowing constraints explains close to half of the increase in the share of wealth going to the top 1%. Changes in credit conditions are crucial to matching the decrease in the saving rate of the bottom groups. To the best of my knowledge, this is the first attempt to explicitly model this channel in the quantitative literature studying the causes for the increase in wealth concentration after 1980.

There are two aspects in which this study can be improved in future work: (1) incorporate the observed capital gains by wealth class, as this would improve the fit of the model with respect to short and medium term movements caused by asset price fluctuations; (2) explicitly modeling the role of housing in the relaxation of the borrowing constraints as this would provide a micro founded justification to movements in the ability to borrow. The results of this paper also has public policy implications: even when the decrease in the progressivity of taxation and the concentration of labor earnings do explain an important fraction of the increase in concentration, it is key to consider the role of the forces that decreased the saving
rate of the bottom groups. According to our simulations, increasing the progressivity of taxation and reducing labor earnings inequality to levels similar to the ones observed in 1980 would only be able to reduce about half of the observed increase in wealth concentration. Any effective policy to reduce wealth concentration must incorporate features that increase the saving rate of the bottom groups.
References


6 Appendix

6.1 Modeling the Current Account Deficit

In this section the main analysis presented in Section 4 is expanded to consider the current account deficit observed in the data. This robustness check is introduced because a large influx of foreign savings was one of the main drivers behind changes in the credit supply along with a loosening of credit conditions as pointed out by Bernanke (2005). We first study the evolution of the personal saving rate, the current account and the real interest rate after 1980. Then, we compare how well our baseline model replicates the behavior of these variables relative to a model that explicitly incorporates a current account deficit like the one observed in the data.

As we can see in Figure 14, the current account as a fraction of GDP went from an average of zero between 1960 - 1980 to -4.8% in 2007. The saving rate went from 11.1% in 1980 to 3.7% in 2007, and the real interest rate displayed a decreasing trend reaching a maximum of 8.5% in 1981 and a minimum of 1.13% in 2011. Our baseline model assumes a current account that is equal to zero, and, under this assumption, the relaxation of the borrowing

Figure 14: CA/GDP, Personal Saving Rate, Real Interest Rate: 1980 - 2012.
Sources: FRED and NIPA.
constraints contracts the overall supply of savings in the economy generating an increase in the equilibrium real interest rate, a result that is at odds with the data. Along the same lines, the model predicts an increase in the saving rate of the top 1% after 1985 while in the data there saving rate of this group has not displayed an increasing trend (see Figure 5).

Under our baseline assumption of a close economy, the supply of savings from households is equal to the demand of capital by firms. If we relax this assumption and allow an exogenous current account deficit, the equilibrium interest rate would be lower as the supply of capital would come both from the domestic country and the rest of the world. Following this logic, adding a current account deficit would potentially allow us to better replicate the observed decrease in the real interest rate and the savings of the top 1%, especially after 1985. The equations that allow us to pin down the equilibrium real interest rate under an exogenous current account are the following:

\[
F_t + \int_{\beta} \int_{p} \int_{a} a \, dG^*(a, p, \beta) \, da \, dp \, d\beta = F_t + A_t = K_t \tag{17}
\]

\[
\frac{F_t}{F(K_t, 1))} = L_t \tag{18}
\]

In the previous equations, \(F_T\) and \(L_T\) represent the net foreign savings in the domestic economy and the ratio Current Account over GDP observed in the data.

In this context, we compare the long run saving rate, real interest rate, and the distributions of wealth of our baseline model with one in which we the long run ratio of current account over GDP is equal to -4.17% which is the average of this variable between 2002 and 2012. In table 4, we observe that the open economy version of the model does a better job at replicating both the decrease in the overall saving rate and also the decrease in the equilibrium interest rate.\(^{11}\) It does, however, generate less concentration. This result is driven by the following mechanism: a lower equilibrium real interest rate generated by an influx of foreign resources reduces

\(^{11}\) \(s_{SD}\) and \(r_{SD}\) refer to the long run values for the saving rate and the real interest rate respectively. The top wealth shares presented in the table are the ones corresponding to the long run stationary distribution implied for the model in the year 2070 absent any policy change.
the incentive to accumulate wealth for all agents in the economy. However, individuals at
the borrowing constraint cannot reduce their asset holdings, while individuals close to the
borrowing constraint, and actively trying to avoid it, have an additional saving motive with
respect to agents at the top. The overall result of an lower real interest rate, is a mild reduction
in wealth concentration as individuals at the top diminish their holdings more than individuals
at the bottom. Summarizing, explicitly modeling the current account deficit help the model to
better replicate the behavior of the real interest rate, the overall saving rate and, particularly,
the one of the top 1%. It does, however, generates a small contraction in wealth concentration,
indicating that the loosening in borrowing constraints is needed to match the observed facts.

