Wealth Inequality and the Ergodic Hypothesis: Evidence from the United States

Yonatan Berman* Ole Peters† Alexander Adamou‡ §

September 10, 2020

Abstract

Many studies of wealth inequality make the ergodic hypothesis that rescaled wealth converges rapidly to a stationary distribution. Changes in distribution are expressed through changes in model parameters, reflecting shocks in economic conditions, with rapid equilibration thereafter. Here we test the ergodic hypothesis in an established model of wealth in a growing and reallocating economy. We fit model parameters to historical data from the United States. In recent decades, we find negative reallocation, from poorer to richer, for which no stationary distribution exists. When we find positive reallocation, convergence to the stationary distribution is slow. Our analysis does not support using the ergodic hypothesis in this model for these data. It suggests that inequality evolves because the distribution is inherently unstable on relevant timescales, regardless of shocks. Studies of other models and data, in which the ergodic hypothesis is made, would benefit from similar tests.

Keywords: Wealth inequality, Wealth dynamics, Ergodic hypothesis, Stochastic processes

JEL Codes: C1, D3, G5

---

*London Mathematical Laboratory y.berman@lml.org.uk

†London Mathematical Laboratory and Santa Fe Institute, o.peters@lml.org.uk

‡London Mathematical Laboratory, a.adamou@lml.org.uk

§We dedicate this study to the memory of Tony Atkinson, who provided invaluable comments to us in Oxford in October 2016. His encouragement was crucial to our work. We thank Jean-Philippe Bouchaud for sharing his expertise on the solutions of Equation (2.7), Tobias Morville, and two anonymous referees, as well as numerous seminar and conference participants, for helpful discussions and comments. The code to generate all figures is available at https://github.com/LMLhub/WealthInequalityCodes.

1 Introduction

Paul Samuelson identified the ergodic hypothesis as "a belief in unique long-run equilibrium independent of initial conditions." (Samuelson, 1968, pp. 11-12). Traditionally, the equilibrium is studied and the transient phenomena preceding it are ignored. This makes models tractable and simplifies analysis by confining it to steady states. As a result, the ergodic hypothesis is made widely. Here we present the first empirical test of its validity as applied to the distribution of wealth. We find it inconsistent with historical wealth data from the United States (US).

Studies of wealth distributions treat individual wealth as a growing quantity, which is trivially non-ergodic. Dividing individual wealth by the population average produces a rescaled wealth, which has the potential to be ergodic and is hypothesized to be so.[1] Formally, the distribution of rescaled wealth is assumed to converge to a unique and time-independent distribution. For example, Benhabib, Bisin and Zhu (2011, p. 130) "impose assumptions [...] that guarantee the existence and uniqueness of a limit stationary distribution." Similar formulations appear across the literature.[2] With models thus constrained, stationary distributions can be analyzed, *e.g.* to examine the effects of shocks in policies and conditions, encoded as changes in model parameters.

Use of the ergodic hypothesis carries a second, implicit assumption: that convergence is fast. The stationary distribution is practically relevant only if approached on a timescale shorter than that between shocks. For example, the top rate of income tax in the US was cut to 70% in 1965 and then to 50% in 1982. Here convergence would be fast, and distributional analysis useful, if the model distribution approached its asymptotic form after, say, 5 years. Convergence would be slow, and the stationary distribution irrelevant, if it took, say, 100 years. As Atkinson (1969, p. 137) argued, "the speed of convergence makes a great deal of difference to the way in which we think about the model."

Recent research has moved away from studying stationary distributions in isolation. Gabaix et al. (2016); Berman, Ben-Jacob and Shapira (2016); Kaymak and Poschke (2016); Berman and Shapira (2017); Benhabib, Bisin and Luo (2019) study the dynamics of wealth and income, how they are distributed asymptotically, and (in some cases) how rapidly they converge.

[1] This is an example of a more general strategy in science of finding ergodic transformations of non-ergodic observables to simplify their analysis. For instance, Peters and Gell-Mann (2016) show that utility theory can be viewed as the construction of ergodic growth rates from non-ergodic wealth changes. Expectation values, which would otherwise be misleading, then quantify the long-run growth of wealth.

[2] See, for example, Stiglitz (1969); Bewley (1977); Piketty and Saez (2013); De Nardi (2015); De Nardi, Fella and Yang (2015); Jones (2015).

There is, however, an elephant in the room. To our knowledge, the validity of the ergodic hypothesis for rescaled wealth has never been tested empirically. Here we present such a test. For the test to be possible, a model is required that does not assume ergodicity from the outset. It must have regions of parameter space ("regimes") in which the ergodic hypothesis does and does not hold.

Our model of wealth, Reallocating Geometric Brownian Motion (RGBM), satisfies this condition. In it, individual wealth undergoes noisy exponential growth and all individuals pool and share their wealth at a fixed rate. It is a minimal model of wealth in a growing economy whose participants interact, at which several authors have arrived independently (Marsili, Maslov and Zhang, 1998; Bouchaud and Mézard, 2000; Liu and Serota, 2017). To a first approximation, the sign of the reallocation rate parameter determines whether rescaled wealth is ergodic: if positive (reallocation from richer to poorer), then it is; if negative (from poorer to richer), then it is not.

RGBM is known to reproduce several important stylized facts. When the reallocation rate is positive, the distribution of rescaled wealth converges to a stationary distribution with a Pareto tail (Pareto, 1897). Larger fluctuations and slower reallocation lead to a fatter tail. Typical fitted parameters are consistent with observed tail indices (Vermeulen, 2018).

We test the ergodic hypothesis by fitting RGBM parameters to historical wealth shares in the US population (Saez and Zucman, 2016). The sign of the fitted reallocation rate tells us whether a stationary distribution exists in the model. When convergence is possible, we evaluate convergence times to see whether the ergodic hypothesis is usable in practice.

Our results do not validate the ergodic hypothesis. The fitted reallocation rate is not robustly positive, and has been consistently negative since the 1980s. We cannot overstate our surprise at this finding. Initially, we imagined a model with no reallocation as an extreme and unrealistic model of an advanced economy. We expected to infer consistently positive reallocation rates from data and our aim was to quantify these rates. Instead, we find the opposite: from the 1980s onwards, the US economy is best described in our model as one where wealth is systematically reallocated from poorer to richer, and where the distribution of rescaled wealth never equilibrates

Even when fitted reallocation rates are positive, convergence to the stationary distribution takes decades or centuries. Thus, our analysis does not support use of the ergodic hypothesis, either because there is no convergence or because convergence is slow.

Our contribution is twofold. First, we make a general contribution to the modeling of observables with distributions. We identify a modeling assumption, the ergodic hypothesis, made widely in studies of wealth and income. In an established model of wealth, we find that historical data from a large and developed economy do not support its use. We do not deduce from this that the ergodic hypothesis is unjustified in all models and for all data, but we sound an alarm. It should be tested empirically in studies that make it. Validation would offer reassurance about findings reported and policies crafted. Failure to validate would open the door to new analyses. This study belongs to a growing body of work, known as "ergodicity economics". Asking whether observables are ergodic has yielded insights in various sub-disciplines of economics, see (Peters, 2019) and references therein.

Our second contribution is specific to the model and data used. RGBM is a minimal and realistic model of wealth in a growing economy whose participants interact. With positive reallocation, it generates stable distributions of rescaled wealth with historically realistic tails and inequality levels. With negative reallocation, it predicts unstable distributions and growing negative wealth, consistent with recent observations. That US wealth data are best described with poorer-to-richer reallocation turns our understanding of the dynamics of wealth distributions on its head. Quantitatively, RGBM provides parameter regimes in which rescaled wealth distributions are inherently unstable. Qualitatively, it suggests that social institutions, commonly assumed to reallocate from richer to poorer, may not function as envisaged. Both insights warrant further study, for which RGBM may be a helpful tool.

The paper is organized as follows. Section 2 describes the model we study in this paper. In Sec. 3 we discuss the data used for the empirical analysis, detailed in Sec. 4. We conclude and discuss the results in Sec. 5.

2 Model

Here we introduce our model, Reallocating Geometric Brownian Motion (RGBM). Individual wealth undergoes random multiplicative growth, modeled as Geometric Brownian Motion (GBM), and is reallocated among individuals by a simple pooling and sharing mechanism. Thus everyone's wealth is coupled to the total wealth in the economy. We view RGBM as a null model of an exponentially growing economy with socio-economic structure. Its three parameters represent common economic growth, random shocks to individual wealth, and socio-economic interaction among agents.

Models similar to RGBM, in which agents face random multiplicative idiosyncratic shocks, have been widely studied in the literature on inequality.[3] Several theoretical justifications

[3]See, for example, Champernowne (1953); Stiglitz (1969); Piketty and Zucman (2015); Toda and Walsh (2015); Gabaix et al. (2016); Hubmer, Krusell and Smith Jr. (2019).

exist, including different saving tastes (Piketty and Zucman, 2015), investments in human capital (Gabaix et al., 2016), and returns to wealth (Fagereng et al., 2020; Bach, Calvet and Sodini, 2020). We focus on the properties of RGBM related to the ergodic hypothesis to construct an empirical test of its validity.

RGBM has both ergodic and non-ergodic regimes, characterized by the sign of the reallocation rate parameter. Reallocation from richer to poorer yields an ergodic regime, in which wealths are positive, distributed with a Pareto tail, and confined around their mean value. Reallocation from poorer to richer produces a non-ergodic regime, in which the population splits into two classes, characterized by positive and negative wealths which diverge away from the mean. If the reallocation rate is zero, RGBM reduces to GBM, in which individual wealths grow independently and no socio-economic structure is represented.

2.1 Model definition

We denote by $x_i(t)$ the wealth at time t of the i^{th} member of a population of N individuals. Without reallocation, wealth obeys the stochastic differential equation for GBM, which describes noisy exponential growth:

$$dx_i = x_i \left[\mu dt + \sigma dW_i(t)\right] . \tag{2.1}$$

μ is the drift parameter, σ the volatility parameter, and $dW_i(t)$ the random increment in a Wiener process, normally distributed with zero mean and variance dt. This is our basic model for the evolution of individual wealth. It says that, without interactions, fractional changes in individual wealth are composed of:

- μdt – a common, constant part, representing economic growth; and
- $\sigma dW_i(t)$ – an idiosyncratic, random part, representing individual growth.

We introduce the rescaled wealth,

$$y_i(t) \equiv \frac{x_i(t)}{\langle x(t)\rangle_N} , \tag{2.2}$$

as individual wealth divided by the population average,

$$\langle x(t)\rangle_N \equiv \frac{1}{N}\sum_{i=1}^{N} x_i(t) . \tag{2.3}$$

For a sufficiently large population, the population average wealth is well approximated by the expected wealth, which grows exponentially at growth rate μ, *i.e.*

$$\langle x(t) \rangle_N \approx \langle x(t) \rangle \propto e^{\mu t} . \tag{2.4}$$

Under this approximation, rescaled wealth follows GBM with zero drift,

$$dy_i = \sigma y_i dW_i(t) , \tag{2.5}$$

where the common growth has been scaled out and only idiosyncratic growth remains. $y_i(t)$ has a time-varying lognormal distribution and is, therefore, non-ergodic in the sense of Samuelson (1968).[4]

To place this basic model of wealth in the context of existing work, we note its similarity to the model of Meade (1964), summarized by De Nardi, Fella and Yang (2015, p. 9) in their literature review as: "a simple accounting framework [in which] individuals are endowed with a, possibly individual-specific, fraction of the contemporaneous, aggregate capital stock. The only source of income in the economy is an idiosyncratic rate of return on individual wealth."

Wealth evolves multiplicatively under GBM. There are no additive changes akin to labor income and consumption. This is unproblematic for large wealths, where additive changes are dwarfed by capital gains. For small wealths, however, wages and consumption are significant. Indeed, empirical distributions exhibit different regularities for low and high wealths. As a robustness test, we treat additive earnings explicitly in Appendix B. We find that including them in a less parsimonious model of wealth accumulation does not alter fundamentally our conclusions.

We note that t in our framework denotes time, rather than the age of an agent. As Meade (1964, p. 41) puts it, our agents "do not marry or have children or die or even grow old." Therefore, the individual in our setup is best imagined as a household or a family, *i.e.* some long-lasting unit into which personal events are subsumed.

Finally, we add to the model a reallocation mechanism that is absent from many other models in the literature. The aim is to represent redistributive socio-economic structure. We imagine, as the simplest social interaction, that each individual pays a fixed proportion of its wealth, $\tau x_i dt$, into a central pot ("contributes to society") and gets back an equal

[4] By "sufficiently large" we mean $2 \ln N/\sigma^2 t > 1$, which is the self-averaging condition for wealth under GBM (see Peters and Adamou (2018) and Appendix A). This condition is satisfied for typical national populations with realistic parameters over typical observation times.

share of the pot, $\tau \langle x \rangle_N \, dt$, ("benefits from society"):

$$dx_i = x_i \left[(\mu - \tau) dt + \sigma dW_i(t) \right] + \tau \langle x \rangle_N \, dt \, . \tag{2.6}$$

A more complex model would treat the economy as a system of agents that interact with each other through a network of relationships. These relationships include trade in goods and services, employment, paying taxes, receiving welfare, using centrally-organized infrastructure (roads, schools, hospitals, a legal system, scientific research, and so on), insurance schemes, wealth transfers through inheritance and gifts, and everything else that constitutes a socio-economic network. It would be a hopeless task to produce an exhaustive list of all these interactions, let alone include them as testable model components. Instead we introduce a single parameter – the reallocation rate, τ – to represent their net effect. If τ is positive, the direction of net reallocation is from richer to poorer (from individuals whose wealth is above the average wealth to individuals whose wealth is below it). If negative, it is from poorer to richer.

