



LUND UNIVERSITY

Capital Taxation and Investment

Matching 100 Years of Wealth Inequality Dynamics

Boehl, Gregor; Fischer, Thomas

2017

Document Version:

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Boehl, G., & Fischer, T. (2017). *Capital Taxation and Investment: Matching 100 Years of Wealth Inequality Dynamics*. (Working Papers; Vol. 2017, No. 8). Department of Economics, Lund University.

Total number of authors:

2

Creative Commons License:

Unspecified

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Working Paper 2017:8

Department of Economics
School of Economics and Management

Capital Taxation and Investment: Matching 100 Years of Wealth Inequality Dynamics

Gregor Boehl
Thomas Fischer

June 2017



LUND
UNIVERSITY

Capital Taxation and Investment: Matching 100 Years of Wealth Inequality Dynamics

(Latest version available at: <https://gregorboehl.com/live/boehl-fischer2017.pdf>)

Gregor Boehl* and Thomas Fischer†

June, 2017

Abstract

Using a parsimonious, analytically tractable dynamic model, we are able to explain up to 100 years of the available data on the dynamics of top-wealth shares for several countries. We build a micro-founded model of heterogeneous agents in which - in addition to stochastic returns on investment - individuals disagree marginally on their expectations of future returns and thus hold different asset positions. We show that, given a positive tax on capital gains, the distribution converges to a double Pareto distribution for which the degree of wealth inequality decreases with the tax rate. Closed-form solutions confirm that without government intervention there is infinite inequality. Moreover, transition dynamics are shown to increase with the tax rate. We discuss the model's ability to match the measured wealth inequality for the US, the UK, Sweden, and France, both in levels and transitions. The heterogeneous development in the different countries and across time can be traced back to different tax regimes.

Keywords: Wealth inequality, capital taxation, stochastic simulation, heterogeneity

JEL No. D31, H23, C63, G11

*IFMS, Goethe University Frankfurt, Germany. Email: G.A.Boehl@uva.nl

†Department of Economics, Lund University and Knut Wicksell Centre of Financial Studies, Sweden. Email: thomas.fischer@nek.lu.se

¹We are grateful to Christiane Clemens, Alex Clymo, Herbert Dawid, Cees Diks, David Edgerton, and participants of several conferences and workshops for discussions and helpful comments on the contents of this paper.

1 Introduction

*“All animals are equal
But some animals are more equal than others”*

— George Orwell, *Animal Farm* (1946)

In November 2016 investor George Soros lost about one billion US\$ when betting against the market during the market race that greeted the election of Donald Trump to become the 45th president of the United States.¹ Meanwhile, others like the billionaire Warren Buffett have put billions into stock markets, betting on the increase in government investment and the accompanying effects on stock prices. Thus, the disagreement in expectations about the evolution of the asset market has a major impact on the overall wealth of individuals and also influences the distribution of wealth. Given the fact that the top wealth holders are heavily invested, even small disagreements will have large impacts and finally shape the distribution of wealth.

In this work we develop a formal model that explicitly incorporates the role of disagreement in expectations and shows its ability to match the dynamics of the upper tail of the empirical distribution of wealth. Our second key contribution is not only to match the correct levels of wealth, but we are also able to replicate the transition dynamics for four countries in the long run – the US, UK, Sweden, and France – by only using time series of income and capital gains taxes as a model input.

The interest in the distribution of wealth gained momentum after the publication of Piketty (2014) in which he documents an increasing concentration of wealth starting in the 1980s. For the important case of the United Kingdom and the USA, Piketty (2014) argues that the tax policy under Margaret Thatcher respectively Ronald Reagan was one major factor leading to the increase in wealth inequality. Yet, the book does not feature a formal model to discuss the claim. Using the novel data evidence, both Kaymak and Poschke (2016) and Hubmer et al. (2016) employ extended Bewley type models in order to explain the observed wealth inequality dynamics and, in particular, the fat tails in the wealth distribution. They confirm that the evolution of wealth inequality can to a large extent be explained by changes in the tax system. In particular, Hubmer et al. (2016) argue that other potential mechanisms such as the witnessed increased income inequality (the key mechanism generating wealth inequality in Bewley models), the falling labor share, or the $r > g$ story of Piketty (2014) all fall short of accounting for the data. We argue that these models are highly complex and require numerical solutions and thus might obscure the fundamental underlying mechanism.

The model employed in this work can be classified as being a *random growth* model in the tradition of Benhabib et al. (2011) which has been recently calibrated to US data in Aoki and Nirei (2017). In our work, we present a closed-form solution of both the stationary distribution and the transition dynamics as emphasized in Gabaix et al. (2016). The key underlying mechanism in these type of models is idiosyncratic investment risk creating the fat tails in the distribution of wealth. As formally shown in Benhabib et al. (2015) – in which the idiosyncratic investment risk is introduced into an otherwise standard Bewley economy – the investment risk is important to explain the top-wealth inequality. In general, Bewley models inherit the wealth inequality from the income inequality. As it is well-known from the empirical evidence the top

¹As stated by news agencies such as *CNBC*:
<http://www.cnn.com/2017/02/17/george-soros-loaded-up-with-big-bets-against-the-stock-market.html>

inequality of the stock measure of wealth is far higher than for the flow measure of income making Bewley models incapable of producing top wealth inequality.² Usually, it is argued that the idiosyncratic shocks can be generated by idiosyncratic asset holdings. This work extends the literature by also including expectation disagreement. We show that wealth (approximately) follows a symmetric double Pareto distribution in line with the empirical evidence of Benhabib et al. (2014). Without expectation disagreement the distribution would be characterized by a fat tail only in the right part of the distribution (the *rich*) and proper calibration to also match the transition speed would overestimate idiosyncratic risk.

We show furthermore that taxation is a key ingredient to understanding wealth inequality. In particular, without taxation inequality there would be no finite inequality.³ Note that we focus on the top-shares - i.e. the right tail of the wealth distribution. It is well-known that the low share at the low end of the Lorenz curve can be explained by different forms of (financial) frictions - most prominently borrowing constraints.⁴ In this work, we abstract from these frictions. Moreover, by assuming constant relative risk aversion individuals have identical risk exposure and identical marginal propensities to consume. In such a world, given that traders share identical beliefs, there would be no trade. Instead, we let them make random and considerably small forecasting errors that are normally distributed around the Rational Expectations solution. Thus - and in contrast to the partial equilibrium model employed in Benhabib et al. (2011) - this work features a closed economy.

With this micro-founded - yet simple - model we are able to capture both the level as well as the dynamics⁵ of wealth inequality for the USA, UK, France, and Sweden for a very long time period ranging up to 100 years.⁶ To our best knowledge we are the first to present a comprehensive cross-country view of inequality as the existing literature has focused solely on the USA (also being the country with the highest increase in wealth inequality among the covered countries). The different levels of taxation are key to understanding the differences in top wealth shares across countries and in time.

The remainder of this work is structured as follows. In Section 2 we provide an overview of the empirical and theoretical literature on wealth inequality with a focus on recent papers trying to fit the empirical evidence. In the following section, we present the micro-foundations for our formal model and discuss analytic statistic properties in Section 4. Section 5 compares these analytical results with the empirical findings. Section 6 uses the model to generate forecasts about the future evolution of wealth inequality and presents robustness checks. Finally, Section 7 wraps up.

²Other extensions proposed in the literature to match the top wealth share problem is the consideration of entrepreneurs in the tradition of Quadrini (2000) or Cagetti and Nardi (2009a) having different savings patterns than the rest of the population.

³A similar point was made in Fernholz and Fernholz (2014). This work, also featuring a slightly different model, only relied on simulations, while we can provide a closed-form solution.

⁴Other factors are different access to financial markets (entry barriers for the poor, lower transaction costs for the wealthy) or heterogeneous access to important information and expertise. We abstract from these systematic frictions in favor of the wealthy. Sometimes the role of portfolio structuring (in particular a utility function with decreasing relative risk implying higher return portfolios for wealthy), different consumption patterns (heterogeneous marginal propensity to consume) and wealth transmission are emphasized.

⁵As argued in Gabaix et al. (2016) the problem with this type of model is that the convergence dynamics are usually very slow.

⁶The first three countries are the only three countries with long-run data for wealth inequality, as documented in the *World Wealth & Income Database*. Sweden is covered extensively in the work of Lundberg and Waldenström (2017).

2 Literature

Following the major public debate surrounding the publication of the work of Piketty (2014), the interest in distributional measures has recently increased. In particular, the empirical evidence regarding inequality - especially for the flow measure of income - has substantially improved. Cross-country evidence is assembled and made freely available on the *World Income & Wealth Database* maintained by the collaborative effort of many researchers.⁷ Despite this effort the data availability of consistent and long-run measures of wealth inequality is still highly limited. The database provides long-run data for both the United States of America and the United Kingdom. The US data was updated recently by Saez and Zucman (2016). The latest data update for the UK was conducted by Alvaredo et al. (2017). The quality of the French data (especially from the 1970s onward) was recently substantially improved by Garbinti et al. (2017). Evidence for Sweden is compiled by Daniel Waldenström and his collaborators (Lundberg and Waldenström, 2017).⁸ A recent comprehensive survey on the overall empirical evidence regarding wealth inequality is given in Roine and Waldenström (2015). The discussion about the distribution of wealth is not an end in itself, but also contains important policy implications as it impacts on the conduct of monetary policy (Kaplan et al., 2016) and also influences economic growth (Clemens and Heinemann, 2015).

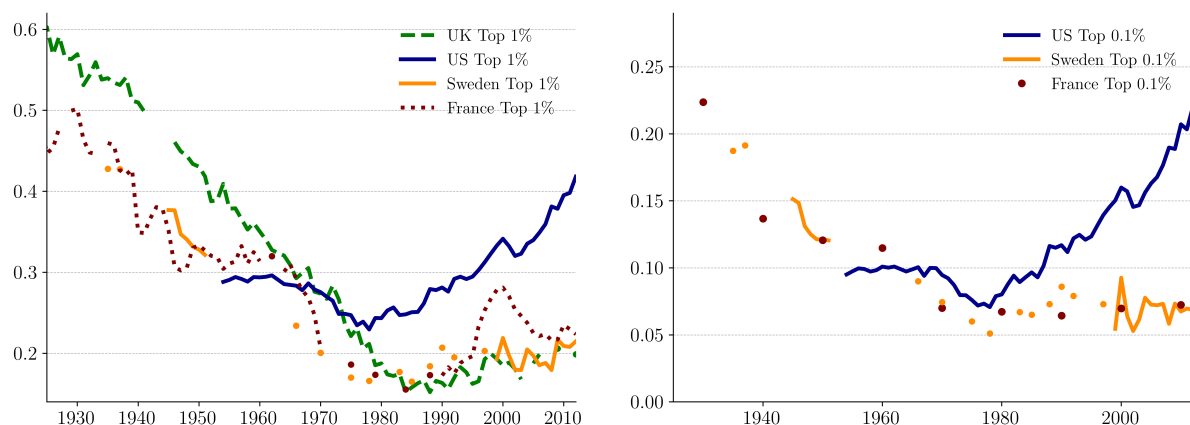


Figure 1: Top wealth shares. Data source: *wid.world* and Lundberg and Waldenström (2017) for Sweden.

Figure 1 present evidence on the top-shares for the Anglo-Saxon countries - the USA and the UK - and the European Countries - France and Sweden - in the long run.⁹ While there is an overall decrease in all countries until the 1980s, inequality has subsequently increased. This increase is modest in the European countries, but highly pronounced in the USA. In particular, it emerges for the top wealth holders. In general, inequality is higher in the USA. As we will discuss in the course of this work this can be traced back to the different taxation systems.

Different theoretical models compete in order to explain the witnessed degree of inequality. Usually, models in the Bewley-type tradition are considered in order to discuss inequality

⁷The data is available at *wid.world*.

⁸The data is freely available on his homepage.

⁹Note that there is no top 0.1% data available for the United Kingdom. Thus, the graph only displays the USA, France and Sweden.

(Bewley, 1977; Huggett, 1993; Aiyagari, 1994). Yet, it has been formally shown by Benhabib et al. (2011) that these types of models - build around the notion of additive idiosyncratic labor income risk - will fail to generate the fat tails in the wealth distribution and thus match the shares of the top wealth holders. Benhabib et al. (2011) propose a model with multiplicative idiosyncratic capital income risk in order to replicate the current state of wealth inequality in the USA. They follow an argument laid out as early as Wold and Whittle (1957), building on random growth. Thus, this type of literature is often referred to as *random growth models* (Gabaix et al., 2016). Taxation of capital (income) plays a crucial role in these models. Using simulations, Fernholz and Fernholz (2014) show that wealth inequality does explode without redistribution in a standard model with idiosyncratic investment risk. We build a similar model, yet introduce dispersion of individual opinion about the prospects of an investment in order to match the shares of the top wealthy individuals.

