Displacement and the Rise in Top Wealth Inequality*

Matthieu Gomez †

First version: November 2016

This version: June 2019

[latest version]

Abstract

I show that the growth of top wealth shares can be decomposed into two terms: (i) a within term, driven by the *average* wealth growth of households in top percentiles relative to the economy and (ii) a displacement term, driven by all *higher-order moments* of their wealth growth. After mapping this decomposition to the data, I find that the displacement term accounts for more than half of the rise in top wealth shares in the United States since 1983. This finding has important implications for the relationship between wealth inequality and economic growth, as well as for wealth mobility.

*I am grateful to my advisors Valentin Haddad, Atif Mian, Wei Xiong, and Motohiro Yogo for their help and support. I am also grateful to Laurent Bach, Jarda Borovička, Jack Favilukis, Gustavo Grullon, Bernard Herskovic, Leonid Kogan, Wojciech Kopczuk, Erik Loualiche, Ben Moll, Stavros Panageas, Jonathan Parker, Dejanir Silva, Annette Vissing-Jorgensen and Gabriel Zucman for helpful discussions and comments.

†Columbia University, mg3901@columbia.edu

1 Introduction

What drives the recent rise in top wealth shares? One common view is that households in top percentiles grow faster than the economy, i.e. that "rich are getting richer" (Piketty (2014), Hubmer et al. (2016)). This view implicitly assumes that the composition of households in top percentiles remains constant over time. In reality, however, less than 10% of the households in the 1983 Forbes list of the 400 richest households in the United States were still on the list in 2017. This paper examines the role of these composition changes for the growth of top wealth inequality.

I show that the growth of the wealth share of a top percentile can always be decomposed into two terms. The first term ("within" term) corresponds to the difference between the average wealth growth of households in the top percentile and the growth of the economy. The second term ("displacement" term) corresponds to the wealth of households entering the top percentile minus the wealth of households that exit the top.

The displacement term is driven by the dispersion of wealth shocks among rich households. When the wealth of rich households follows a diffusion process (i.e. log-normal innovations), the displacement term equals $1/2(\zeta - 1)\nu^2$, where ν denotes the idiosyncratic volatility of wealth growth and ζ denotes the local tail index of the wealth distribution. Thus, displacement increases with the dispersion of wealth shocks ν and decreases with the level of wealth inequality (i.e. increases with the tail exponent ζ). Intuitively, the higher the level of wealth inequality, the bigger the gap between the rich and the rest of the population, and, therefore, the harder it is for households to displace the existing rich.

The rapid rise of a few entrepreneurs at the top of the distribution suggests that higher-order cumulants may play an important role in the rise in top wealth inequality. To examine this hypothesis, I also consider a model in which the wealth of rich households follows a jump-diffusion process (i.e. with non log-normal innovations). In this case, the displacement term equals $\sum_2^{+\infty} 1/j!(\zeta^{j-1} - 1)\kappa_j$ where κ_j denotes the j-th cumulant of wealth growth and ζ denotes the tail index of the wealth distribution. For instance, skewness κ_3 increases the growth of top wealth shares by $1/6(\zeta^2 - 1)\nu^3\kappa_3$ while excess kurtosis κ_4 increases the growth of top wealth shares by $1/24(\zeta^3 - 1)\nu^4\kappa_4$.

After applying this framework to the share of wealth owned by the 400 wealthiest in the U.S., I document three facts about the role of displacement in the recent rise of top wealth inequality. First, displacement accounts for more than half of the increase in top wealth inequality since 1983. While the within term accounts for an annual growth of the wealth share of the top 400 of 1.9%, the displacement term accounts for an annual growth of 2.3%. The magnitude of this displacement term is well explained by a simple diffusion model (i.e. log normal innovations). Indeed, with a measured tail index $\zeta \approx 1.5$ and a measured annual idiosyncratic volatility of wealth $\nu \approx 27\%$, the diffusion model predicts a displacement term around $1/2(\zeta - 1)\nu^2 \approx 2.0\%$ per year, which is close to the actual displacement term 2.3%.

Second, displacement has steadily declined over time, from 3.2% in the 1980s to 1.5% in the 2010s. Using the theoretical model discussed above, this change can be decomposed into a change in the idiosyncratic volatility ν and a change in the tail index ζ. Half of the decrease of the displacement term is driven by a decrease in the idiosyncratic volatility of household wealth from $\nu_{1983-1993} = 28\%$ to $\nu_{2005-2015} = 23\%$, which follows from a similar decrease of the cross-sectional standard deviation of firm-level returns. This is consistent with a recent literature documenting a recent decline in business dynamism (Decker et al. (2016a), Decker et al. (2016b)). The other half is driven by a fattening of the wealth distribution over time, from $\zeta_{1980s} = 1.8$ to $\zeta_{2010s} = 1.4$. Intuitively, as wealth inequality increased during the time period, it became gradually harder for households with positive shocks to reach the top.

Third, technological innovation is an important driver of displacement. In the cross-section, households entering top percentiles tend to innovate more than the households that they displace. In the time series, displacement is correlated with aggregate measures of innovation. I also find that most of displacement happens within industries rather than between industries. Overall, these findings suggest that, far from being a symptom of a stalling economy, the rise in wealth inequality in the 1980s and 1990s reflects the rapid pace of innovation during that period.

I then use the diffusion model to estimate the role of displacement in the wealth share of the top 1%, 0.1%, and 0.01% over the 20th century, for which we lack panel data. I estimate the idiosyncratic volatility ν of wealth at these top percentiles by interacting the share of wealth invested in equity with the yearly cross-sectional standard variance of firm-level returns. This allows me to obtain a model-implied displacement term without using panel data. Overall, displacement matches the inverted U-shape of top wealth inequality over the 20th century. It first peaked during the Great Depression, remained low during the World War and the postwar economic boom, before peaking again during the technological revolutions of the 1980s and 1990s. This suggests that displacement is a driving force of the low-frequency fluctuations in wealth inequality observed during the 20th century.

Finally, I examine the implications of my findings on wealth mobility. I define wealth mobility as the average time a household in a given top percentile remains in the top. I find that, while a rise in the average wealth growth of top households tends to decrease wealth mobility, a rise in the dispersion of their wealth growth tends to increase wealth mobility. This is true even though it increases wealth inequality in the long-run. The importance of displacement in the recent rise of top wealth shares suggests that wealth inequality and wealth mobility go hand in hand.

Related Literature. This paper is related to a recent empirical literature documenting the rise in top wealth shares in the U.S. in the last thirty years (Kopczuk and Saez (2004), Piketty (2014), Saez and Zucman (2016), Piketty and Zucman (2015), Garbinti et al. (2017), and Kuhn et al. (2017)). This literature tends to interpret the rise in top wealth shares as a rise in the wealth growth of households in top percentiles relative to the rest of the economy. In particular, Saez and Zucman (2016) defines a "synthetic saving rate" as the difference between the wealth growth of top wealth shares and the average return of top households. My paper clarifies that this synthetic saving rate is actually the sum of three conceptually different terms: a household saving rate, a "displacement" term due to idiosyncratic wealth shocks, and a "demography" term due to the death of households in top percentiles and population growth.

Recent empirical papers stress the importance of idiosyncratic shocks at the very top. Guvenen et al. (2014) documents the importance of skewness and kurtosis for labor income growth at the top. Bach et al. (2015) stresses the dispersion of wealth growth across households, using administrative data from Sweden. Similarly, Fagereng et al. (2016) documents a large heterogeneity in asset returns at the top using administrative data from Norway. Relative to this literature, my contribution is to (i) present an accounting decomposition to identify the contribution of these idiosyncratic shocks to the growth of top wealth shares, and to (ii) present a theoretical framework to relate this term to the cumulants of wealth growth and to the shape of the wealth distribution. While I focus on the dynamics of top wealth shares in the U.S., these tools could be applied to wide range of settings. Bach et al. (2017) uses the accounting framework presented in this paper to decompose the dynamics of top wealth shares in Sweden.

I relate the dispersion of wealth growth to the dispersion of asset returns. This connects this paper to a large asset pricing literature studying the dispersion of stock market returns. Campbell et al. (2001) documents the rise in idiosyncratic volatility in the 1980s and the 1990s. Herskovic et al. (2016) points out the degree of comovement among the idiosyncratic volatilities of U.S. stocks. Martin (2013) examines the role of higher-order cumulants for the equity premium. Bessembinder (2018) focuses on skewness in firm-level returns: more than half of stocks have returns lower than one month Treasuries. Oh and Wachter (2018) examines the role of this skewness in the distribution of firm size.

This work also contributes to a more theoretical literature that studies wealth inequality through the lens of random growth models (Wold and Whittle (1957), Acemoglu and Robinson (2015), Jones (2015)). In particular, Luttmer (2012), Gabaix et al. (2016) and Jones and Kim (2016) recently developed tools to study the *dynamics* of the wealth density over time. My contribution is to extend these tools to characterize directly the dynamics of top shares.[1] I also develop a new accounting framework that allows me to map directly random growth models to the data. This new method reveals the importance of the dispersion of wealth shocks in the recent rise in inequality.

A recent macroeconomic literature examines the drivers of top wealth inequality in general equilibrium models. For instance, Benhabib et al. (2011) and Benhabib et al. (2015b) examine the stationary wealth distribution in an economy with idiosyncratic returns. More recently, Benhabib et al. (2015a) and Hubmer et al. (2016) develop equilibrium models to match the recent rise in top wealth shares. Compared to these papers, I stress the role of the rise in idiosyncratic shocks in the recent rise in top wealth inequality. While I take a reduced-form approach, the decomposition developed in my paper can be used to further discipline these types of models.

This paper contributes to a growing literature documenting the relationship between inequality and technological innovation. Kogan et al. (Forthcoming) shows that the skewness of consumption in CEX correlates with economy-wide innovation. In the cross-section, Aghion et al. (2015) documents a positive relationship between innovation and top income inequality across U.S. states. In a contemporaneous paper, Gârleanu and Panageas (2017) stresses the growth of self-made billionaires compared to pre-existing billionaires, using data from Forbes 400. Papanikolaou et al. (2018) documents a relationship between innovation and the rise in income inequality. This paper is also related to a literature stressing the role of entry and exit for aggregate growth. In particular, Melitz and Polanec (2015) decomposes the growth of aggregate productivity into a term due to the productivity growth of existing firms and a term due to the entry and exit of firms in the economy. My focus is different since I decompose the aggregate growth of a top percentile, rather than the aggregate growth of the economy: what I call displacement corresponds to the flow of agents in and out top percentiles, rather than the flow of agents in and out the economy.

Campbell (2016) proposes an alternative way to decompose the growth of wealth inequality. In the paper, the change in the variance of the distribution of log wealth is decomposed into a term

[1] To examine the impact of an increase of idiosyncratic volatility on top wealths shares, Gabaix et al. (2016) uses the Kolmogorov Forward Equation to simulate the dynamics of wealth density, and then integrate the simulated path of the density to obtain the dynamics top wealths shares. By contrast, this paper shows how to examine directly the dynamics of top wealth shares.

due to differences in expected wealth growth and a term due to differences in unexpected wealth shocks. Compared to this paper, my decomposition is *local*, in the sense that it decomposes the rise in wealth inequality at any top percentile of the distribution. Moreover, my focus on top wealth shares fits more directly with the empirical literature on inequality, that focuses on top wealth share to describe the evolution of wealth inequality.

Roadmap. The rest of my paper is organized as follows. In Section 2, I derive in continuous-time the law of motion of top wealth shares to the law of motion of household wealth. In Section 3, I present an accounting framework to decompose the growth of top wealth shares into a within term, a displacement term, and a demography term using panel data. In Section 4, I apply this framework to decompose the growth of Forbes 400. In Section 5, I examine the role of displacement for the top 1%, 0.1%, and 0.01% over the 20th century. In Section 6, I discuss the implication of my findings for technological innovation and wealth mobility.

2 Theory

In this section, I present the main theoretical contribution of this paper: I derive a formula relating the growth of top wealth shares to the dynamics of individual wealth. Section 2.1 first presents the result in the simple case where wealth follows a simple diffusion process. Section 2.2 extends the result to more realistic wealth dynamics, that account for households heterogeneity, jumps, population death, and population growth.

Denote g_t the density of normalized wealth in the economy. For a given top percentile p, denote q_t the normalized wealth of an household at the percentile threshold.[2] The wealth share owned by the top percentile p, S_t, can be written as the total normalized wealth of households above the threshold q_t:[3]

$$S_t = \int_{q_t}^{+\infty} w g_t(w) dw \tag{1}$$

2.1 Baseline Model

Law of Motion of Relative Wealth Denote w_{it} the normalized wealth of an individual i, i.e. the ratio between individual wealth and aggregate wealth. Assume that, in the upper tail of the

[2] Formally, q_t corresponds to the $1 - p$ quantile.

[3] Here and for the rest of the paper, I assume that the top wealth share is finite, i.e. that the tail index of the distribution is higher than one.

wealth distribution, the normalized wealth of household i, w_{it}, follows a diffusion process:

$$\frac{dw_{it}}{w_{it}} = \mu_t dt + \nu_t dB_{it} \tag{2}$$

where B_{it} is an idiosyncratic Brownian Motion. The average growth of normalized wealth, μ_t, corresponds to the difference between the average growth rate of individuals in the right tail and the growth rate of total wealth in the economy. The idiosyncratic volatility, ν_t, models the dispersion in the wealth growth of individuals. For now, I assume that the geometric drift μ_t and the geometric volatility ν_t do not depend on wealth, at least in the upper tail.

Law of Motion of Top Wealth Share I now give a heuristic derivation for the law of motion of top wealth share.[4] During a short period of time dt, two things happen. First, individuals in the top percentile grow by μ_t relative to the economy. Holding the composition of households at the top fixed, this increases the top wealth share S_t by μ_t. This is the "within" term.

Second, individuals experience idiosyncratic shocks in their wealth. By changing the composition of households in the top percentile, this changes the top wealth share S_t due to a "displacement" term. First, as shown in Figure 1a, some households below the percentile threshold, with a positive wealth shock, enter the top percentile. Because population size in the top percentile is held constant, each entering household displaces a household at the lower percentile threshold, with wealth q_t. Overall, this increases wealth in the top percentile by

$$\int_{q_t/(1+\nu_t\sqrt{dt})}^{q_t} ((1+\nu_t\sqrt{dt})w - q_t)\frac{1}{2}g_t(w)dw = \frac{1}{4}g_t(q_t)q_t^2\nu_t^2 dt + o(dt)$$

Second, as shown in Figure 1b, some households above the percentile threshold, with a negative wealth shock, exit the top percentile. Each exiting household is replaced by a household at the lower percentile threshold, with wealth q_t. Overall, this increases wealth in the top percentile by

$$\int_{q_t}^{q_t/(1-\nu_t\sqrt{dt})} (q_t - (1-\nu_t\sqrt{dt})w)\frac{1}{2}g_t(w)dw = \frac{1}{4}g_t(q_t)q_t^2\nu_t^2 dt + o(dt)$$

Summing the term due to entry and exit gives the following proposition.