Fitting τ to data will allow us to answer questions such as:

- what is the net reallocating effect of socio-economic activity on the wealth distribution?
- are observations consistent with the ergodic hypothesis that the distribution of rescaled wealth converges to a stationary form?
- if so, how long after a change in conditions does this convergence take?

To be clear, we do not want to separate different processes that affect wealth, as would be useful if we wanted to understand the effect of a specific process, say income tax, on economic inequality. We aim for the opposite: a model that summarizes in as few parameters as possible everything that affects the rescaled wealth distribution and, crucially, without assuming ergodicity. If, for example, effects of death and inheritance were separated out explicitly with additional parameters, rather than being included in τ, the fitted τ values would no longer tell us if the system as a whole is in an ergodic regime.

2.2 Model behavior

Equation (2.6) is our model for the evolution of wealth with socio-economic structure and the basis for the empirical study that follows. It is instructive to write it as

$$dx_i = \underbrace{x_i \left[\mu dt + \sigma dW_i\left(t\right)\right]}_{\text{Growth}} \; \underbrace{- \;\; \tau(x_i - \langle x \rangle_N)dt}_{\text{Reallocation}} \,. \tag{2.7}$$

This can be thought of as GBM plus a mean-reverting term like that of Uhlenbeck and Ornstein (1930) in physics and Vasicek (1977) in finance. This representation exposes the importance of the sign of τ. We discuss the two regimes in turn.

Positive τ For $\tau > 0$, wealth, x_i, reverts to the population average, $\langle x \rangle_N$. The large-sample approximation, $\langle x\left(t\right)\rangle_N \propto e^{\mu t}$, is valid and yields a stochastic differential equation for rescaled wealth,

$$dy_i = \sigma y_i dW_i\left(t\right) - \tau(y_i - 1)dt \,, \tag{2.8}$$

in which the common growth rate, μ, has been scaled out.[5]

The distribution of $y_i\left(t\right)$ can be found by solving the corresponding Fokker-Planck equation (also known as the Kolmogorov forward equation). A stationary distribution exists with a Pareto tail, see Appendix C. It is known as the inverse gamma distribution and has probability density function,

$$\mathcal{P}\left(y\right) = \frac{\left(\zeta - 1\right)^{\zeta}}{\Gamma\left(\zeta\right)} e^{-\frac{\zeta-1}{y}} y^{-(1+\zeta)} \,, \tag{2.9}$$

where $\zeta = 1 + 2\tau/\sigma^2$ is the Pareto tail index, $\Gamma\left(\cdot\right)$ is the gamma function, and the person index i has been dropped.

The usual stylized facts are recovered. The larger σ (more randomness in the returns) and the smaller τ (slower reallocation), the smaller the tail index and the fatter the tail of the distribution (higher inequality). Moreover, the fitted τ values we obtain in Sec. 4 give typical ζ values between 1 and 2 for the different datasets analyzed, consistent with observed tail indices between 1.2 to 1.6 (Klass et al., 2006; Gabaix, 2009; Brzeziński, 2014; Vermeulen, 2018). Thus, not only does RGBM predict a realistic functional form for the distribution of rescaled wealth, but also it admits fitted parameters which match observed

[5]Strictly speaking, the large-sample approximation and resulting rescaled-wealth process, Equation (2.8), hold only for $\tau > \tau_c$. However, $\tau_c \approx 0$ for realistic model parameters, and fits to data do not allow us to distinguish it from zero. Nonetheless, the derivation of τ_c is instructive, see Appendix A.

tails. The inability to do the latter is a known shortcoming in models of earnings-based wealth accumulation, see Benhabib, Bisin and Luo (2017) and Appendix B.

In this regime, the model also recovers a relationship between wealth mobility and wealth inequality. The larger τ is, the lower the steady-state inequality is. At the same time, wealth mobility would also be higher. This resembles the relationship coined "the Great Gatsby curve" in the context of income (Krueger, 2012; Corak, 2013). A thorough analysis of wealth mobility in this model is left for future work.

Equation (2.8) and extensions of it have received much attention in statistical mechanics (Marsili, Maslov and Zhang, 1998; Bouchaud and Mézard, 2000; Bouchaud, 2015b; Liu and Serota, 2017). As a combination of GBM and an Ornstein-Uhlenbeck process, it is a simple and analytically tractable stochastic process. Liu and Serota (2017) provide an overview of the literature and known results.

Negative τ For $\tau < 0$ the model exhibits mean repulsion rather than reversion. The ergodic hypothesis is invalid and no stationary rescaled wealth distribution exists. The population splits into those above the mean and those below the mean. Whereas in RGBM with non-negative τ it is impossible for positive wealth to turn negative, negative τ leads to negative wealth. No longer is total economic wealth a limit to the wealth of the richest individual, because the poorest develop large negative wealth. The wealth of the rich in the population increases exponentially away from the mean, and the wealth of the poor becomes negative and exponentially large in magnitude, see Figure 1(D). Qualitatively, this echoes the findings that the rich experience higher growth rates of their wealth than the poor (Piketty, 2014; Wolff, 2014) and that the cumulative wealth of the poorest 50% of the US population was negative during 2008–2013 (Rios-Rull and Kuhn, 2016; Saez and Zucman, 2016).

Such splitting of the population is a common feature of non-ergodic processes. If rescaled wealth were an ergodic process, then individuals would, over long enough time, experience all parts of its distribution. People would spend 99% of their time as "the 99 percent" and 1% of their time as "the 1 percent". The social mobility implied by models that assume ergodicity might not exist in reality if that assumption is invalid. That inequality and immobility have been linked (Corak, 2013; Liu et al., 2013; Berman, 2019) may be unsurprising when both are viewed as consequences of non-ergodic wealth or income.

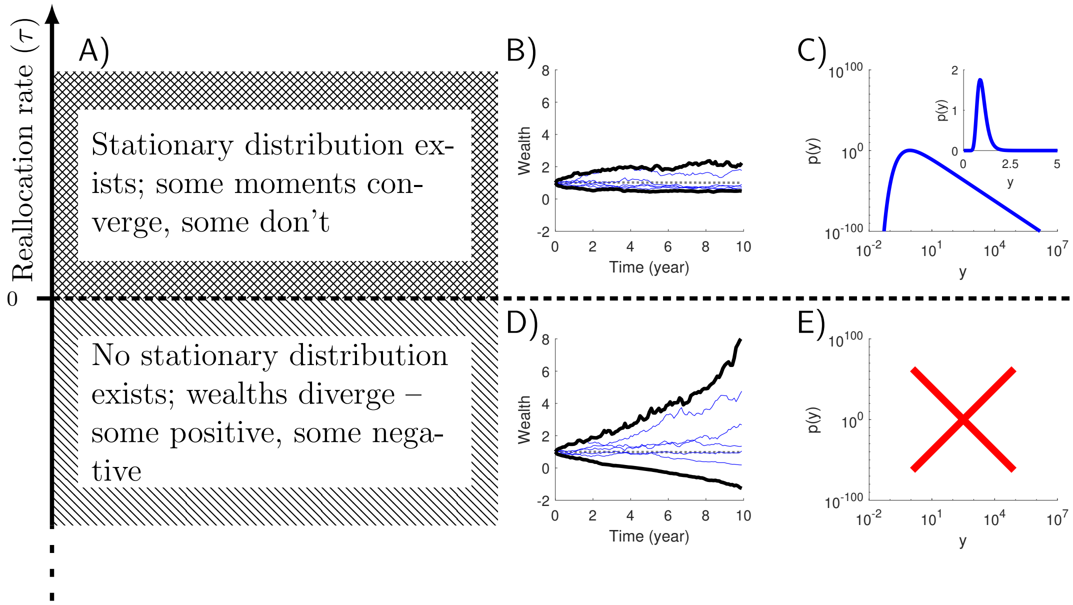


Figure 1: Regimes of RGBM. A) $\tau = 0$ separates the two regimes of RGBM. For $\tau > 0$, a stationary rescaled wealth distribution exists. For $\tau < 0$, no stationary rescaled wealth distribution exists and wealths diverge. B) Simulations of RGBM with $N = 1000$, $\mu = 0.021\ \text{year}^{-1}$ (presented after rescaling by $e^{\mu t}$), $\sigma = 0.14\ \text{year}^{-1/2}$, $x_i(0) = 1$, $\tau = 0.15\ \text{year}^{-1}$. Thick black lines: largest and smallest wealths. Thin blue lines: five randomly chosen wealth trajectories. Dotted grey line: sample mean. C) The stationary distribution to which the system in B) converges. Inset: same distribution on linear scales. D) Similar to B), with $\tau = -0.15\ \text{year}^{-1}$. E) in the $\tau < 0$ regime, no stationary rescaled wealth distribution exists.

3 Data

3.1 Wealth share data

To estimate the model parameters we analyze the wealth shares of the top quantiles of the US population. We consider net wealth, *i.e.* assets minus liabilities, and allow it to be negative. Wealth shares are estimated using different methods. Our baseline results rely on the estimates of Saez and Zucman (2016) using the capitalization method, in which capital income data from individual income tax returns is used to estimate the underlying stock of wealth. The idea is that if "we can observe capital income $k = rW$, where W is the underlying value of an asset and r is the known rate of return, then we can estimate wealth based on capital income and capitalization factor $1/r$ defined using the appropriate choice of rate of return" (Kopczuk, 2015, p. 54). We consider annual data for the top 10, 5, 1, 0.5, 0.1 and 0.01 percent wealth shares in the US between 1913 to 2014. Appendix D describes additional data sources based on other estimation methods. It also tests the robustness of our conclusions to differences between sources. We find that using different data sources does not alter qualitatively our conclusions.

3.2 Wealth Growth Rate

We find numerically that the results of our analysis do not depend on μ. This is because wealth shares depend only on the distribution of rescaled wealth and, for $\tau > 0$, it is possible to scale out μ completely from the wealth dynamic to obtain Equation (2.8) for rescaled wealth. The fitted $\tau < 0$ values we find are not large or persistent enough to make our simulations significantly μ-dependent. However, formally, since we allow negative τ, we must simulate Equation (2.7) and not Equation (2.8). This requires us to specify a value of μ, which we estimate as $\mu = 0.021 \pm 0.001$ year^{-1} by a least squares fit of historical per capita private wealth in the US (Piketty and Zucman, 2014) to an exponential growth curve.

3.3 Volatility

We must also specify the volatility parameter, σ, in Equation (2.7). In principle, this can vary with time. We have no access to real individual wealth trajectories, so we resort to estimating $\sigma(t)$ in two ways: by fitting our model to observed wealth shares; and by assuming volatility in personal wealth is coupled to stock market volatility. Specifically, we do the following:

1. We estimate $\sigma(t)$ and $\tau(t)$ together to fit the top wealth shares ($q = 10$, 5, 1, 0.5, 0.1 and 0.01 percent) from Saez and Zucman (2016) (see Appendix E for more details).

2. We estimate $\sigma(t)$ assuming that the volatility in individual wealths tracks the volatility in the values of the companies that constitute the commercial and industrial base of the national economy. Therefore, for each year, we estimate $\sigma(t)$ as the standard deviation of daily logarithmic changes of the Dow Jones Industrial Average (Quandl, 2016), annualized by multiplying by $(250/\text{year})^{1/2}$.

In both cases the values of σ change in time. In the first case, the values change within the range 0.02–0.56 year$^{-1/2}$, averaging at 0.2 year$^{-1/2}$ and with a median of 0.16 year$^{-1/2}$. In the second case, the annual volatility is within the range 0.06–0.53 year$^{-1/2}$, averaging at 0.16 year$^{-1/2}$ and with a median of 0.14 year$^{-1/2}$ for the years 1913–2012.

Despite the different sources of the two estimates, both yield a similar range of values for $\sigma(t)$ and similar typical values. Running our analysis with constant σ in this range has little effect on the results, see the robustness tests in Appendix E. Therefore, for simplicity we present our analysis using $\sigma(t) = 0.16 \text{ year}^{-1/2}$ for all t. Fixing σ means that we have only one model parameter – the effective reallocation rate, $\tau(t)$ – to fit to the historical wealth shares.

4 Empirical Analysis

The goal of the empirical analysis is to estimate $\tau(t)$ from the historical wealth data, using RGBM as our model. This estimation allows us to address two main questions:

1. Is it valid to make the ergodic hypothesis for rescaled wealth in the US? For the hypothesis to be valid, fitted values of $\tau(t)$ must be robustly positive.

2. If $\tau(t)$ is positive, does the rescaled wealth distribution converge to its asymptotic form quickly enough for the ergodic hypothesis to be useful?

We fit a time series, $\tau(t)$, that reproduces the annually observed wealth shares in Saez and Zucman (2016). The wealth share, S_q, is defined as the proportion of total wealth, $\sum_i^N x_i$, owned by the richest fraction q of the population, *e.g.* $S_{10\%} = 80\%$ means that the richest 10 percent of the population own 80 percent of the total wealth.