While heterogeneous portfolios are often motivated by different degrees of risk aversion (and marginal propensities to consume), this argument does not hold for the very rich. In such models, the highly wealthy should have a relatively similar portfolio structure. This, however, does not take into account the vast variety of different asset classes and almost infinite supply of similar assets within classes. Moreover, due to heterogeneous individual expectations about future prospects, agents will hold different positions in identical assets. For that reason we motivate heterogeneous portfolios by marginal disagreement on future returns, of which a considerable degree is documented by Greenwood and Shleifer (2014) in a survey over six data sets on investor expectations of future stock market returns. Furthermore, evidence from the lab has shown that individuals generally do not to a good job when forming expectations and these expectations are furthermore largely heterogeneous.¹⁰ If agents are heterogeneous in their projections, they also invest in different assets.

Thus, compared to the literature in the tradition of Benhabib et al. (2011) assuming some exogenous random noise in a partial equilibrium setting, our model is closed in a general equilibrium tradition by allowing individuals to trade with each other. Moreover - and compared to for example Fernholz and Fernholz (2014) - our work goes beyond pure numerical simulation and is able to not only quantify the stationary distribution but the whole dynamics of the top wealth shares. Similar to the existing literature we are able to match the steady state of wealth inequality.

The second aim of this work is to take the available data and test whether the model can match in the transition. This is of particular importance as it was recently argued in Gabaix et al. (2016) that while these models capture the steady state of inequality well, the dynamics in these models are far too slow compared to empirical evidence. Both closed form solution and simulations confirm that our model also matches the dynamics of inequality well.

Of course, using the new data evidence, similar projects have been undertaken. Most prominently, Kaymak and Poschke (2016) use the evidence for the United States from 1960 to the most present date to present a calibrated model in the Bewley tradition. Using the modification of Castaneda et al. (2003), allowing for extreme *superstar* shocks producing high levels of income inequality, the authors are able to match the data. With a very detailed modeling of the US-tax system (including income, corporate, and estate taxes as well as the pension system) the authors identify the contributing factors. They argue that the (exogenous) increase in income inequality and - for the distribution of wealth - the structure of the taxation and transfer system are highly

¹⁰For a recent overview on the heterogeneous expectations hypothesis see for example Hommes (2013).

important in order to explain the evolution of inequality.

A more comparable approach to this work is presented in Aoki and Nirei (2017), featuring a rich model in continuous time. In line with empirical evidence and due to idiosyncratic firm shocks the distribution of heterogeneous firms is given by Zipf's law.¹¹ The firm's income translates into income for private households, implying a realistic distribution of both income and wealth for private households. Combining this with tax rates, they are able to match both the dynamics and the state of inequality in the USA from the 1970s to the most recent years. Note that we focus on the distributional impact on taxation and do not make a statement about the macroeconomic impact of the wealth tax or even its optimal level.¹²

Most closely connected to this work, Hubmer et al. (2016) extend an otherwise standard Bewley-type model with heterogeneous rates of time preference β_i in the tradition of Krusell and Smith (1998), with Pareto tails in the income distribution and idiosyncratic investment risk following Benhabib et al. (2011). They are able to quantitatively reproduce wealth inequality dynamics for the USA from the 1970s. As outlined before, they conclude that the most important mechanism driving the increases in wealth inequality is the change in the taxation system since the 1980s which is also central in our work. Compared to the existing literature focusing solely on the USA, we compare different countries and also use longer data evidence starting before World War II.

3 Model

We assume an economy with a large number n of individuals indexed by i . Their only income consists of investment returns and they are free to choose between a risk-free asset paying constant gross return R and a continuum of ex-ante identical risky assets of which each pays an idiosyncratic, stochastic dividend $d_{i,t}$ every period t . To maximize their intertemporal consumption over an infinite time horizon the agents accumulate wealth $w_{i,t}$. Hence, each agent i faces the question of which amount $c_{i,t}$ to consume and which amount $x_{i,t} = z_{i,t}w_{i,t}$ of the risky asset to purchase. In this case $z_{i,t}$ relates the demand for risky assets as a share of individual wealth $w_{i,t}$. Assuming log-preferences, the individual problem is then given by

$$\max_{c,z} \sum_{t=0}^{\infty} \beta^t \ln c_{i,t}$$

subject to the two constraints

$$\begin{aligned} c_{i,t} &= (1 - s_{i,t})w_{i,t}, \\ w_{i,t} &= (R + [d_{i,t} + p_t - Rp_{t-1}]z_{i,t-1})s_{i,t-1}(1 - \tau)w_{i,t-1}. \end{aligned}$$

Here we denote by β the intertemporal discount rate and $s_{i,t}$ the savings rate. The value τ captures a tax on the stock level of wealth.¹³ p_t is the price for a risky asset in t and

¹¹The latter is a Power-law with an exponent $\alpha = 1$.

¹²As shown in the seminal contribution of Judd (1985) in a standard model the optimal tax on the stock value of wealth is zero. Yet, it is well-known that in Bewley-type models this result fails to hold and optimal taxes are positive in order to counteract excessive savings (Aiyagari, 1995). More recent contributions discussing the welfare impact in both the state and the transition of wealth taxes in the broader sense (including capital gain and inheritance taxes) are (among others) Castaneda et al. (2003), Domeij and Heathcote (2004), and Cagetti and Nardi (2009b).

¹³Note that in the empirical part it is very important that taxes vary in time. For the sake of readability we, however, suppress the time index in this section.

$d_{i,t} = d + \epsilon_{i,t}^d$ its dividend with an idiosyncratic stochastic term $\epsilon_{i,t}^d \sim N(0, \sigma_d)$.¹⁴ This problem does not directly entail a closed form solution, but can be separated into two stages that are relatively standard in the literature. Let us first solve the consumption problem.

Levhari and Srinivasan (1969) show that for log-utility, which is a particular case of Constant Relative Risk Aversion (CRRA) preferences¹⁵, in equilibrium agents consume $1-\beta$ of their wealth at the end of each period, i.e. $s_{i,t} = \beta \forall i, t$.¹⁶ It is important to point out that this result holds despite the tax rate. Due to the exact offsetting of income and substitution effects for log-utility the savings rate is not distorted by the tax rate.¹⁷ Note that the assumption of CRRA also explicitly avoids inequality dynamics induced by a different marginal propensity to consume. The fact that agents are not subject to borrowing constraints simplifies the model considerably, but is not realistic in the context of the lower 50% share of wealth holders. We hence focus on the upper tail of the wealth distribution. Then the law of motion for each individual's wealth follows

$$w_{i,t} = (1 - \tau)\beta \{R + (d_{i,t} + p_t - Rp_{t-1})z_{i,t-1}\} w_{i,t-1} \quad (3.1)$$

$$= (1 - \tau)\beta R_t^z(z_{t-1})w_{i,t-1}, \quad (3.2)$$

for which $R_t^z(z_{i,t-1})$ summarizes the individuals gross return on investment.

For the second stage, in which we solve for the optimal demand for risky asset $x_{i,t}$, let us use Equation (3.1) to rewrite the maximization problem as

$$\max_z \sum_{t=0}^{\infty} \beta^t \ln\{(1 - \beta)w_{i,t}\} \quad \text{s.t.} \quad w_{i,t} = (1 - \tau)\beta R_t^z(z_{i,t-1})w_{i,t-1}$$

which is equivalent to

$$\max_z \sum_{t=0}^{\infty} \beta^t \ln \left\{ (1 - \beta)\beta w_{i,0} (1 - \tau)^t \prod_{k=0}^t R_{i,k}^z(z_{i,k-1}) \right\}.$$

Due to the logarithmic laws, the term $\ln\{\prod_{k=0}^t R_{i,k}(z_{i,k-1})\}$ can be separated and is the only part that depends on z_t . Since we can rewrite $\prod_{k=0}^t R_{i,k}^z(z_{i,k-1}) = R_{i,t}^z(z_{i,t-1}) \prod_{k=0}^{t-1} R_{i,k}^z(z_{i,k-1})$ this portfolio problem can be well-approximated by mean-variance maximization as laid out in Pulley (1983). The optimal demand for the risky asset $x_{i,t}$ is then, up to a second order approximation, given by

$$x_{i,t} = z_{i,t}w_{i,t} = (E_t[d_{t+1} + p_{t+1}] - Rp_t) w_{i,t} / \sigma_d^2. \quad (3.3)$$

Note that – identical to the optimal consumption plan – the portfolio structure is independent of the wealth tax. As presented in Stiglitz (1969) for Constant Relative Risk Aversion preferences –

¹⁴Since those assets are ex-ante identical, their price is likewise the same.

¹⁵It is easy to show that for the special case $\gamma \rightarrow 1$ the CRRA utility function boils down to log-utility $\lim_{\gamma \rightarrow 1} \frac{c^{1-\gamma}-1}{1-\gamma} = \ln(c)$ for which income and substitution effects exactly cancel each other out.

¹⁶ If for gross return $1 + r$, r is given by a normal distribution with mean μ and standard deviation σ , Levhari and Srinivasan (1969) show that the optimal value is given by $s = (\beta \exp(\mu + 0.5\sigma^2)^{1-\gamma} \exp(-\gamma(1-\gamma)0.5\sigma^2))^{\frac{1}{\gamma}}$. It is easy to see that for $\gamma = 1$ we have $s = \beta$. Note that here we abstract from the taxes that would modify both μ and σ . Given log-utility the latter, however, has no impact.

¹⁷The interested reader is also referred to Lansing (1999) showing that the seminal result of a zero optimal tax rate as proposed in Judd (1985) fails to hold with log-utility. For a more recent and general approach the reader is referred to Straub and Werning (2014).

of which the assumed log-utility is a special case – wealth taxation does not lead to a restructuring of the portfolio.

Market clearing requires $\sum_{i \in n} x_{i,t} = X_t$, with X_t being the total supply of the risky asset. Without loss of generality we can fix supply and normalize X_t to unity for all periods.¹⁸

We want to assume that return expectations are heterogeneous and each agent's expectation is a draw from the normal distribution around the rational expectation of future returns. Thus, the rational expectation operator E is replaced with a noise individual expectation operator $\hat{E}_{i,t}$, giving

$$\hat{E}_{i,t}[d_{t+1} + p_{t+1}] = d + E_t p_{t+1} + \epsilon_{i,t}^E, \quad \epsilon_{i,t}^E \sim N(0, \sigma_E).$$

Assume furthermore that no single person is rich enough or has an $\epsilon_{i,t}^E$ large enough to influence the price.¹⁹ Taking market clearing into account, it is clear that all wealth is owned by all agents and $\sum_{i \in n} x_{i,t} = 1$, i.e. aggregate demand for the risky asset equals one every period.

We want to assume that our taxation is redistributive transferring a lump-sum value T to all individuals $T_t = \frac{\tau \sum_{i \in n} w_{i,t}}{n} = \frac{\tau W_t}{n}$ for an aggregate wealth $W_t = \sum_{i \in n} w_{i,t}$. Thus, no wealth is lost in the act of taxation. For our specific assumption, for which the aggregate wealth (being stationary) is identical to the number of agents $W = n$, this boils down to $T_t = \tau_t$.

Keeping this in mind and the aggregating over Equation (3.3) and (3.1) yields:

$$p_t = R^{-1}(E_t p_{t+1} + d - \sigma_d^2 W_t^{-1}) \quad (3.4)$$

$$W_t = \beta(RW_{t-1} + d + p_t - Rp_{t-1}), \quad (3.5)$$

which is the law of motion for prices and aggregated wealth W_t . Since n is large, due to the law of large numbers idiosyncratic disturbances level out and aggregate wealth $W_t = W$ is constant in the absence of aggregate shocks. The levels of prices and aggregated wealth thus reflect the detrended steady growth path.²⁰ Then we can also normalize the price to unity without explicitly accounting for market clearing. The steady state versions of (3.4) and (3.5) is

$$\begin{aligned} \sigma_d^2/W &= d + 1 - R \\ W(\beta^{-1} - R) &= d + 1 - R. \end{aligned}$$

This implies that, given the normalization of prices,

$$W = \frac{\sigma_d}{\sqrt{\beta^{-1} - R}} \quad (3.6)$$

$$d + 1 - R = \sqrt{\beta^{-1} - R} \sigma_d. \quad (3.7)$$

¹⁸This implies that the market clearing price is defined by:

$$p_t = \{p : \sum_{i \in n} x_{i,t}(p, w_{i,t}) = 1\}.$$

Note that the optimal individual demand depends on individual wealth, which in turn also depends on the current price.

¹⁹This is indeed satisfied by the law of large numbers. The additional advantage of this assumption is, without loss of generality, that we can provide analytic results for the law of motion of individual wealth, aggregated wealth, and prices.

²⁰This assumption implies that all growth in aggregate wealth can be attributed to increases in output, which constitutes a further simplification for the sake of simplicity. Since the key targets in this paper are wealth shared, these are unaffected by the overall level of wealth.