Proposition 1 (Dynamics of Top Wealth Share)**.** *Assume that, in the upper tail of the distribution, the law of motion for wealth is given by (2).*[5] *Then the top wealth share S_t follows the law of motion:*

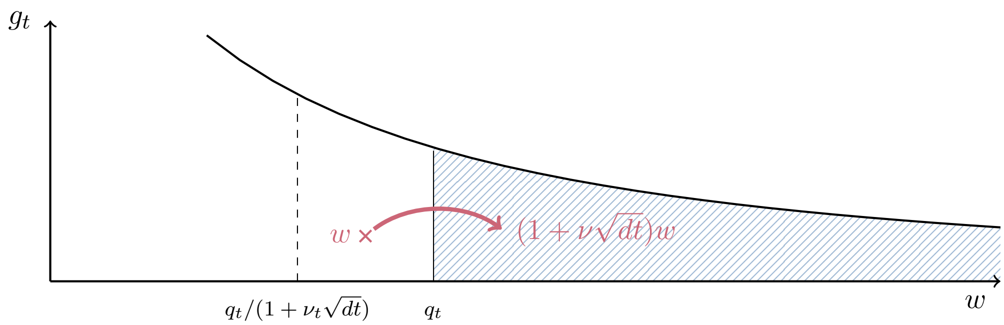
$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt}_{dr_{within}} + \underbrace{\frac{g_t(q_t)q_t^2}{2S_t}\nu_t^2 dt}_{dr_{displacement}} \tag{3}$$

[4] A formal proof is given in Appendix A.

[5] This assumption will be made more precise in Proposition 3.

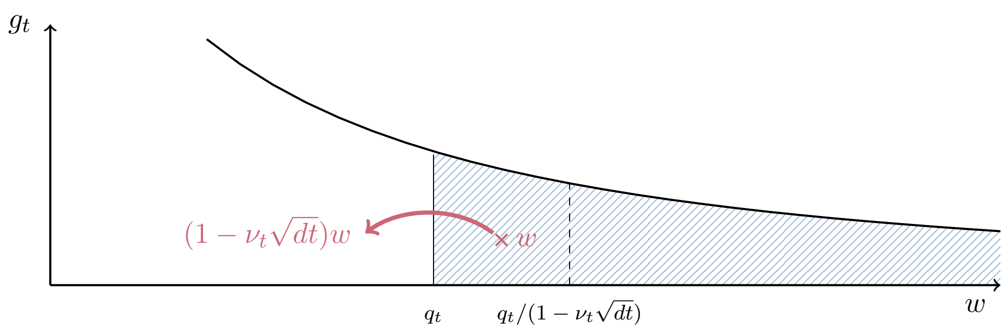
Figure 1: Heuristic Derivation of Displacement Term

(a) Displacement due to Entry



$$dS_t^{\text{entry}} = \int_{q_t/(1+\nu_t\sqrt{dt})}^{q_t} ((1+\nu_t\sqrt{dt})w - q_t)\frac{1}{2}g_t(w)dw = \frac{1}{4}g_t(q_t)q_t^2\nu_t^2 dt + o(dt)$$

(b) Displacement due to Exit



$$dS_t^{\text{exit}} = \int_{q_t}^{q_t/(1-\nu_t\sqrt{dt})} (q_t - (1-\nu_t\sqrt{dt})w)\frac{1}{2}g_t(w)dw = \frac{1}{4}g_t(q_t)q_t^2\nu_t^2 dt + o(dt)$$

This formula expresses the growth rate of the top wealth share as the sum of two terms: the average growth rate of individuals in the top percentile (a "within" term) and a term due to the idiosyncratic volatility of wealth (a "displacement" term).

As seen in the heuristic derivation, the displacement term can be written as product of the mass of households that enter or exit the top percentile during a short period of time dt, relative to the mass of existing households in the top percentile, $\frac{g_t(q_t)q_t}{p}\nu_t\sqrt{dt}$, times the average increase of wealth in the top percentile per entry or exit, relative to the average wealth of households in the top percentile, $\frac{1}{2}\frac{q_t p}{S_t}\nu_t\sqrt{dt}$.

Define the local tail index of the wealth distribution, $\zeta_t(q_t)$, as one minus the elasticity of the top wealth share to the top quantile at q_t:

$$\zeta_t(q_t) = 1 - \frac{\partial \ln S_t}{\partial \ln q_t}(q_t) = 1 + \frac{g_t(q_t)q_t^2}{S_t} \tag{4}$$

The law of motion of the top wealth share can be rewritten in terms of this local tail index:[6]

$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt}_{dr_{\text{within}}} + \underbrace{\frac{\zeta_t(q_t)-1}{2}\nu_t^2 dt}_{dr_{\text{displacement}}} \tag{5}$$

The displacement term depends on two easily measurable quantities: the idiosyncratic variance of wealth growth ν^2 and the local tail index of the wealth distribution at the top quantile q_t, $\zeta_t(q_t)$.

For a Pareto distribution (i.e. $S(q) = Cq^{1-\zeta}$), the local tail index is constant across the wealth distribution, i.e. $\zeta_t(q_t) = \zeta$. For a distribution with a thick tail, (i.e. $S(q) = L(q)q^{1-\zeta}$ where $L(q)$ is a slowly varying function[7]), the local tail index is constant in the right tail of the distribution, i.e. $\zeta_t(q_t) \to \zeta$ as $q_t \to +\infty$.

The displacement term decreases as the local tail index ζ decreases, i.e. as the right tail of the wealth distribution becomes thicker. This is for two reasons. First, as ζ decreases, the distribution becomes more spread out, and there are less households near the lower percentile threshold ; therefore, the mass of households that enter and exit the top percentile during a short time period dt decreases.[8] Second, as ζ decreases, the wealth of these entering households becomes smaller

[6]This is because the first and second derivative of the top wealth share with respect to the top percentile are related to the wealth density and the quantile, $q_t = -\partial_p S_t$ and $g_t(q_t(p)) = -1/\partial_{pp}S_t$.

[7] A slowly varying function L is defined as

$$\lim_{w\to+\infty} \frac{L(tw)}{L(w)} \to 1 \tag{6}$$

[8]As seen in the heuristic derivation of Proposition 1, the mass of households that enter the percentile during a short period of time dt, relative to the mass of existing households in the top percentile, is $\frac{g_t(q_t)q_t}{p}\nu_t\sqrt{dt} = \zeta\nu_t\sqrt{dt}$.

compared to the average wealth of households above the percentile ; therefore, the average change in the top wealth share per entry or exit decreases.[9] As the tail exponent ζ converges to 1 (Zipf's law), the displacement term converges to zero.

Stationary Case. When μ_t and ν_t are constant over time, under certain conditions, the wealth distribution converges to a stationary wealth distribution.[10] In this case, Equation (3) characterizes the tail index of the stationary wealth distribution ζ: it is the index such that the (positive) displacement term exactly compensates the (negative) within term:

$$0 = \mu dt + \frac{\zeta - 1}{2}\nu^2 dt \tag{7}$$

that is, $\zeta = 1 - 2\mu/\nu^2$. This is a well-known formula, that is usually derived from the law of motion of the wealth density (Kolmogorov-Forward equation).[11] Deriving it from the law of motion of top wealth shares allows one to interpret it as a balance equation for top wealth shares.

2.2 Extensions

Equation (3) was derived under the simplifying assumption that wealth followed the same simple diffusion process for all households in the economy. I now focus on four deviations from this model that correspond to more realistic wealth dynamics: scale dependence, household heterogeneity, jumps, and demographic forces. In each case, I derive analytical expressions for the displacement term in terms of the dynamics of individual wealth.

Scale Dependence. Equation (3) assumed the law of motion of households wealth was linear in wealth. While this is a natural assumption, we may expect saving and investment decisions to be heterogeneous across the wealth distribution. This may be due to non homothetic preferences (Roussanov (2010), Wachter and Yogo (2010)), credit constraints (Wang et al. (2016)), stochastic labor income (see Carroll and Kimball (1996)), or heterogeneous investment opportunities at different levels of wealth.

Formally, I assume that the law of motion of normalized wealth depends on the wealth level w,

[9] As seen in the heuristic derivation of Proposition 1, the average increase of wealth per entry or exit, relative to the average wealth of households in the top percentile, corresponds to $\frac{1}{2}\frac{q_t p}{S_t}\nu_t\sqrt{dt} \approx \frac{1}{2}\left(1 - \frac{1}{\zeta}\right)\nu_t\sqrt{dt}$.

[10] One needs to impose that $\mu \leq 0$ and that there is some lower bound on wealth, see Gabaix et al. (2016). Alternatively, one could augment the model with a positive death rate, see Section 2.2.

[11] See for instance Gabaix (2009).

that is

$$\frac{dw_{it}}{w_{it}} = \mu_t(w_{it})dt + \nu_t(w_{it})dB_{it} \tag{8}$$

where μ and ν are differentiable functions of wealth w_{it}.

Proposition 2 (Dynamics of Top Wealth Share with Scale Dependence)**.** *Assume that the law of motion for wealth is given by (8). Then the top wealth share S_t follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mathrm{E}^{gw}[\mu_t(w)|w \geq q_t]dt}_{dr_{within}} + \underbrace{\frac{\zeta_t(q_t) - 1}{2}\nu_t(q_t)^2 dt}_{dr_{displacement}} \tag{9}$$

where E^{gw} denotes the wealth-weighted cross-sectional average along the wealth distribution.

The within term is the wealth-weighted average of the drift in the top percentile. It simply corresponds to the instantaneous growth of total wealth of individuals in the top percentile.

The displacement term depends exclusively on the idiosyncratic variance of households at the lower percentile threshold, $\nu_t(q_t)^2$. This is because, as seen in the heuristic derivation of Proposition 1, only households near the lower percentile threshold enter or exit the top percentile during a short period of time dt. The key assumption is that wealth is a continuous process (i.e. no jumps), which will be relaxed below.

Heterogeneity. Equation (3) was derived under the assumption that all households have the same process for normalized wealth. In reality, different households may have different average wealth growth or different idiosyncratic volatility.

To model this heterogeneity in a parsimonious way, I assume that households can belong to one of $1 \leq n \leq N$ groups. The law of motion of the normalized wealth of households in group n is:

$$\frac{dw_{nt}}{w_{nt}} = \mu_{nt}dt + \nu_{nt}dB_{it} \tag{10}$$

Proposition 3 (Dynamics of Top Wealth Share with Heterogeneity)**.** *Assume that the law of motion for wealth is given by (10). Then the top wealth share S_t follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mathrm{E}^{gw}[\mu_{nt}|w_{it} \geq q_t]dt}_{dr_{within}} + \underbrace{\frac{\zeta_t(q_t) - 1}{2}\mathrm{E}^{gw}[\nu_{nt}^2|w_{it} = q_t]}_{dr_{displacement}} \tag{11}$$

where E^{gw} denotes the wealth-weighted cross-sectional average along the wealth distribution.

Similarly to the case of scale dependence, the within term corresponds to the total wealth growth of households in the top percentile. The displacement term now depends on the average of idiosyncratic variance at the threshold $\mathrm{E}^{gw}[\nu_{nt}^2|w_{it} = q_t]$.

Aggregate Risk. Different households may also have different exposures to aggregate risks. To model this heterogeneity, as above, assume that households can belong to one of $1 \le n \le N$ groups. The law of motion of the normalized wealth of households in group n is

$$\frac{dw_{nt}}{w_{nt}} = \mu_{nt}dt + \sigma_{nt}dZ_t + \nu_{nt}dB_{it} \tag{12}$$

where Z_t is a $d-$dimensional aggregate Brownian motion. Each group n has a different exposure to aggregate risk given by σ_{nt}. Because the aggregate Brownian motion is multidimensional, this setup includes the situation in which households are differently exposed to the same aggregate risk, or in which households are exposed different aggregate risks (such as different industries).

Proposition 4 (Dynamics of Top Wealth Share with Aggregate Risk)**.** *Assume that the law of motion for wealth is given by* (12)*. Then the top wealth share S_t follows the law of motion:*

$$\begin{aligned}\frac{dS_t}{S_t} =& \underbrace{\mathrm{E}^{gw}[\mu_{nt}|w_{it} \ge q_t]dt + \mathrm{E}^{gw}[\sigma_{nt}|w_{it} \ge q_t]dZ_t}_{dr_{within}} \\ &+ \underbrace{\frac{\zeta_t(q_t) - 1}{2}\left(\mathrm{E}^{gw}[\nu_{nt}^2|w_{it} = q_t] + \mathrm{Var}^{gw}[\sigma_{nt}|w_{it} = q_t]\right) dt}_{dr_{displacement}}\end{aligned} \tag{13}$$

where E^{gw} *denotes the wealth-weighted cross-sectional average along the wealth distribution, and* Var^{gw} *denotes the wealth-weighted cross-sectional variance along the wealth distribution.*

The within term is the sum of a deterministic term, the wealth-weighted average wealth growth of households in the top percentile, and a stochastic term. The exposure of the top wealth share to aggregate risk is given by the wealth-weighted exposure of households in the top percentile.

The displacement term is the sum of two terms. The first term is due to the average idiosyncratic variance at the lower percentile threshold. The second term is due to the variance of risk exposures for households at the lower percentile threshold. Heterogeneous exposures to aggregate risks is another component of displacement. The displacement term can still be interpreted as the cross-sectional variance of the wealth growth of individuals at the wealth threshold q_t.

Jumps. The preceding analysis assumed that household wealth followed a diffusion process. This implied that the wealth process of households at the top was continuous. In reality, we observe that some entrepreneurs appear to reach top percentiles almost instantaneously. These large jumps in wealth may come from jumps in asset valuations, as in Aït-Sahalia et al. (2009).

Formally, I assume that normalized wealth is the sum of a Brownian Motion and a Compound Poisson-Process, i.e.:

$$\frac{dw_{it}}{w_{it-}} = \mu_t dt + \nu_t dB_{it} + (e^{J_{it}} - 1)dN_{it} \tag{14}$$

where N_{it} is an idiosyncratic jump process with intensity λ_t. The innovations J_{it} are drawn from an exogenous distribution such that $E^f[e^{J_{it}}] = 1$, where E^f denotes the expectation with respect to the jump density f_t.[12]

Define $\phi_t(\theta)$ the instantaneous cumulant generating function of log wealth, i.e. $\phi_t(\theta) = E_t[dw_{it}^{\theta}/w_{it-}^{\theta}]$. Define κ_{jt} the instantaneous cumulant of log wealth, i.e. $\kappa_{jt} = \phi_t^{(j)}(0)$. Applying Ito's lemma on (14), one obtains $\kappa_{2t} = \nu_t^2 + \lambda_t E^f[J^2]$ and $\kappa_{jt} = \lambda_t E^f[J^j]$ for $j > 2$.

I focus here on the case in which the wealth distribution has a Pareto tail at time t with tail index ζ.[13]

Proposition 5 (Dynamics of Top Wealth Share with Jumps)**.** *Suppose that the wealth distribution is Pareto with tail index ζ, and that the law of motion of normalized wealth w_{it} is given by* (14)*. Then the top wealth share S_t follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt}_{dr_{within}} + \underbrace{\sum_{j=2}^{+\infty} \frac{\zeta^{j-1} - 1}{j!} \kappa_{jt} dt}_{dr_{displacement}} \tag{15}$$

The displacement term depends on all higher-order cumulants of wealth growth. It can be written as

$$\begin{aligned} dr_{displacement} = & \frac{\zeta - 1}{2}\kappa_{2t} dt + \frac{\zeta^2 - 1}{6}\kappa_{2t}^{3/2} \cdot \text{skewness}_t \cdot dt + \frac{\zeta^3 - 1}{24}\kappa_{2t}^2 \cdot \text{excess kurtosis}_t \cdot dt \\ & + \text{higher-order cumulants} \ldots \end{aligned} \tag{16}$$

When the process for wealth follows a diffusion, all cumulants for $j \geq 3$ are equal to zero, and the displacement term only depends on the idiosyncratic variance, as in the baseline model. A negative skewness tends to decrease the displacement term while a positive kurtosis tends to increase the displacement term.