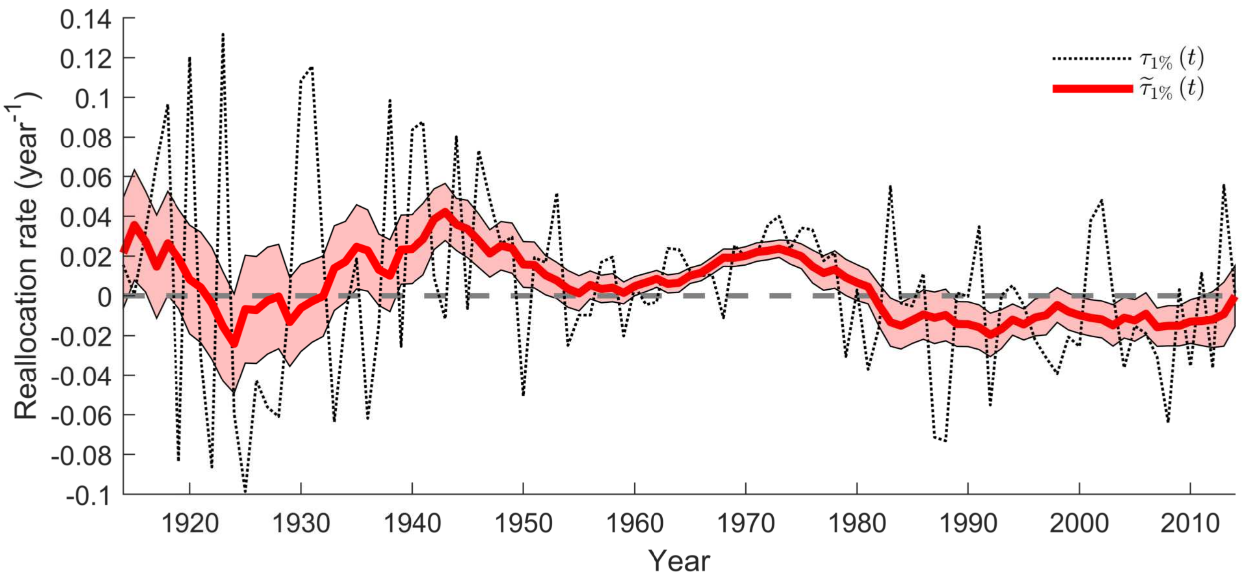
For an empirical wealth share time series, $S_q^{\text{data}}(t)$, we proceed as follows:

Step 1 Initialize N individual wealths at time t_0, $\{x_i(t_0)\}$, as random variates of the inverse gamma distribution with parameters chosen to match $S_q^{\text{data}}(t_0)$.

Step 2 Propagate the N individual wealths, $\{x_i(t)\}$, according to Equation (2.7) from time t to $t + \Delta t$, finding the value of τ that minimizes the difference between the wealth share in the modeled population, $S_q^{\text{model}}(t + \Delta t, \tau)$, and the observed wealth share, $S_q^{\text{data}}(t + \Delta t)$. We perform the minimization using the Nelder-Mead algorithm (Nelder and Mead, 1965).

Step 3 Repeat Step 2 until the end of the time series.

We consider historical wealth shares of the richest $q = 10$, 5, 1, 0.5, 0.1 and 0.01 percent and obtain time series of fitted effective reallocation rates, $\tau_q(t)$, shown in Figure 2 for $q = 1\%$. For each value of q we perform a run of the simulation for $N = 10^8$. Since, in practice, the $\{dW_i(t)\}$ are chosen randomly, each run of the simulation will result in slightly different $\tau_q(t)$ values. However, we find that the differences between such calculations are negligible.



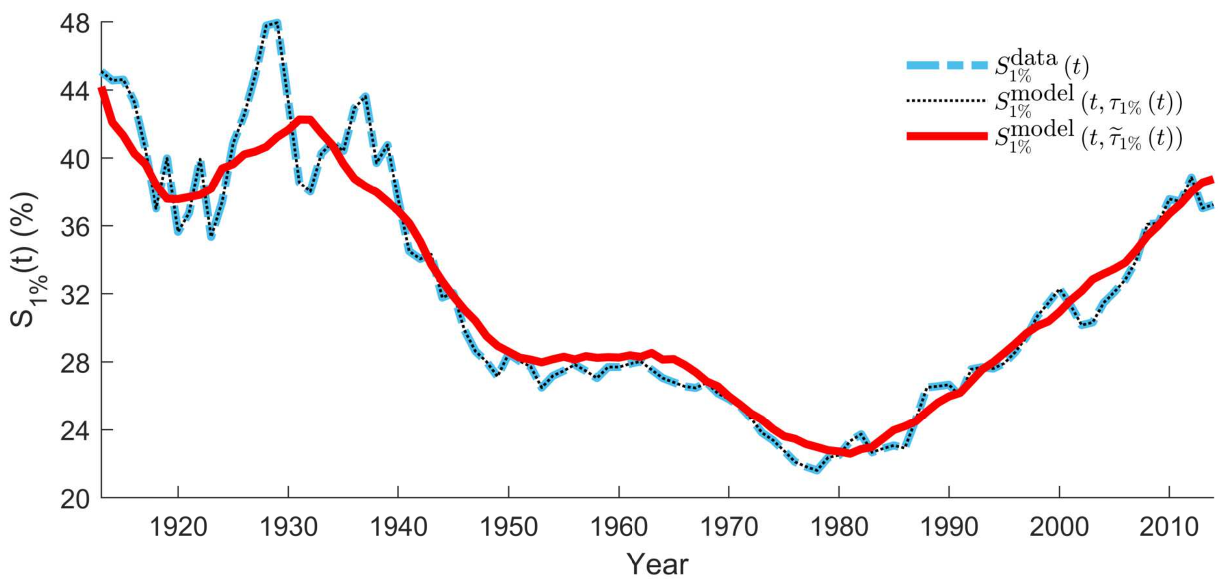


Figure 2: Fitted effective reallocation rates. Calculations done using $\mu = 0.021$ year^{-1} and $\sigma = 0.16$ year$^{-1/2}$. Top: $\tau_{1\%}(t)$ (dotted black) and $\widetilde{\tau}_{1\%}(t)$ (solid red). Translucent envelopes indicate one standard error in the moving averages. Bottom: $S_{1\%}^{\text{data}}$ (dash-dotted blue), $S_{1\%}^{\text{model}}$ based on the annual $\tau_{1\%}(t)$ (dotted black), based on the 10-year moving average $\widetilde{\tau}_{1\%}(t)$ (red).

Figure 2 (top) shows the fitted values of $\tau_{1\%}(t)$. These exhibit large annual fluctuations, whereas we are interested in longer-term changes in reallocation driven by structural economic and political changes. To elucidate these, we smooth the data by taking a central 10-year moving average, $\widetilde{\tau}_{1\%}(t)$, where the window is truncated at the ends of the time series. To ensure the smoothing does not introduce artificial biases, we reverse the procedure and use

$\widetilde{\tau}_{1\%}(t)$ to propagate the initial $\{x_i(t_0)\}$ to determine the modeled wealth shares, $S_{1\%}^{\text{model}}(t)$. Their close agreement with $S_{1\%}^{\text{data}}(t)$ in Figure 2 (bottom) suggests that the smoothed $\widetilde{\tau}_{1\%}(t)$ series is meaningful.

We find that the fitted reallocation rate, $\widetilde{\tau}_{1\%}(t)$, is not robustly positive for the time period studied, as would be required for the ergodic hypothesis to be valid. Indeed, it is consistently negative, *i.e.* reallocation from poorer to richer, from the 1980s onwards. This result holds for all of the top wealth shares considered.[6]

The fitted RGBM model predicts that poor households accumulate large and growing negative wealth. This is not unrealistic: the cumulative wealth of the poorest half of the US population turned negative in recent years; and the wealth required to exit the bottom wealth decile has been negative during the last fifty years, with its magnitude growing exponentially from −\$3650 to −\$16500 (in 2018 US Dollars) between 1962 and 2012 (The World Inequality Database, 2016). Thus, the existence of large negative wealth under poorer-to-richer reallocation in the RGBM model accords well with observations.

In the ergodic regime of RGBM, $\tau > 0$, it is possible to calculate how fast the wealth shares of different quantiles converge to their asymptotic value. We do this numerically in Appendix F and find a quantitative relationship between τ, σ, and the convergence time. Using the fitted $\widetilde{\tau}_{1\%}(t)$ values, we present the evolution of the implied convergence times in Figure 3. When they exist, convergence times for wealth shares are long, ranging from a few decades to several centuries. They are longer than the times between policy changes and other shocks that are of interest to economists.

The evidence we present does not support using the ergodic hypothesis in RGBM to model US wealth data from the last hundred years. We find either negative reallocation, for which the rescaled wealth distribution never stabilises, or slow positive reallocation, for which stabilisation takes decades or centuries.

[6] It is also possible to estimate the effective reallocation rate $\tau_{\text{all}\%}(t)$ that best fits the data for all top shares ($q = 10$, 5, 1, 0.5, 0.1 and 0.01) simultaneously. In addition, since σ was fixed at 0.16 year$^{-1/2}$ in our analysis, we can estimate both $\tau_{\text{all}\%+\sigma}(t)$ and $\sigma(t)$ together to best fit the data for all six top shares, for robustness. These results are presented in Appendix E. They show that although fitting τ to a specific top wealth quantile matters quantitatively, the results are qualitatively robust. Letting $\sigma(t)$ and $\tau(t)$ change together to fit all top wealth shares does not change the results substantially. In addition, similar qualitative observations are obtained when fitting the effective reallocation rate to the reported top wealth shares in other data sources, see Appendix D.

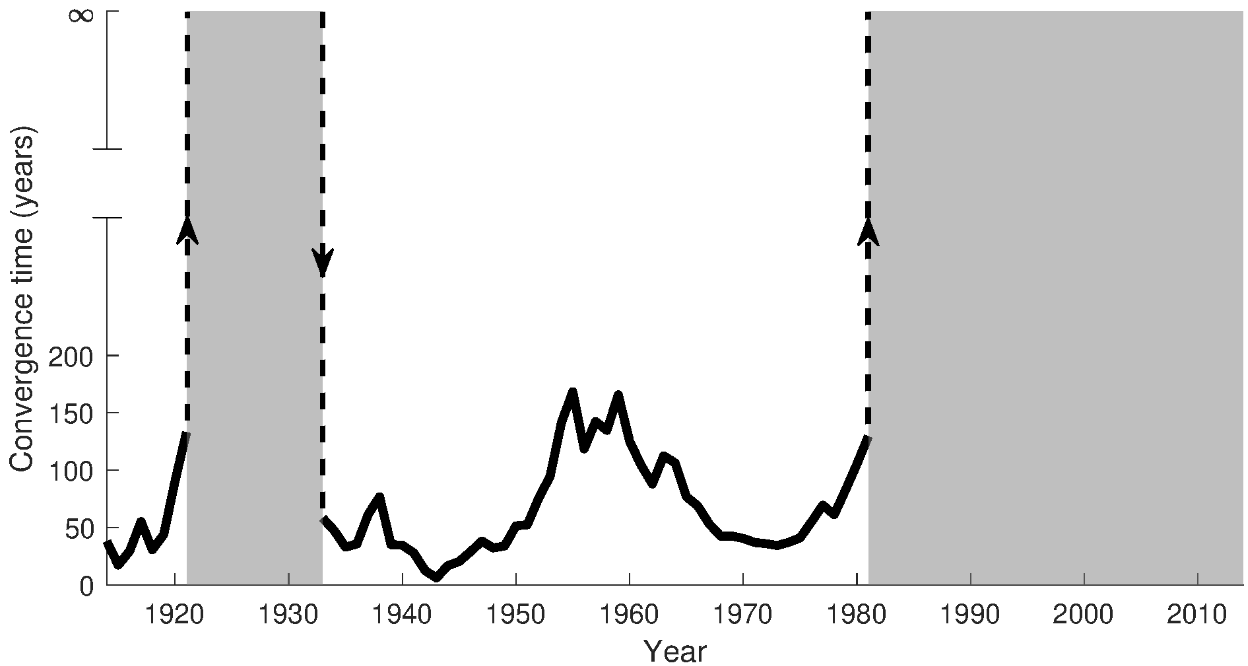


Figure 3: Time series of wealth share convergence times. The convergence times (black line) are based on the fitted $\widetilde{\tau}_{1\%}(t)$ values in Figure 2. The grey shaded areas indicate time periods when $\widetilde{\tau}_{1\%}(t) < 0$, the non-ergodic regime of RGBM where no stationary distribution for rescaled wealth exists. In these periods, the distribution never stabilises and the convergence time is undefined.

5 Discussion

Many studies of economic inequality assume ergodicity of rescaled wealth. This assumption often goes under the headings of equilibrium, stationarity, or stability (Adamou and Peters, 2016). Specifically, it is assumed that:

1. rescaled wealth equilibrates, *i.e.* a stationary distribution exists to which the observed distribution converges in the long-time limit; and
2. it equilibrates quickly, *i.e.* the observed distribution approaches the stationary distribution over a time shorter than other relevant timescales, such as the time between policy changes.