<table>
<thead>
<tr>
<th>Model</th>
<th>$s_{SD}/s_{1980}$</th>
<th>$r_{SD}/r_{1980}$</th>
<th>Top 1%</th>
<th>Top 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA=0</td>
<td>.77</td>
<td>1.1</td>
<td>49.2%</td>
<td>83.5%</td>
</tr>
<tr>
<td>CA=-4.17%</td>
<td>.41</td>
<td>.62</td>
<td>46.9%</td>
<td>79.8%</td>
</tr>
</tbody>
</table>

Table 5: Model Comparison: Baseline & Open Economy

6.2 Approximation to the Value Function and Policy Function

Each agent maximizes its expected utility given a sequence for the tax scheme, the wage rate
and the real return on assets $\{w_t, r_t, \tau_t(\cdot)\}^{\infty}_{t=0}$. In a stationary distribution, $w_t, r_t, \tau_t(\cdot)$ are
fixed over time and the consumer problem can be written recursively as follows:

$$
V(a, p, \beta) = \max \left\{ u(c) + \beta E\{V(a', \beta', p') \mid (\beta, p)\} \right\}
$$

(19)

And $(c, a')$ must satisfy the budget constraint:

$$
c + a' = [w_t l_t(p_t, v_t) + r_t a_t][1 - \tau_t(I_t)] + T_t + a
$$

(20)

For a given $w_t, r_t, \tau_t(\cdot)$, there are three state variables $(a, p, \beta)$ and one control variable $a'$.
Let $V_0$ be an arbitrary $R^3 \rightarrow R$ function and let the sequence $\{V_t\}^{\infty}_{t=0}$ be defined by:
\[ V_{t+1}(a, p, \beta) = \max \left\{ u(c) + \beta E\{V_t(a', \beta', p') \mid (\beta, p)\} \right\} \]  

(21)

Given that \( u(\cdot) \) is an strictly concave function, we know that for any well defined function, \( V_0 : R^3 \rightarrow R \), the contraction mapping theorem guarantees that the sequence converges to the function \( V^* \) and that \( V^* = V \). To numerically approximate \( V^* \), I first discretize the values of the state and variables \( p \), and \( \beta \) using the method developed by Rouwenhorst (1995). The number of points in the approximation of \( p \), \( n_p \), must be sufficiently large so that individuals at the top of the unconditional distribution of \( p \) [top 10\%, top 1\%, .1\%, .01\%] are considered. This is important, because one of the shocks I consider is changes in the mass of wage income going to the very top. The values of \( \{\beta_1, \beta_2, \ldots, \beta_{n_{\beta}}\} \) are also approximated using the Rouwenhorst method with \( n_\beta = 3 \). This is equivalent to the seminal example studied by Krusell, Smith, and Jr. (1998). To discretize \( a \), I use a log-spaced grid with 100 points, a lower bound equal to \( a \) and a top bound, \( a_{n_a} \) equal to a large multiple of average wealth.

The process to obtain an approximation to \( V^* \) is straight forward: i) propose a matrix \( V_0 \), ii) obtain the sequence \( \{V_t\}_{t=0}^\infty \) implied by equation 21 , iii) the approximation of \( V^* \) is reached when \( \|V_t - V_{t-1}\| \) is less than or equal than 1e-4. To speed up the convergence process, I use the Endogenous Grid-point Method (EGM) developed by Carroll (2005). In the traditional solution method, an optimal level for \( a' \) associated to each state vector \((a, p, \beta)\) is obtained by solving equation 21. This process is computationally costly as it requires to numerically optimize for each \((a, p, \beta)\) and at every iteration step. The idea behind the EGM is the following: instead of using the grid for the asset on \( a \), denoted \( A^G = \{a_1, a_2, \ldots, a_{n_A}\} \) we use it on \( a' \) and then find the \( a \) that “supports” the election of \( a' \) as the optimum: take \((p, \beta)\) and \( a' \) such that all are elements of their respective grids, \((p, \beta, a') \in P^G \times \beta^G \times A^G \) at the optimum it must be that:

\[ u'(c^*(p, \beta; a')) = \beta E\{V_{a,t}(a', \beta', p') \mid (\beta, p)\} \]  