As the tail index of the wealth distribution ζ increases, the importance of higher-order cumulants increases relative to the variance component. To take a simple example, going from a distribution

[12] The assumption that jumps average to one could be easily relaxed, at the cost of notational complexity.

[13] In contrast to the case of a diffusive process, the displacement term depends on the shape of the wealth distribution beyond the top percentile p. Lemma 3 derives the displacement term for an arbitrary wealth distribution.

with $\zeta = 1.5$ (the approximate tail index of the wealth distribution) to a distribution with $\zeta = 2.5$ (the approximate tail index of the labor income distribution), the term due to the variance of wealth shocks is multiplied by 4, the term due to skewness is multiplied by 6, and the term due to kurtosis is multiplied by 9. Intuitively, the lower the level of wealth inequality, the more entry and exit there is from households far from the percentile threshold (i.e. due to higher-order cumulants) rather than from households close to the percentile threshold (i.e. due to the variance of wealth growth).

2.3 Demography

For simplicity, the preceding analysis assumed away any change in top wealth shares due to demography. In reality, households at the top die, which changes the composition of households at the top. Moreover, due to population growth, the total number of households in a given percentile also increases over time. I now augment the framework of Section 2.1 to account for these two demographic forces.

Households in the top percentile p die with a hazard rate δ_t that can vary over time. I model inheritance as follows. When a household in the top percentile dies, it is replaced by their offspring, who is born with a fraction $\chi_t \in [0, 1]$ of their initial wealth. Other newborns are born below the top percentile. The parameter χ_t controls the extent to which top fortunes are able to maintain themselves. When $\chi_t = 100\%$ (i.e. perfect inheritance), households that die are directly replaced by their offspring: death has no impact on top wealth shares. At the other end of the spectrum, when $\chi_t = 0\%$ (i.e. no inheritance) there is no transmission of wealth across generations. Economically, the fraction of wealth that cannot be passed to offspring, $1 - \chi_t$, is related to the average estate tax. Finally, I assume that population grows with rate η_t, that can also vary over time.

I focus here on the case in which the wealth distribution is Pareto at time t.

Proposition 6 (Dynamics of Top Wealth Share with Death and Population Growth)**.** *Suppose that the wealth distribution is Pareto with tail index ζ, and that the instantaneous law of motion of normalized wealth w_{it} is given by (2), with death rate δ_t, inheritance parameter χ_t, and population growth η_t. Then the top wealth share S_t follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt}_{dr_{within}} + \underbrace{\frac{\zeta - 1}{2}\nu_t^2 dt}_{dr_{displacement}} + \underbrace{\overbrace{\frac{\chi_t^{\zeta} - 1}{\zeta}\delta_t dt}^{dr_{death}} + \overbrace{\left(1 - \frac{1}{\zeta}\right)\eta_t dt}^{dr_{pop.\ growth}}}_{dr_{demography}} \tag{17}$$

Due to death and population growth, a new "demography" term appears in the growth of top wealth share S_t, which is the sum of a term due to death and a term due to population growth.

The term due to death depends on only three parameters: ζ, the tail index of the wealth distribution, δ_t, the death rate of top households, and χ_t, the fraction of wealth that can be passed to offspring. The term due to population growth depends on only two parameters: ζ, the tail index of the wealth distribution, and η, the population growth rate.

Like the displacement term, the demography term increases in ζ. As ζ decreases (i.e. as wealth inequality decreases), the ratio between the wealth of households at the lower percentile threshold and the average wealth of households in the top percentile decreases. Therefore both the death term and the population growth term decrease.

The term due to death is always negative. It increases with the degree of inheritance χ_t. The term due to population growth is always positive.[14] As population grows, a given top percentile p includes more and more households, which increases the top wealth share S_t. Overall, the demography term has an ambiguous sign.

Quantitatively, we can expect the demography term to be small. To take realistic parameters, for a distribution with a tail index $\zeta \approx 1.5$, a death rate $\delta_t \approx 1.5\%$, a fraction of wealth passed to offspring $\chi_t \approx 60\%$,[15] and a population growth rate $\eta_t \approx 1\%$, one obtains that the demography increases the growth rate of top wealth shares by $dr_{\text{demography}} \approx -0.2\%$ per year.

3 Accounting Framework

A natural question to ask is whether the law of motion of top wealth shares presented in the previous section can be mapped to the data. In this section, I present an accounting framework that does exactly this. I show how to decompose empirically the growth of top wealth shares into a within term, a displacement term, and a demography term using panel data.

Case without Death or Population Growth. I first present the accounting decomposition in the case without demographic forces, i.e. without death or population growth. This makes it easier

[14] Alternatively, one could define the within term as the wealth growth of households in the top percentile relative to the wealth growth of *existing households*, rather than the wealth growth of the economy. In this case, denoting θ_t the ratio between the wealth of a newborn household relative to the average wealth a household in the economy, the within term would equal $\mu_t dt + \theta_t \eta_t dt$ while the population growth term would equal $\left(1 - \frac{1}{\zeta} - \theta_t\right) \eta_t dt$. In any case, the displacement term remains unchanged.

[15] This corresponds to an average estate tax of 40%.

to understand the intuition behind the decomposition. Assume that the econometrician observes a representative sample of households in the economy at time t and at time $t+\tau$. The growth of the top wealth share S_t of a given top percentile p between t and $t+\tau$ is given by:

$$\frac{S_{t+\tau}-S_t}{S_t}=\frac{\sum_{i\in\mathcal{T}'} w_{it+\tau}}{\sum_{i\in\mathcal{T}} w_{it}}-1$$

where $\mathcal{T}$ denotes the set of households in the top percentile at time t and $\mathcal{T}'$ the set of households in the top percentile at time $t+\tau$. Denoting $\mathcal{X}$ the set of households that exit the top percentile between t and $t+\tau$, and $\mathcal{E}$ the set of households that enter the top percentile between t and $t+\tau$, we can write $\mathcal{T}'=(\mathcal{T}\cup\mathcal{E})\setminus\mathcal{X}$. The growth of the top wealth share S_t between t and $t+\tau$ can be decomposed into a term due to the average wealth growth of households at the top, and a term due to entry and exit:

$$\frac{S_{t+\tau}-S_t}{S_t}=\underbrace{\frac{\sum_{i\in\mathcal{T}} w_{it+\tau}}{\sum_{i\in\mathcal{T}} w_{it}}-1}_{R_{within}}+\underbrace{\frac{\sum_{i\in\mathcal{E}} w_{it+\tau}-\sum_{i\in\mathcal{X}} w_{it+\tau}}{\sum_{i\in\mathcal{T}} w_{it}}}_{R_{displacement}} \tag{18}$$

The first term ("within" term) is the wealth change for households in the top percentile at time t, whether or not they drop out of the top between t and $t+\tau$. The displacement term $R_{\text{displacement}}$ is the difference between the wealth of households that enter the top percentile between t and $t+\tau$ and the wealth of households that exit the top percentile between t and $t+\tau$.

To separate the role of entry and exit in the growth of the top wealth share S_t, it is useful to rewrite the displacement term as the sum of a term due to entry and a term due to exit:

$$R_{\text{displacement}}\equiv\frac{\sum_{i\in\mathcal{E}}(w_{it+\tau}-q_{t+\tau})}{\sum_{i\in\mathcal{T}} w_{it}}+\frac{\sum_{i\in\mathcal{X}}(q_{t+\tau}-w_{it+\tau})}{\sum_{i\in\mathcal{T}} w_{it}} \tag{19}$$

where $q_{t+\tau}$ is the wealth of the last household in the top at time $t+\tau$. Intuitively, when Mark Zuckerberg entered the Forbes 400 list in 2008, he displaced the last household in Forbes 400, that became the 401th wealthiest household. The net increase of Forbes 400 wealth share due to this entry is the difference between his wealth and the wealth of this last household. Conversely, when, say, Elizabeth Holmes dropped out of the Forbes 400 list in 2016, this caused the 401th wealthiest household to enter Forbes 400 list. The net increase in Forbes 400's wealth share due to this exit, relative to the within term, is the difference between the wealth of this last household and her new wealth.

Case with Death and Population Growth. I now extend this decomposition to account for demographic forces, i.e. death and population growth, which also generate entry and exit in the top percentile. Denoting $\mathcal{X}_\mathcal{D}$ the set of households in the top percentile that die between t and $t+\tau$, and $\mathcal{E}_\mathcal{D}$ the set of their offsprings that enter the top percentile after inheriting, we can write $\mathcal{T}' = (\mathcal{T} \cup \mathcal{E} \cup \mathcal{E}_\mathcal{D}) \setminus (\mathcal{X} \cup \mathcal{X}_\mathcal{D})$.[16]

Proposition 7 (Accounting Decomposition)**.** *The growth of the top wealth share S_t between t and $t+\tau$ can be decomposed as follows:*

$$\frac{S_{t+\tau} - S_t}{S_t} = R_{within} + R_{displacement} + R_{demography} \tag{20}$$

where the within term R_{within} is defined as

$$R_{within} = \frac{\sum_{i \in \mathcal{T} \setminus \mathcal{X}_\mathcal{D}} w_{it+\tau}}{\sum_{i \in \mathcal{T} \setminus \mathcal{X}_\mathcal{D}} w_{it}} - 1 \tag{21}$$

the displacement term $R_{displacement}$ is defined as

$$R_{displacement} \equiv \frac{\sum_{i \in \mathcal{E}} (w_{it+\tau} - q_{t+\tau})}{\sum_{i \in \mathcal{T}} w_{it}} + \frac{\sum_{i \in \mathcal{X}} (q_{t+\tau} - w_{it+\tau})}{\sum_{i \in \mathcal{T}} w_{it}} \tag{22}$$

and the demography term $R_{demography}$ is defined as[17]

$$R_{demography} \equiv \underbrace{\frac{\sum_{i \in \mathcal{E}_\mathcal{D}} w_{it+\tau} + (|\mathcal{X}_\mathcal{D}| - |\mathcal{E}_\mathcal{D}|)\, q_{t+\tau} - \sum_{i \in \mathcal{X}_\mathcal{D}} (1 + R_{within}) w_{it}}{\sum_{i \in \mathcal{T}} w_{it}}}_{R_{death}} + \underbrace{\frac{(|\mathcal{T}'| - |\mathcal{T}|) q_{t+\tau}}{\sum_{i \in \mathcal{T}} w_{it}}}_{R_{pop.\ growth}} \tag{23}$$

The first term ("within" term) is the average wealth growth of households in the top at time t, that do not die between t and $t+\tau$. The new demography term $R_{\text{demography}}$ is the sum of a term due to death R_{death} and a term due to population growth $R_{\text{pop. growth}}$. The term due to death is the difference between the wealth of the households that replace deceased households in the top percentile (the wealth of their offspring, or, in absence of offspring, the wealth of the last household in the top percentile) and the wealth of deceased households.[18] The term due to population growth is the wealth of the last household in the top percentile times the number of households that enter the top percentile due to population growth.

As shown in Appendix B,[19] when the wealth of households in the top percentile follows the law of motion (2), the decomposition identifies theoretical decomposition presented in (17) as the time period τ tends to zero.

[16] Here, $\mathcal{X}$ denotes the set of households that exit the top percentile for reasons other than death, and $\mathcal{E}$ denotes the set of households that enter the top percentile for reasons other than inheritance.

[17] For a set Ω, $|\Omega|$ denotes the number of its elements

[18] The offspring can refer to one or multiple children.

[19] See the proof of Proposition 7.

4 Empirical Analysis using Forbes 400

In this section, I apply the accounting framework presented in the previous section to decompose the growth of the wealth share of Forbes 400. I present the Forbes 400 data in Section 4.1. I discuss the results of the decomposition in Section 4.2. Finally, in Section 4.3, I examine the robustness of the decomposition with respect to measurement error.

4.1 Data

I focus on the list of the wealthiest 400 Americans constructed by Forbes Magazine every year since 1983. The list is created by a dedicated staff of the magazine, based on a mix of public and private information.[20] Because Forbes nominatively identifies the 400 wealthiest individuals in the U.S, one can track the wealth of the same individuals over time,[21] which is key to measure displacement. By contrast, other data sources used to track the level of wealth inequality in the U.S. rely on repeated cross-sections.[22] Using data from Forbes and Execucomp, I also match individual to the firms they own. Firm-level stock returns are obtained through CRSP.

I focus on the wealth share owned by a percentile that includes the richest 400 U.S. households in 2017.[23] To obtain the wealth share of this percentile, I divide the total wealth of households as reported by Forbes 400 by the aggregate wealth of U.S. households from the Financial Accounts (Flow of Funds). While this top percentile accounts for a small percentage of the total U.S. population, it accounts for a substantial share of total U.S. wealth (almost 3% in 2017).

Figure 2 plots the cumulative growth of the share of wealth owned by this top percentile since 1983, as well as the cumulative growth of the wealth share of the top 0.01%, 0.01%, 1%, and 10% from Saez and Zucman (2016). Most of the increase of top wealth inequality during the period is concentrated in the top 0.01%. Moreover, the rise in Forbes 400 wealth share tracks very well the

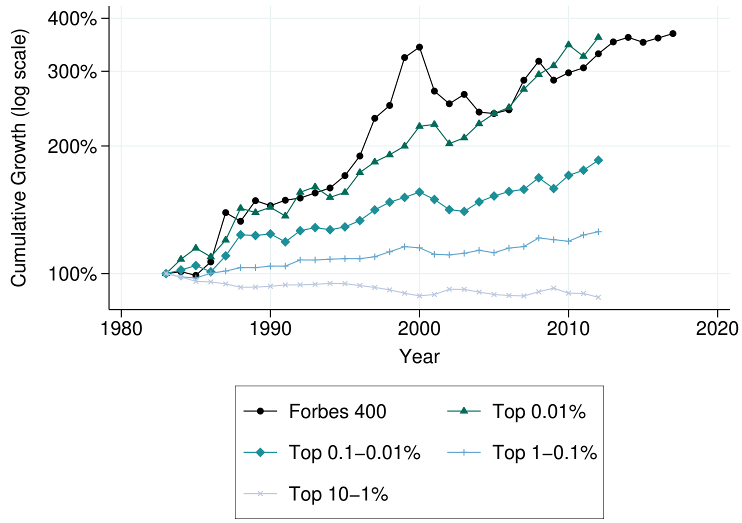
[20] Forbes Magazine reports: "We pored over hundreds of Securities Exchange Commission documents, court records, probate records, federal financial disclosures and Web and print stories. We took into account all assets: stakes in public and private companies, real estate, art, yachts, planes, ranches, vineyards, jewelry, car collections and more. We also factored in debt. Of course, we don't pretend to know what is listed on each billionaire's private balance sheet, although some candidates do provide paperwork to that effect."

[21] I extend the construction from Capehart (2014) for the last five years. In Appendix C.1, I describe how I obtain the wealth of individuals that exit the top percentile.

[22] The three main datasets on the wealth distribution in the U.S. are the Survey of Consumer Finances, Estate Tax Returns (see Kopczuk and Saez (2004)) and Income Tax Returns (see Saez and Zucman (2016)), which all correspond to repeated cross-sections.

[23] It corresponds to approximatively 0.0003% of U.S. population. Due to population growth, it includes 264 households in 1983. Data on household population is from the U.S. Census Bureau.

Figure 2: Cumulative Growth of Wealth Share Top 0.01% Tracks Forbes 400



Notes. The figure plots the cumulated growth of top wealth shares for groups defined in the top Forbes percentile, which includes 400 households in 2017. Data for the top 10%, 1%, 0.1%, 0.01% is from Saez and Zucman (2016).

rise in the wealth share of the top 0.01%. This suggests that understanding the wealth growth of Forbes 400 can shed light on the overall rise in top wealth inequality during this period.