The second assumption is usually left unstated, but it is necessary for the stationary distribution of the model to be a useful proxy for the empirical distribution. This matters because the stationary distribution is a key object of study. Model parameters are found by fitting the stationary distribution to observed inequality, and the effects of changes in these parameters on the stationary distribution are explored. While in some cases this strategy is useful, it may be misleading if the ergodic hypothesis is unjustified. We echo the warning of Cowell (2014, p. 708) that "the long-run analysis of policy should not presume that there is an inherent tendency for the wealth distribution to approach equilibrium."

We do not assume ergodicity. Fitting $\tau(t)$ in RGBM allows the data to speak without constraint as to whether the ergodic hypothesis is valid. We find it is not because:

A. We observe negative τ values in all datasets analyzed, most notably using the capitalization method and especially since the 1980s. The rescaled wealth distribution is non-stationary and inequality increases while these conditions prevail.

B. When we observe positive τ, the associated convergence times are typically decades or centuries, see Figure 7 and Figure 11 (bottom). They are much longer than the periods over which economic policies and conditions change. They are the timescales of history rather than of politics.

Observation A corresponds to reallocation that moves wealth systematically from poorer to richer individuals. This dramatic finding is inconsistent with the ergodic hypothesis for rescaled wealth in RGBM. Worryingly, therefore, it would not be made in such a study under the conventional strategy of assuming ergodicity at the outset. The most recent US wealth data are best described by RGBM with $\tau < 0$ (Saez and Zucman, 2016; The World Inequality Database, 2016) or $\tau \approx 0$ (Bricker et al., 2016). Under such conditions, each time we observe the rescaled wealth distribution, we see a snapshot of it either in the process of diverging or far from its asymptotic form. It is much like a photograph of an explosion in space: it will show a fireball whose finite extent tells us nothing of the future distances between pieces of debris.

Important economic phenomena, such as diverging inequality, social immobility, and the emergence of negative wealth, are difficult to reproduce in a model that assumes ergodicity. In our simple model, this is easy to see: with $\tau > 0$, our model cannot reproduce these phenomena at all. We might be tempted to conclude that their existence is a sign of special conditions prevailing in the real world. However, when we admit the possibility of non-ergodicity, $\tau \leq 0$, it becomes clear that these phenomena can emerge in an economy that does not actively guard against them.

As well as shedding light on the dynamics of wealth inequality in the US, our findings have two major implications for the modeling of distributions. First, existing models that make the ergodic hypothesis explicitly, or are built such that the existence and uniqueness of a stationary distribution are guaranteed, should have the validity of their assumptions tested empirically. If a steady state does exist, it is important to confirm that convergence happens over shorter times than those between studied changes in conditions. This is particularly relevant when policy recommendations are based on such models. We encourage researchers to make such checks. Second, our findings demonstrate that models of wealth (or, indeed, income) must contain non-ergodic regimes, if they are to capture the full spectrum of possible economic phenomena. We hope this insight will aid future modeling efforts.

References

Adamou, Alexander, and Ole Peters. 2016. "Dynamics of Inequality." Significance, 13(3): 32–35.

Atkinson, Anthony B. 1969. "The Timescale of Economic Models: How Long is the Long Run?" Review of Economic Studies, 36(2): 137–152.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini. 2020. "Rich Pickings? Risk, Return, and Skill in Household Wealth." American Economic Review. *Forthcoming.*

Benhabib, Jess, Alberto Bisin, and Mi Luo. 2017. "Earnings Inequality and Other Determinants of Wealth Inequality." American Economic Review, 107(5): 593–597.

Benhabib, Jess, Alberto Bisin, and Mi Luo. 2019. "Wealth Distribution and Social Mobility in the US: A Quantitative Approach." American Economic Review, 109(5): 1623–1647.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2011. "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents." Econometrica, 79(1): 123–157.

Berman, Yonatan. 2019. "On the Link Between Intergenerational Mobility and Inequality: Are They Truly Distinct?" Mimeo.

Berman, Yonatan, and Yoash Shapira. 2017. "Revisiting r > g—The Asymptotic Dynamics of Wealth Inequality." Physica A, 467: 562–572.

Berman, Yonatan, Eshel Ben-Jacob, and Yoash Shapira. 2016. "The Dynamics of Wealth Inequality and the Effect of Income Distribution." PLOS ONE, 11(6): e0199999.

Bewley, Truman. 1977. "The Permanent Income Hypothesis: A Theoretical Formulation." Journal of Economic Theory, 16(2): 252–292.

Bouchaud, Jean-Philippe. 2015a. "Note on Mean-Field Wealth Models and the Random Energy Model." Mimeo.

Bouchaud, Jean-Philippe. 2015b. "On Growth-Optimal Tax Rates and the Issue of Wealth Inequalities." arXiv:1508.00275.

Bouchaud, Jean-Philippe, and Marc Mézard. 2000. "Wealth Condensation in a Simple Model of Economy." Physica A, 282(4): 536–545.

Bricker, Jesse, Alice Henriques, Jacob Krimmel, and John Sabelhaus. 2016. "Measuring Income and Wealth at the Top Using Administrative and Survey Data." Brookings Papers on Economic Activity, 261–331.

Brzeziński, Michał. 2014. "Do Wealth Distributions Follow Power Laws? Evidence from 'Rich Lists'." Physica A, 406: 155–162.

Champernowne, David G. 1953. "A Model of Income Distribution." The Economic Journal, 63(250): 318–351.

Corak, Miles. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." Journal of Economic Perspectives, 27(3): 79–102.

Cowell, Frank A. 2014. "Piketty in the Long Run." The British Journal of Sociology, 65(4): 708–720.

De Nardi, Mariacristina. 2015. "Quantitative Models of Wealth Inequality: A Survey." Working Paper 21106, NBER.

De Nardi, Mariacristina, Giulio Fella, and Fang Yang. 2015. "Piketty's Book and Macro Models of Wealth Inequality." Working Paper 21730, NBER.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2020. "Heterogeneity and Persistence in Returns to Wealth." Econometrica, 88(1): 115–170.

Gabaix, Xavier. 2009. "Power Laws in Economics and Finance." Annual Review of Economics, 1(1): 255–294.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2016. "The Dynamics of Inequality." Econometrica, 84(6): 2071–2111.

Hubmer, Joachim, Per Krusell, and Anthony A. Smith Jr. 2019. "Sources of U.S. Wealth Inequality: Past, Present, and Future." Mimeo.

Jones, Charles I. 2015. "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality." Journal of Economic Perspectives, 29(1): 29–46.

Kaymak, Barış, and Markus Poschke. 2016. "The Evolution of Wealth Inequality Over Half a Century: The Role of Taxes, Transfers and Technology." Journal of Monetary Economics, 77: 1–25.

Klass, Oren S., Ofer Biham, Moshe Levy, Ofer Malcai, and Sorin Solomon. 2006. "The Forbes 400 and the Pareto Wealth Distribution." Economics Letters, 90(2): 290–295.

Kopczuk, Wojciech. 2015. "What Do We Know about the Evolution of Top Wealth Shares in the United States?" Journal of Economic Perspectives, 29(1): 47–66.

Kopczuk, Wojciech, and Emmanuel Saez. 2004. "Top Wealth Shares in the United States, 1916–2000: Evidence from Estate Tax Returns." National Tax Journal, 57(2): 445–487.

Krueger, Alan. 2012. "The Rise and Consequences of Inequality." Presentation Made to the Center for American Progress, January 12th 2012.

Liu, Kang, Nicholas Lubbers, William Klein, Jan Tobochnik, Bruce Boghosian, and Harvey Gould. 2013. "The Effect of Growth on Equality in Models of the Economy." arXiv:1305.0794.

Liu, Z., and Rostislav A. Serota. 2017. "Correlation and Relaxation Times for a Stochastic Process with a Fat-tailed Steady-state Distribution." Physica A, 474: 301–311.

Marsili, Matteo, Sergei Maslov, and Yi-Cheng Zhang. 1998. "Dynamical Optimization Theory of a Diversified Portfolio." Physica A, 253: 403–418.

Meade, James E. 1964. Efficiency, Equality and The Ownership of Property. London, UK:Allen & Unwin.

Nelder, John A., and Roger Mead. 1965. "A Simplex Method for Function Minimization." The Computer Journal, 7(4): 308–313.

Pareto, Vilfredo. 1897. Cours d'Économie Politique. Lausanne, Switzerland:F. Rouge.

Peters, Ole. 2019. "The Ergodicity Problem in Economics." Nature Physics, 15: 1216–1221.

Peters, Ole, and Alexander Adamou. 2018. "The Sum of Log-Normal Variates in Geometric Brownian Motion." arXiv:1802.02939.

Peters, Ole, and Murray Gell-Mann. 2016. "Evaluating Gambles using Dynamics." Chaos, 26: 23103.

Peters, Ole, and William Klein. 2013. "Ergodicity Breaking in Geometric Brownian Motion." Physical Review Letters, 110(10): 100603.

Piketty, Thomas. 2014. Capital in the Twenty-First Century. Cambridge, MA:Harvard University Press.

Piketty, Thomas, and Emmanuel Saez. 2013. "A Theory of Optimal Inheritance Taxation." Econometrica, 81(5): 1851–1886.

Piketty, Thomas, and Gabriel Zucman. 2014. "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700–2010." Quarterly Journal of Economics, 129(3): 1255–1310.

Piketty, Thomas, and Gabriel Zucman. 2015. "Wealth and Inheritance in the Long Run." In Handbook of Income Distribution. , ed. Anthony B. Atkinson and François Bourguignon, 1303–1368. Elsevier.

Pinkovskiy, Maxim, and Xavier Sala-i-Martin. 2009. "Parametric Estimations of the World Distribution of Income." Working Paper 15433, NBER.

Quandl. 2016. "Dow Jones Industrial Average." Accessed: 04/19/2016.

Rios-Rull, Jose-Victor, and Moritz Kuhn. 2016. "2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics." Quarterly Review, 1–75.

Saez, Emmanuel, and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." Quarterly Journal of Economics, 131(2): 519–578.

Saez, Emmanuel, and Gabriel Zucman. 2019. "Progressive Wealth Taxation." Brookings Papers on Economic Activity.

Samuelson, Paul A. 1968. "What Classical and Neoclassical Monetary Theory Really Was." The Canadian Journal of Economics, 1(1): 1–15.

Stiglitz, Joseph E. 1969. "Distribution of Income and Wealth Among Individuals." Econometrica, 37(3): 382–397.

The World Inequality Database. 2016. "USA Top 10% and 1% and Bottom 50% Wealth Shares, 1913–2014." http://wid.world/data/, Accessed: 12/26/2016.

Toda, Alexis A., and Kieran Walsh. 2015. "The Double Power Law in Consumption and Implications for Testing Euler Equations." Journal of Political Economy, 123(5): 1177–1200.

Uhlenbeck, George E., and Leonard S. Ornstein. 1930. "On the Theory of the Brownian Motion." Physical Review, 36(5): 823–841.

Vasicek, Oldrich. 1977. "An Equilibrium Characterization of the Term Structure." Journal of Financial Economics, 5(2): 177–188.

Vermeulen, Philip. 2018. "How Fat is the Top Tail of the Wealth Distribution?" Review of Income and Wealth, 64(2): 357–387.

Wolff, Edward N. 2014. "Household Wealth Trends in the United States, 1983–2010." Oxford Review of Economic Policy, 30(1): 21–43.

A The Derivation of Self-Averaging Time t_c

In transforming the differential equation for wealth, $x(t)$, into a differential equation for rescaled wealth, $y(t)$, we use the approximation $\langle x(t) \rangle_N \approx \langle x(t) \rangle \propto e^{\mu t}$. In other words, we assume that the wealth averaged over the finite population grows like the expectation value of wealth.

It is known that this approximation is invalid for long times (Peters and Klein, 2013). Specifically, the population average $\langle x(t) \rangle_N$ grows exponentially at rate $\mu - \sigma^2/2$ in the long-time limit, whereas the expectation value $\langle x(t) \rangle$ grows exponentially at rate μ at all times.

This raises the question for how long our approximation is valid. The answer depends on the population size N, as it must because the expectation value is just the $N \to \infty$ limit of the population average. To assess whether fluctuations in $\langle x(t) \rangle_N$ are important, we compute a quantity known as the relative variance, which is the variance of $\langle x(t) \rangle_N$ divided by the square of its expectation value. If the relative variance is less than one, then fluctuations are small enough to make the approximation acceptable. If this is not the case, then we cannot use the approximation.

The calculations that use the approximation relate to properties of the stationary distribution of rescaled wealth in RGBM. This exists for τ above some positive threshold, *i.e.* with sufficiently fast reallocation. The coupling of wealth trajectories through reallocation lengthens the timescale over which the population average resembles the expected wealth. Therefore, we are on safe ground if we can show that the timescale on which the approximation is valid when $\tau = 0$ is longer than practically relevant timescales. This is a sufficient condition for the approximation to be valid when $\tau > 0$.