Plugging Equation (3.3) into Equation (3.1), integrating individual forecast errors and setting prices to the steady state:

$$w_{t,i} = \beta \left\{ R + \left(d + \epsilon_t^d + 1 - R \right) \left(d + \epsilon_{t,i}^E + 1 - R \right) \sigma_d^{-2} \right\} (1 - \tau) w_{t-1,i} + T_t.$$

Using Equation (3.7) and some algebra, the law of motion (LOM) for individual wealth can be written as:

$$w_{i,t} = \beta \left\{ \beta^{-1} + \sqrt{\beta^{-1} - R} (\epsilon_{i,t}^d + \epsilon_{i,t}^E) \sigma_d^{-1} + \epsilon_{i,t}^d \epsilon_{i,t}^E \sigma_d^{-2} \right\} (1 - \tau) w_{i,t-1} + T_t.$$

For our simulations we use a quarterly calibration, so let $\beta = 0.99$ and the real interest rate R be 1.02. As a result we have $\sqrt{\beta^{-1} - R} \approx 0.05$, which is negligibly small.²¹ Defining $\gamma \equiv \beta \frac{\sigma_E}{\sigma_d}$ and $\varepsilon_{i,t} \equiv \epsilon_{i,t}^1 \epsilon_{i,t}^2$ to be the product of two independent random variables that follow a standard normal distribution and recalling that $T_t = \frac{\tau_t W_t}{n}$, the final law of motion can be further simplified to:

$$w_{i,t} = (1 + \gamma \varepsilon_{i,t}) (1 - \tau) w_{i,t-1} + \tau_t,$$

which leaves γ as the only free parameter of our model.

4 Analytic Results

This section aims to enrich our understanding of the process that generates the wealth distribution by finding a closed form solution for the stationary distribution as well as for the transition dynamics. In order to do so, we have to overcome some technical obstacles.

The portfolio returns, a product of two standard normal variables, follow a so-called *product-normal distribution*. To obtain a closed form solution, we have to transfer this distribution to another distribution that is easier to handle analytically.

Proposition 1. The first three moments of the product normal distribution and the Laplace distribution with shape parameter of $\lambda = \sqrt{0.5}$ are equal.

Proof. See Appendix A.1. □

The Laplace distribution is very handy in our context for identifying a closed-form solution. The individual law of motion (LOM) has to be rewritten in continuous time in order to solve the Fokker-Planck equations which allows us to identify the cross-sectional distribution in terms of the free parameters γ and τ . It would read as

$$dw_{i,t} = \tau(\bar{w} - w_{i,t})dt + (1 - \tau)\gamma w_{i,t}dNP, \quad (4.1)$$

for which NP is the noise following the product-normal distribution and $\bar{w} = W/N$ the mean level of wealth. In order to retrieve a closed-form solution we transform this to the Laplace distribution using the scaling factor λ which we just introduced. The equation thus reads

$$dw_{i,t} = \frac{1}{\lambda} \tau(\bar{w} - w_{i,t})dt + \frac{1}{\lambda} (1 - \tau) \gamma w_{i,t}dL, \quad (4.2)$$

for which L signifies Laplace distributed noise.

²¹Note that a positive demand for the risky assets requires $R < 1 + d$. Stationarity of aggregate wealth furthermore demands for $R < \beta^{-1} < 1 + d$ i.e., $\sqrt{\beta^{-1} - R} > 0$ but small. Furthermore, the variance of $\epsilon_{i,t}^E$ is already relatively small. Rewriting $\epsilon_{i,t}^d$ in terms of a standard normal reveals that the term is relatively small.

Proposition 2. Using Itô's lemma as a second-order approximation, disregarding redistribution for the very rich, and solving the Fokker-Planck equation, the stationary cross-sectional distribution can be expressed as a double Pareto distribution with an inverse shape parameter α :

$$\alpha = 1 + \frac{\sqrt{2}\tau}{\gamma^2(1-\tau)^2}. \quad (4.3)$$

Proof. For the assumption of overall wealth W being identical to the number of agents N , the mean wealth is $\bar{w} = \frac{W}{N} = 1$. In fact, this process is a mean reversion process and the redistributive taxation makes individuals converge to the overall mean if noise is absent. The pace of convergence increases with the level of taxation τ . Let us define the log of wealth $\hat{w}_{i,t} = \log(w_{i,t})$ and apply Itô's lemma. Thus the equation reads

$$d\hat{w}_{i,t} = \left(-\frac{\tau}{\lambda} + \frac{\tau\bar{w}}{\hat{w}_{i,t}\lambda} - 0.5(1-\tau)^2\frac{\gamma^2}{\lambda^2} \right) dt + \frac{1}{\lambda}(1-\tau)\gamma dL.$$

Unfortunately, a complete closed-form solution is not feasible. We want to focus on the rich, whose (log) wealth $\hat{w}_{i,t}$ is far higher than the mean wealth ($\hat{w}_{i,t} \gg \bar{w}$). Thus, we can disregard the effect of redistribution. The equation simplifies to

$$d\hat{w}_{i,t} \approx \left(-\frac{\tau}{\lambda} - 0.5\frac{\gamma^2}{\lambda^2}(1-\tau)^2 \right) dt + \frac{1}{\lambda}(1-\tau)\gamma dL = -\mu dt + \delta dL,$$

with a diffusion term $\delta \equiv \frac{1}{\lambda}(1-\tau)\gamma$ and a drift $\mu \equiv \frac{\tau}{\lambda} + 0.5\frac{\gamma^2}{\lambda^2}(1-\tau)^2 = \frac{\tau}{\lambda} + 0.5\delta^2$. As shown in Toda (2012), the Laplace distribution can be modeled by

$$dL = -\kappa \text{sign}(L)dt + \delta dB$$

with B being the standard Brownian motion and sign representing the sign function. Using this relationship we can rewrite the overall equation as

$$d\hat{w}_{i,t} = \begin{cases} \mu dt + \delta dB & \hat{w}_{i,t} < 0 \\ -\mu dt + \delta dB & \hat{w}_{i,t} > 0. \end{cases} \quad (4.4)$$

The cross-sectional distribution can be found by solving the so-called Fokker-Planck equation²²

$$\frac{\partial f(\hat{w}, t)}{\partial t} = -\frac{\partial}{\partial \hat{w}} (\mu f(\hat{w}, t)) + 0.5 \frac{\partial^2}{\partial \hat{w}^2} (\delta^2 f(\hat{w}, t)).$$

We consider the stationary distribution ($\frac{\partial f(\hat{w}, t)}{\partial t} \stackrel{!}{=} 0$). The solution is well-known (Karlin and Taylor, 1981, p. 221) and given by

$$f(\hat{w}) = 0.5\alpha \exp(-\alpha|\hat{w}|). \quad (4.5)$$

For our case we have

$$\alpha = \frac{2\mu}{\delta^2} = 1 + \frac{\sqrt{2}\tau}{\gamma^2(1-\tau)^2}. \quad (4.6)$$

²²The latter is frequently referred to as Kolmogorov forward equation. The terms can, however, be used interchangeably.

As shown in Toda (2012), it is easy to transfer the Laplace distribution to a symmetric double Pareto distribution. In fact, if \hat{w} follows the described Laplace distribution, wealth $w = \exp(\hat{w})$ is given by the probability density function:

$$f(w) = \begin{cases} 0.5\alpha w^{-\alpha-1} & w \geq M \\ 0.5\alpha w^{\alpha-1} & 0 < w < M. \end{cases}$$

The mode of this double Pareto distribution is given by $M = 1$.²³ □

In fact, the complete distribution is characterized by the single value α . Thus, other measures regarding inequality can be derived starting from this assumption.

Proposition 3. The stationary ($t \rightarrow \infty$) share $s^x(\tau, \infty)$ of the top x (e.g. the top 1% implying $x = 0.01$) wealth holders is given by

$$s^x(\tau, \infty) = (1 + \alpha^{-1})0.5^{1/\alpha}x^{1-1/\alpha}, \quad (4.7)$$

for which α , as above, is implicitly a function of taxes τ and $\frac{\partial}{\partial \alpha}s^x(\tau, \infty) < 0$ for realistic case of $\alpha > 1$.

Proof. See Appendix A.1.1. □

The same rationale can also be used to derive a closed form for the Gini coefficient.²⁴ This implies that high α is accompanied by low inequality. This very neat result has some strong implications. First of all, without taxation $\tau = 0$ the tail-coefficient is $\alpha = 1$, identical to Zipf's law. In fact, the Gini coefficient then takes the value of $Gini(w) = 1$ and $s_x(\tau) = 1$ for all $x \in (0, 1]$, implying total inequality. Thus, in a laissez-faire economy without government intervention, there is no finite level of inequality. In general, inequality increases (α decreases) with γ while decreasing with taxation τ . For the extreme case of $\tau \rightarrow 1$ - which can be thought of as a socialist society - we would have $\alpha \rightarrow \infty$, and thus have a Lorenz-curve identical to the 45-degree line and thus no inequality at all.

Note that our proof heavily relies on second-order approximations.²⁵ This is, however, not problematic for realistic values of $\alpha < 2$, for which only the first two moments exists.²⁶

We can also make a statement about the convergence speed.

²³The cumulative probability distribution takes the following form:

$$F(w) = \begin{cases} 1 - 0.5w^{-\alpha} & w \geq 1 \\ 0.5w^{\alpha} & 0 < w < 1. \end{cases}$$

²⁴The closed-form value for the Gini coefficient is given by

$$Gini(w) = \frac{3\alpha}{4\alpha^2 - 1}, \quad (4.8)$$

and decreasing with α for realistic case of $\alpha > 1$. The proof can be found in Appendix A.1.2.

²⁵In fact, the Fokker-Planck equation is also only a second-order approximation to the more general Master equation. Moreover, the use of Itô's lemma is only a second-order approximation. For noise generated by a Brownian motion (rather than the product-normal distribution) it would still hold exactly.

²⁶As a result, we can only model degrees of inequality with $s^{0.01} > s^{0.01}(a = 2) \approx 10.6\%$ using the closed-form solutions.

Proposition 4. The convergence to the stationary distribution is given by

$$||f(w, t) - f(w, \infty)|| \sim \exp(-\phi t), \quad (4.9)$$

with an average convergence speed of

$$\phi = (0.5\gamma(1 - \tau)\alpha)^2. \quad (4.10)$$

Proof. The convergence speed is given by $\phi = \frac{\mu^2}{2\delta^2}$. See proposition 1 of Gabaix et al. (2016). \square

This implies a half-life of $t_{0.5} = \frac{\ln(2)}{\phi}$. In fact the effective taxation τ not only decreases steady-state inequality, but also increases the speed of convergence to the latter. This also means that there is an asymmetry in the convergence. The increase of inequality for low taxes is slower than the decrease after high tax rates. Thus, the positive message for the policy maker is that it is faster to come down to lower inequality rather than to increase the level of inequality.

Proposition 5. The top-shares evolve according to an autoregressive process of first-order with

$$s_t^x = \rho_t s_{t-1}^x + (1 - \rho_t) s^x(\tau_t, \infty), \quad (4.11)$$

and $\rho_t = \exp(-\phi_t)$ for the average convergence speed $\phi_t = \phi(\tau_t)$ as defined in Equation 4.10.

Proof. See Appendix A.1.3. \square

Hence, the share owned by the fraction x of the population is a linear combination of last period's share and the share of the top x of the stationary distribution given the tax rate τ_t at each time t .

5 An Empirical Application

In this section we feed empirical time series for taxes into our model's LOM in Equation (3) and attempt to match the data on wealth inequality. We start our investigation in the year 1900. From this year on (Piketty, 2014, chapter 14) provides data for the top income and top inheritance taxes.²⁷ Our approach is less rich than, for example Hubmer et al. (2016) or Kaymak and Poschke (2016), yet allows for analytical tractability.²⁸ Rather than only focusing on the USA, we have a broader view by including the UK as well and - in order to contrast the Anglo-Saxon evolution - also consider France and Sweden.

We take the available tax series on capital gains taxes or, if these are unavailable, top marginal income taxes as a proxy for the wealth taxes introduced in Section 3. For the United States, tax data on capital gains taxes are available from the *Tax Policy Center*²⁹, while for the other countries we have to rely on the marginal top income taxes by Piketty (2014). Moreover, we are only provided with the top capital income tax rate but not with the average tax rate.

²⁷Note that wealth inequality data for France is already available from 1807. Yet, the data frequency is very low and moreover there is little variation in time.

²⁸Note that the approach of Aoki and Nirei (2017) is also analytically tractable to a high extent.

²⁹The data is available at <http://www.taxpolicycenter.org/statistics/historical-capital-gains-and-taxes>.