4.2 Results

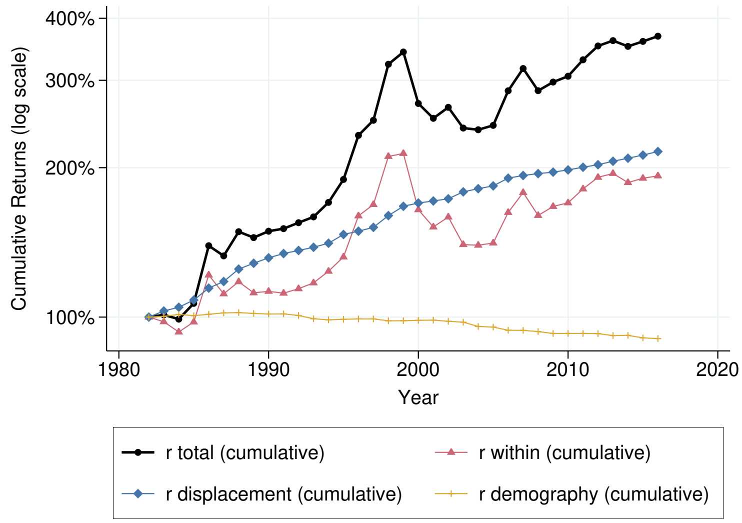
Fact 1: Displacement accounts for half of the rise in top wealth inequality. Table 1 reports the result of the accounting decomposition. The first line reports each term geometrically averaged over the entire time period. I find that the displacement term is responsible for half of the increase of the top wealth share. More precisely, the 3.9% yearly growth of the top wealth share can be decomposed into a within term equal to 1.9%, a displacement term equal to 2.3%, and a demography term equal to -0.3%.

Figure 3 plots the cumulative sum of the terms since 1983. Business-cycle fluctuations in top shares are driven by fluctuations in the within term, rather than fluctuations in the displacement or the demography term. This is not surprising: as seen in the theoretical section above, when top households are particularly exposed to aggregate risks, the instantaneous variance of the top wealth share is entirely driven by the within term.[24]

I examine the displacement term through the lens of the theoretical framework laid out in

[24]See Equation (13).

Figure 3: Decomposing the Cumulative Growth of Forbes 400 Wealth Share



Notes. The figure plots the growth of the wealth share of the top percentile, as well as its accounting decomposition using Equation (18). It plots the cumulative log terms, i.e. the sum of log terms from 1983 to t. The plots for the within term, the displacement term, and the demography term approximately sum up to the total growth of the top wealth share. Data from Forbes 400.

Section 2.[25] When wealth follows a diffusion process (i.e. normal shocks), Section 2 predicts that the displacement term equals $1/2(\zeta - 1)\nu^2$ where ζ is the tail index of the wealth distribution and ν^2 is the idiosyncratic variance of wealth growth. To compare the prediction of this model with the actual displacement term, I estimate the tail index of the wealth distribution ζ as well as the standard deviation of wealth growth ν in Table 2. I obtain a model-predicted displacement term equal to 2.0%, which comes from an average ζ equal to 1.5 and an average ν equal to 27%. This is very close to the actual displacement term, which averages 2.3%.

What is the role of higher-order cumulants for displacement? When wealth follows a jump-diffusion process (i.e. non-normal shocks), Section 2 shows that the displacement term equals $\sum_2^{+\infty} \frac{\zeta^{j-1}-1}{j!}\kappa_{jt}$ where κ_{jt} denotes the $j-$th cumulant of wealth growth.[26] To examine whether the wealth growth of top households displays non-normality, I estimate the skewness and kurtosis of wealth growth in Table 2. The average skewness is negative around -0.3 (i.e. more downward realizations compared to the log-normal distribution), while the average excess kurtosis is positive around 5 (i.e. more extreme realizations compared to the log-normal distribution). Combined with

[25]The dynamics of the within term and the demography term are relegated in Appendix C.

[26]While it allows for jumps, the model with jumps assumes that the distribution of wealth is exactly Pareto and that the law of motion of wealth is the same at any level of wealth.

a tail index around $\zeta \approx 1.5$, this implies that skewness decreases the displacement term by 0.2%, while kurtosis increases the displacement term by 0.3% annually (see Table 3). In other words, the effect of higher-order cumulants on the displacement term is small. This comes from the fact that ζ is close to one. Intuitively, wealth inequality is so high that most of the entry in the top percentile is driven by households already close to the percentile threshold, rather than entrepreneurs from the bottom of the distribution with extremely high wealth realization.[27]

To examine the effect of higher-order cumulants at yearly frequency, Figure 4 plots the actual displacement term, the displacement term predicted by the diffusion model, as well as the term predicted using by the jump-diffusion model. While the term predicted by the diffusion model tracks the actual displacement term very well, it misses the rise of the displacement term in 1986 and 1998, as well as the decline of the displacement term during the burst of the tech bubble. Accounting for the skewness and kurtosis of wealth shocks is important to match these fluctuations.

I now examine the role of firm-level returns in driving the dispersion of wealth shocks for households in the top percentile. I regress the variance of household-level wealth growth on the equal weighted variance of firm-level returns in column (1) of Table 5. If households split their wealth in n uncorrelated firms, the idiosyncratic volatility of their wealth equals $\nu_{\text{stocks}}/\sqrt{n}$, where ν_{stocks} denotes the idiosyncratic volatility of firm-level returns. The estimate for the slope is 0.18, which can be interpreted as an average number of distinct firms owned by top households $n = 5$. The estimate for the intercept is close to zero, which means that the number n, identified purely from time-series variation, also accounts for the level of the idiosyncratic volatility of wealth. This suggests that the idiosyncratic volatility of wealth growth is almost entirely driven by the idiosyncratic volatility of firm-level returns.

Fact 2: Displacement has steadily declined over time. To examine low-frequency changes in the decomposition since 1983, Table 1 reports the terms averaged across three time periods of equal duration since 1983. Each time period covers an entire business cycle. The first period covers 1983-1993, which includes the 1990-1991 recession. The second period covers 1994-2004, which includes the 2001 recession. The third period covers 2005-2016, which includes the 2007-2009 recession. I find that the displacement term has substantially decreased over time: it goes from 3.0% in the first part of the sample (1983-1993), to 2.5% in the second part of the sample (1994-2004), and finally to 1.4% in the third part of the sample (2005-2016). Table A1 in the

[27] Consistent with this result, Bessembinder (2018) stresses that lognormality implies a large skewness in the distribution of individual *level* stock returns.

Figure 4: Displacement Term: Data vs Theory

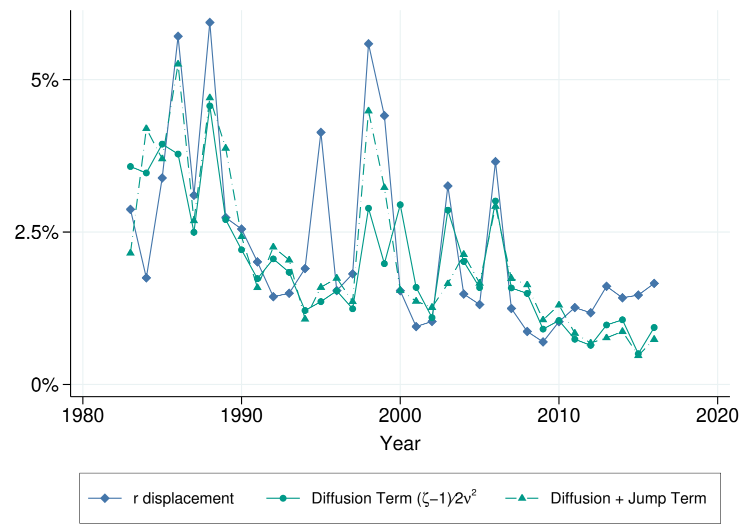


Figure 4 plots the displacement term (defined in Proposition 7) as well as the term predicted by the random-growth model for a diffusion model (normal shocks) and a jump diffusion (non-normal shocks). Following Equation (5), the local tail index ζ is given by $\zeta = 1 + g_t(q_t)q_t^2/S_t$, where the density $g_t(q_t)$ is estimated from the mass of households with a wealth between q_t and $1.3q_t$. The idiosyncratic variance of wealth growth is estimated using the corresponding sample moments of the log wealth growth among households in the top in a given year. The term with all higher-order cumulants is computed as $\log \mathrm{E}[\tilde{R}^{\zeta}]^{\frac{1}{\zeta}}$ where $\tilde{R}$ denotes the normalized wealth growth of households at the top. Data from Forbes 400.

appendix formally regresses the terms obtained in the accounting decomposition on year trends, showing that the decrease of the displacement term over time is statistically significant.

What explains the decline of the displacement term over time? To answer this question, I use the displacement term predicted by the diffusion model $1/2(\zeta - 1)\nu^2$ to decompose the decline of the displacement term into a decline in the idiosyncratic volatility of wealth shocks ν and a decline in the shape of the wealth distribution ζ in Figure 4. Half of the decrease of the displacement term is due to the decrease of the dispersion of wealth shocks from $\nu_{1980s} \approx 28\%$ to $\nu_{2010s} \approx 23\%$, which follows a similar decline in the cross-sectional variance of firm-level returns. The slow-down of the displacement term in the last two decades is therefore related to the general decline in the pace of business dynamism. As documented by Decker et al. (2016a), much of the decline occurs within industry, firm-size and firm-age categories.

The remaining half is due to the decrease of the tail index from $\zeta_{1980s} \approx 1.8$ to $\zeta_{2010s} \approx 1.4$. Intuitively, following the rapid rise in idiosyncratic volatility at the end of the 20th century, wealth inequality increased so much that households with high wealth shocks now have a harder time entering the top.[28]

Fact 3: Technological Innovation Drives Displacement Households that enter the top percentile innovate more than the households that they displace. To prove this, I regress a measure of firm innovation on a dummy that is equal to one if the household enters the top during the year, and zero if the household is already at the top in Table 6. As a proxy for the innovation of each firm in a given year, Column (1) uses the number of patents issued during the year, Column (2) uses the number of their citations, and Column (3) uses their value using Kogan et al. (2017). I find that, compared to the households already at the top, households that enter the top percentile in a given year tend to own firms that file twice the number of patents, with three times the total number of citations, and with twice the economic value. To compare households within the same year and industry, regressions are done with year and industry fixed effects.

I then study the relationship between displacement and aggregate innovation over time. I proxy for aggregate innovation using the economic value of patents issued during the year. This measure is constructed by Kogan et al. (2017) by aggregating the value of all patents every year, normalized by the total market capitalization in the economy. In Column (1) of Table 7, I regress the variance of the log wealth growth of households in the top on aggregate patent activity. The estimate for

[28] Formally, the tail index of the stationary wealth distribution decreases with the idiosyncratic volatility of wealth, as shown in Equation (7). Appendix D explores the dynamics of the tail index after changes in idiosyncratic volatility.

Figure 5: Contribution of ζ and ν to Displacement

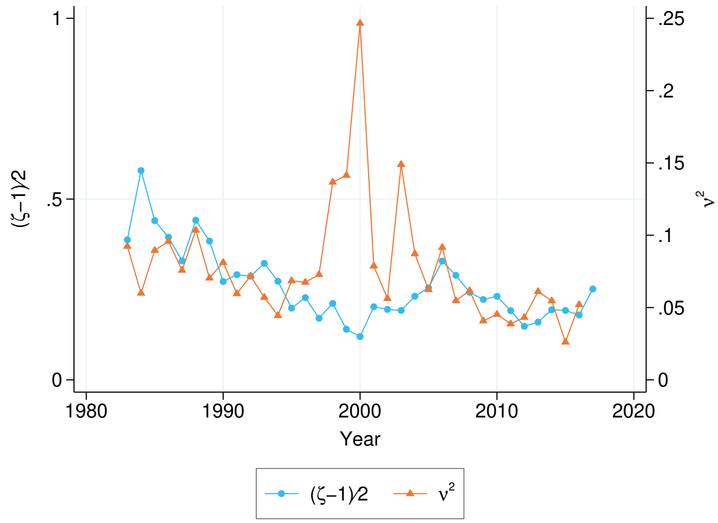


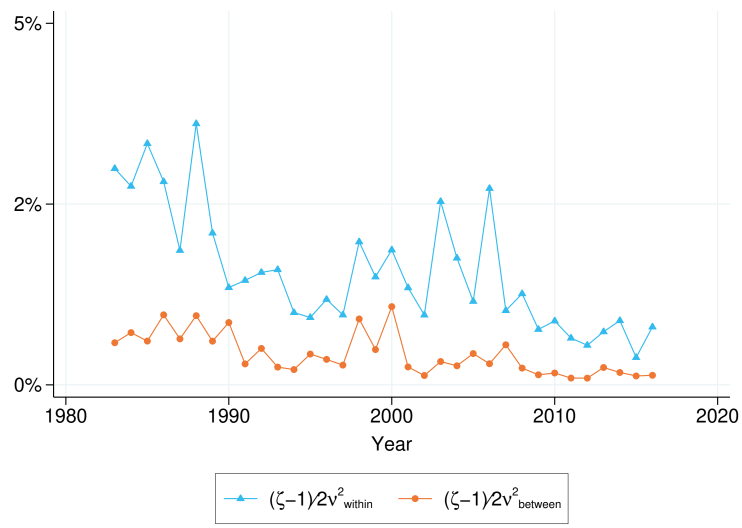
Figure 5 decomposes the model-predicted term into its two components, $(\zeta-1)/2$ and ν^2. Following Equation (4), the local tail index of the wealth distribution ζ is given by $\zeta = 1 + g_t(q_t)q_t^2/S_t$, where the density $g_t(q_t)$ is estimated from the mass of households with a wealth between q_t and $1.3q_t$. The idiosyncratic variance of wealth growth ν^2 is estimated using the corresponding sample moments of the log wealth growth among households in the top in a given year. The product of the term equals the model-predicted term with normal shocks $1/2(\zeta-1)\nu^2$. Data from Forbes 400.

the slope is positive and strongly significant, with a R^2 equal to 36%. In Column (2) of Table 7, I replace the variance of log wealth growth by the displacement term divided by $(\zeta-1)/2$. This alternative measure potentially reflects the effect of innovation on displacement through higher-order cumulants. The coefficient increases to 0.12. Overall, this suggests that a 10% increase of patent innovation increases the growth of top wealth shares due to displacement by 0.3 percentage points ($= 1/2(1.5-1) \times 0.12 \times 0.1$).

The effect of innovation on displacement weakens when using rougher proxies for the dispersion of wealth shocks. Column (3) regresses directly the displacement term on aggregate patent activity. The estimate is only significant at the 10% level. This is because regressing the displacement term on innovation is misspecified, due to low-frequency changes in ζ.[29] In Column (4), I regress directly the growth of top wealth share on aggregate patent activity. The coefficient is not significant. A researcher that simply regresses the growth on top wealth shares on innovation would not find any relation between inequality and innovation. This is because the within term, which is very volatile masks the relationship between the displacement term and innovation.

[29] Since innovation is serially correlated, high innovation today is correlated with high innovation in previous years, and therefore with a low tail index of the wealth distribution ζ. See Appendix D.