This means we can work with a population of N independent trajectories of GBM, for which computation of the relative variance of $\langle x(t) \rangle_N$ is straightforward. For simplicity, we assume all trajectories start at the same value, $x(0) = 1$. After time t, the outcome of a single trajectory, $x(t)$, is lognormally distributed,

$$\ln x(t) \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right) . \tag{A.1}$$

The expectation value is

$$\langle x(t) \rangle = e^{\mu t} , \tag{A.2}$$

and the variance is

$$V[x(t)] = e^{2\mu t}\left(e^{\sigma^2 t} - 1\right) . \tag{A.3}$$

Since the wealth trajectories are independent, the variance of the population average, $\langle x(t)\rangle_N$, is simply $1/N$ times the variance of the individual trajectory. Thus, the relative variance is

$$\frac{V\left[\langle x(t)\rangle_N\right]}{\langle\langle x(t)\rangle_N\rangle^2} = \frac{e^{\sigma^2 t}-1}{N}\,. \tag{A.4}$$

If this is less than one, then $\langle x(t)\rangle$ is a good approximation for $\langle x(t)\rangle_N$. Rearranging and taking $N \gg 1$ gives an upper bound on the time for which this approximation holds:

$$t < t_c \equiv \frac{\ln N}{\sigma^2}\,. \tag{A.5}$$

For realistic parameter values, $N = 10^8$ and $\sigma = 0.16\,\text{year}^{-1/2}$, we find $t_c \approx 700$ years. The minimum value of τ required for rescaled wealth to have a stationary distribution in RGBM is proportional to the inverse of this timescale (Bouchaud, 2015a). Specifically, rescaled wealth is ergodic if

$$\tau > \tau_c \equiv \frac{\sigma^2}{2\ln N}\,. \tag{A.6}$$

The intuition behind this is that the inequality-increasing effects of multiplicative growth drive wealths apart on the timescale t_c, whereas the inequality-reducing effects of reallocation drive wealths back together on the timescale $1/\tau$. Thus, τ_c marks the point at which reallocation overcomes the forces that drive wealths apart, leading to a stationary distribution of rescaled wealth. For the parameters in our model $\tau_c \approx 0.0007$ year^{-1}, which is practically indistinguishable from zero. Therefore, for simplicity in the main text, we treat $\tau = 0$ as the critical reallocation rate, which separates the ergodic and non-ergodic regimes of RGBM.

The timescale t_c was derived assuming that everyone starts out equally, which is not generally the case. If the initial distribution of wealth is very unequal, then the variance of $\langle x\rangle_N$ will be dominated by the fluctuations experienced by the wealthiest individuals, and the approximation $\langle x\rangle_N \approx \langle x(0)\rangle_N\, e^{\mu t}$ will become invalid more quickly. We confirmed numerically that the effect is negligible for our study: our results are indistinguishable whether we simulate dx (which requires an estimate for μ) or dy (where μ does not appear but the aforementioned approximation is made).

B The Effect of Earnings

In a 2017 review, Benhabib, Bisin and Luo (2017, p. 593) remark that "the literature has largely emphasized the role of earnings inequality in explaining wealth inequality" and point to empirical failures of this approach. The most popular models have agents accumulate wealth through stochastic earnings and precautionary savings. These models predict a strong correlation between earnings and wealth inequality, which is not observed in cross-country data, and struggle "to reproduce the thick right tail of the wealth distribution observed in the data" (Benhabib, Bisin and Luo, 2017, p. 593). The latter failure is also noted by Hubmer, Krusell and Smith Jr. (2019). Benhabib, Bisin and Luo (2017, p. 595) conclude that "other factors, like stochastic idiosyncratic returns on wealth" must be at play.

Consistent with these observations, RGBM, the model studied in this paper, models wealth accumulation as a multiplicative process. Changes in individual wealth are composed of terms proportional to either individual wealth or average wealth, in effect generating stochastic idiosyncratic returns. Additive changes akin to labor income and consumption are not treated explicitly. Instead their effects are wrapped into the reallocation rate, τ. While this parsimony makes the model tractable – in that we can write down the rescaled wealth distribution, Equation (2.9), for positive τ – it makes it impossible to disentangle the various real-world drivers of the observed increase in wealth inequality. It is, therefore, legitimate to ask whether wealth is inherently unstable (because it is reallocated negatively) or whether the increase in wealth inequality is really due to changes in the earnings distribution.

We check this by adding to Equation (2.7) a term representing earnings, for which data are available. We find that earnings have had a small effect on the dynamics of the rescaled wealth distribution over the last century. Since about 2000 this may have contributed to the increase in wealth inequality, but in general the effect has been stabilizing. This implies that the purely multiplicative dynamics of wealth, *i.e.* excluding additive earnings, have exhibited greater negative reallocation than fits to RGBM might suggest.

We separate earnings out as follows, in a model we call Earnings Geometric Brownian Motion (EGBM):

$$dx_i = \underbrace{x_i \left[\mu\left(t\right)dt + \sigma\left(t\right)dW_i\left(t\right)\right]}_{\text{Growth}} \underbrace{- \ \tau^{\text{EGBM}}\left(t\right)\left(x_i - \langle x\rangle_N\right)dt}_{\text{Reallocation}} \underbrace{+ \ \epsilon_i\left(t\right)dt}_{\text{Earnings}} . \tag{B.1}$$

Equation (B.1) is similar to Equation (2.7) for RGBM, with the addition of the individual earnings term $\epsilon_i\left(t\right)dt$. We consider earnings after spending and after tax, as they directly change wealth. This changes the role of the reallocation rate, as reflected in the notation:

τ^{EGBM} reflects changes in wealth for reasons other than additive income and consumption.

Using the same techniques as for the RGBM fits, we fit a time series, $\tau^{\text{EGBM}}(t)$, that reproduces annually observed wealth shares in the US (for the capitalization method dataset). This requires us to find individual earnings in Equation (B.1), which we do as follows. We assume that the earnings are lognormally distributed (see for example Pinkovskiy and Sala-i-Martin (2009)) and then fit the two parameters of the distribution every year. This is done using the mean earnings (calculated as total disposable national income multiplied by the personal savings rate and divided by the population size, as reported by Piketty and Zucman (2014)) and the top 1% after-tax income share data (as reported in The World Inequality Database (2016)).

Since the joint wealth-earnings distribution is not known fully, we analyze two bounding scenarios. First, a "correlated" scenario in which earnings are perfectly correlated with wealth. Every year, the richest individual has the highest earnings, the second richest individual has the second highest earnings, and so on. This is the least stabilizing possibility. Second, an "anti-correlated" scenario in which the poorest individual has the highest earnings, the second poorest individual has the second highest earnings, and so on. This is the most stabilizing possibility. We fit $\tau^{\text{EGBM}}(t)$ for each scenario, and compare the results to the fitted $\tau(t)$ values for which the effects of earnings have not been disentangled from other effects. See Figure 4.

In general, disentangling earnings has a small effect on the fitted reallocation rates. For the correlated scenario, the fitted reallocation rates $\tau^{\text{EGBM}}(t)$ are within the statistical error of $\tau(t)$, into which the effects of earnings have been subsumed. This is the more realistic scenario of the two, because the correlation between wealth and earnings is known to be positive (Rios-Rull and Kuhn, 2016).

When earnings have a stabilizing effect, the fitted reallocation rate is lower under EGBM than under RGBM. Even in the correlated scenario, which is the least stabilizing, earnings have a stabilizing effect for years 1930–2000. During 2000–2010, we find that earnings destabilize the rescaled wealth distribution in the correlated scenario. However, since the difference between the fitted values is small, we can conclude that the contribution of earnings to the reported increase in wealth inequality from 1985 to 2010 is negligible. Moreover, the earnings contribution cannot explain the invalidity of the ergodic hypothesis in our model.

We suspect the main reason for the small effect of earnings is the low ratio of earnings to wealth. Figure 5 shows the ratio of average earnings (after spending) to average wealth in the US during the period under study. Typically this ratio is around 1%, at which level earnings play only a secondary role in the dynamics of wealth. This echoes the observations

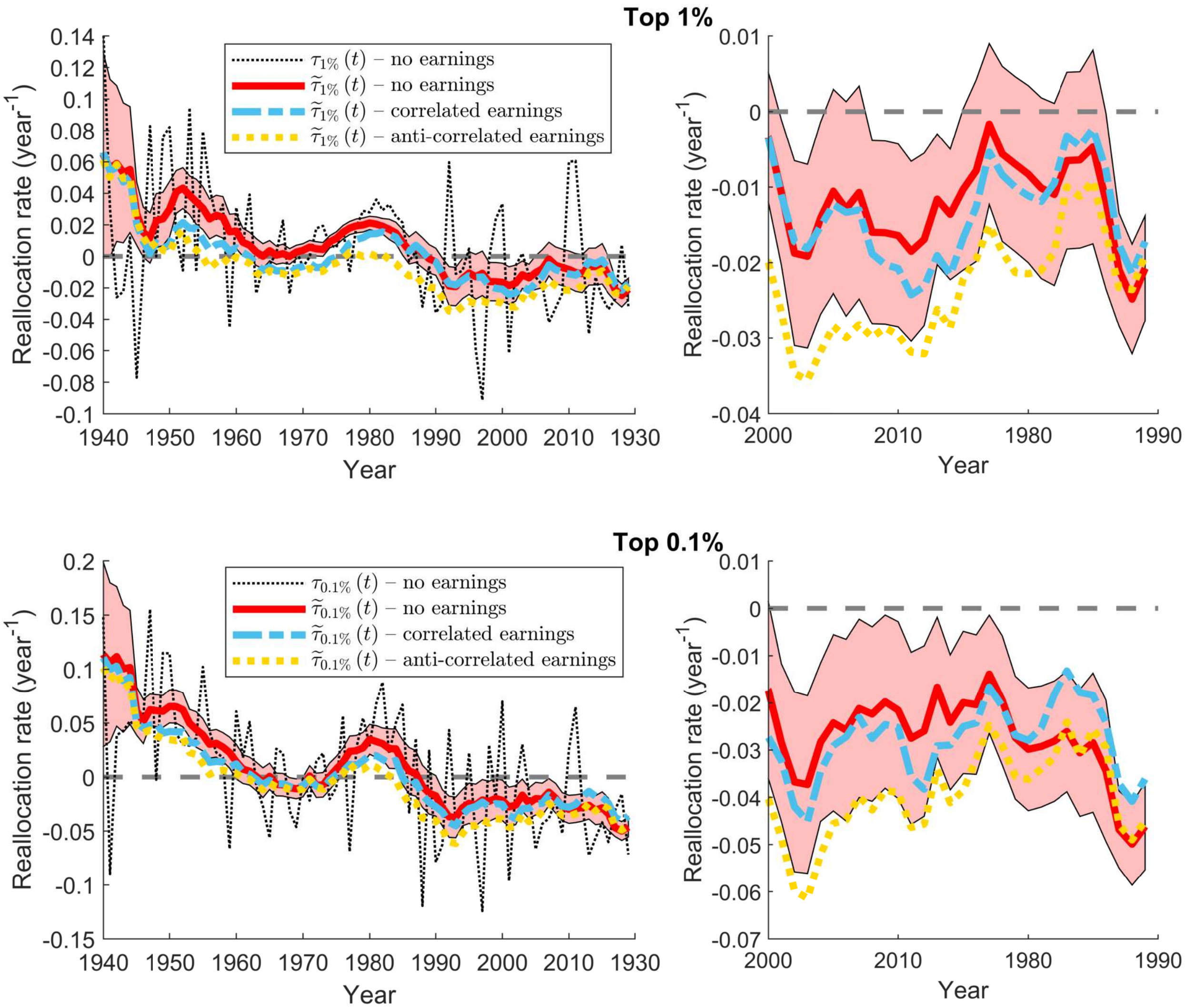


Figure 4: The effective reallocation rates for the RGBM and EGBM models. The red translucent envelopes indicate one standard error in the moving averages of the RGBM fitted reallocation rates.

of Piketty (2014); Piketty and Zucman (2014); Berman and Shapira (2017). Under such conditions, it is unsurprising that models reliant on earnings as the primary mechanism of wealth accumulation fail to resemble reality (Benhabib, Bisin and Luo, 2017; Hubmer, Krusell and Smith Jr., 2019). By including a multiplicative growth mechanism for wealth, RGBM avoids these pitfalls. Not only does it predict the Pareto tail of the stationary rescaled wealth distribution, when it exists, but also it allows for fitted tail exponents that reproduce the observed tail thickness (see Sec. 2.2).