Proposition 6. Usually capital taxes only apply to the additional, positive returns on wealth and not on wealth itself or on losses. We approximate the gross-wealth tax τ - given a capital gains tax θ_r - by finding a τ such that the expected value of after tax returns from a capital gains tax and after tax returns of a gross-wealth tax are equal. Gross-wealth taxes are then given by

$$\tau = \frac{1}{2}\theta_r\gamma\lambda. \quad (5.1)$$

Proof. Dropping time subscripts for taxes, the after-tax returns given the capital income tax θ_r are

$$\bar{R}_{\theta_r} = 1 + \begin{cases} (1 - \theta_r)\gamma\varepsilon_{i,t} & \text{if } \varepsilon_{i,t} > 0 \\ \gamma\varepsilon_{i,t} & \text{if } \varepsilon_{i,t} \leq 0. \end{cases}$$

To use the LOM in Equation (4.11) we approximate τ given θ_r by finding a τ such that the expected value of \bar{R}_τ equals the expected value of \bar{R}_{θ_r} . Then, given that $\varepsilon_{i,t}$ approximately follows a Laplace distribution with scale $\lambda = \sqrt{0.5}$, the expected value $E[\varepsilon_{i,t}|\varepsilon_{i,t} \leq 0]$ is the mean of an exponential distribution with inverse scale λ , which is again λ . Then

$$\begin{aligned} E\bar{R}_\tau &= E\bar{R}_{\theta_r} \\ E\{(1 - \tau)(1 + \gamma\varepsilon_{i,t})\} &= 1 + \gamma P(\varepsilon_{i,t} \leq 0)E[\varepsilon_{i,t}|\varepsilon_{i,t} \leq 0] + (1 - \theta_r)\gamma P(\varepsilon_{i,t} > 0)E[\varepsilon_{i,t}|\varepsilon_{i,t} > 0] \\ 1 - \tau &= 1 - 0.5\gamma\lambda + 0.5(1 - \theta_r)\gamma\lambda \\ \tau &= \frac{1}{2}\theta_r\gamma\lambda, \end{aligned}$$

where $P(\varepsilon_{i,t} > 0)$ denotes the probability that $\varepsilon_{i,t}$ is positive.³⁰ □

Unfortunately, time series on capital or wealth taxation are unavailable for the UK, France and Sweden, which is why we focus on the data on income taxation as a proxy. In order to account for income instead of capital taxation, we cannot directly use the available data as $\theta_{r,t}$, as we do in the case for the US simulations. Using the top marginal income taxes directly will potentially imply too much taxation for various reasons. Frequently, income taxes are far more complicated than the capital gains taxes considered here. Income taxes are progressive to a varying degrees while the above approximation applies the top marginal tax rate to all agents independent of the magnitude of their capital gains or income level. Also, capital gains taxes are, for most countries, considerably lower than income taxes due to higher mobility of capital (as compared to labor) giving countries an incentive to offer lower capital gain taxes in the international tax competition. Furthermore, both income and capital gains taxes provide different forms of exemptions that can not be modeled in detail if we want to make use of our closed form solution. Lastly, we do not adjust for the possibility of tax evasion and tax avoidance.³¹

For this reason we scale down the available series of top income taxes by multiplying it with a scale parameter $0 < \eta \leq 1$ to adjust for progressivity, exemption levels, the difference between capital gains taxes and income taxes, and tax avoidance. We define

$$\tau_t = 0.5\lambda\gamma\eta\theta_{r,t}. \quad (5.2)$$

³⁰In fact, this also approximates the conversion of annual tax rates to a quarterly calibration.

³¹Analyzing evidence from leaked tax evasion documents, Alstadsaeter et al. (2017) show an increase of tax evasion with wealth rising up to 30% for the top 0.01% of the wealth distribution.

Our second free parameter is $\gamma = \beta \frac{\sigma_E}{\sigma_d}$. The finance literature usually uses a value of σ_d in the range from 0.08 to 0.3 (Campbell and Viceira, 2002). Estimates for return disagreement are hard to gather. Greenwood and Shleifer (2014) present values ranging between 1% and 3%. Combining these values implies an estimate for γ in the range from 0.1 to 0.375.³²

To obtain some intuition, let us plug in Equation 5.2 into the closed-form solution from Equation (4.11), and approximate $(1 - \tau) \approx 1$. Then

$$s_t^x = e^{-\phi_t} s_{t-1}^x + (1 - e^{-\phi_t})(1 + 1/\alpha_t)0.5^{1/\alpha_t} x^{1-1/\alpha_t},$$

with

$$\begin{aligned}\alpha(\theta_{r,t}) &\approx 1 + \frac{0.5\theta_{r,t}}{\gamma}, \\ \phi(\theta_{r,t}) &\approx \frac{1}{4}\gamma^2 + \frac{1}{4}\theta_{r,t}\gamma + \frac{1}{16}\theta_{r,t}^2,\end{aligned}$$

for which follows that *in the steady state*

$$\frac{\partial s_t^x}{\partial \gamma} > 0 > \frac{\partial s_t^x}{\partial \theta_{r,t}} \quad \text{and} \quad \frac{\partial |\Delta s_t^x|}{\partial \theta_{r,t}}, \frac{\partial |\Delta s_t^x|}{\partial \gamma} > 0.$$

The weight on the most recent value s_{t-1}^x decreases in the transition speed, which depends positively on taxes $\theta_{r,t}$ and dispersion γ . This means that in terms of inequality dynamics, an increase in the standard deviation of portfolio returns is a complement to an increase in taxes and will speed up dynamics. However, in terms of inequality levels these two have opposing effects: an increase in taxes $\theta_{r,t}$ decreases the stationary level of inequality, while a higher value of γ will increase it.

We use this LOM (without the approximation) to estimate γ (and η if necessary) for each country by using numerical optimization routines to minimize the squared difference between a time series created by the LOM in Equation (4.11) and the available data. Denote as $s_t^{0.01}(\{\tau\}_0^t, S)$ the 1% share at t of a series initialized with $S \in (0, 1)$ and fed with the tax series $\{\tau\}_0^t$. This model-generated time series is compared to the empirical top 1% share at time t , denoted as $\hat{s}_t^{0.01}$. Then

$$\{\gamma, \eta\} = \arg \min \sum_{t=0}^T \left[s_t^{1\%}(\{\tau\}_0^t, \hat{s}_0^{1\%}) - \hat{s}_t^{1\%} \right]^2. \quad (5.3)$$

The resulting estimates are summarized in Table 1 and the (analytic) trajectories can be found in Figure 2.³³ We will discuss these results in detail below.

As reference, we also show simulations using Equations (3) and (5.2) and given the empirical tax rate at time t . These simulations are important in so far, that they allow us to identify the 95% confidence interval of our model, which is useful for evaluating its empirical fitness. Since

³²Keep in mind that for our quarterly calibration we set $\beta = 0.99 \approx 1$.

³³For further comparison, we also quantified the fitness using mean absolute percentage errors (MAPE) and the Pearson correlation coefficient (PCC):

	US	UK	FR	SE
MAPE	0.0639	0.0782	0.0904	0.0858
PCC	0.8777	0.9948	0.9579	0.9827

	US	UK	SE	FR
γ	0.3523	0.0398*	0.2012	0.1341
η	–	0.2484	0.3709	0.2372

Table 1: Parameter estimates used for closed-form low-of-motion and simulations. For the UK the analytic solution underestimates γ , which is why for the simulations $\gamma = 0.082$ is used.

the real world is a stochastic environment with a finite number of agents but the closed form assumes an infinite number of agents, we can assess the degree of discrepancy.

For simulations, we created an initial distribution Pareto distribution that satisfies $W = \sum w_0 = N = 50,000$ and matches the top 1 and top 0.1 percentiles from the WID database for the respective moment in time.³⁴ Starting from this distribution, the series follows the model outlined above while feeding in the tax series as explained above. Robustness checks for the variation of the free parameter γ are presented in Appendix B.

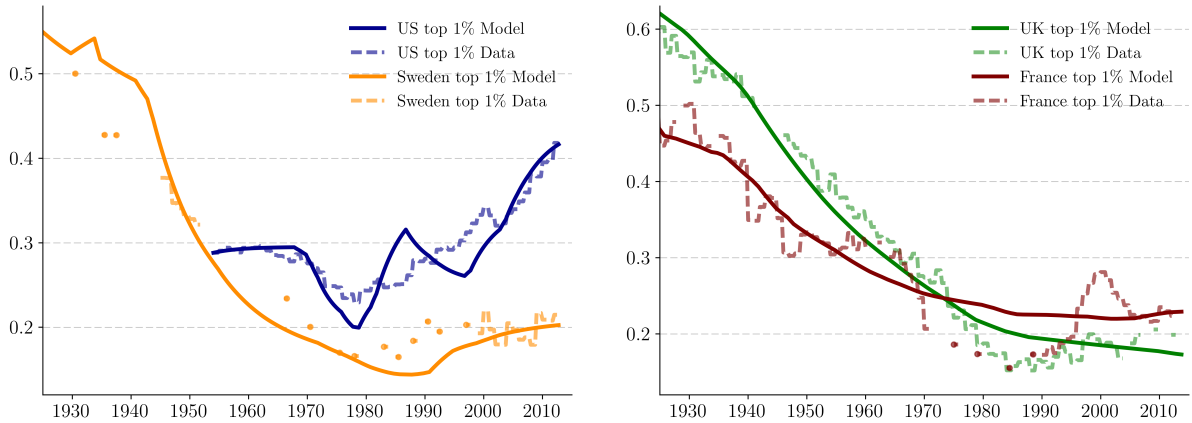


Figure 2: Trajectories using the law-of-motion in closed form (Equation 4.11) and the parameters from Table 1. Each time series is initialized with the top-share from the empirical data and then uses the available series of (capital) income taxes as the single input.

5.1 The United States

In the United States, individuals generally pay income tax on the net of their capital gains. There are, however, a reasonable number of exemptions, depending on investment duration, net-worth and status. The series in use here represents the maximum tax rate on long-term gains obtained from the US Tax Foundation. This has the great advantage that it is a tax explicitly and only on capital gains, which spares the need to use (and estimate) η to scale down top marginal income taxes. Yet, we note that reducing a complicated system of tax progression to just one number bears the risk of misalignment.

³⁴This can be done easily by using the well-known relationship with the share of the top $x\%$ given by $s^x = x^{1-1/a}$. We can then estimate the Pareto coefficient to fit the top 1% and the top 0.1%, create a distribution and check the fit with the empirical values. We do not intend to match the lower 50% interval since our model is one of investment and frictionless financial markets. These assumptions do not hold for the bottom of the wealth distribution.

In Figure 3, we plot our closed form solution for the 1% top shares and the US calibration. The dashed line stands for the analytic result of the final cross-sectional distribution after convergence and given the tax rate $\theta_{r,t}$ at time t . The solid line is the analytic LOM. For this reason the final analytic top-shares *jump* with each change in the tax rate and the numerical simulation follows more slowly, in line with the data.

The results, together with the series of taxes, are shown in Figure 4. The dashed blue/orange line depicts the original data while we add the median as well as the 95% confidence intervals from our simulations. Median and intervals from 200 simulations are calculated starting from the same distribution and using different random seeds. We initialize our simulation for the US in 1954 because from this year on capital gains taxes are available on an annual basis. In Appendix D we presents histograms of the overall simulated data and compare it to the closed-form solutions.

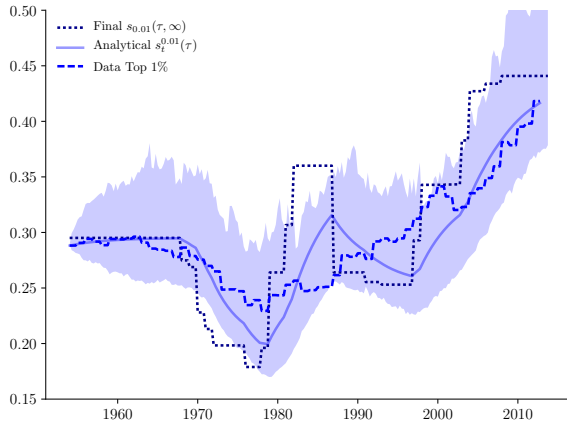


Figure 3: Closed form LOM $s_t^{0.01}(\tau_t)$ (solid line) and closed form stationary share $s^{0.01}(\tau_t, \infty)$ (dotted line) taking tax series and the calibration for the USA together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.

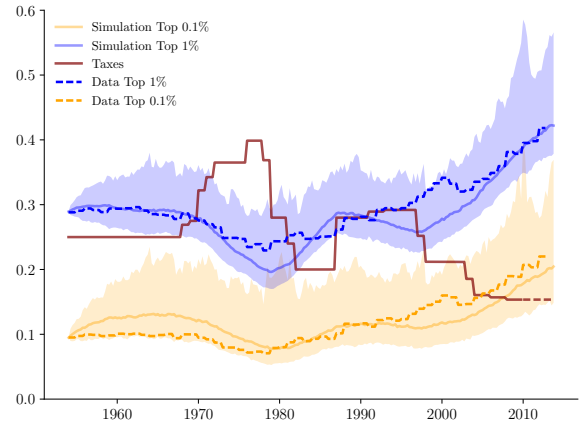


Figure 4: Simulations starting from the wealth distribution in 1954 for the USA, using only the respective time series of US capital gain taxes as input ($\gamma = 0.3523$).