Figure 6: Displacement Within and Between Industries



Notes. The table decomposes the model-predicted displacement term $1/2(\zeta - 1)\nu^2$ into a displacement "within" industries $1/2(\zeta - 1)\nu^2_{\text{within}}$ and a displacement "between" industries $1/2(\zeta - 1)\nu^2_{\text{between}}$. The decomposition follows from the law of total variance: the variance of wealth growth ν^2 is the sum of the average variance of wealth growth within industry groups ν^2_{within} and the variance of wealth growth between groups ν^2_{between}. Industries are defined using the Fama-French 49 industry classification. Data from Forbes 400.

How important is the rise and fall of certain industries (i.e. software v.s. oil) for the dynamics of top wealth shares? To answer this question, I use the displacement term predicted by the diffusion model $1/2(\zeta - 1)\nu^2$ to decompose the displacement term into a displacement within industries and a displacement between industries. This decomposition uses the fact that the cross-sectional variance of wealth shocks can always be decomposed into the average variance within industry and the variance of average wealth growth between industries.[30] Table 4 reports that the displacement term within industries averages to 1.6% whereas the displacement term between industries averages to 0.4%. In other words, displacement within industries is much more important than displacement term between industries.[31] Figure 6 plots the two terms over time: the only time when the displacement between industries is quantitatively important is the height of the dot-com bubble.[32]

[30]This decomposition mirrors the theoretical decomposition in Equation (13).

[31]This finding is consistent with Campbell et al. (2001), who find that the variance of firm-level returns within industries is much higher than the variance across industry portfolio returns.

[32]In Appendix C.3, I use a similar method to decompose the displacement term into the variance within families and the variance between families. The variance within families is negligible compared to the variance between families.

4.3 Measurement Error

The wealth of individuals at the top is inevitably measured with errors. I conclude this section by assessing the effect of measurement error on the displacement term, as measured in the accounting decomposition Proposition 7.

The first concern is that Forbes may systematically underestimate or overestimate the wealth of top 400 households. Along these lines, Atkinson (2008) argues the magazine may give inflated values of the wealth of top households, because debts are harder to track than assets. Empirically, Raub et al. (2010) document that the wealth of deceased households reported for on estate tax returns is approximately half of the wealth estimated by Forbes. However, this measurement error in level does not impact the growth of top wealth shares.

A related concern is that Forbes measures the wealth of top households with noise. If the measurement error is completely persistent, as noted in Luttmer (2002), this leads Forbes to overestimate the level of top wealth shares, without affecting the growth of top wealth shares, nor the accounting decomposition. If, however, the measurement error is non completely persistent, it may generate artificial entry and exit in the top percentile. While this does not change the growth of top wealth shares, this leads the econometrician to underestimate the within term and to overestimate the displacement term.

I deal with this potential bias in three ways. First, Forbes usually report the reasons households drop off the list. Less than 5% of these exits are due to the fact that the previously reported wealth was inflated.[33] I simply remove these households from the sample. Second, I estimate the importance of transitory measurement errors in the remaining sample. Table 8 reports that the autocorrelation of wealth growth at the individual level is close to zero, which suggests that there is little mean-reversion in wealth growth. Formally, I show in Appendix C.2 that the relative bias in the displacement term is well approximated by -2ρ, where ρ is the AR(1) coefficient of wealth growth. With an estimated $\rho \approx 0.01$, this suggests that transitory measurement error accounts for only 4 basis points in the displacement term. Third, if measurement error was important, we would expect the regression of the variance of household wealth growth on the variance of firm-level returns to have a large positive intercept. As shown in Table 5, the intercept is fairly small, suggesting that measurement error does not play a significant role in driving the dispersion of wealth growth.

A final concern is that Forbes 400 coverage may become more and more precise over time, and therefore, that the magazine gradually discovers rich households that were not reported earlier

[33]This includes in particular Donald Trump.

(Piketty (2014)). This would lead the econometrician to overestimate the displacement term as well as the growth of top wealth shares. If this were an important driver of top wealth shares, the observed displacement term would be higher than the term predicted by the dispersion of wealth among existing households. This is not the case, as seen in Table 3.

5 Displacement Along the Wealth Distribution

Measuring the displacement term as the wealth of households entering the top minus the wealth of households exiting the top requires panel data. However, most of the data on wealth inequality beyond Forbes 400 is based on repeated cross-sections.[34] In this case, however, the empirical results of the previous section suggests that the term predicted by the diffusion model approximates well the displacement t.

Methodology. In this section, I proxy for the displacement term for the top $1\%, 0.1\%$, and 0.01% from 1916 to 2012 following Equation (9), i.e. as $1/2(\zeta(q_t) - 1)\nu^2(q_t)$ where $\zeta(q_t)$ denotes the local tail index of the wealth distribution around the percentile threshold q_t and $\nu^2(q_t)$ denotes the idiosyncratic volatility of households at the percentile threshold q_t.

I first estimate the local tail index of the wealth distribution $\zeta(q_t)$ at top percentiles $1\%, 0.1\%$, and 0.01% using data on wealth thresholds, and top wealth shares from Kopczuk and Saez (2004) for 1916-1962, and Saez and Zucman (2016) for 1962-2012.[35] Table 9 reports the estimated $\zeta(q_t)$ for top percentiles. Over the time period, the local tail index ζ equals 1.5 for the Top 1% and 1.7 for the Top 0.01%. The estimate ζ does not change much across top percentiles, which reflects the fact that the wealth distribution is close to Pareto.

I estimate the idiosyncratic volatility at each percentile by interacting the share of wealth invested in equity, using data from Kopczuk and Saez (2004) for 1916-1962, and Saez and Zucman (2016) for 1962-2012, with the cross sectional standard deviation of firm-level returns, using data from CRSP. I scale this product so that the idiosyncratic volatility of the top 0.01% matches the idiosyncratic volatility of Forbes 400 in 1983-2012. Table 9 reports the estimated ν for top percentiles. Over the time period, ν equals 14% for the Top 1%, and 21% for the top 0.01%. The fact that ν increases in the right tail of the distribution reflects the fact that top percentiles tend to invest more in equity.

[34] See Footnote 22.

[35] I estimate the density around a top percentile from the difference in wealth threshold in the neighborhood of the percentile.

Results. Figure 7 plots the model-predicted displacement term $1/2(\zeta - 1)\nu^2$ for the top 1%, the top 0.1%, and the top 0.01% from 1916 to 2012. The displacement term roughly follows a U-shape for all top percentiles. The displacement term for the top 0.01% peaked at 2% during the Great Depression, then steadily decreased, reaching its minimum in 1945. The displacement term again increased starting in 1960, and reached its maximum at the height of the dot-com bubble. Overall, the displacement term was roughly twice as high in 1983-2012 as it had been for the rest of the century.

To understand better what drives the displacement term over time, Figure 8 plots separately the term due to the wealth distribution $1/2(\zeta - 1)$ and the term due to the idiosyncratic variance of wealth ν^2 for the top 0.01%. Most of the fluctuations in the model-predicted displacement term arise from fluctuations in the idiosyncratic variance of wealth rather than from fluctuations in the tail index of the wealth distribution. This is because the right tail of the distribution tends to move slowly, as shown in Gabaix et al. (2016).

According to Saez and Zucman (2016), the yearly growth rate of the wealth share of the top 0.01% in 1982-2012 averaged to 4.3%, while the yearly growth rate of the top 1% averaged to 1.9%, i.e. a difference of 2.4% per year. The results of Table 9 suggest that the differences in displacement between the two percentiles can explain almost half of this gap.[36]

6 Wealth Mobility

How does a rise in wealth inequality impacts wealth mobility? In this section, I show that whether a rise in wealth inequality is driven by a rise in the average wealth growth of households at the top (within term) or a rise in the dispersion of wealth shocks (displacement term) has opposite effects on mobility.

While a rise in the wealth growth of households at the top unambiguously decreases wealth mobility, the effect of a rise in the dispersion of wealth shocks on mobility is ambiguous. On the one hand, the higher the dispersion of wealth shocks of households at the top, the more likely it is for their wealth to decrease, which tends to increase mobility. On the other, the higher the dispersion of wealth shocks, the more unequal the wealth distribution in the long run, and, therefore, the higher the typical distance between a household in the top percentile and the lower percentile threshold.

To examine the overall effect of an increase in the dispersion of wealth shocks on mobility, I focus on the average time a household in the top percentile remains in the top. The advantage of

[36] A more detailed discussion is given in Appendix D.

Figure 7: Model-Predicted Displacement Term

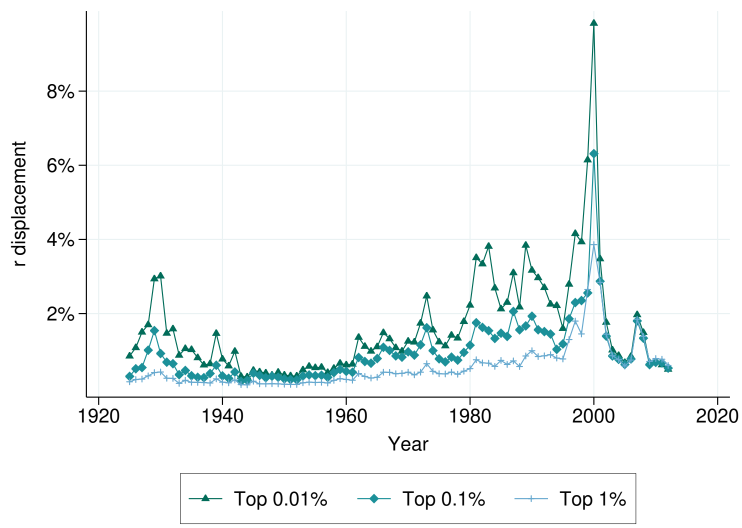


Figure 7 plots the model-predicted displacement term $(\zeta - 1)/2\nu^2$ for the Top 0.01%, 0.1%, and 1%. The local tail index ζ at each percentile p is estimated as $1/(1 - \frac{q_t}{S_t/p})$. The idiosyncratic volatility of wealth ν is estimated by interacting the share of wealth invested in equity at each percentile with half of the idiosyncratic volatility of firm-level returns. Data from Kopczuk and Saez (2004) and Saez and Zucman (2016).

this notion of "downward" mobility is that it only depends on the wealth dynamics of individuals in the right tail of the distribution.[37] Formally, for a household with wealth w, denote $T_q(w)$ the average time the household remains above the wealth threshold q (also called the "average first passage time"), i.e.

$$T_q(w) \equiv \mathrm{E}[\inf\{\tau \text{ s.t. } w_{it+\tau} \leq q \text{ or } i \text{ dies }\}|w_{it} = w] \tag{24}$$

For the remainder of this section, I assume that the law of motion of wealth is given by

$$\frac{dw_{it}}{w_{it}} = \mu dt + \nu dB_{it} \tag{25}$$

with death rate $\delta > 0$. Having a positive death rate ensures that the average first passage time is always finite.

Lemma 1 (Average First Passage Time)**.** *When wealth follows the law of motion* (25)*, the average first passage time for* $w \geq q$ *is:*[38]

[37] In particular, compared to a notion of "upward" mobility, it allows me to abstract from the role of labor income or government programs.

[38] The average first passage time of a Brownian Motion is a classic result, for instance see Karlin and Taylor (1981). This formula simply generalizes it to the case of a process with Brownian Motion with death probability.

Figure 8: Contribution of ζ and ν to Model-Predicted Displacement

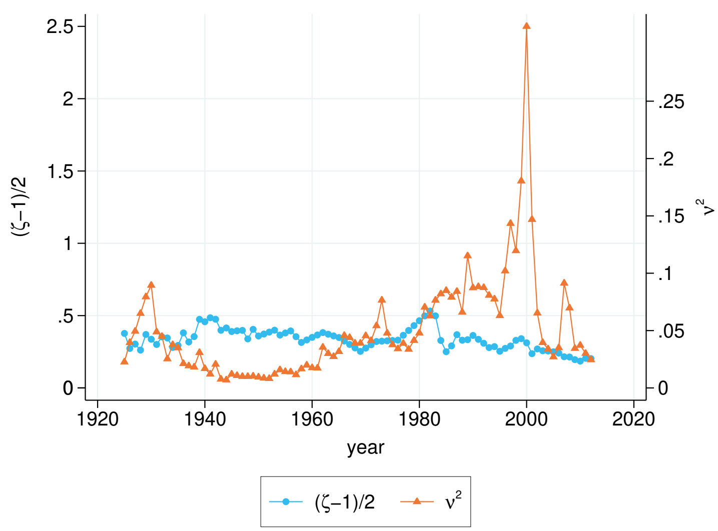


Figure 8 decomposes the model-predicted displacement term into its two components, $(\zeta - 1)/2$ and ν^2. Following (4), the local tail index of the wealth distribution ζ is given by $\zeta = 1 + g_t(q_t)q_t^2/S_t$, where the density $g_t(q_t)$ is estimated from the mass of households with a wealth between q_t and $1.3q_t$. The idiosyncratic volatility at each percentile ν is estimated by interacting the share of wealth invested in equity at each percentile with the idiosyncratic volatility of firm-level returns. The product of the term equals the model-predicted term with normal shocks $^1/_2(\zeta - 1)\nu^2$. Data from Kopczuk and Saez (2004) and Saez and Zucman (2016).

1. If the death rate is $\delta = 0$,[39]

$$T_q(w) = \frac{1}{-\mu + \frac{1}{2}\nu^2} \log \frac{w}{q} \tag{26}$$

2. If the death rate is $\delta > 0$

$$T_q(w) = \frac{1}{\delta}\left(1 - \left(\frac{w}{q}\right)^{\zeta_-}\right) \tag{27}$$

where ζ_- is the negative zero of $\zeta \to \mu\zeta + \frac{\zeta(\zeta-1)}{2}\nu^2 - \delta$.

This lemma gives a closed-form formula for the average time a household with initial wealth w remains above a wealth threshold q. Naturally, the first passage time increases in w/q. As the household wealth w converges to q, this time converges to zero. As w converges to infinity, this time converges to $1/\delta$. The first passage time is a power law in w/q. The exponent ζ_- captures how fast the first passage time increases as the household wealth increases.

The average first passage time $T_q(w)$ increases in the average wealth growth of individuals μ but decreases in the idiosyncratic volatility ν.[40] Intuitively, the higher the dispersion of wealth shocks, the more likely it is to have a negative wealth shock, and therefore the more likely it is for the wealth of an household to drop below q.

While an increase in idiosyncratic volatility decreases the average first passage time at a given wealth level, it also increases in the long run the typical distance between individuals. To determine the overall effect of idiosyncratic volatility on mobility, one needs to take this long run adjustment into account. Instead of considering the average first passage time for a household with given wealth level, I examine the average first passage time for an average household in a top percentile p, denoted $T(p)$. Formally,

$$T(p) \equiv \mathrm{E}^g[T_q(w_{it})|w_{it} \geq q] \tag{28}$$

where q denotes the wealth at the lower threshold of the top percentile p and E^g denotes the cross-sectional average with respect to the wealth density g.

Proposition 8 (Average First Passage Time for an Average Household)**.** *Consider the stationary distribution in an economy where household wealth follows the process given in (14) with death rate δ, inheritance parameter χ, and population growth η. Then, the average time someone in the top percentile p remains in the top percentile is:*

[39]This is assuming that $\mu - \nu^2/2 = \frac{\mathrm{E}[d\log(w)]}{dt} > 0$. Otherwise, the average passage time is infinite.

[40]See the proof in Appendix E.

1. If the death rate $\delta = 0$

$$T(p) = \frac{1}{-\mu + \frac{1}{2}\nu^2}\frac{1}{\zeta_+} \tag{29}$$

2. If the death rate $\delta > 0$

$$T(p) = \frac{1}{\delta}\frac{\zeta_-}{\zeta_- - \zeta_+} \tag{30}$$

where ζ_+ denotes the Pareto tail of the stationary wealth distribution[41] *and ζ_- is defined in Lemma 1.*

This formula characterizes in closed-form the average passage time for a household in the top percentile p. Strikingly, the average first passage time of an average household in the top percentile p, $T(p)$, does not depend on the top percentile p.