(22)

\( c^*(p, \beta; a') \) is defined as the optimal level of consumption given that the state variables are
\((p, \beta)\) and that the optimal choice of the asset for the next period is \(a'\). Inverting the function \(u(\cdot)\) allow us to obtain \(c^*(p, \beta; a')\) as:

\[
c^*(p, \beta; a') = u'^{-1}\left\{\beta E\{V_{a,t}(a', \beta', p') \mid (\beta, p)\}\right\}
\]

With the value of \(c^*(p, \beta; a')\) and the budget constraint it is possible to obtain the value of \(a\) that supports the election of \(a'\) as optimal given \(p\) and \(\beta\):

\[
c^*(p, \beta; a') + a' = [w_t l_t(p_t, v_t) + r_t a_t][1 - \tau_t(I_t(a))] + a + T_t
\]

The left hand side of equation 25 is a real number and the right hand side is a function of \(a\) that change slope when the tax bracket changes. The value of \(a\) that solves equation 25 is denoted \(a^*(p, \beta; a')\) and its solved using the Newton-Raphson algorithm. The values of \((a^*, p, \beta)\) and \(a'\) determine the value of \(V_t\):

\[
V_t(a^*(p, \beta; a'), p, \beta) = u(c^*(p, \beta; a')) + \beta E\{V_{t-1}(a', \beta', p') \mid (\beta, p)\}
\]

Notice that \(V(a^*, p, b)\) is not necessarily defined at the points at the grid of \(a\) and for that reason it is necessary to use a cubic interpolation to approximate its values at every \(a_i \in A^G\).

To obtain the policy function associated to the value function \(V^*\), a similar process is followed using a finer grid for the value of the assets \((n^A = 1,000)\). For every \((p, \beta) \in P^G \times \beta^G\) and for every \(a' \in A^G\), a value \(a^*(p, \beta; a')\) is obtained using the EGM. The policy function is then defined as \(g : (a^*(p, \beta; a'), p, \beta) \rightarrow a'\) for every \((p, \beta; a')\). Then, this function is interpolated to obtain its values at \(a_i \in A^G\), \(g : (a_i, p, \beta)\).
6.3 Computing the Stationary Distribution

Let $G_t$ be defined as the cumulative distribution function over the amount of assets for a given pair $(p, \beta)$ at period $t$:

$$G_t(a^*, p^*, \beta^*) = \text{Probability of } \left( a \leq a^*; p = p^*; \beta = \beta^* \right) \quad (26)$$

$G_{t+1}$ is a function of $G_t$ and the policy function $g : (a, p, \beta) \rightarrow a'$. Consider an arbitrary element of $A^G \times P^G \times \beta^G, (a^*, p^*, \beta^*)$:

$$G_{t+1}(a^*, p^*, \beta^*) = \sum_{\beta_i \in \beta^G} \sum_{p_i \in P^G} \left[ G_t(g^{-1}(a^*; p_i, \beta_i), p_i, \beta_i) \Pi_{p_i, p^*} \Pi_{\beta_i, \beta^*} \right] \quad (27)$$

$\Pi_{\beta_i, \beta^*}$ denotes the probability of getting $\beta^*$ in $t+1$ given that $\beta_t = \beta_i$; $\Pi_{p, p^*}$ denotes the probability of getting $p^*$ in $t+1$ given that $p_t = p_i$ and $g^{-1}(a^*; p_i, \beta_i)$ is the inverse of the policy function given $(p_i, \beta_i)$. For a given initial cdf $G_0$, iterating on equation 27 yields a sequence $\{G_t\}_{t=0}^{\infty}$. The limit of this sequence, $G^*$, is the stationary distribution of assets. To numerically approximate $G^*$, I first propose $G_0$ such that the unconditional distribution of $\beta$ and $p$ are already at its long run values to speed up the convergence process. For all $(p, \beta)$ and for all $i = 1, \ldots, n^{A_2}$:

$$G_0(a_i, p_i, \beta_i) = \frac{a_i - a_0}{a_n A^2 - a_0} \Pi(p_i) \Pi(\beta_i)$$