The value of γ is 0.3523, which implies that the standard deviation of forecasting errors is yet very small compared to the real standard deviations of returns and hence also lies in a sensible range. Thus, our value of $\gamma = \beta \frac{\sigma_E}{\sigma_D}$ for the USA would be in line with relatively high disagreement variance and low return variance. Our simulation matches the data very well with the time series of the data, always lying within the 95% interval of our simulation and the median sticking close to the real time series.

In the late 1950s the series seems to be stable around the steady state given constant taxes. After successively increasing capital gain taxes starting in the late 1960s, inequality decreases up until the Reagan period, in which taxes return to the previous level. In particular, the levels as well as the responses to this tax increase are well matched while simultaneously providing realistic transition dynamics. The relaxation of taxes is accompanied by an immediate increase in wealth inequality both in the data and predicted by our model. While, after a short dip, taxes return to their level from the 1950s in the late 1980s, inequality stabilizes at the same postwar-

level in the data as well as in the model. Finally, the tax decreases in the late 1990s and the early 2000s initiate convergence to a steady state that is yet unreachable. It is furthermore noteworthy that, even though the 95% intervals indicate a considerable variation among simulations, the amount of total variation remains almost constant and the movement of the outer boundaries coincides with the movements in the data.

5.2 The United Kingdom

Let us now turn to the United Kingdom. The tax system in place was introduced in 1965. However, we start our investigation back in 1913. Thus, we rely on the only available tax series dating back that far, which is the series on top marginal income taxes provided by Piketty (2014). In fact, the actual capital gains tax rate in 2010 (depending on income) ranges from 18% up to 28% while allowing for several exemptions. This, again, leaves a considerable degree of freedom when attempting to summarize these tax regimes by one number per period, supposed to address only the top shares.

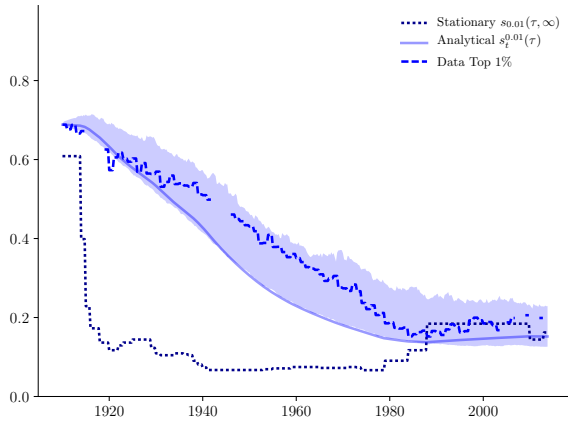


Figure 5: Closed form LOM $s_t^{0.01}(\tau_t)$ (solid line) and closed form stationary share $s^{0.01}(\tau_t, \infty)$ (dotted line) taking tax series and the calibration for the UK together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.

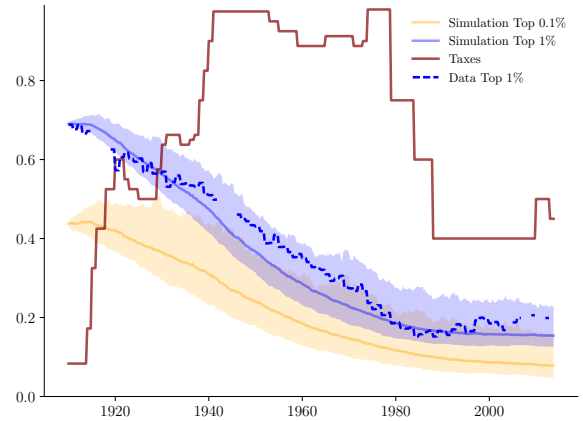


Figure 6: Simulations starting from the wealth distribution in 1913 for the UK, using only the respective time series of UK top income taxes as input. $\gamma = 0.082$ and η as in Table 1.

Figures 5 and 6 show our model results and the simulation for the UK starting in 1904 and document the income taxation for each period, taken from Piketty (2014). Compared to the US we opt for a γ of 0.082.³⁵ The ratio fits in line with the reported evidence for expectation disagreement (Greenwood and Shleifer, 2014) and return variance (Campbell and Viceira, 2002). Compared to the USA the value of γ is lower, implying less disagreement among traders. In

³⁵For the US and France we used the same values for our simulation that were identified by using the closed form solution. This is not feasible for the UK, since the second-order approximations used for our analytic results becomes imprecise when γ is very small. In particular, for $\alpha > 2$ the closed-form results fail to hold. Moreover, our closed-form solution slightly underestimates γ because it understates inequality when taxes are relatively high. For this reason, for the UK and Sweden, we repeated the estimation procedure using the median of the simulation instead of the analytic LOM.

fact, in the sample of the countries considered in this study the United Kingdom is the one with the lowest wealth inequality.

In accordance with the data, the model predicts an overall decrease in wealth inequality until the Thatcher years, in which - similar to the USA - taxes decreased considerably. This is reflected by a gradual increase and rise in inequality, both in the data and predicted in the model. Again, not only the levels but also the slow transitions are well matched. As shown in Figure 5 the stationary level of inequality in the periods of the tax hike until the 1980s is substantially below the transition level owing to low transition dynamics. This is especially surprising because one would expect to obtain a worse fit given that marginal top income taxes are not an optimal proxy for capital income taxes. Yet, after initiation of the distribution given only one relatively narrow bounded parameter, our model provides a good fit for over 100 years of data.

5.3 Sweden

Let us have a look at Sweden. The usual calibration procedure produces a value of $\eta = 0.3709$ higher than in the other considered cases. The estimate for γ is 0.2186, lower than the value for the USA, but higher than the value for the UK. The results are presented in Figures 7 and 8. Similar to the Anglo-Saxon economies, there was an increase in taxes and thus a decrease in wealth inequality until the 1980s in which the trend reversed. Roine and Waldenström (2015) discuss the evolution of wealth inequality in Sweden in detail. The strong increase in taxation after the Second World War was a reaction of the ruling Social Democratic Party to crowd out the nascent Communist movement in Sweden. Sweden had a wealth tax in place from 1911 until 2007, when it was abolished when a center-right government came to take over the government responsibilities. Due to this fact a pure wealth tax is not available for the complete sample. Thus - and therefore in line with our approach for the UK and France - we approximate the wealth tax by the income tax.

As documented in Du Rietz and Henrekson (2014) the top marginal wealth tax rate increased until the 1980s, decreasing afterwards in line with the evolution of top income tax rates. In general, Du Rietz and Henrekson (2014) presents a detailed description about the evolution of taxation in Sweden including capital gain and wealth taxes. Yet, for reasons of consistency with the data on wealth inequality we use the tax data provided by Lundberg and Waldenström (2017).

Once again we are able to match the decrease in wealth inequality until the 1980s (cf. figure 8). After a substantial tax decrease top wealth inequality is on the rise again. As furthermore presented in the closed form solution depicted in figure 7 it is, however, already close to its stationary value. As before, the transition dynamics are well in line with empirical evidence.

5.4 France

We also want to compare the two Anglo-Saxon economies and Sweden that underwent deep regulatory changes in the 1980s with France as an example for a central European economy. In France, capital gains taxes, in addition to taxes on sales of financial instruments, are generally subject to the marginal tax rate plus some social contributions. Certain deductions might be feasible depending again on holding duration and other individual indicators. Thus, we again keep in mind that these deductions, together with potential changes in tax progression, can not be captured well by the one-number-summary of each periods tax regime.

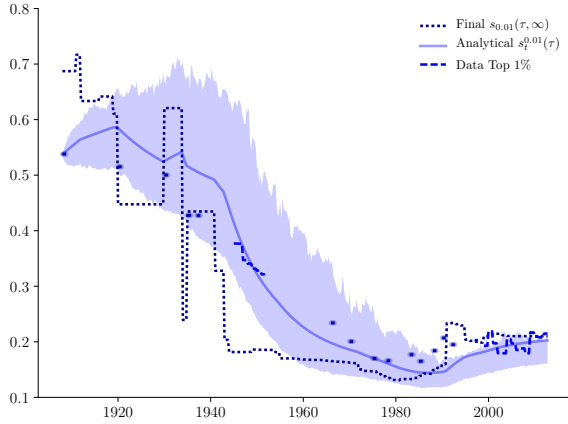


Figure 7: Closed form LOM $s_t^{0.01}(\tau_t)$ (solid line) and closed form stationary share $s^{0.01}(\tau_t, \infty)$ (dotted line) taking tax series and the calibration for Sweden together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.

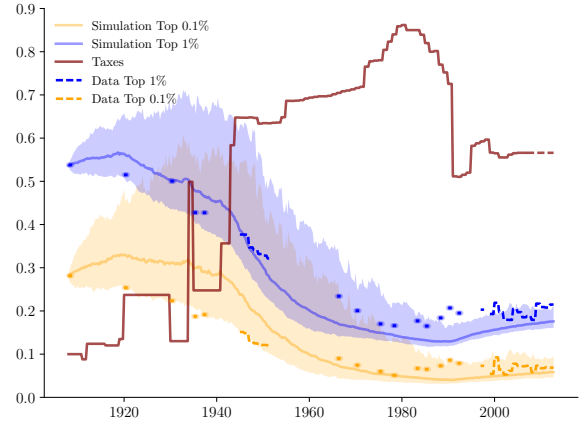


Figure 8: Simulations starting from the wealth distribution in 1908 for Sweden, using only the respective time series of Swedish top income taxes as input. Parameters as in Table 1.

We again start our investigation in the year 1917. From this year on data for the top income and top inheritance taxes are available (Piketty, 2014, chapter 14). Similar to the UK, the empirical time series on capital gain taxes is not available. The fitted value of γ amounts to 0.1341, slightly lower than for the case of Sweden, but higher than the value for the UK, while η is estimated to be 0.2372. Basically, we see an increase in taxes until World War II. Top income taxes leveled out afterwards only to slightly decrease after the 1980s. Yet, the decrease in taxes is not as pronounced as in the other countries covered, which explains the lower degree of wealth inequality.

Figure 10 presents simulation results. While in the beginning our model is able to match the data fairly well, the fit from the 1970s until the 1990s is rather loose. However, in this period only a few data points are provided and most of the series still remains within our 95% interval. It is furthermore noteworthy that the empirical series shows relatively large fluctuations, which is also reflected in relatively abrupt changes in the tax rate. The main driving force of this result is that the distribution starts off from a steady state with very low taxation, and then after the introduction of relatively high tax rates converges slowly to the lower steady state.

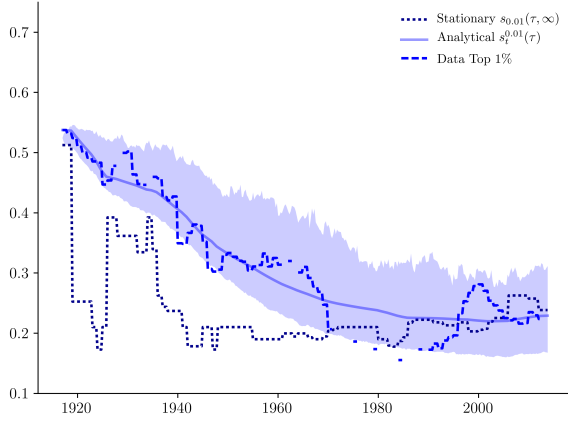


Figure 9: Closed form LOM $s_t^{0.01}(\tau_t)$ (solid line) and closed form stationary share $s^{0.01}(\tau_t, \infty)$ (dotted line) taking tax series and the calibration for France together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.

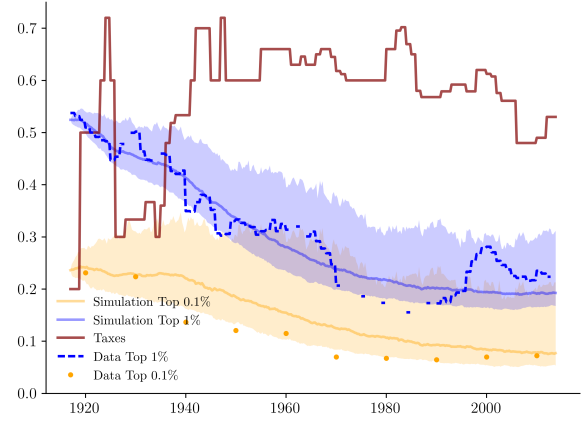


Figure 10: Simulations starting from the wealth distribution in 1917 for France, using only the respective time series of French top income taxes as input. Parameters as in Table 1.

6 Validation and Projections 2040

We validate our model through out-of-sample testing. Instead of estimating the free parameters γ (and potentially η) for each country given the full available dataset, we only use a shorter sample from the history to determine γ and then initiate our model into an episode that does not lie within the sample. For this purpose we make use of our closed-form solution from Equation (4.11) and feed in only the available tax series. The model simulations can then be compared with the empirical data for the periods after the sample phase.

The quality of our out-of-sample predictions depends crucially on the properties of the sample period. For the US, our sample starts in 1956 as above. The years thereafter stand out as a period of very stable inequality. As we have argued above, γ has two effects on inequality dynamics. First, a lower (higher) γ is associated with a lower (higher) steady state level of inequality. Second, a lower (higher) γ is linked with slower (faster) transitions. Hence, if the sample contains a stationary period, the respective value of γ can be easily identified through the first criterion. The sensitivity and robustness with respect to this parameter is further discussed in Appendix B.