The average first passage time depends on the ratio between ζ_+ and ζ_-. Intuitively, $-\zeta_-$ controls the average first passage time from a given distance to the threshold, while ζ_+ corresponds to the tail index of the right tail of the stationary wealth distribution. Both statistics matter to determine the first passage time for an average household in the top percentile.

As the average wealth growth of top households μ increases, T increases (i.e. mobility decreases). This is due to two reasons. First, the average first passage time at a given wealth level increases ($-\zeta_-$ increases). Second, in the long run, the wealth distribution becomes more unequal, which increases the typical distance between a household in the top percentile and the lower percentile threshold (ζ_+ decreases). These two forces combine to decrease mobility.

In contrast, as the idiosyncratic volatility of wealth ν increases, T tends to decrease (i.e. mobility increases). On the one hand, as ν increases, the average first passage time from a given wealth level decreases, which tends to increase mobility ($-\zeta_-$ decreases). On the other hand, in the long run, the wealth distribution becomes more unequal, which increases the typical distance between a household in the top percentile and the lower percentile threshold (i.e. ζ_+ decreases). For realistic parameters, this long-run effect on the wealth distribution is not strong enough to compensate for the first force. Overall, mobility increases.

This formula allows to relate changes in fundamental parameters to changes in mobility. First consider the pre-1980 economy, with normalized wealth growth of top households $\mu = 0\%$, idiosyncratic volatility $\nu = 10\%$, death rate $\delta = 2\%$, inheritance parameter $\chi \approx 50\%$, and population

[41] It can be defined as the positive zero of $\zeta \to \mu\zeta + \frac{\zeta(\zeta-1)}{2}\nu^2 + (\chi^\zeta - 1)\delta + (\zeta - 1)\eta$

growth rate 1.5%.[42] According to Proposition 8, the average time a top household remains in a top percentile is $T \approx 25$ years in this economy. Now, consider a change in parameters that corresponds to the rise in top wealth shares during the period. In particular, the normalized wealth growth of top households increases to $\mu = 2\%$ and the idiosyncratic volatility of wealth growth increases to $\nu = 27\%$, which follows the empirical evidence discussed above.[43] Applying Proposition 8, I obtain that the average time a top household remains at the top becomes $T \approx 20$ years. Even though wealth inequality increases between these two states, wealth mobility *increases*.[44]

7 Conclusion

This paper stresses the importance of composition effects on the dynamics of inequality. I document that half of the rise of the wealth share of the top 400 is driven by displacement, i.e. the entry and exit of households in top percentiles. This empirical result contradicts the "rich getting richer" hypothesis, which posits that the rise in top wealth shares is exclusively due to the average wealth growth of households in top percentiles. I show that the growth of top wealth shares due to displacement is well approximated $1/2(\zeta - 1)\nu^2$, where ζ denotes the tail index of the wealth distribution and ν denotes the idiosyncratic volatility of wealth. This formula is useful to understand the drivers displacement, as well as the role of displacement in setups where panel data is not available. Finally, I document a positive relationship between displacement and technological innovation. In particular, the slow-down of displacement in the last two decades seems to reflect the recent decline in business dynamism documented in Decker et al. (2016b).

The implications of my analysis extend beyond the literature on wealth inequality. Economic studies have recently documented rising concentrations in other areas, such as in the distribution of labor income (Piketty and Saez (2003)) or in the distribution of firms' market shares (Autor et al. (2017)). The tools developed here could easily be adapted to these other settings. In particular, the comovement of the dispersion of wealth shocks and the dispersion of firm-level returns suggests a deep link between the recent rise in wealth concentration and the recent rise in firm concentration. Understanding better this connection is an important direction for future research.

[42] I choose the death rate, inheritance parameter and population growth to match the demography term of Forbes 400 in 1983-2917 (see Table A4). I choose the idiosyncratic volatility to target the average idiosyncratic volatility for the top 0.01% in 1960-1980 from Section 5. Finally, I choose the drift μ so that the Pareto tail of the stationary wealth distribution is 1.8.

[43] The value of μ and ν are taken from the average accounting decomposition (see Table 1).

[44] This mechanism may be also explain the empirical findings of Kopczuk et al. (2010), which find that, even though labor inequality increased at the end of the 20th century, labor mobility remained constant.

Table 1: Decomposing the Growth of Top Wealth Share

Period	**Total** (%)	R_{within} (%)	$R_{\text{displacement}}$ (%)	$R_{\text{demography}}$ (%)
All Years	3.9	1.9	2.3	−0.3
1983-1993	4.3	1.5	3.0	−0.1
1994-2004	3.7	1.6	2.5	−0.3
2005-2016	3.7	2.7	1.4	−0.5

Notes. The table reports the geometric average of the growth of the wealth share of the top 0.0003% R_{total}, as well as the geometric average of the within term R_{within}, the displacement term $R_{displacement}$, and the demography term $R_{demography}$, as defined in Proposition 7. All terms in percentage. Data from Forbes 400.

Table 2: Tail Index and Cumulants of Wealth Growth

Period	Tail-Index (ζ)	Volatility (ν)	Skewness (sk)	Excess Kurtosis (kurt)
All Years	1.5	0.27	−0.35	4.70
1983-1993	1.8	0.28	−0.24	4.10
1994-2004	1.4	0.31	−0.37	4.90
2005-2016	1.4	0.23	−0.44	5.07

Notes. The table reports summary statistics on the tail index of the wealth distribution and higher-order cumulants of log wealth growth. The local tail index of the wealth distribution ζ is estimated yearly as $1 + g_t(q_t)q_t^2/S_t$, where the density $g_t(q_t)$ is estimated from the mass of households with a wealth between q_t and $1.3q_t$. The variance, skewness and kurtosis of wealth growth are estimated yearly using the corresponding sample moments of the log wealth growth among households in the top in a given year. Data from Forbes 400.

Table 3: Displacement Predicted by Diffusion-Jump Model

Year	$R_{\text{displacement}}$						
	Total (%)	Displacement Predicted by All Cumulants $\sum_2^{+\infty} \frac{\zeta^j-1}{j!}\kappa_j$ (%)					ϵ (%)
		Total	Variance Term	Skewness Term	Kurtosis Term	Higher-Order Term	
			$\frac{\zeta-1}{2}\nu^2$	$\frac{\zeta^2-1}{6}\nu^3 \cdot \text{sk}$	$\frac{\zeta^3-1}{24}\nu^4 \cdot \text{kurt}$	$\sum_5^{+\infty} \frac{\zeta^j-1}{j!}\kappa_j$	
All Years	2.3	2.1	2.0	−0.2	0.3	0.0	0.2
1983-1993	3.0	3.2	2.9	−0.2	0.5	0.0	−0.2
1994-2004	2.5	1.9	1.9	−0.3	0.4	0.0	0.6
2005-2016	1.4	1.2	1.2	−0.1	0.1	0.0	0.2

Notes. The table reports the geometric average of the displacement term, the average displacement term predicted by the diffusion-jump model, as well as their difference ϵ. The term predicted by diffusion-jump is split into different cumulants. The local tail index of the wealth distribution ζ is estimated yearly as $1 + g_t(q_t)q_t^2/S_t$, where the density $g_t(q_t)$ is estimated from the mass of households with a wealth between q_t and $1.3q_t$. The cumulants are estimated yearly using the cross-section of the log wealth growth of households in the top. Data from Forbes 400.

Table 4: Role of Industry Shocks for Displacement

Year	$R_{\text{displacement}}$				
	Total (%)	Displacement Predicted by Variance $1/2(\zeta-1)\nu^2$ (%)			ϵ (%)
		Total	Within Industries	Between Industries	
All Years	2.3	2.0	1.6	0.4	0.3
1983-1993	3.0	2.9	2.3	0.6	0.0
1994-2004	2.5	1.9	1.5	0.4	0.6
2005-2016	1.4	1.2	1.0	0.2	0.2

Notes. The table decomposes the model-predicted displacement term $1/2(\zeta-1)\nu^2$ into a displacement "within" industries $1/2(\zeta-1)\nu^2_{\text{within}}$ and a displacement "between" industries $1/2(\zeta-1)\nu^2_{\text{between}}$. The decomposition follows from the law of total variance: the variance of wealth growth ν^2 is the sum of the average variance within groups ν^2_{within} and the variance between groups ν^2_{between}. Groups are defined by industry appurtenance of households, using the Fama-French 49 industry classification. Data from Forbes 400.

Table 5: Regressing the Variance of Wealth Growth on the Variance of Stock Returns

	ν^2
	(1)
Variance of Firm-Level Returns	0.18^{***}
	(0.05)
Constant	0.02
	(0.02)
R^2	0.47
Period	1983-2016
N	34

Notes. The table reports the results of the regression of the cross-sectional variance of wealth growth for households at the top percentile ν on the cross-sectional variance of firm-level returns.
Estimation via OLS. Standard errors in parentheses and estimated using Newey-West with 3 lags. $^{*},^{**},^{***}$ indicate significance at the 0.1, 0.05, 0.01 levels, respectively. Data from Forbes 400 and CRSP.

Table 6: Patent Activity of Entrants

	Patent Activity of Firms		
	Number of Patents	Number of Patent Citations	Market-Value of Patents
	(1)	(2)	(3)
Entry	0.008^{***}	0.197^{***}	0.097^{***}
	(0.003)	(0.056)	(0.030)
$E[Y]$	0.008	0.120	0.180
R^2	0.31	0.24	0.53
Period	1983-2009	1983-2009	1983-2009
FE	Year, Industry	Year, Industry	Year, Industry
N	690	690	690

Notes. The table reports the results of regressions of firm-level patent activity on an entry dummy on the sample of the firms in the top at time t or entering the top. Measures of patent activity are respectively the number of patents, the number of total citations, and the market-value of patents, divided by the firm market value.
Estimation via OLS. Standard errors in parentheses. $^{*},^{**},^{***}$ indicate significance at the 0.1, 0.05, 0.01 levels, respectively. Data from Kogan et al. (2017) and Forbes 400.

Table 7: Regressing Measures of Wealth Inequality on Aggregate Patent Activity

	ν^2	$R_{\text{displacement}}/\left(1/2(\zeta-1)\right)$	$R_{\text{displacement}}$	R_{total}
	(1)	(2)	(3)	(4)
Aggregate Patent Activity	0.07***	0.12***	0.01*	0.07
	(0.02)	(0.03)	(0.01)	(0.08)
Constant	0.71***	1.23***	0.13**	0.74
	(0.23)	(0.30)	(0.06)	(0.74)
R^2	0.36	0.42	0.07	0.06
Period	1983-2008	1983-2008	1983-2008	1983-2008
N	26	26	26	26

Notes. The table reports the result of measures of wealth inequality on aggregate patent activity. Aggregate patent activity is defined as the log-ratio between the total market value of patents issued in a given year and the total market capitalization of U.S. firms, as constructed in Kogan et al. (2017). The dependent variables are: the yearly cross-sectional variance of wealth shocks (first column), the displacement term divided by $(\zeta-1)/2$ (second column), the displacement term (third column), and the net growth of top wealth share (fourth column).
Estimation via OLS. Standard errors in parentheses and estimated using Newey-West with 3 lags. *,**,*** indicate significance at the 0.1, 0.05, 0.01 levels, respectively. Data from Kogan et al. (2017) and Forbes 400.

Table 8: Wealth Growth is Serially Uncorrelated

	Future Wealth Growth
	(1)
Current Wealth Growth	−0.01
	(0.01)
Constant	0.04***
	(0.00)
R^2	0.20
Period	1983-2016
FE	Individual
N	11,453

Notes. The table reports the result of a regression of future wealth growth on current wealth growth, i.e. denoting w_{it} the wealth of household i at time t,

$$\log\left(\frac{w_{it+2}}{w_{it+1}}\right) = \alpha_i + \beta \log\left(\frac{w_{it+1}}{w_{it}}\right) + \epsilon$$

Estimation via OLS. Standard errors in parentheses and estimated using Newey-West with 3 lags. *,**,*** indicate significance at the 0.1, 0.05, 0.01 levels, respectively. Data from Forbes 400.

Table 9: Displacement Along the Wealth Distribution

	Top 1%	Top 0.1%	Top 0.01%	Top 400
Panel A: 1926-2012				
Power-law ζ	1.5	1.6	1.7	
Idiosyncratic Volatility ν (%)	0.14	0.17	0.21	
Displacement Term $1/2(\zeta-1)\nu^2$ (%)	0.5	0.9	1.4	
Panel B: 1983-2012				
Power-law ζ	1.5	1.5	1.6	1.5
Idiosyncratic Volatility ν (%)	0.20	0.23	0.28	0.28
Displacement Term $1/2(\zeta-1)\nu^2$ (%)	1.1	1.5	2.2	2.1

Notes. The local tail index ζ is estimated as $1/(1-\frac{q_t}{S_t/p})$. The idiosyncratic volatility of wealth ν is estimated by interacting the share of wealth invested in equity at each percentile with half of the idiosyncratic volatility of firm-level returns. Data from from Kopczuk and Saez (2004) and Saez and Zucman (2016).

A Appendix for Section 2

I start with a lemma given the dynamics of top wealth shares in terms of the dynamics of the wealth density:

Lemma 2. *The dynamics of the top wealth share S_t is given by*

$$dS_t = \int_{q_t}^{\infty} (w - q_t) dg_t(w) dw \tag{A1}$$

The intuition is as follows. During a small time period dt, a net mass $\int_{q_t}^{+\infty} dg_t(w)dw$ of households enter the top percentile. Because population size in the top percentile is held constant, an equal mass of households at the threshold must exit the top percentile, with a wealth q_t. The formula expresses that the total change in S_t is given by the difference between the wealth change due to entry and the wealth change due to exit.