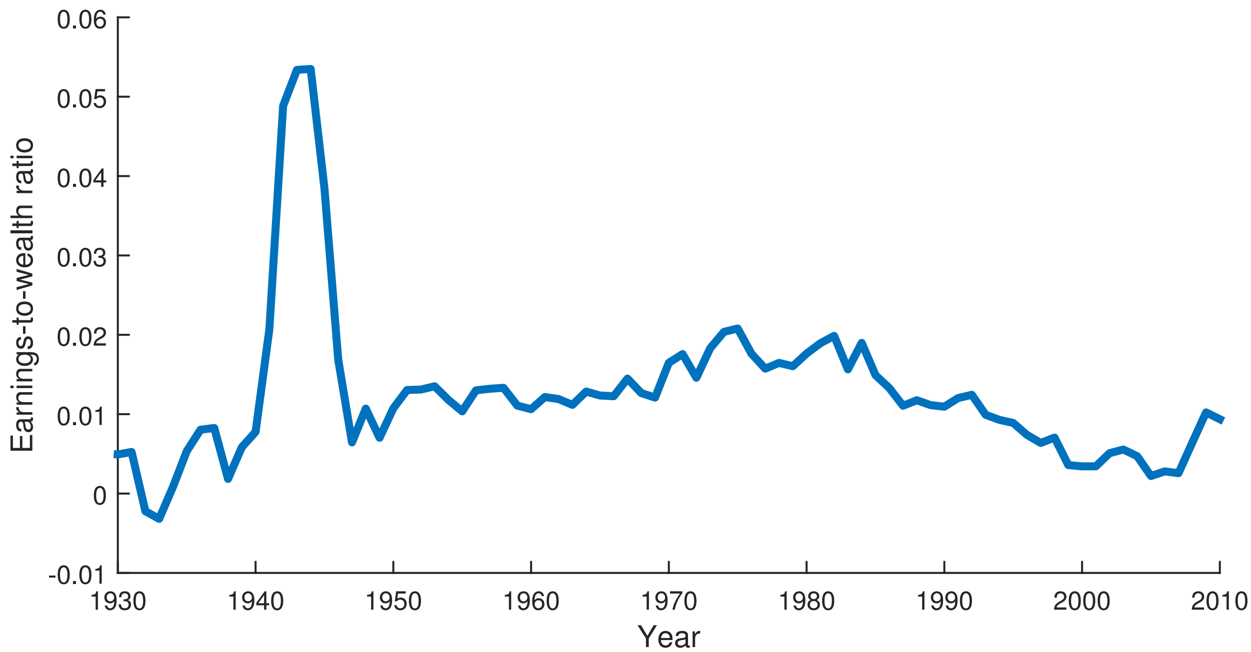


Figure 5: Time series of the ratio of average earnings after spending to average private wealth in the US. Data from Piketty and Zucman (2014).

C The Derivation of the Stationary Distribution

We start with the SDE for the rescaled wealth,

$$dy = \sigma\, y\, dW - \tau\,(y-1)\, dt\,. \tag{C.1}$$

This is an Itô equation with drift term $A = \tau\,(y-1)$ and diffusion term $B = y\sigma$.

Such equations imply ordinary second-order differential equations that describe the evolution of the pdf, called Fokker-Planck equations. The Fokker-Planck equation describes the change in probability density, at any point in (rescaled wealth) space, due to the action of the drift term (like advection in a fluid) and due to the diffusion term (like heat spreading). In this case, we have

$$\frac{dp\,(y,t)}{dt} = \frac{\partial}{\partial y}\left[Ap\,(y,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[B^2 p\,(y,t)\right]\,. \tag{C.2}$$

The steady-state Fokker-Planck equation for the pdf $p\,(y)$ is obtained by setting the time derivative to zero,

$$\frac{\sigma^2}{2}\left(y^2 p\right)_{yy} + \tau\left[(y-1)\,p\right]_y = 0\,. \tag{C.3}$$

Positive wealth subjected to continuous-time multiplicative dynamics with non-negative reallocation can never reach zero. Therefore, we solve Equation (C.3) with boundary condition $p\,(0) = 0$ to give

$$p\,(y) = \frac{(\zeta-1)^{\zeta}}{\Gamma\,(\zeta)} e^{-\frac{\zeta-1}{y}} y^{-(1+\zeta)}\,, \tag{C.4}$$

where

$$\zeta = 1 + \frac{2\tau}{\sigma^2} \tag{C.5}$$

and $\Gamma\,(\zeta) = \int_0^\infty x^{\zeta-1} e^{-x}\, dx$ is the gamma function. The distribution has a power-law tail as $y \to \infty$, resembling Pareto's often confirmed observation that the frequency of large wealths tends to decay as a power law. The exponent of the power law, ζ, is called the Pareto parameter and is one measure of economic inequality.

D Data Sources for US Top Wealth Shares

To estimate the model parameters, we analyze the wealth shares of the top quantiles of the US population. The wealth shares can be estimated using different methods. Our baseline results presented above rely on the estimates of Saez and Zucman (2016), using the capitalization method. Here we describe the various data sources and methods, and the differences between them:

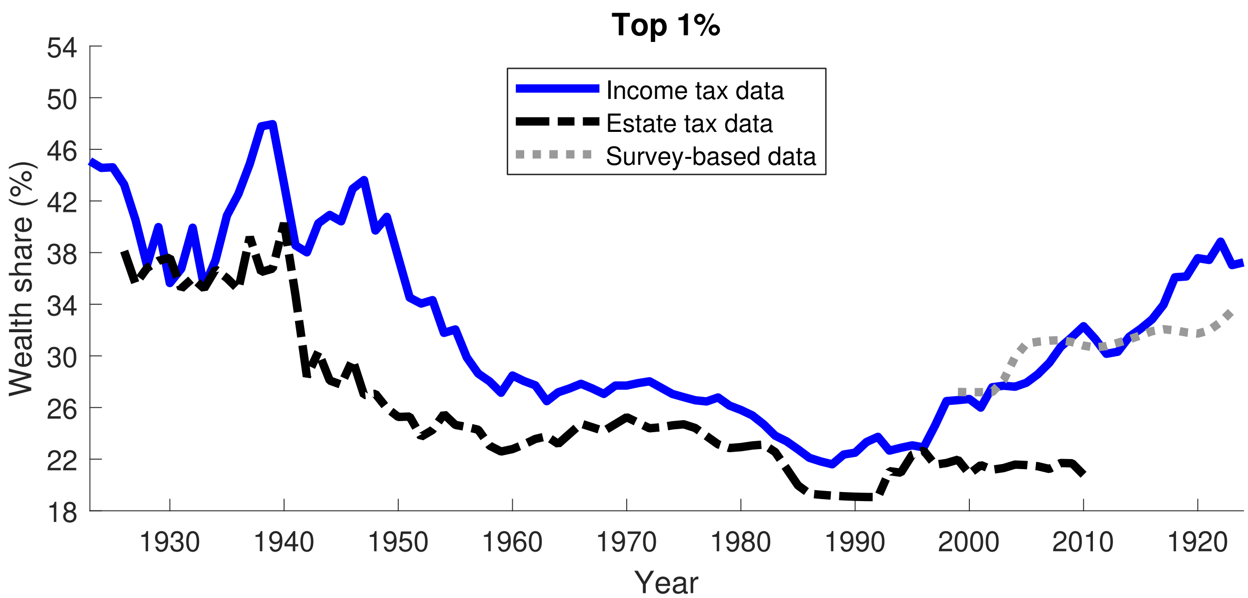
- The capitalization method uses information on capital income from individual income tax returns to estimate the underlying stock of wealth (Saez and Zucman, 2016; The World Inequality Database, 2016). "If we can observe capital income $k = rW$, where W is the underlying value of an asset and r is the known rate of return, then we can estimate wealth based on capital income and capitalization factor $1/r$ defined using the appropriate choice of rate of return" (Kopczuk, 2015, p. 54). Data availability: the wealth shares of the top 5, 0.5, 0.1 and 0.01 percent for 1917–2012 and of the top 10 and 1 percent for 1913–2014 (annually).

- The estate multiplier method uses data from estate tax returns to estimate wealth for the upper tail of the wealth distribution (Kopczuk and Saez, 2004). "The basic idea is to think of decedents as a sample from the living population. The individual-specific mortality rate m_i becomes the sampling rate. If m_i is known, the distribution for the living population can be simply estimated by reweighting the data for decedents by inverse sampling weights $1/m_i$, which are called 'estate multipliers"' (Kopczuk, 2015, p. 53). Data availability: the wealth shares of the top 1, 0.5, 0.25, 0.1, 0.05 and 0.01 percent for 1916–2000 (annually, with several missing years).

- The survey-based method uses data from the Survey of Consumer Finances (SCF) conducted by the Federal Reserve, plus defined-benefit pension wealth, plus the wealth of the members of the Forbes 400 (Bricker et al., 2016). Data availability: the wealth shares of the top 1 and 0.1 percent for 1989–2013 (for every three years).

These sources are based on different datasets and for different time periods. In the overlapping periods, they sometimes report markedly different wealth share estimates (see Figure 6).

Kopczuk (2015) reviewed the advantages and disadvantages of the different methods (see also the comment by Kopczuk on Bricker et al. (2016)). He observed that "the survey-based and estate tax methods suggest that the share of wealth held by the top 1 percent has not increased much in recent decades, while the capitalization method suggests that it has" (Kopczuk, 2015, p. 48). However, recent evidence suggests that the different sources may



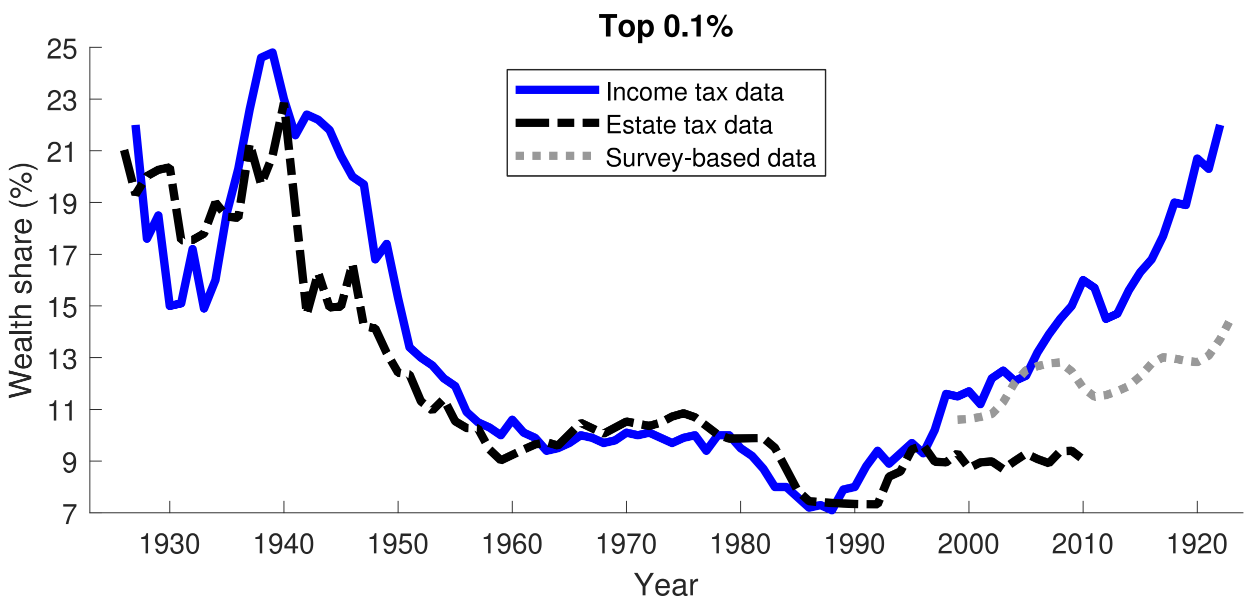
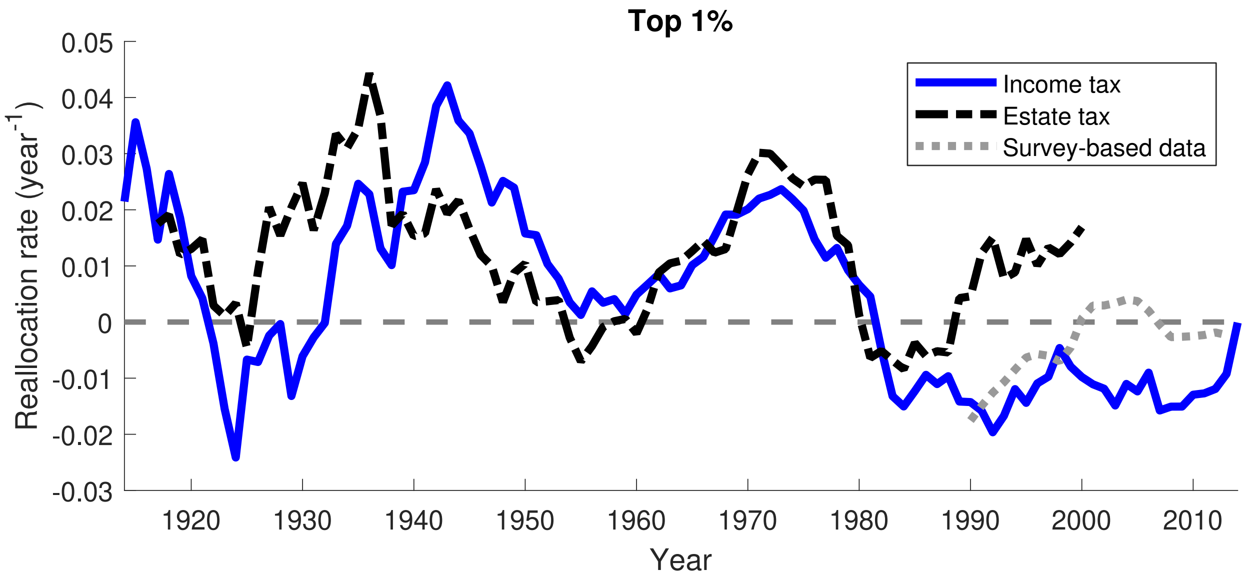


Figure 6: Top 1% and 0.1% wealth shares in the US, 1913–2014. Sources – Saez and Zucman (2016); The World Inequality Database (2016) (solid blue); Kopczuk and Saez (2004) (dash-dotted black); Bricker et al. (2016) (dotted grey).

lead to closer estimates of top wealth shares than previously thought (Saez and Zucman, 2019).