For the case of the USA, proper identification is even easier because the stationary period is - after an increase in taxes in the late 1960s - followed by a significant decrease in inequality. Furthermore, there is no need to estimate η since capital taxes are readily provided. For this reason our US dataset allows a good approximation with $\gamma = 0.3497$ when using only a sample of 10 years (1956 - 1966). Increasing the sample size does not then have a significant effect on the estimate of γ . In Figure 11 (left) we show the fit of this simulation for different samples for the USA. For each sample, the fit is very close to our benchmark results from the previous section since the associated γ is very close to its value for the whole available dataset.

Unfortunately, such periods of stable inequality are absent in the data for the other countries.

Since for the UK, France, and Sweden wealth inequality is constantly decreasing until the 1980s, given any subsample of this data it is only possible to identify that inequality is at a transient path, but this is insufficient to narrow down our parameters. This can be readily understood by the same argument as before. If inequality is high and taxes are high too, inequality will decrease until the stationary level is reached where convergence speed and final level both depend on γ . If the value of γ is low, the final level will be low, but convergence speed will be low as well. If the sample only includes a decreasing path of inequality, it is not possible to identify whether the increase is due to slow convergence to a low level, or fast convergence to a high level, as long as the final level is not part of the sample data. Further, for the mentioned three countries it is also necessary to identify η , which reflects a further complication since even if a period of stable inequality would be given, the estimation would be singular. The parameter η has a contrasting effect compared to γ . An increase (decrease) also leads to an increase (decrease) in transition speed, but to a lower (higher) level of inequality.

We focus on the parameter estimation for Sweden here and redirect estimation and discussion for the UK and France to the Appendix. In short, while our procedure shows mild success for the UK given the limitations discussed above, the erratic nature of the data for France does not allow a precise and liable estimation of the parameters. We also adjust our estimation procedure slightly, by iteratively estimating only one parameter at a time while using the latest estimate of the others. This procedure is repeated until changes in the parameter estimates become marginal. This is to avoid the singularity problem of estimating both parameters simultaneously given a very small sample.³⁶

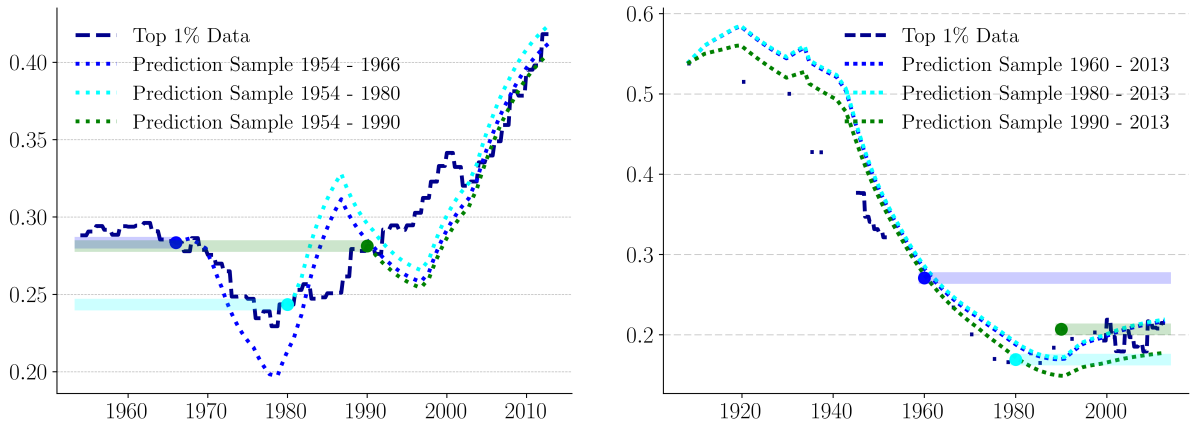


Figure 11: Out-of-sample tests for the US (left) and Sweden (right). The fat bars indicate the sample period.

Keeping in mind our discussion above, Sweden does not provide a stationary period either. To find a value for γ and η that captures the dynamics as precisely as in the previous section, the period that follows 1990 is essential. If we, however, for the reasons outlined above, only use this period and initialize the comparing simulation at $t = 0$, we are able to estimate γ and η such that they provide a fairly good fit for the out-of-sample period. In Figure 11 (center) we show out-of-sample predictions for Sweden in different sample sets ending in 2013.

³⁶In fact, this procedure then explicitly searches for local minima only, a compromise that we are willing to make in this context.

As discussed the out-of-sample predictions work well for the USA implying that we can also make reasonable future forecasts for a long and very long run horizon. The same holds true for the last 20 years of Swedish data featuring a rather constant level of inequality helping us to identify the parameters. Using the calibration from the previous section and starting from a Pareto distribution, we first simulate the last 20 years for which we know the empirical shares and the time series of taxes. This then also serves as a second test for our calibration, which was targeted to match a longer horizon.

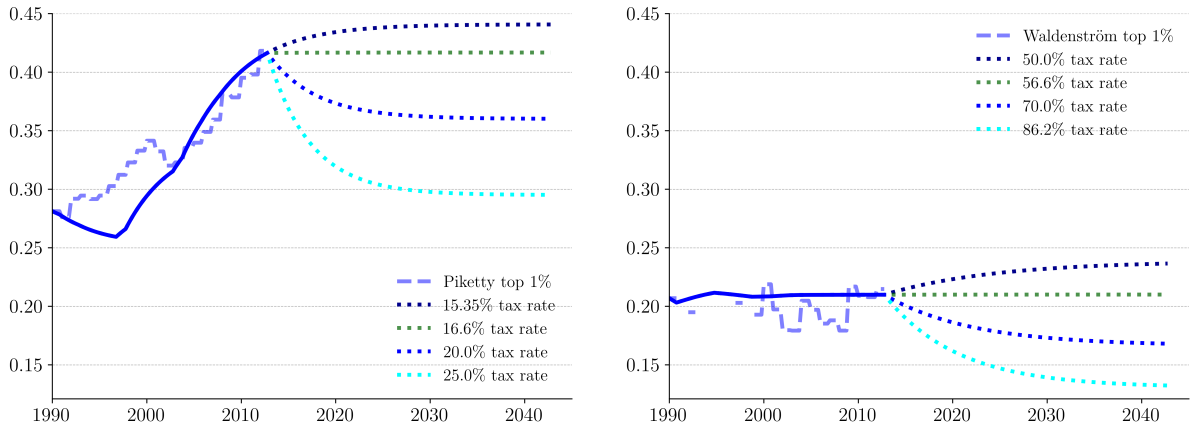


Figure 12: At the crossroads: projections for different tax regimes. For the US (left) the top dashed line projects the current tax rate, the second line shows that a rate of at least 16.6% would be necessary to keep inequality at the present level. A rate of 20% would partly reverse the trend. The bottom line shows a 25% rate. For Sweden a rate of 56.6% is required to keep inequality at the present level.

In Figure 12 (left) we show analytic forecasts given the US estimation and different tax regimes. A considerable increase in the capital gains tax rate to 25%, as for example currently prevailing in Germany would be necessary to reverse the trend and bring inequality back to the level of the 1990s. For the other countries, simulations and closed form solutions imply that, given timely tax rates, the stationary level of inequality is almost reached. Figure 12 (right) displays the (exemplary) behavior of Sweden, showing that for a tax rate of 56.6% wealth inequality is at a steady state level. Lower, respectively, higher tax rates will increase (decrease) the level of wealth inequality.

7 Conclusion

The main purpose of this work is to develop a simple, yet micro-founded model to explain the dynamics of wealth inequality given empirical tax series. Although a quite straightforward approach, this stands in contrast to the majority of the theoretical literature on wealth inequality which takes income inequality as a starting point.

We apply this method to the USA, the UK, France, and Sweden. For the USA we use capital taxes, while for the UK, Sweden and France marginal top income taxes are taken as a proxy. Due to the parsimonious nature of our model the degree of freedom to match the data is very limited. Nevertheless, our model matches the data surprisingly well, both in levels and also

in transition speed. Our analytic results emphasize that the level and the transition speed of wealth inequality depend crucially on the degree of capital taxation, which is well in line with our numerical results quantitatively and qualitatively. We conclude that the given tax series have a very high explanatory power regarding the dynamics of wealth distribution over the last 70-100 years.

This also implies that the answer to the question of how to approach – and potentially reverse – the recent increase in wealth inequality that can be observed in developed economies is considerably simple. An increase in capital gains taxes, or alternatively a gross tax on wealth, as suggested in Piketty (2014), will very likely mitigate the issue and has the potential to upturn these trends. Our projections predict that, in particular for the USA - continuing on the present path of capital taxation - the gap between the rich and poor is expected to further increase.

There are two implications for future research. Although our model fits the data quite well, there are periods where it falls short of accounting for the data. First, we consider it important to identify whether the loose fit is due to bad tax data or to reasons that are exogenous to our model. Second, if these reasons are exogenous it is crucial to investigate them further.

Of course, the quality of the model's result severely hinges on the quality of the data. In particular, a better availability of data on wealth dispersion at higher frequencies would give better means to testing and improving our model and improve the understanding of the issue of wealth inequality in the 21st century.

References

- Aiyagari, S. R., 1994. Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics* 109 (3), 659–84.
URL <http://ideas.repec.org/a/tpr/qjecon/v109y1994i3p659-84.html>
- Aiyagari, S. R., 1995. Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting. *Journal of Political Economy* 103 (6), 1158–75.
URL <http://ideas.repec.org/a/ucp/jpolec/v103y1995i6p1158-75.html>
- Alstadsaeter, A., Johannesen, N., Zucman, G., 2017. Tax Evasion and Inequality. Working Paper.
- Alvaredo, F., Atkinson, A. B., Morelli, S., 2017. Top Wealth Shares in the UK Over More than a Century. Dept. of Economics Research Paper Series No. 01/WP/2017, University Ca' Foscari of Venice.
- Aoki, S., Nirei, M., 2017. Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S. *American Economic Journal: Macroeconomics* forthcoming.
- Benhabib, J., Bisin, A., Zhu, S., 2011. The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents. *Econometrica* 79 (1), 123–157.
URL <http://ideas.repec.org/a/ecm/emetrp/v79y2011i1p123-157.html>
- Benhabib, J., Bisin, A., Zhu, S., 2014. The distribution of wealth in the blanchard–yaari model. *Macroeconomic Dynamics* FirstView, 1–16.
URL <http://journals.cambridge.org/article.S1365100514000066>
- Benhabib, J., Bisin, A., Zhu, S., 2015. The wealth distribution in bewley economies with capital income risk. *Journal of Economic Theory* 159, Part A, 489 – 515.
URL <http://www.sciencedirect.com/science/article/pii/S0022053115001362>
- Bewley, T., 1977. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory* 16 (2), 252–292.
URL <https://ideas.repec.org/a/eee/jetheo/v16y1977i2p252-292.html>
- Cagetti, M., Nardi, M. D., 2009a. Estate Taxation, Entrepreneurship, and Wealth. *American Economic Review* 99 (1), 85–111.
URL <https://ideas.repec.org/a/aea/aecrev/v99y2009i1p85-111.html>
- Cagetti, M., Nardi, M. D., 2009b. Estate Taxation, Entrepreneurship, and Wealth. *American Economic Review* 99 (1), 85–111.
URL <https://ideas.repec.org/a/aea/aecrev/v99y2009i1p85-111.html>
- Campbell, J. Y., Viceira, L. M., 2002. Strategic Asset Allocation: Portfolio Choice for Long-Term Investors. No. 9780198296942 in OUP Catalogue. Oxford University Press.
URL <https://ideas.repec.org/b/oxp/obooks/9780198296942.html>
- Castaneda, A., Diaz-Gimenez, J., Rios-Rull, J.-V., 2003. Accounting for the U.S. Earnings and Wealth Inequality. *Journal of Political Economy* 111 (4), 818–857.
URL <http://ideas.repec.org/a/ucp/jpolec/v111y2003i4p818-857.html>