Proof of Lemma 2. Applying Ito's lemma on $p = \int_{q_t}^{+\infty} g_t(w)dw$ gives the law of motion of the quantile q_t

$$0 = -g_t(q_t)\frac{dq_t}{dt} + \int_{q_t}^{+\infty} \frac{dg_t(w)}{dt} dw \tag{A2}$$

Applying Ito's lemma on $S_t = \int_{q_t}^{+\infty} wg_t(w)dw$ gives the law of motion of S_t:

$$dS_t = -q_t g_t(q_t) dq_t + \int_{q_t}^{\infty} w dg_t(w) dw \tag{A3}$$

Injecting the law of motion for q_t into (A4), we obtain the law of motion for S_t:

$$dS_t = \int_{q_t}^{\infty} (w - q_t) dg_t(w) dw \tag{A4}$$

□

Proof of Proposition 1. This follows from Proposition 2 □

Proof of Proposition 2. The Kolmogorov Forward equation corresponding to the wealth process is

$$dg_t = -\partial_w(\mu_t(w)dtwg_t) + \partial_w^2(\nu_t^2(w)dtw^2 g_t/2) \tag{A5}$$

Substituting this equation into Lemma 2, one obtains:

$$dS_t = \int_{q_t}^{+\infty} (w - q_t)(-\partial_w(\mu_t(w)dtwg_t) + \partial_w^2(\nu_t^2(w)dtw^2 g_t/2))dw$$

Integrating by parts twice, one obtains:

$$dS_t = \int_{q_t}^{\infty} \mu_t(w)wdtg_t(w)dw + \frac{1}{2}g_t(q_t)q_t^2\nu_t^2(q_t)dt \tag{A6}$$

□

Proof of Proposition 3. Denote π_n the population share of group n and g_{nt} the density of wealth within group n. The density of wealth in the economy is the sum of the wealth density within each group:

$$g_t = \sum_{1 \le n \le N} \pi_n g_{nt} \tag{A7}$$

Therefore, the law of motion of g_t in terms of the law of motion of g_{nt} is:

$$dg_t = \sum_{1 \le n \le N} \pi_n dg_{nt} \tag{A8}$$

Applying the law of motion for S_t in terms of g_t (A13) and the Kolmogorov Forward Equation for g_{nt} for $1 \le n \le N$ (A16), we obtain:

$$dS_t = \int_{q_t}^{\infty} (w - q_t) \sum_{1 \le n \le N} \pi_n \left(-\partial_w(\mu_{nt} dtw g_{nt}(w)) + \partial_w^2(\nu_{nt}^2 dtw^2 g_{nt}(w)/2) \right) dw \tag{A9}$$

Integrating by parts, we obtain the law of motion of top wealth share S_t:

$$dS_t = \int_{q_t}^{+\infty} \sum_{1 \le n \le N} \mu_{nt} dtw \pi_n g_{nt}(w) dw + \frac{g_t(q_t) q_t^2}{2} \sum_{1 \le n \le N} \frac{\pi_n g_{nt}(q_t)}{g_t(q_t)} \nu_{nt}^2 \tag{A10}$$

□

Proof of Proposition 4. Applying Ito's lemma on the implicit definition of quantile $p = \int_{q_t}^{+\infty} g_t(w) dw$ gives the law of motion of the quantile q_t

$$0 = -g_t(q_t) \frac{dq_t}{dt} + \int_{q_t}^{+\infty} \frac{dg_t(w)}{dt} dw - \sigma_t[dg_t(q_t)] \sigma_t[dq_t] \tag{A11}$$

where $\sigma_t[dg_t(q_t)]$ and $\sigma_t[dq_t]$ denote the exposure of $g_t(q_t)$ and q_t to aggregate shocks.

Applying Ito's lemma on $S_t = \int_{q_t}^{+\infty} g_t(w) dw$ gives the law of motion of S_t:

$$dS_t = -q_t g_t(q_t) dq_t + \int_{q_t}^{\infty} w dg_t(w) dw - q_t \sigma_t[dg_t(q_t)] \sigma_t[dq_t] dt - \frac{1}{2} g_t(q_t) \sigma_t[dq_t]^2 dt \tag{A12}$$

Injecting (A11) into (A12), we obtain:

$$\begin{aligned} dS_t &= \int_{q_t}^{\infty} (w - q_t) dg_t(w) dw - \frac{1}{2} g_t(q_t) \sigma_t[dq_t]^2 dt \\ &= \int_{q_t}^{\infty} (w - q_t) dg_t(w) dw - \frac{1}{2} \frac{1}{g_t(q_t)} (\int_{q_t}^{\infty} \sigma_t[dg_t(w)] dw)^2 dt \end{aligned} \tag{A13}$$

Denote π_n the population share of group n and g_{nt} the density of wealth within group n. The density of wealth in the economy is the sum of the wealth density within each group:

$$g_t = \sum_{1 \le n \le N} \pi_n g_{nt} \tag{A14}$$

Therefore, the law of motion of g_t in terms of the law of motion of g_{nt} is:

$$dg_t = \sum_{1 \le n \le N} \pi_n dg_{nt} \tag{A15}$$

The Kolmogorov Forward equation for g_{nt} gives:

$$dg_{nt} = -\partial_w(\mu_{nt} dtwg_t + \sigma_{nt} wdZ_t) + \partial_w^2((\sigma_{nt}^2 + \nu_{nt}^2)dtw^2 g_{nt}/2) \tag{A16}$$

Applying the law of motion for S_t in terms of g_t (A13) and the Kolmogorov Forward Equation for g_{nt} for $1 \leq n \leq N$ (A16), we obtain:

$$\begin{aligned} dS_t = & \int_{q_t}^{\infty} (w - q_t) \sum_{1 \leq n \leq N} \pi_n \left(-\partial_w((\mu_{nt} dt + \sigma_{nt} dZ_t) w g_{nt}(w)) + \partial_w^2((\nu_{nt}^2 + \sigma_{nt}^2) dtw^2 g_{nt}(w)/2)\right) dw \\ & - \frac{1}{2} \frac{1}{g_t(q_t)} (\int_{q_t}^{+\infty} \sum_{1 \leq n \leq N} \pi_n \partial_w (\pi_n \sigma_{nt} w g_{nt}(w) dw))^2 dt \end{aligned} \tag{A17}$$

Integrating by parts, we obtain the law of motion of top wealth share S_t:

$$\begin{aligned} dS_t = & \int_{q_t}^{+\infty} \sum_{1 \leq n \leq N} (\mu_{nt} dt + \sigma_{nt} dZ_t) w \pi_n g_{nt}(w) dw \\ & + \frac{g_t(q_t) q_t^2}{2} \left(\sum_{1 \leq n \leq N} \frac{\pi_n g_{nt}(q_t)}{g_t(q_t)} \nu_{nt}^2 + \sum_{1 \leq n \leq N} \frac{\pi_n g_{nt}(q_t)}{g_t(q_t)} \sigma_{nt}^2 - \left(\sum_{1 \leq n \leq N} \frac{\pi_n g_{nt}(q_t)}{g_t(q_t)} \sigma_{nt} \right)^2 \right) dt \end{aligned} \tag{A18}$$

□

Lemma 3 (Dynamics of Top Wealth Share with Jumps (General Distribution))**.** *Assume that the law of motion for wealth is given by* (14)*. Then the top wealth share S_t follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt}_{dr_{within}} + \underbrace{\frac{g_t(q_t) q_t^2}{2 S_t} \nu_t^2 dt + \frac{\lambda_t dt}{S_t} \mathrm{E}^f [\int_{q_t e^{-J}}^{q_t} (e^J w - q_t) g_t(w) dw]}_{dr_{displacement}} \tag{A19}$$

Proof of Lemma 3. The Kolmogorov Forward equation corresponding to the wealth process is

$$dg_t = -\partial_w(\mu_t dtwg_t) + \partial_w^2(\nu_t^2 dtw^2 g_t/2) + \lambda_t dt \mathrm{E}^f[e^{-J} g_t(we^{-J}) - g_t(w)]. \tag{A20}$$

Substituting the law of motion for dg_t from the Kolmogorov Forward equation into Lemma 2 and integrating by parts:

$$\begin{aligned} dS_t = & \int_{q_t}^{\infty} (w - q_t)(-\partial_w(\mu_t dtwg_t(w)) + \partial_w^2(\nu_t^2 dtw^2 g_t(w)/2) \\ & + \lambda \int_{q_t}^{+\infty} (w - q_t)(\mathrm{E}^f[e^{-J} g_t(we^{-J})] - g_t(w)) dw \\ = & - \int_{q_t}^{+\infty} (-\mu_t dtwg_t(w) + \partial_w(\nu_t^2 dtw^2 g_t(w)/2)) dw \\ & + \lambda \int_{q_t}^{+\infty} (w - q_t)(\mathrm{E}^f[e^{-J} g_t(we^{-J})] - g_t(w)) dw \\ = & \mu_t S_t dt + \frac{g_t(q_t) q_t^2}{2} \nu_t^2 dt + \lambda_t dt \int_{q_t}^{+\infty} (w - q_t)(\mathrm{E}^f[e^{-J} g_t(we^{-J})] - g_t(w)) dw \end{aligned} \tag{A21}$$

The jump term can be rewritten as a summation with respect to the wealth pre-jump. Denoting f_t the density function of jump sizes J_{it}, we have:

$$
\begin{aligned}
\int_{q_t}^{+\infty}(w-q_t)(\mathrm{E}^f[e^{-J}g_t(we^{-J})]-g_t(w))dw &= \int_{q_t}^{+\infty}(w-q_t)\int_{J=-\infty}^{+\infty}(e^{-J}g_t(we^{-J})-g_t(w))f_t(J)dJ \\
&= \int_{J=-\infty}^{+\infty}f_t(J)dJ\int_{q_t}^{+\infty}(w-q_t)(e^{-J}g_t(we^{-J})-g_t(w))dw \\
&= \int_{J=-\infty}^{+\infty}f_t(J)dJ(\int_{q_t}^{+\infty}(w-q_t)e^{-J}g_t(we^{-J})dw-\int_{q_t}^{+\infty}(w-q_t)g_t(w)dw) \\
&= \int_{J=-\infty}^{+\infty}f_t(J)dJ(\int_{q_te^{-J}}^{+\infty}(e^Jw-q_t)g_t(w)dw-\int_{q_t}^{+\infty}(w-q_t)g_t(w)dw) \\
&= \int_{J=-\infty}^{+\infty}f_t(J)dJ(\int_{q_te^{-J}}^{q_t}(e^Jw-q_t)g_t(w)dw+\int_{q_t}^{\infty}(e^J-1)wg_t(w)dw) \\
&= \mathrm{E}^f[\int_{q_te^{-J}}^{q_t}(e^Jw-q)g_t(w)dw].
\end{aligned}
$$

This concludes the proof. $\square$

Proof of Proposition 5. When the distribution is Pareto, the growth of top wealth shares due to jumps in Lemma 3 can be written

$$
\begin{aligned}
\frac{1}{S_t}\mathrm{E}^f\left[\int_{q_te^{-J}}^{q_t}(e^Jw-q_t)g_t(w)dw\right] &= \frac{1}{S_t}\mathrm{E}^f\left[\int_{q_te^{-J}}^{q_t}(e^Jw-q_t)Cw^{-\zeta-1}dw\right] \\
&= \frac{1}{q_t^{1-\zeta}/(\zeta-1)}\mathrm{E}^f\left[[e^J\frac{x^{1-\zeta}}{1-\zeta}]_{q_te^{-J}}^{q_t}+[q_t\frac{x^{-\zeta}}{\zeta}]_{q_te^{-J}}^{q_t}]\right] \\
&= \frac{1}{q_t^{1-\zeta}/(\zeta-1)}\mathrm{E}^f\left[e^J\frac{q_t^{1-\zeta}}{1-\zeta}-e^Je^{(-J)(1-\zeta)}\frac{q_t^{1-\zeta}}{1-\zeta}+\frac{q_t^{1-\zeta}}{\zeta}-e^{(-J)(-\zeta)}\frac{q_t^{1-\zeta}}{\zeta}\right] \\
&= \mathrm{E}^f\left[e^{J\zeta}-e^J+\left(1-\frac{1}{\zeta}\right)(1-e^{J\zeta})\right] \\
&= \mathrm{E}^f\left[\frac{e^{J\zeta}-1}{\zeta}-(e^J-1)\right] \\
&= \frac{\mathrm{E}^f[e^{J\zeta}]-1}{\zeta}
\end{aligned}
$$

Note that, by Ito's lemma, we have:

$$
\frac{\phi_t(\zeta)}{\zeta}=\mu_tdt+\frac{\zeta-1}{2}\nu_t^2dt+\lambda_tdt\frac{\mathrm{E}^f[e^{J_{it}\zeta}]-1}{\zeta} \tag{A22}
$$

Therefore we obtain

$$
\frac{dS_t}{S_t}=\frac{\phi_t(\zeta)}{\zeta}
$$

Plugging the definition for κ_{jt} gives (15).

Note that κ_{jt} corresponds to the derivative of the $j-$th cumulant of wealth growth between t and $t+\tau$, at $\tau=0$. To see this, note that:

$$
\mathrm{E}_t\left[\frac{dw_{it}^\zeta}{w_{it}^\zeta}\right]=\left(\lim_{\tau\to 0}\frac{1}{\tau}\log E_t\frac{w_{it+\tau}^\zeta}{w_{it}^\zeta}\right)dt=\sum_{j=1}^{+\infty}\frac{\zeta^j}{j!}\left(\lim_{\tau\to 0}\frac{\kappa_{jt}(\tau)}{\tau}\right)dt \tag{A23}
$$

where $\kappa_{jt}(\tau)$ denotes the $j-$th cumulant of wealth growth between t and $t+\tau$. $\square$

Lemma 4 (Dynamics of Top Wealth Share with Death and Population Growth (General Distribution))**.** *When wealth follows the law of motion* (2)*, with death rate* δ_t*, inheritance parameter* χ_t*, and population growth* η_t*, the top wealth share* S_t *follows the law of motion:*

$$\frac{dS_t}{S_t} = \mu_t dt + \underbrace{\frac{\zeta_t(q_t)-1}{2}\nu_t^2 dt}_{dr_{displacement}} + \overbrace{\underbrace{\left(\frac{\chi_t S_t(\pi_t p) + (1-\pi_t)q_t p}{S_t} - 1\right)\delta_t dt}_{dr_{death}} + \underbrace{\frac{q_t p}{S_t}\eta_t dt}_{dr_{pop.\ growth}}}^{dr_{demography}} \tag{A24}$$

where $\pi_t \equiv \mathbb{P}(\chi_t w \geq q_t | w \geq q_t)$ *is the proportion of households that die with a wealth high enough that their offspring enters the top percentile.*

Proof of Lemma 4. **Heuristic Derivation.** The death term can be derived as follows. Between t and dt, a mass $\delta_t p dt$ of households in the top die, which decreases total wealth in the top percentile by $S_t \delta_t dt$. A proportion π_t of these households has a wealth high enough that their newborn offspring enters the top percentile. Since the average wealth of this offspring is $\chi S_t(\pi_t p)/(p\pi_t)$, this increases total wealth in the top percentile by $\chi S_t(\pi_t p)$. The other households are simply replaced by households that enter at the lower percentile threshold, with wealth q_t. The total increase of S_t due to death is $(\chi_t S_t(\pi_t p) + (1-\pi_t)q_t p - S_t)\delta_t dt$.

I now turn to the term due to population growth. Between t and dt, a new mass $\eta_t p dt$ of households enters the top percentile. Since the wealth of these households equals to q_t, this increases total wealth in the top percentile by $\eta_t p q_t dt$. The total increase of S_t due to population growth is $\eta_t p q_t dt$.

Formal Derivation. The Kolmogorov Forward Equation is

$$dg_t = -\partial_w(\mu_t dt w g_t) + \partial_w^2(\nu_t^2 dt w^2 g_t/2) - \eta_t dt g_t(w) + \delta_t dt\left(\frac{g(w/\chi)}{\chi} - g(w)\right) \tag{A25}$$

Plugging the law of motion of the density in Lemma 2 and integrating by parts, one obtains the dynamics of the top wealth share S_t.