Which method best reflects the recent trends in wealth inequality is a matter of ongoing debate. Each method suffers from bias. For example, the survey-based method suffers from some underrepresentation of families who belong to the top end of the distribution. The capitalization method suffers from some practical difficulties – "not all categories of assets generate capital income that appears on tax returns. [...] Owner-occupied housing does not generate annual taxable capital income." (Kopczuk, 2015, p. 54) The estate tax method suffers from the need to accurately estimate mortality rates for the wealthy, known to be lower than those for the rest of the population conditional on age. We refer the reader to Kopczuk (2015); Bricker et al. (2016) for a thorough discussion.

As a robustness test, we fit $\tau(t)$ for different top wealth shares based on each dataset separately, see Figure 7. We observe briefer periods in which $\widetilde{\tau}(t) < 0$ for the survey-based wealth shares (Bricker et al., 2016). The same is true for the estate tax data (Kopczuk and Saez, 2004). When $\tau(t)$ is positive, relevant convergence times are long compared to the timescales of policy changes, namely at least several decades. Therefore, none of the datasets support making the ergodic hypothesis for rescaled wealth in RGBM.



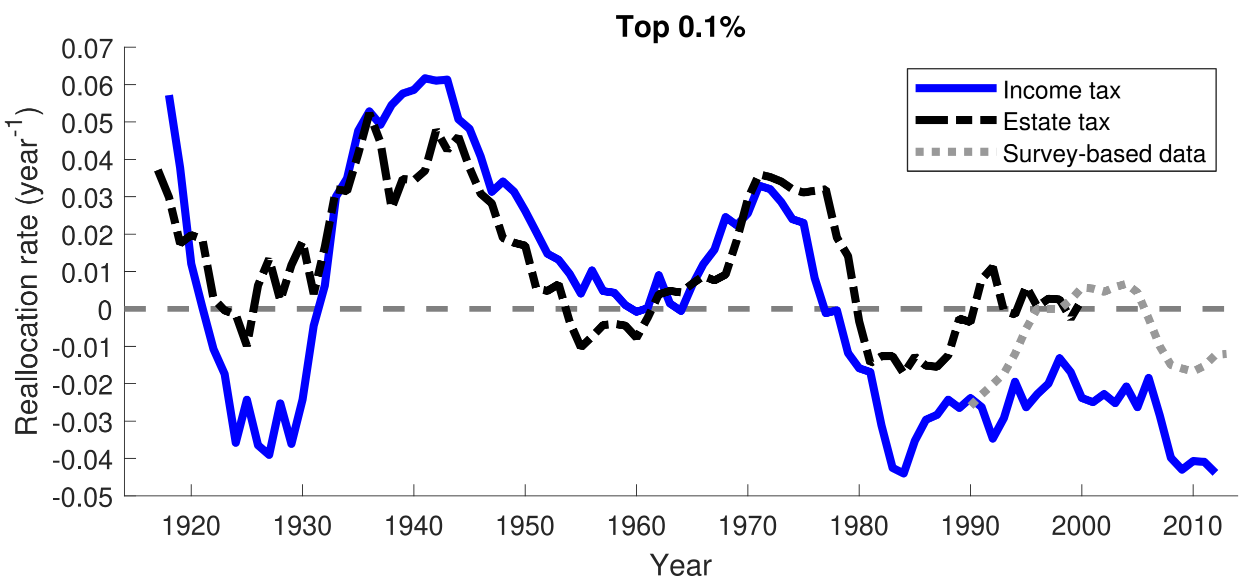


Figure 7: Effective reallocation rates for different datasets.

E The Effect of Fixed Versus Time-Varying σ

Our analysis requires us to set the volatility parameter, σ, in the RGBM model. Our baseline estimates (see Sec. 4) are done with a fixed $\sigma = 0.16$ year$^{-1/2}$. In this appendix we show that letting $\sigma(t)$ change in time has only a limited quantitative impact on the results.

First, for each year in the analysis, we define $\sigma(t)$ as the standard deviation of the daily logarithmic changes in the Dow Jones Industrial Average (DJIA) for that year (Quandl, 2016), annualized by multiplying by $(250/\text{year})^{1/2}$. We present this in Figure 8. In order to show that using the time-varying $\sigma(t)$ has little effect on the estimated reallocation rates, a comparison is presented in Figure 9.

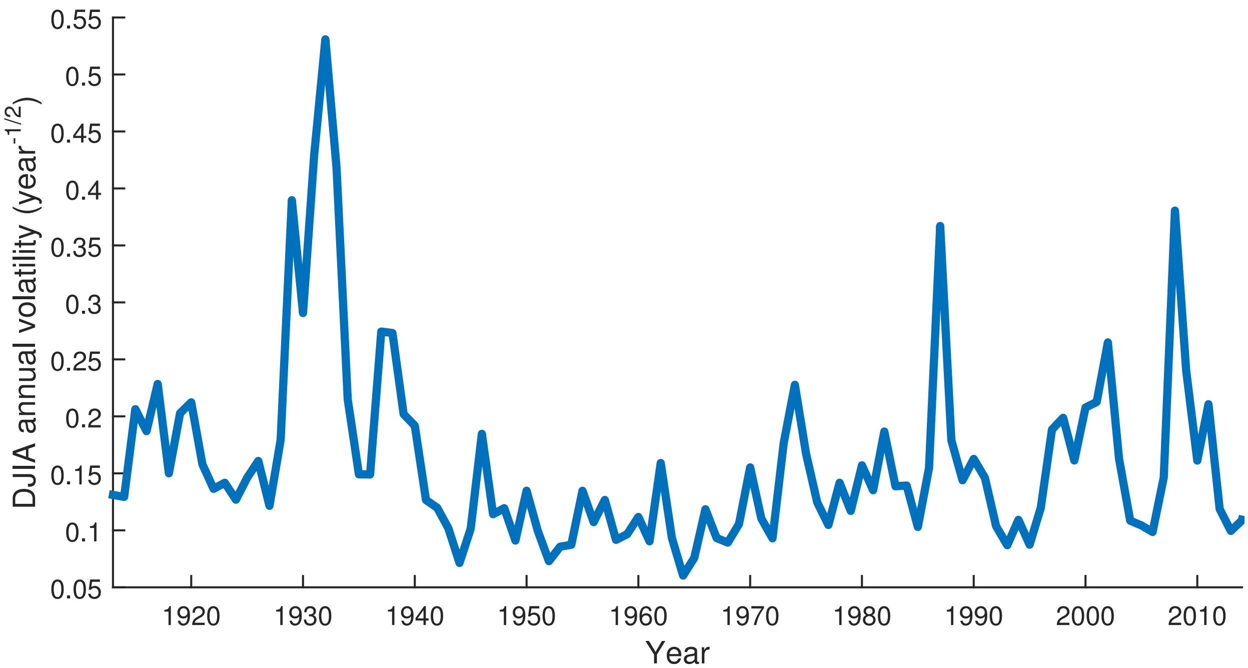


Figure 8: Annualized $\sigma(t)$ estimated from daily changes in the DJIA. Data taken from Quandl (2016).

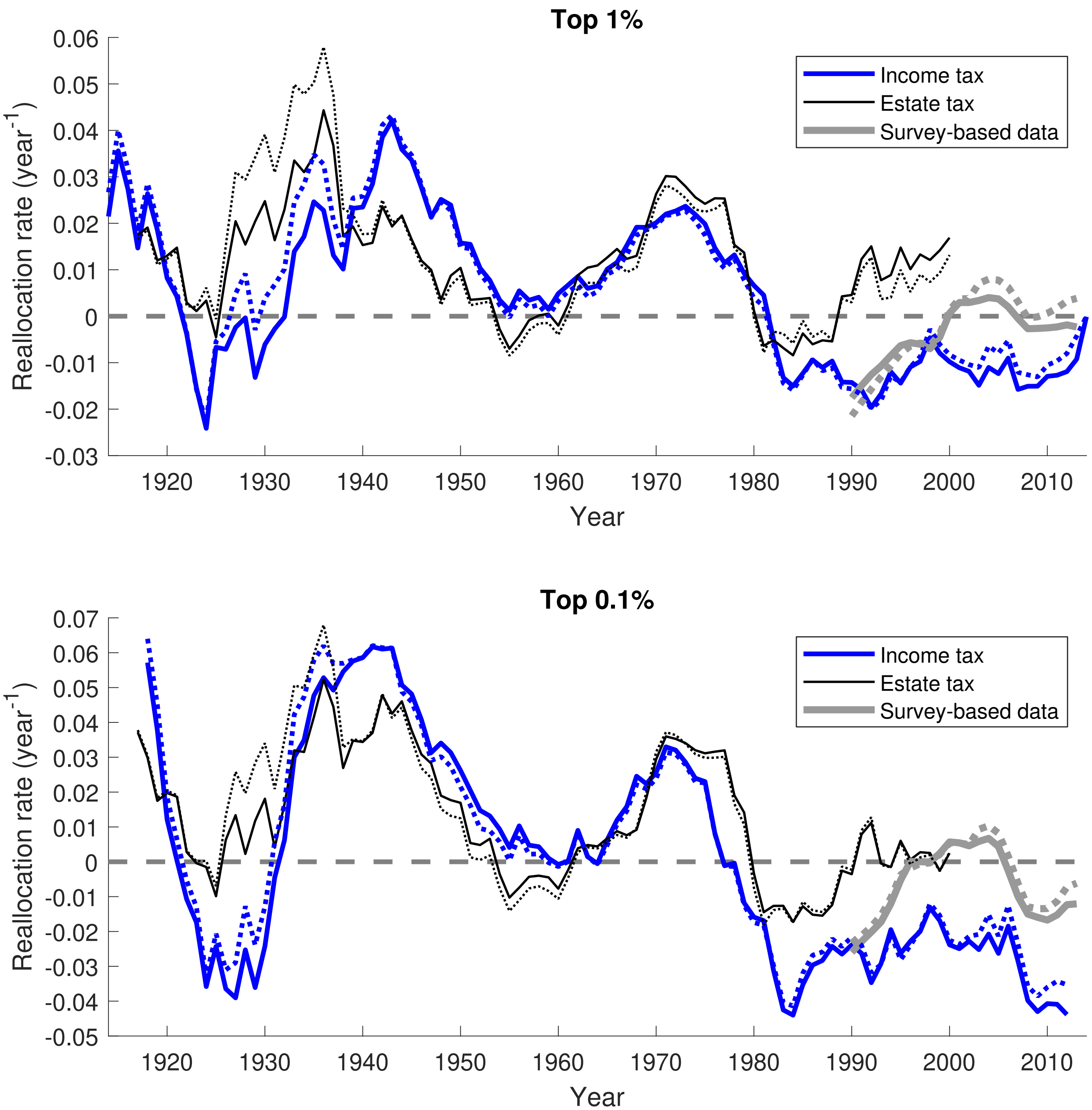


Figure 9: Smoothed effective reallocation rates with fixed and time-varying σ. Solid lines – fixed σ; Dotted lines – time-varying σ.

E.1 Robustness of Fitting τ and σ Together

It is also possible to estimate the effective reallocation rate $\tau_{\text{all\%}}(t)$ that best fits the data for all top shares together (q = 10, 5, 1, 0.5, 0.1 and 0.01 percent), and not to a specific share only. In addition, instead of fixing σ at 0.16 year$^{-1/2}$ as we did in our analysis, we can estimate both $\tau_{\text{all\%}+\sigma}(t)$ and $\sigma(t)$ together to best fit the data for all six top shares, for robustness. These results are presented in Figure 10. They show that although fitting τ

to a specific top wealth quantile matters quantitatively, the results are qualitatively robust. Also, letting $\sigma(t)$ and $\tau(t)$ change together to fit all top wealth shares does not change the results substantially.