- Clemens, C., Heinemann, M., 2015. Endogenous growth and wealth inequality under incomplete markets and idiosyncratic risk. *Journal of Macroeconomics* 45 (C), 300–317.
URL <https://ideas.repec.org/a/eee/jmacro/v45y2015icp300-317.html>
- Craig, C. C., 03 1936. On the frequency function of xy . *Ann. Math. Statist.* 7 (1), 1–15.
URL <http://dx.doi.org/10.1214/aoms/1177732541>
- Domeij, D., Heathcote, J., 2004. On The Distributional Effects Of Reducing Capital Taxes. *International Economic Review* 45 (2), 523–554.
URL <https://ideas.repec.org/a/ier/iecrev/v45y2004i2p523-554.html>
- Du Rietz, G., Henrekson, M., 2014. Swedish Wealth Taxation (1911–2007). Working Paper Series 1000, Research Institute of Industrial Economics.
URL <https://ideas.repec.org/p/hhs/iuiwop/1000.html>
- Fernholz, R., Fernholz, R., 2014. Instability and concentration in the distribution of wealth. *Journal of Economic Dynamics and Control* 44 (C), 251–269.
URL <http://ideas.repec.org/a/eee/dyncon/v44y2014icp251-269.html>
- Gabaix, X., Lasry, J.-M., Lions, P.-L., Moll, B., 2016. The dynamics of inequality. *Econometrica* 84 (6), 2071–2111.
URL <http://dx.doi.org/10.3982/ECTA13569>
- Garbinti, B., Goupille-Lebret, J., Piketty, T., 2017. Accounting for wealth inequality dynamics: Methods, estimates and simulations for france (1800-2014). CEPR Discussion Papers 11848, C.E.P.R. Discussion Papers.
URL <https://ideas.repec.org/p/cpr/ceprdp/11848.html>
- Greenwood, R., Shleifer, A., 2014. Expectations of Returns and Expected Returns. *Review of Financial Studies* 27 (3), 714–746.
URL <https://ideas.repec.org/a/oup/rfinst/v27y2014i3p714-746..html>
- Hommes, C., 2013. Behavioral rationality and heterogeneous expectations in complex economic systems. Cambridge University Press.
- Hubmer, J., Krusell, P., Anthony A. Smith, J., 2016. The Historical Evolution of the Wealth Distribution: A Quantitative-Theoretic Investigation. NBER Working Papers 23011, National Bureau of Economic Research, Inc.
URL <https://ideas.repec.org/p/nbr/nberwo/23011.html>
- Huggett, M., 1993. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control* 17 (5-6), 953–969.
URL <http://ideas.repec.org/a/eee/dyncon/v17y1993i5-6p953-969.html>
- Judd, K. L., 1985. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics* 28 (1), 59–83.
URL <http://ideas.repec.org/a/eee/pubeco/v28y1985i1p59-83.html>
- Kaplan, G., Moll, B., Violante, G. L., 2016. Monetary Policy According to HANK. NBER Working Papers 21897, National Bureau of Economic Research, Inc.
URL <https://ideas.repec.org/p/nbr/nberwo/21897.html>

- Karlin, S., Taylor, H., 1981. *A Second Course in Stochastic Processes*. No. 2. Academic Press.
- Kaymak, B., Poschke, M., 2016. The evolution of wealth inequality over half a century: The role of taxes, transfers and technology. *Journal of Monetary Economics* 77 (C), 1–25.
URL <https://ideas.repec.org/a/eee/moneco/v77y2016icp1-25.html>
- Krusell, P., Smith, T., 1998. Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy* 106 (5), 867–896.
URL <http://ideas.repec.org/a/ucp/jpolec/v106y1998i5p867-896.html>
- Lansing, K. J., 1999. Optimal redistributive capital taxation in a neoclassical growth model. *Journal of Public Economics* 73 (3), 423–453.
URL <http://ideas.repec.org/a/eee/pubeco/v73y1999i3p423-453.html>
- Levhari, D., Srinivasan, T. N., 1969. Optimal savings under uncertainty. *The Review of Economic Studies* 36 (2), 153–163.
- Lundberg, J., Waldenström, D., 2017. Wealth inequality in sweden: What can we learn from capitalized income tax data? *Review of Income and Wealth* (77(1)), n/a–n/a.
URL <http://dx.doi.org/10.1111/roiw.12294>
- Piketty, T., 2014. *Capital in the twenty-first century*. Harvard University Press, Boston.
- Pulley, L. B., 1983. Mean-variance approximations to expected logarithmic utility. *Operations Research* 31 (4), 685–696.
- Quadrini, V., 2000. Entrepreneurship, Saving and Social Mobility. *Review of Economic Dynamics* 3 (1), 1–40.
URL <https://ideas.repec.org/a/red/issued/v3y2000i1p1-40.html>
- Roine, J., Waldenström, D., 2015. Chapter 7 - long-run trends in the distribution of income and wealth. In: Atkinson, A. B., Bourguignon, F. (Eds.), *Handbook of Income Distribution*. Vol. 2 of *Handbook of Income Distribution*. Elsevier, pp. 469 – 592.
URL <http://www.sciencedirect.com/science/article/pii/B9780444594280000084>
- Saez, E., Zucman, G., 2016. Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data. *The Quarterly Journal of Economics* 131 (2), 519–578.
URL <https://ideas.repec.org/a/oup/qjecon/v131y2016i2p519-578..html>
- Shorrocks, A., Davies, J. B., Lluberas, R., 2016. *Global Wealth Databook 2016*. Research institute, Credit Suisse.
- Stiglitz, J. E., 1969. The Effects of Income, Wealth, and Capital Gains Taxation on Risk-Taking. *The Quarterly Journal of Economics* 83 (2), 263–283.
URL <https://ideas.repec.org/a/oup/qjecon/v83y1969i2p263-283..html>
- Straub, L., Werning, I., 2014. Positive Long Run Capital Taxation: Chamley-Judd Revisited. NBER Working Papers 20441, National Bureau of Economic Research, Inc.
URL <http://ideas.repec.org/p/nbr/nberwo/20441.html>

Toda, A. A., 2012. The double power law in income distribution: Explanations and evidence. *Journal of Economic Behavior & Organization* 84 (1), 364 – 381.

URL <http://www.sciencedirect.com/science/article/pii/S0167268112000923>

Wold, H. O. A., Whittle, P., 1957. A model explaining the pareto distribution of wealth. *Econometrica* 25 (4), 591–595.

URL <http://www.jstor.org/stable/1905385>

Appendix

A Proofs of the Propositions

A.1 Proof of Proposition 1

The product-normal distribution is treated extensively in Craig (1936). The probability distribution function is given by

$$f(z_{PN}) = \frac{1}{\pi} K_0(|z_{PN}|), \quad (\text{A.1})$$

with $z_{PN} \equiv \epsilon^1 \epsilon^2$ with $\epsilon^i \sim N(0, 1)$ and K_0 being the modified Bessel-function of the second kind. The function is symmetric around the mean of zero and exhibits leptokurtic behavior. It is more appealing to write this using the Moment-Generating Function (MGF), which in this case is given by

$$M_{Z_{PM}}(t) = \frac{1}{\sqrt{1-t^2}}. \quad (\text{A.2})$$

Using this it is easy to show that the mean and skewness are zero, while the standard deviation is given by

$$SD(z_{PN}) = 1. \quad (\text{A.3})$$

This distribution is highly comparable to the Laplace distribution. For a zero-mean the probability density function of the latter is given by

$$f(z_L) = \frac{1}{2\lambda} \exp\left(-\frac{|z_L|}{\lambda}\right) \quad (\text{A.4})$$

for shape parameter $\lambda > 0$, having both a mean and a skewness of zero. The standard deviation of Laplace is

$$SD(z_L) = \sqrt{2}\lambda. \quad (\text{A.5})$$

The Laplace distribution is also very appealing as each half takes the form of an exponential function. The moment generating function of the Laplace distribution is

$$M_{Z_L}(t) = \frac{1}{1-\lambda^2 t^2}. \quad (\text{A.6})$$

Comparing this with the MGF of the product-normal distribution it becomes obvious that the two are not identical. In fact, the sum of two product-normal variables follows a Laplace distribution.³⁷

³⁷Using the MGF it is easy to show that if there are four independently distributed normal shocks with zero mean $X_i \sim N(0, \sigma_i)$ and we have $\sigma_1 \sigma_2 = \sigma_3 \sigma_4$ then $X_1 X_2 + X_3 X_4$ follows a Laplace distribution with zero mean and $\lambda = 1$.

As a reasonable approximation we replace the product-normal with the Laplace distribution. To obtain the shape parameter λ that best approximates the standard normal product distribution we equalize the second order Taylor expansions of both MGFs around $t = 0$, which in fact is equivalent to choosing λ to match the first two moments of the function. This yields

$$\begin{aligned} \sum_{n=0}^2 \frac{\partial^n M_{Z_{PM}}(0)}{n! \partial t^n} (t-0)^n &= \sum_{n=0}^2 \frac{\partial^n M_{Z_L}(0)}{n! \partial t^n} (t-0)^n \\ 1 + \frac{t}{2} &= 1 + t^2 \lambda^2 \\ \lambda &= \sqrt{\frac{1}{2}} \approx 0.707. \quad \square \end{aligned}$$

A.1.1 Proof of Proposition 3

From the CDF given by

$$F(w) = \begin{cases} 1 - 0.5w^{-\alpha} & w \geq 1 \\ 0.5w^\alpha & 0 < w < 1, \end{cases} \quad (\text{A.7})$$

we can compute the Lorenz-curve with

$$L(F) = \frac{\int_0^F w(F') dF'}{\int_0^1 w(F') dF'}. \quad (\text{A.8})$$

This requires the inverse of the CDF which takes the following form

$$w(F) = \begin{cases} 2^{-\frac{1}{\alpha}}(1-F)^{-\frac{1}{\alpha}} & F \geq 0.5 \\ 2^{\frac{1}{\alpha}}F^{\frac{1}{\alpha}} & 0 < F < 0.5. \end{cases} \quad (\text{A.9})$$

With integration by parts we can write

$$L(F) = \frac{\int_0^{0.5} w(F') dF' + \int_{0.5}^F w(F') dF'}{\int_0^{0.5} w(F') dF' + \int_{0.5}^1 w(F') dF'}, \quad (\text{A.10})$$

for the relevant part $F > 0.5$ (the *rich*). Integrating in the boundaries leads to

$$L(F) = \frac{\frac{\alpha}{\alpha+1}0.5 - \frac{\alpha}{1-\alpha}0.5 + 2^{-\frac{1}{\alpha}}\frac{\alpha}{1-\alpha}(1-F)^{1-\frac{1}{\alpha}}}{\frac{\alpha}{\alpha+1}0.5 - \frac{\alpha}{1-\alpha}0.5}. \quad (\text{A.11})$$

Let us define the denominator $z \equiv \frac{\alpha}{\alpha+1}0.5 - \frac{\alpha}{1-\alpha}0.5 = -\frac{\alpha^2}{1-\alpha^2}$. Thus, we have

$$L(F) = 1 + \frac{2^{-\frac{1}{\alpha}}\frac{\alpha}{1-\alpha}(1-F)^{1-\frac{1}{\alpha}}}{z} \quad (\text{A.12})$$

Following the logic of the Lorenz-curve the share s^x of the top x is given by

$$s^x = 1 - L(1-x). \quad (\text{A.13})$$

Using the definition we can write

$$s^x = -\frac{2^{-\frac{1}{\alpha}} \frac{\alpha}{1-\alpha} (1-F)^{1-\frac{1}{\alpha}}}{z} = -\frac{2^{-\frac{1}{\alpha}} \frac{\alpha}{1-\alpha} (x)^{1-\frac{1}{\alpha}}}{z} = \frac{1+\alpha}{\alpha} 2^{-\frac{1}{\alpha}} x^{1-\frac{1}{\alpha}}. \quad \square \quad (\text{A.14})$$

It is important to point out the relationship to the standard Pareto-distribution, for which the share is given by $s^x = x^{1-\frac{1}{\alpha}}$. For the symmetric double Pareto it amounts to $s = Kx^{1-\frac{1}{\alpha}}$ with $K = \frac{1+\alpha}{\alpha} 2^{-\frac{1}{\alpha}} \geq 1$. Note that for the special case of $\alpha = 1$ respectively the inverse case $\alpha \rightarrow \infty$ we have $K = 1$. In fact, K is a u-shaped function with a local maximum at $\alpha = \frac{\ln(2)}{1-\ln(2)} \approx 2.26$ amounting to $K = \frac{2}{e \ln(2)} \approx 1.06$. Thus, the top share is slightly higher than in the standard Pareto-case. Yet, the standard Pareto-approach fares quite well.

A.1.2 Proof of analytic expression for the Gini coefficient

Using the algebraic expression of the Lorenz-curve we can also derive a value for the Gini coefficient defined as

$$Gini = 1 - 2 \int_0^2 L(F) dF = 1 - 2 \left[\int_0^{0.5} L(F) dF + \int_{0.5}^1 L(F) dF \right]. \quad (\text{A.15})$$

The overall Lorenz-curve is defined as

$$L(F) = \begin{cases} \frac{1-\alpha}{\alpha} 2^{\frac{1}{\alpha}} F^{\frac{1}{\alpha}+1} & 0 < F < 0.5 \\ 1 - \frac{1+\alpha}{\alpha} 2^{-\frac{1}{\alpha}} (1-F)^{1-\frac{1}{\alpha}} & F \geq 0.5. \end{cases} \quad (\text{A.16})$$

With the integration by parts we have

$$Gini = 1 - 1 - 0.5 \left(\frac{\alpha-1}{1+2\alpha} + \frac{1+\alpha}{1-2\alpha} \right) = \frac{3\alpha}{4\alpha^2-1}. \quad \square \quad (\text{A.17})$$

A more general result for a non-symmetric Laplace distribution is given in Toda (2012). Note that the Gini coefficient for a double Pareto distribution is always larger than for a standard Pareto (for which it amounts to $\frac{1}{2\alpha-1}$) for a reasonable $\alpha > 1$. Note that the symmetric double Pareto distribution underestimates overall inequality due to not accounting for the close to zero or even negative net worth for low income households. In order to account for the top 1% share of approx. 60% (as being the peak of wealth inequality in the UK) we would require $\alpha \approx 1.3$, implying a Gini of approx. 0.68. Recent empirical evidence for more moderate inequality suggest a value of larger than 0.7 for the considered countries (Shorrocks et al., 2016). For $\alpha \approx 1.3$ the share of the bottom 50% would be given by $s^{0.5} = 11.5\%$. For the countries discussed this share is always below 10%.