$\Pi(p_i)$ and $\Pi(\beta_i)$ are the probability of $p_i$ and $\beta_i$ given by their unconditional distributions. I then approximate the inverse of the policy function using a piece-wise cubic Hermite interpolating polynomial to obtain $G_1$ using equation 27. This process is repeated until $\|G_t - G_{t-1}\|$ is less than or equal than $1e-8$ and then set $G_t = G^*$. Notice that $G^*$ is associated to the policy function $g(\cdot)$ which is a function of $(w, r)$. So, each $(w, r)$ has a different $G^*$ associated to it. Furthermore, the overall stock of assets held by households is given by:
\[ A = \int_\beta \int_p \int_a^\infty a dG^*(a, p, \beta) \, da \, dp \, d\beta \]  

(28)

6.4 Equilibrium

In a stationary equilibrium, there is a stationary distribution over assets and the stock of assets held by the households, \( A \), is equal to the stock of capital required by the firm, \( K \). In equilibrium, \( r \) is such that \( A = K \). The process for computing the equilibrium is the following:

1. Set an initial range for \( K \), \([K, \bar{K}]\). The lower bound is equal to half of the level of capital in the complete markets case.

2. Set \( K_1 = [K, \bar{K}] / 2 \).

3. Let \( w_1 = F^L(K_t, 1) \) and \( r_1 = F^K(K_1, 1) - \delta \).

4. Compute \( A \) associated to \((w_1, r_1)\)

   (a) If \(|A - K_1| \leq 1e-3\), then \( r_1 \) is the equilibrium interest rate.

   (b) If \(|A - K_1| > 1e-3\), then:

   i. If \( K > A \) then set \( K_1 = \bar{K} \) and repeat step 1 to 4.

   ii. If \( K \leq A \) then set \( K_1 = K \) and repeat step 1 to 4.

6.5 Transitions

The original stationary equilibrium is calibrated to match the distribution of wealth observed in 1980 and is a function of the deep parameters \((\mu^\beta, \rho^\beta, \sigma^\beta, \delta, \sigma)\), the properties of the labor income process \((\rho^p, \sigma^p, \sigma^v, \kappa)\) and the taxation and transfers system given by \( \tau(\cdot) \). The stationary distribution associated with the 1980 values for the labor income process and the tax system is denoted as \( G^*_0 \). The values for \( \sigma^p, \sigma^v, \kappa, \tau(\cdot) \) and \( a \) between 1980 and 2012 are taken from the data (or chosen to match certain aspects of the data) and then are assumed
to remain constant at their last observed value. Let $G_T^*$ denote the stationary distribution associated with the last observed values of $\sigma^p, \sigma^v, \kappa, \tau(\cdot)$ and $\underline{a}$. We know that the economy will eventually converge from $G_0^*$ to $G_T^*$ in $T$ periods and we are interested in the transition.

To compute the transition, I first guess the number of periods that it will take to reach it, $T$, and a path for the aggregate level of capital $\{K_t\}_{t=1}^{T-1}$. $K_0$ and $K_T$ are the ones associated to the equilibrium interest rate in the original and new stationary distributions. Note that for every $K_t$ there is a real interest rate and wage associated to it $(w_t, r_t)$ such that $w_t = F^L(K_t, 1)$ and $r_t = F^K(K_t, 1) - \delta$.

For every $t = 1, \ldots, T-1$ the value function $V_t$ is computed as:

\begin{equation}
V_{t+1}(a, p, \beta) = \max \left\{ u(c) + \beta E\left\{ V_t(a', \beta', p') \mid (\beta, p) \right\} \right\} \tag{29}
\end{equation}

Subject to the borrowing constraint, $a_t$, the current value for the labor income process $(\sigma^p_t, \sigma^v_t, \kappa_t)$, the tax system $\tau_t(\cdot)$ and the budget constraint in $t$. $V_0$ and $V_T$ are the value functions associated to the initial and final stationary distributions. In addition to the value functions, I use the EGM to obtain the sequence of policy functions $\{g_t(a, p, \beta) \rightarrow a'\}_{t=1}^{T-1}$. Using the sequence of policy functions and equation 27 is possible to obtain the total stock of assets held by households $\{A_t\}_{t=1}^{T-1}$ during the transition.

If $\max_t |A_t - K_t| \leq 1e-3$ then the original guess $\{K'_t\}_{t=1}^{T-1}$ is accurate and the equilibrium path for the transition has been found. If not, the original path for $K_t$ is updated according to the following rule: $K'_t = \lambda K_t + (1 - \lambda)A_t$ and the original process is repeated with the guess for $K_t$ equal to $\{K'_t\}_{t=1}^{T-1}$ until convergence is reached.