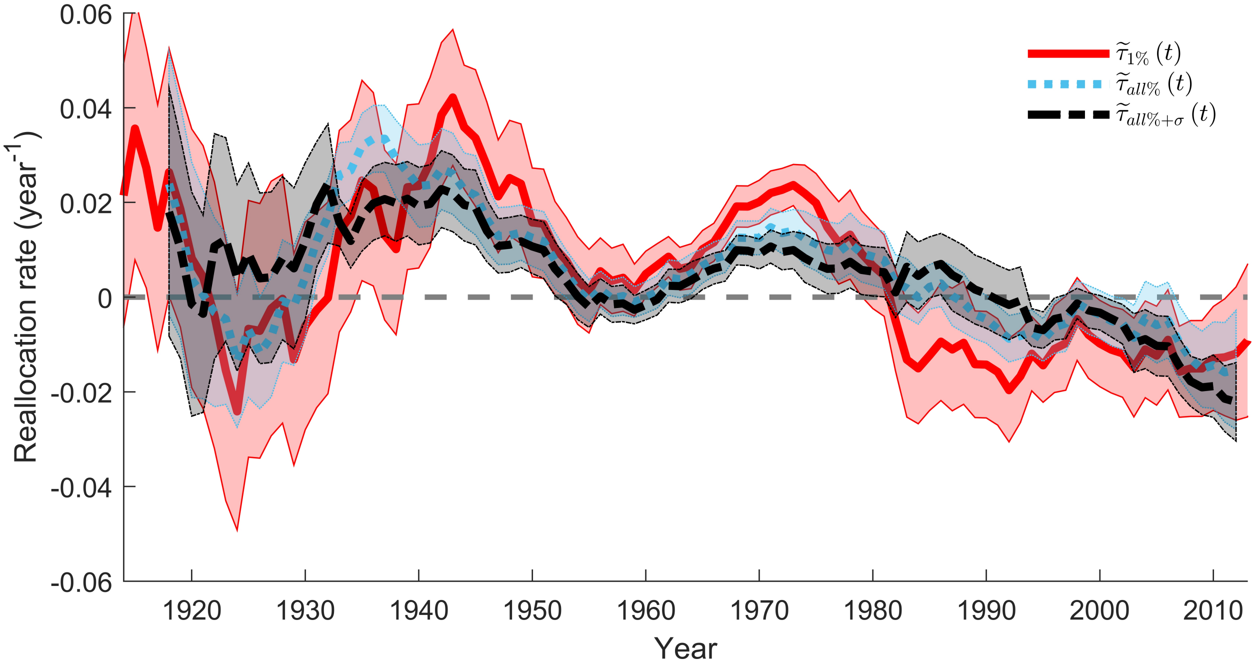


Figure 10: Fitted effective reallocation rates. Calculations done using $\mu = 0.021$ year^{-1} and $\sigma = 0.16$ year$^{-1/2}$ when fitting $\widetilde{\tau}_{1\%}(t)$ (solid red) and $\widetilde{\tau}_{\text{all}\%}(t)$ (dotted blue) and using $\mu = 0.021$ year^{-1} when fitting $\widetilde{\tau}_{\text{all}\%+\sigma}(t)$ (dash-dotted black). Translucent envelopes indicate one standard error in the moving averages.

F The Calculation of Variance Convergence Times

Our key finding is that, under currently prevailing economic conditions, it is unsafe to assume the existence of stationary rescaled wealth distributions in models of wealth dynamics. Nevertheless, we present some results for the regime of our model where a stationary distribution exists.

In this regime it is possible to calculate how fast the wealth shares of different quantiles converge to their asymptotic value. We do this numerically. Starting with a population of equal wealths and assuming $\mu = 0.021$ year^{-1}, $\sigma = 0.16$ year$^{-1/2}$, and $\tau = 0.04$ year^{-1}, we let the system equilibrate for long enough (3000 years in the simulation), so the distribution reaches its asymptotic form to numerical precision. We then create a shock, by changing τ to a different "shock value", and allow the system to equilibrate again for a long enough time (see top panel of Figure 11). Following the shock, the wealth shares converge to their asymptotic values. We fit this convergence numerically with an exponential function and interpret the inverse of the exponential convergence rate as the convergence time. The bottom panel of Figure 11 shows the convergence times versus the shock value of τ. This relationship was used in Figure 3 to present the evolution of the top wealth share convergence times in the US.

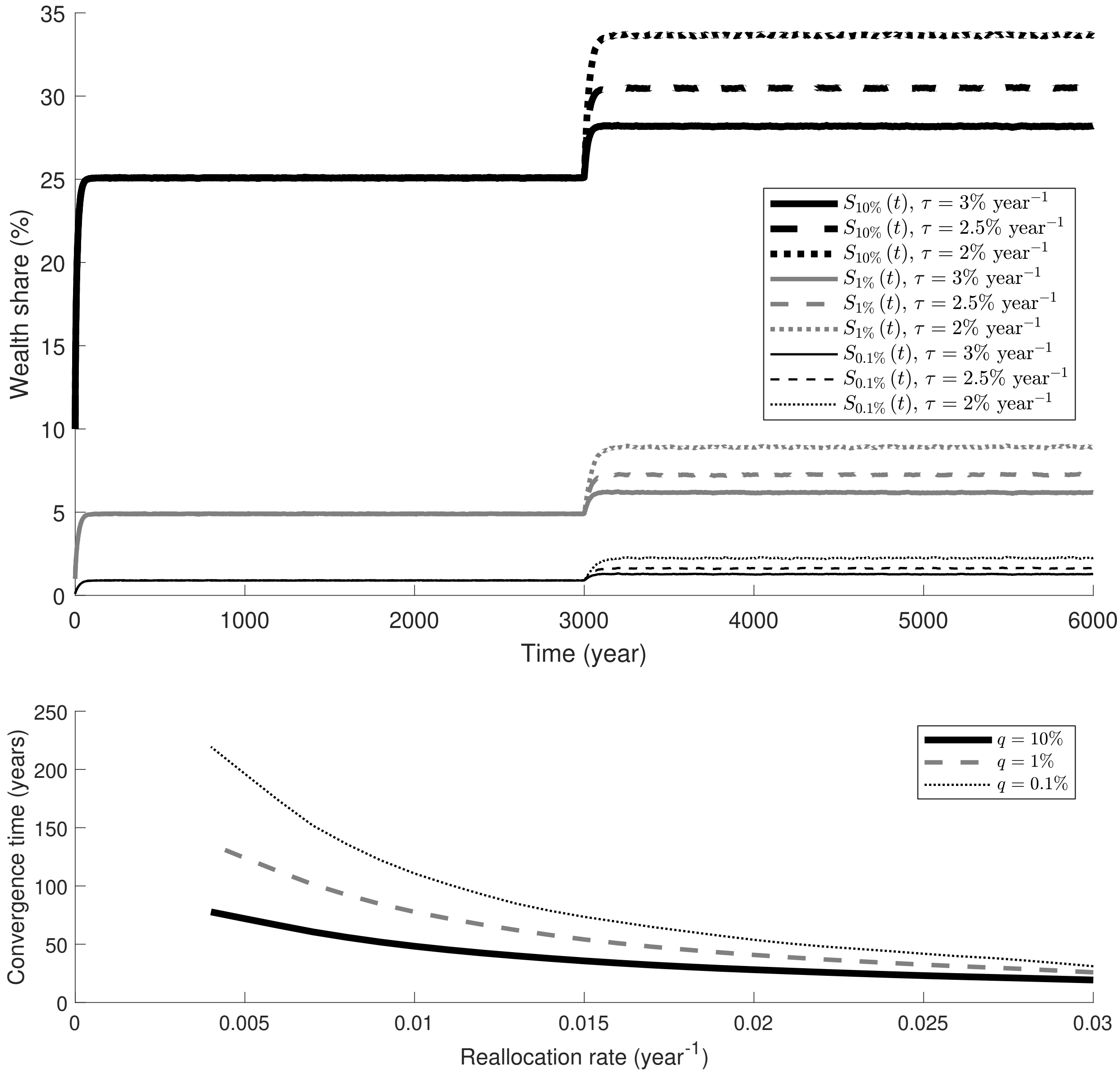


Figure 11: Wealth share convergence time. Top: The convergence of the wealth share for $q = 10\%$ (thick black), $q = 1\%$ (grey) and $q = 0.1\%$ (thin black) following a change in the value of τ from 0.04 year^{-1} to 0.03 year^{-1} (solid), 0.025 year^{-1} (dashed) and 0.02 year^{-1} (dotted). Bottom: The wealth share exponential convergence time for $q = 10\%$ (solid black), $q = 1\%$ (dashed grey) and $q = 0.1\%$ (dotted black) as a function of τ.

F.1 The Derivation of the Variance Convergence Time

In the ergodic regime it is possible to derive analytically the convergence time of the variance of the stationary distribution (and other cumulants and moments of interest).

The variance of y is a combination of the first moment, $\langle y \rangle$ (the average), and the second moment, $\langle y^2 \rangle$:

$$V(y) = \left\langle y^2 \right\rangle - \langle y \rangle^2 \tag{F.1}$$

We thus need to find $\langle y \rangle$ and $\langle y^2 \rangle$ in order to determine the variance. The first moment of the rescaled wealth is, by definition, $\langle y \rangle = 1$.

To find the second moment, we start with the SDE for the rescaled wealth:

$$dy = \sigma y \, dW - \tau (y - 1) \, dt \, . \tag{F.2}$$

This is an Itô process, which implies that an increment, df, in some (twice-differentiable) function $f(y,t)$ will also be an Itô process, and such increments can be found by Taylor-expanding to second order in dy as follows:

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial y} dy + \frac{1}{2} \frac{\partial^2 f}{\partial y^2} dy^2 \, . \tag{F.3}$$

We insert $f(y,t) = y^2$ and obtain

$$d\left(y^2\right) = 2y dy + (dy)^2 \, . \tag{F.4}$$

We substitute dy in Equation (F.4), which yields terms of order dW, dt, dW^2, dt^2, and $dW dt$. The scaling of Brownian motion allows us to replace dW^2 by dt, and we ignore $o(dt)$ terms. This yields

$$d\left(y^2\right) = 2\sigma y^2 \, dW - \left(2\tau - \sigma^2\right) y^2 \, dt + 2\tau y \, dt \, . \tag{F.5}$$

Taking expectations on both sides, and using $\langle y \rangle = 1$, produces an ordinary differential equation for the second moment:

$$\frac{d\langle y^2 \rangle}{dt} = -\left(2\tau - \sigma^2\right) \langle y^2 \rangle + 2\tau \tag{F.6}$$

with solution

$$\langle y(t)^2 \rangle = \frac{2\tau}{2\tau - \sigma^2} + \left(\langle y(0)^2 \rangle - \frac{2\tau}{2\tau - \sigma^2} \right) e^{-\left(2\tau - \sigma^2\right)t} \, . \tag{F.7}$$

The variance $V(t) = \langle y(t)^2 \rangle - 1$ therefore follows

$$V(t) = V_\infty + (V_0 - V_\infty) \, e^{-\left(2\tau - \sigma^2\right)t} \, , \tag{F.8}$$

where V_0 is the initial variance and

$$V_\infty = \frac{2\tau}{2\tau - \sigma^2} - 1\,. \tag{F.9}$$

V converges in time to the asymptote, V_∞, provided the exponential in Equation (F.8) is decaying. This can be expressed as a condition on τ

$$\tau > \frac{\sigma^2}{2}\,. \tag{F.10}$$

Clearly, for negative values of τ the condition cannot be satisfied, and the variance (and inequality) of the rescaled wealth distribution will diverge. In the regime where the variance exists, $\tau > \sigma^2/2$, it also follows from Equation (F.8) that the convergence time of the variance is $1/\left(2\tau - \sigma^2\right)$.

As τ increases, increasingly high moments of the distribution become convergent to some finite value. The above procedure for finding the second moment can be applied to the k^{th} moment, just by changing the second power y^2 to y^k, and any other cumulant can therefore be found as a combination of the relevant moments. For instance Liu and Serota (2017) also compute the third cumulant.

Figure 12 shows the dependence of the variance convergence time on τ for different values of σ.

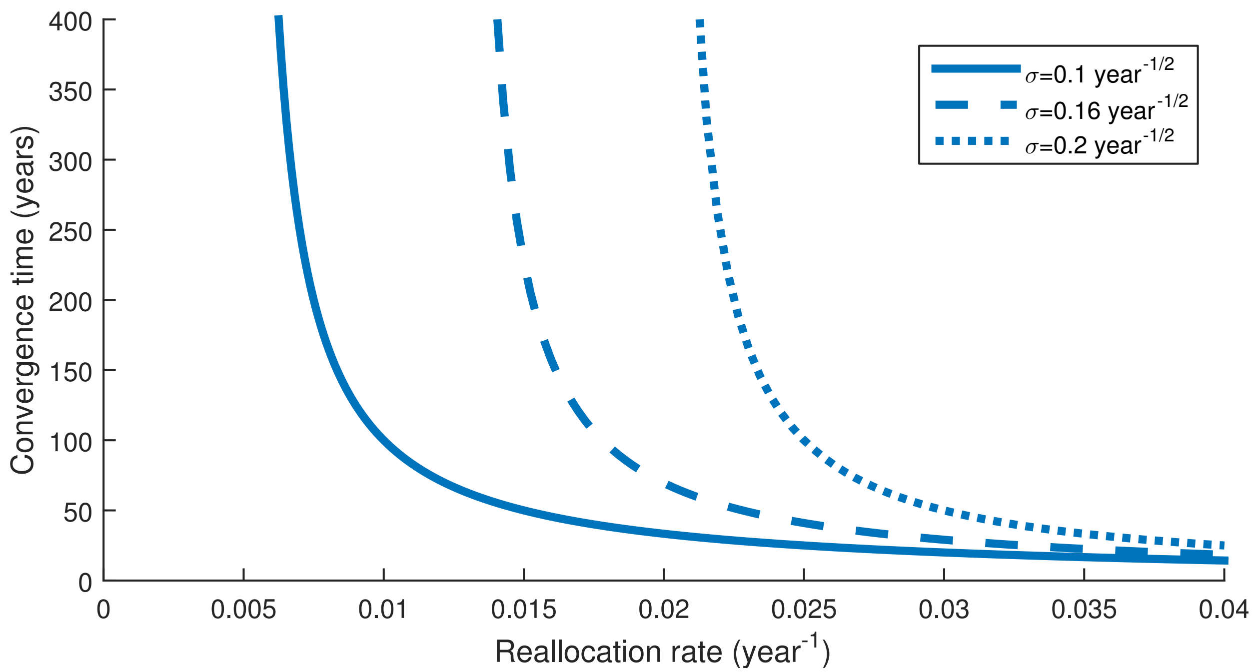


Figure 12: The variance convergence time as a function of the reallocation rate.