A.1.3 Proof of Proposition 5

A change in taxation influences the mean reversion speed $\mu = -\frac{\tau}{\lambda} - \frac{\gamma^2}{\lambda^2} (1-\tau)^2 < 0$. It leads to a jump in μ , changing the form of the Pareto-distribution (cf. figure 2(b) of Gabaix et al. (2016) as well as figure 17). For an increase in inequality (lower taxes) the Pareto-coefficient decreases firstly at the mode of the distribution only to fade out to the tails. The same holds true for the top-share which, given the average convergence speed ϕ , can be described by

$$s_{t+n}^x = \exp(-\phi n) s_t^x + (1 - \exp(-\phi n)) s^x(\tau_t, \infty). \quad (\text{A.18})$$

This would be the case for a one-time shock with $s_{n=0}^x = s_t^x$ and $s_{n \rightarrow \infty}^x = s^x(\tau_t, \infty)$. A continuous time representation of the convergence process is given by the following differential equation

$$\frac{ds^x}{dt} = \phi(s^x(\tau_t, \infty) - s^x), \quad (\text{A.19})$$

which is solved by the aforementioned equation for an initial condition s_t^x . The discrete time equivalent would be a mean-reversion process

$$s_t^x - s_{t-1}^x = (1 - \rho)(s^x(\tau_t, \infty) - s_{t-1}^x) \leftrightarrow s_t^x = \rho s_{t-1}^x + (1 - \rho)s^x(\tau_t, \infty), \quad (\text{A.20})$$

with $\Delta t = 1$. Comparing Equation A.18 with Equation A.20 for $n = \Delta t = 1$ it is easy to see that $\rho = \exp(-\phi)$.

Given the overall time varying nature of the tax rate τ_t the long-run steady state $s^x(\tau_t, \infty)$ is also time varying. Moreover, the reversion speed ϕ_t is also subject to time changes. Thus, the final recursive equation reads

$$s_t^x(\tau_t) = \exp(-\phi_t)s_{t-1}^x + (1 - \exp(-\phi_t))s^x(\tau_t, \infty). \quad \square \quad (\text{A.21})$$

B Robustness with Respect to γ

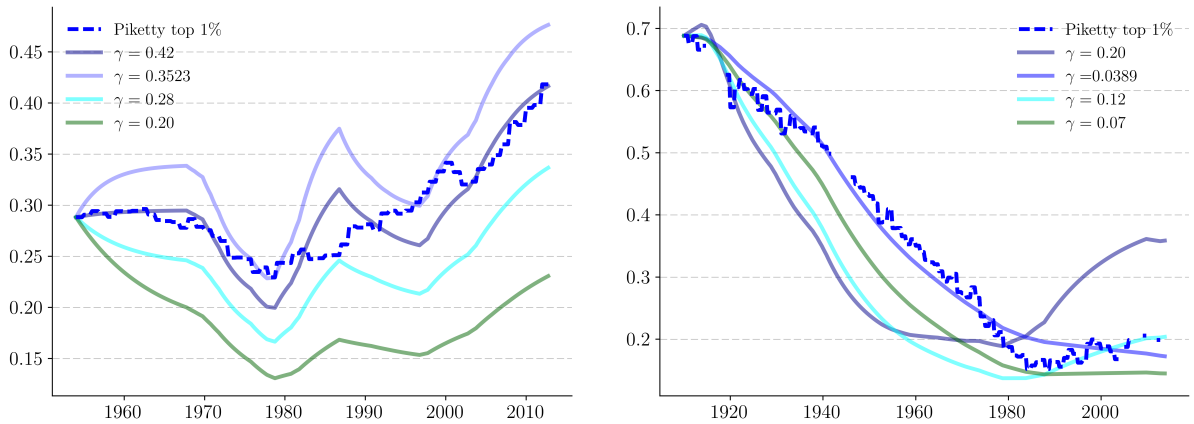


Figure 13: Variations of γ in the closed-form solutions for the USA (left) and the UK (right).

In Figures 13 and 14 we plot the closed-form LOM with different values of γ for all considered countries in order to show that our results are robust even when γ varies considerably. In general - and as formally stated - a lower value of the free parameter γ implies both lower inequality and slower dynamics. This implies a potential effect that while converging to lower equilibrium a low γ can (temporarily) be accompanied by higher inequality (cf. e.g. the case for France with $\gamma = 0.05$). Yet, in the stationary case a lower γ is always accompanied by lower inequality.

Moreover, these simulations imply that in particular for the UK, France and Sweden our results until the 1980s are considerably robust with respect to γ .

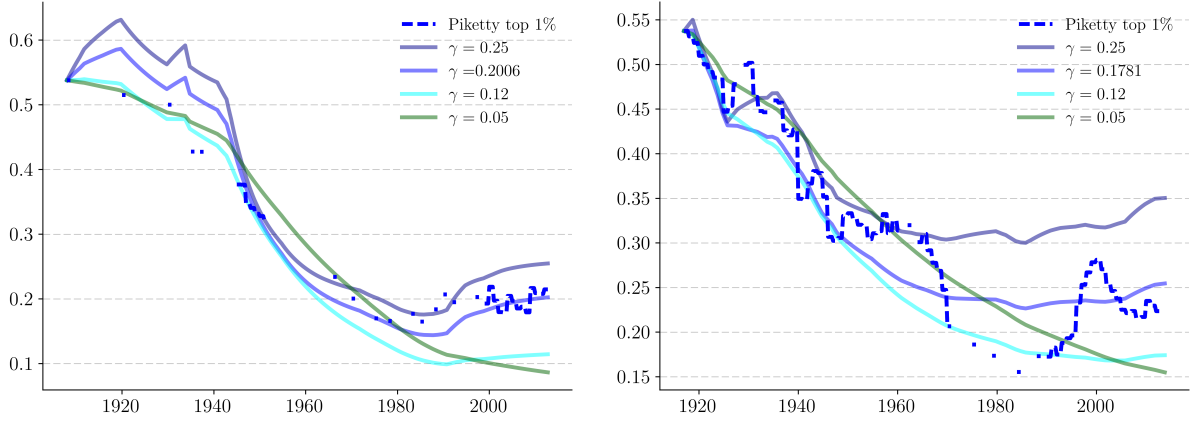


Figure 14: Variations of γ in the closed-form solutions for Sweden (left) and France (right).

C Out-of-sample tests for the UK and France

As discussed in the main body correct identification of γ and η for the UK is difficult if one excludes the stable (final) period from the data. Any period before the 1990s is insufficient to nail down γ . Only when considering the 1990s would we be able to conduct a meaningful prediction, but then we would be short of further datapoints to confirm the quality of our predictions. The results are shown in Figure 15 (left). A further problem is induced by the fact that, for France and Sweden, in a considerable time span only few data points are given.

For the sake of completeness we also perform out-of-sample forecasts for France. While the period from 1950 to 1970 is relatively stable, our above simulations fall short of accounting for the relatively low inequality from 1970 until 1990. Respectively, when excluding the years since 1970 the estimate for γ is around 0.23 regardless of when the calibration period begins. However, the quality of the out-of-sample predictions then crucially depends on the initialization date. If the simulation is then already initialized in the late 1960s, predicted inequality will be higher than documented by the date. Likewise, the level of inequality in the year 2000 is underestimated when we initiate our simulation around 1980. This indicates, again, that the tax changes proceeding the 1980s are very relevant to both forming the level and dynamics of inequality. The predictions are shown in Figure 15 (right). Since our attempt to match the data for France in the previous section has led to rather mixed results, it was not to be expected that the model would perform extraordinarily well in the out-of-sample forecasts.

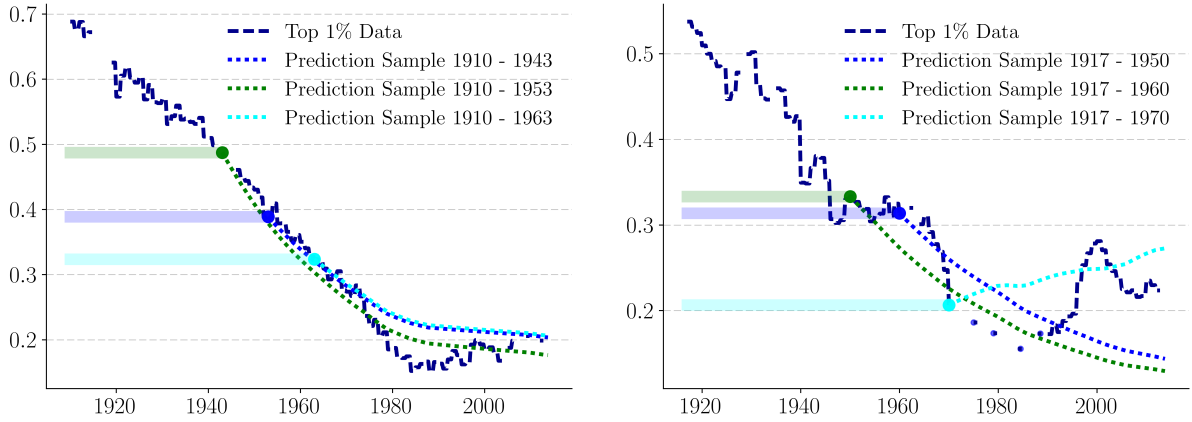


Figure 15: Out-of-sample predictions for the UK (left) and France (right). Fat bars indicate the respective sample period.

D Additional Notes on the Stationary Distribution

In Section 4 we present a closed-form solution for the stationary distribution of our model. Figure 17 presents a close-up look at the right tail with a Pareto-fit. The fit misses the extreme tail of the very rich (less than 0.1% of the population). As discussed in Gabaix et al. (2016), this is due to the slow convergence in the tail.³⁸ The change in inequality - here an increase in inequality - has not yet reached the very right tail.

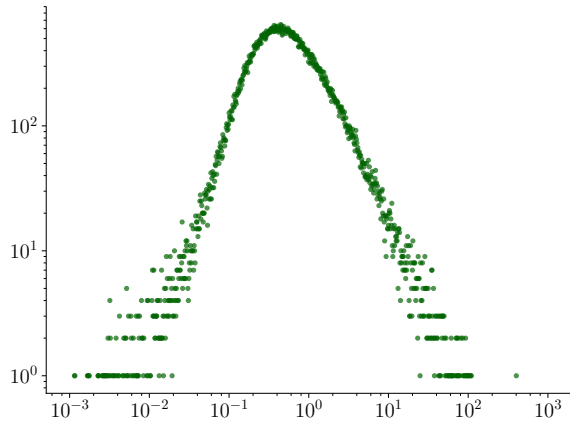


Figure 16: Log histogram of simulation using US calibration after a very long horizon of 50,000 periods. $\gamma = 0.3523$, $\tau = 25\%$ and $N = 100,000$.

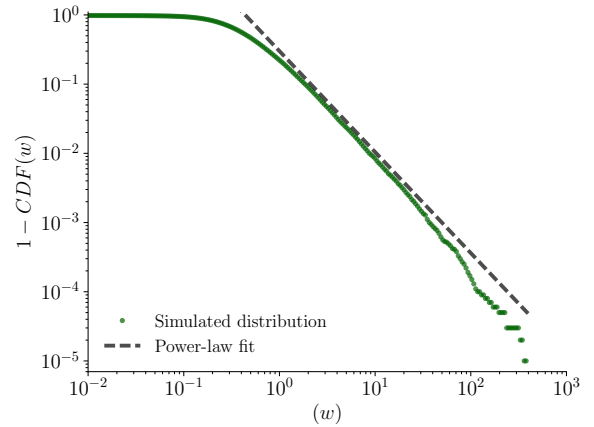


Figure 17: Right tail of the distribution for the last period for exemplary simulation using the US calibration of $\gamma = 0.3523$, $\bar{\omega}_{US}^{-1} = 0$ and $\theta_t = 0.25$ at all t . The dashed line represents a Pareto-fit with an estimated coefficient of 1.459.

Figure 16 presents the histogram of the wealth distribution as created by an exemplary

³⁸The interested reader is referred in particular to Figure 2 of Gabaix et al. (2016).

simulation using the USA calibration and tax time series. As predicted by the computation the distribution is symmetric around the mode of 1.