□

Proof of Proposition 6. When the wealth distribution is Pareto with tail exponent ζ, Lemma 4 can be simplified. We have $\pi_t = \chi_t^\zeta$, and therefore $S_t(\pi_t p) = \chi_t^{\zeta-1} S_t(p)$. Therefore, the term due to death can be written as

$$\begin{aligned} dr_{\text{death}} &= \left(\frac{\chi_t S_t(\pi_t p) + (1-\pi_t)q_t p}{S_t} - 1\right)\delta_t dt + \frac{q_t p}{S_t}\eta_t dt \\ &- ((\chi^\zeta + (1-\chi_t^\zeta)(1-\frac{1}{\zeta}) - 1)\delta_t dt \\ &= \frac{\chi^\zeta - 1}{\zeta}\delta_t dt \end{aligned} \tag{A26}$$

□

Lemma 5 (Dynamics of Quantile)**.** *Assume that the law of motion for wealth is given by* (8)*. Then the top quantile* q_t *follows the law of motion:*

$$\frac{dq_t}{q_t} = \mu_t(q_t)dt - \frac{1}{2}\frac{\partial_w(w^2\nu_t^2(w)g_t(w))}{g_t(q_t)q_t}dt \tag{A27}$$

Proof of Lemma 5. Combining the definition of the quantile (A2) with the Kolmogorov Forward equation (A5), we obtain

$$\begin{aligned} dq_t &= \frac{1}{g_t(q_t)} \int_{q_t}^{+\infty} dg_t(w)dw \\ &= \frac{1}{g_t(q_t)} \int_{q_t}^{+\infty} (-\partial_w(\mu_t(w)dtwg_t(w)) + \frac{1}{2}\partial_w^2(\nu_t^2(w)dtw^2 g_t(w))dw \\ &= \mu_t(q_t)q_t dt - \frac{1}{2}\frac{\partial_w(w^2\nu_t^2(w)g_t(w))}{g_t(q_t)}dt \end{aligned} \tag{A28}$$

□

B Appendix for Section 3

Proof of Proposition 7. The set of households in the top at time $t+\tau$ is $\mathcal{T}' = (\mathcal{T} \setminus (\mathcal{X}_\mathcal{D} \cup \mathcal{X})) \cup (\mathcal{E} \cup \mathcal{E}_\mathcal{D})$. Therefore, the total wealth in the top at time $t+\tau$ can be decomposed as follows:

$$\sum_{i\in\mathcal{T}'} w_{it+\tau} = \sum_{i\in\mathcal{T}\setminus\mathcal{X}_\mathcal{D}} w_{it+\tau} + \sum_{i\in\mathcal{E}\cup\mathcal{E}_\mathcal{D}} w_{it+\tau} - \sum_{i\in\mathcal{X}} w_{it+\tau} \tag{A29}$$

Denote R the net wealth growth of households in the top at time t that do not die, i.e.

$$R_{\text{within}} = \frac{\sum_{i\in\mathcal{T}\setminus\mathcal{X}_\mathcal{D}} w_{it+\tau}}{\sum_{i\in\mathcal{T}\setminus\mathcal{X}_\mathcal{D}} w_{it}} - 1 \tag{A30}$$

Equation (A29) can be rewritten using $r_{t+\tau}$:

$$\sum_{i\in\mathcal{T}'} w_{it+\tau} = (1+R_{\text{within}})(\sum_{i\in\mathcal{T}} w_{it} - \sum_{i\in\mathcal{X}_\mathcal{D}} w_{it}) + \sum_{i\in\mathcal{E}\cup\mathcal{E}_\mathcal{D}} w_{it+\tau} - \sum_{i\in\mathcal{X}} w_{it+\tau} \tag{A31}$$

Adding and subtracting $q_{t+\tau}$ to the wealth of households that enter, exit, or die, and dividing by total wealth at time $t+\tau$, one obtains

$$\begin{aligned} \sum_{i\in\mathcal{T}'} w_{it+\tau} = (1+R_{\text{within}})\sum_{i\in\mathcal{T}} w_{it} + \sum_{i\in\mathcal{X}_\mathcal{D}} (q_{t+\tau} - (1+R_{\text{within}})w_{it}) + \sum_{i\mathcal{E}_\mathcal{D}}(w_{it+\tau} - q_{t+\tau}) \\ + \sum_{i\in\mathcal{E}}(w_{it+\tau} - q_{t+\tau}) + \sum_{i\in\mathcal{X}}(q_{t+\tau} - w_{it+\tau}) + (|\mathcal{T}'| - |\mathcal{T}|)q_{t+\tau} \end{aligned} \tag{A32}$$

Dividing by S_t and rearranging, one obtains the accounting decomposition (20). □

I now relate this accounting decomposition to the theoretical decomposition in Lemma 4. I assume that the panel data is a representative sample of the true underlying continuous distribution. I also consider the model without inheritance (i.e. $\chi = 0$) to simplify the exposition.

Integrating the law of motion (2) for household wealth w_{it}, we get for $\tau > 0$

$$\mathrm{E}[R_{\text{within}}] = e^{\int_t^{t+\tau} \mu_s ds} - 1 \tag{A33}$$

Therefore

$$\frac{d}{d\tau}\mathrm{E}[R_{\text{within}}]\Big|_{\tau=0} = \mu_t \tag{A34}$$

Integrating the law of motion for the top wealth share (A24), we get for $\tau > 0$

$$\mathrm{E}[\frac{S_{t+\tau}}{S_t}] - 1 = e^{\int_t^{t+\tau}\left(\mu_s + \frac{g_s(q_s)q_s^2}{2S_s}\nu_s^2 + (\frac{q_s p}{S_s} - 1)\delta_s + \frac{q_s p}{S_t}\eta_s\right)ds} - 1 \tag{A35}$$

Therefore

$$\frac{d}{d\tau}\mathrm{E}\left[\frac{S_{t+\tau}}{S_t}\right]\Big|_{\tau=0} = \mu_t + \frac{g_t(q_q)q_t^2}{2S_t}\nu_t^2 + \left(\frac{q_t p}{S_t} - 1\right)\delta_t + \frac{q_t p}{S_t}\eta_t \tag{A36}$$

The demography term is

$$\mathrm{E}[R_{\text{demography}}] = \left(1 - e^{-\int_t^{t+\tau}\delta_s ds}\right)\left(\frac{q_{t+\tau}p}{S_t} - e^{\int_t^{t+\tau}\mu_s ds}\right) + e^{\int_t^{t+\tau}\eta_s ds} p \frac{q_{t+\tau}}{S_t} \tag{A37}$$

Therefore

$$\frac{d}{d\tau}\mathrm{E}[R_{\text{demography}}]\Big|_{\tau=0} = (\frac{q_t p}{S_t} - 1)\delta_t + \frac{q_t p}{S_t}\eta_t \tag{A38}$$

We obtain the derivative of the displacement term as a residual:

$$\begin{aligned}\frac{d}{d\tau}\mathrm{E}[R_{\text{displacement}}]\Big|_{\tau=0} &= \frac{d}{d\tau}\mathrm{E}\left[\frac{S_{t+\tau}}{S_t} - R_{\text{within}} - R_{\text{demography}}\right]\Big|_{\tau=0} \\ &= \frac{g_t(q_t)q_t^2}{2S_t}\nu_t^2 \end{aligned} \tag{A39}$$

Therefore, the expectation of the terms in the accounting decomposition are asymptotically equivalent to the terms in Lemma 4 as the time period τ tends to zero. In this sense, the accounting decomposition converges to the theoretical decomposition as the time period τ tends to zero.

C Appendix for Section 4

C.1 Left Censoring

The decomposition in Section 3 requires us to know the wealth of households that drop out of the top percentile. However, Forbes only reports the wealth of individuals in Forbes 400 before 2012.

First, 60% households that drop out of the top percentile actually stay in Forbes 400. Indeed, the top percentile used in this paper is composed of only 264 households in 1983 (indeed, it was chosen so that, with population growth, it includes 400 households in 2017). Because wealth is so concentrated in the top, there is usually a large difference between the last individual in this top percentile and the wealth of the last individual in the top 400. Therefore, most households that drop out of this top percentile stay in the top 400.

I now focus on the remaining 40% of households that drop off Forbes 400. Formally, the problem boils down to estimating the average of a variable (the wealth growth of top households) that is left censored. In this particular setting, the Kaplan and Meier (1958) estimator gives tight bounds to estimate this quantity. The idea is to estimate this quantity using the observed big negative jumps of the top households to infer the negative jumps of the households that drop off Forbes 400. The identifying assumption is that the distribution of negative jumps is the same for households at the very top of the distribution compared to households at the quantile.

More precisely, Kaplan and Meier (1958) insight is that the survival function, i.e. in my setting the probability that wealth growth is lower than a certain threshold $P(w_{t+1}/w_t - 1 \leq x)$, can be estimated even if the data is censored. In turn, this survival function can be used to estimate the conditional expectation of wealth growth, given that it is lower than a certain threshold, i.e. $\mathrm{E}[w_{t+1}/w_t - 1 | w_{t+1}/w_t - 1 \leq x]$. Finally, I use this conditional expectation to impute the wealth growth of each household that drops out of the top.

I check the validity of this imputation method by focusing on years where Forbes reports the wealth of drop-offs. Starting from 2012, Forbes systematically reports the wealth of drop-offs. In these years, I compare the result obtained from the estimated method and the result obtained using the real wealth of drop-offs. The results are reported in Table A5. Column (2) and (3) report the average return of these drop-offs using the imputed method and the actual data reported by wealth. The estimates differ by only 3 percentage points in average (-34.6% vs -31.5%). The fact that the Kaplan-Meier estimator gives such a good result is intuitive: because wealth is very concentrated households at the very top of the distribution hold ten times more wealth than the households at the margin, and therefore I do observe a large part of the distribution of downward jumps.

Column (3) reports that the total wealth owned by these imputed households represents only 2% of the total wealth of households at the top. These imputed households represent a very small share of the total households at the top, which suggests that noise due to imputation will have little impact on the average wealth growth of households at the top.

Columns (4) and (5) report the estimates for $R_{\text{within}} = \mathrm{E}[w_{t+1}/w_t - 1]$ using imputed and real data. The estimates differ by 0.1 percentage points. The bias is small because, as discussed above, the Kaplan-Meier method gives accurate estimates of the wealth growth of imputed households and that the wealth share represented by the imputed households is small to begin with.

C.2 Measurement Error

I study the relation between the persistence of wealth growth and measurement error. Suppose the process for wealth is given by $w_{it+1} = w_{it}e^{r_{it+1}}$ where r_{it+1} is an i.i.d. process independent of wealth. Moreover, suppose the observed wealth $\tilde{w}_{it} = w_{it}e^{\epsilon_{it}}$ where ϵ_{it} is an i.i.d process independent of wealth capturing measurement error and $\mathrm{E}[e^{\epsilon_{it}}] = 1$. Denote ξ the ratio between the variance of measurement error and the variance of wealth growth, i.e.

$$\xi \equiv \frac{var(\epsilon_{it})}{var(r_{it})} \tag{A40}$$

The log change in observed wealth can be written as

$$\log\left(\frac{\tilde{w}_{it+1}}{\tilde{w}_{it}}\right) = r_{it+1} + \epsilon_{it+1} - \epsilon_{it} \tag{A41}$$

A regression of observed wealth growth on the lagged observed wealth growth estimates the slope coefficient ρ:

$$\begin{aligned}
\rho &= \frac{cov(\log(\tilde{w}_{it+1}/\tilde{w}_{it}), \log(\tilde{w}_{it+1}/\tilde{w}_{it}))}{var(\log(\tilde{w}_{it+1}/\tilde{w}_{it}))} \\
&= \frac{cov(r_{it+1} + \epsilon_{it+1} - \epsilon_{it}, r_{it} + \epsilon_{it} - \epsilon_{it-1})}{var(r_{it} + \epsilon_{it} - \epsilon_{it-1})} \\
&= -\frac{var(\epsilon_{it})}{var(r_{it}) + 2var(\epsilon_{it})} \\
&= -\frac{\xi}{1+2\xi}
\end{aligned} \tag{A42}$$

Testing whether ρ is equal to zero is a test on whether ξ is different from zero, i.e. that there is measurement error.

The slope coefficient ρ is negative and decreasing in ξ. Moreover, for ξ close to zero, ρ can be well approximated by the opposite of ξ, i.e. $\rho \approx -\xi$.

I now examine how the displacement term depends on ρ. Even though I cannot reject that ρ is statistically different from zero, this allows me to check that small values of ρ do not have a disproportionate effect on the displacement term. To make further progress, I assume that ϵ_{it} and r_{it} are normal variables.

Since $w_{t+1} = e^{r_{t+1}} w_t$, the displacement term is given by

$$r_{\text{displacement}} = \frac{\zeta - 1}{2} Var(r_{it}) \tag{A43}$$

By contrast, since $\tilde{w}_{t+1} = e^{r_{t+1} + \epsilon_{t+1} - \epsilon_t} \tilde{w}_t$, the observed displacement term is

$$\tilde{r}_{\text{displacement}} = \frac{\zeta - 1}{2} \left(Var(r_{it}) + 2Var(\epsilon_{it})\right) \tag{A44}$$

Therefore, the relative bias between the observed displacement term and the measured displacement term is:

$$\frac{\tilde{r}_{\text{displacement}} - r_{\text{displacement}}}{r_{\text{displacement}}} = -\frac{2\rho}{1+2\rho}$$

In particular, when ρ is close to zero, the relative bias is close to -2ρ. Since ρ is very small, we can conclude that the relative bias is also very small.

C.3 Families

This paper follows the empirical literature on wealth inequality by using households, rather than families, as the unit of observation. A concern is that the displacement term captures reallocation of wealth within families, rather than reallocation of wealth across families.

To examine formally the importance of the reallocation of wealth within families, I decompose the displacement term predicted by the diffusion model $1/2(\zeta - 1)\nu^2$ into a term due to a displacement within families and another term due to the displacement between families. This decomposition relies on the law of total variance: the variance of wealth growth is the sum of the average variance within groups and the variance between groups. Table A6 reports the annual displacement within families and between families: the displacement within families is negligible relative to the displacement between families: it accounts for less than 5% of the overall displacement term.

C.4 Within Term

I now examine the level and the dynamics of the within term. Overall, the within term averages 1.9%, but is very volatile. There is no significant trend in the within term over time, as shown in Table A1.

The within term R_{within} can be approximated by the difference between the wealth growth of top households, denoted $R_{\text{top households}}$, and the wealth growth of the economy, denoted $R_{\text{U.S.}}$:

$$R_{\text{within}} \approx R_{\text{top households}} - R_{\text{U.S.}} \tag{A45}$$

I report both series in real terms in Table A3. The wealth growth of the overall economy is pretty stable over time: most of the dynamics of the within term come from an increase of $R_{\text{households}}$.

To determine the exposure of the wealth of top households to priced factors, I run regressions of their average wealth growth on a variety of factors in Table A2. Because the wealth of top households may contain illiquid assets that are difficult to valuate, one concern is that the true volatility of wealth is higher than the volatility reported by Forbes.[45] To avoid this issue, I estimate the exposure of top households by regressing three-year horizon wealth growth on one year factors returns. After obtaining a beta, I compute a constant term as the average of the difference between the wealth growth of top households and the return predicted by factor exposures. I compute the standard errors of factor exposures and of the constant terms by bootstrapping.[46].

Column (1) reports the results where the only factor is the stock market. The slope coefficient, which reflects the exposure of top households to the stock market, is close to one. Column (2) reports the results for the Fama-French three factor models, that add the value factor and the size factor. The exposure to the size factor is weakly negative, significant at 10%, which reflects the fact that households at the top tend to own bigger firms. The exposure to the value factor is not significant. Similarly, Column (3) reports the results for the Fama-French five factor models, that add profitability and investment factors. Similarly, only the exposure market is significant. Finally, in column (4), I add the excess returns of long-term bonds, as well as the excess returns of corporate bonds. Similarly, only the exposure to the market is significant. Overall,

[45]This problem is known as the “stale pricing” problem in the private equity literature, see for instance Emery (2003).

[46]More precisely, I use block-bootstrap to correct for the serial correlations of the returns across time.

the stock market appears to be the main factor for the average wealth growth of top households. Moreover, the exposure to the market is relatively constant around 1.0 across specifications.

I also compare the within term to a representative portfolio of industries at the top, rather than simply the market in Column (5) of Table A2. I classify households in the Forbes 400 based on the 49 industries of Fama-French. The industries of households in top percentiles are not representative of the market (in particular, Real Estate, Printing and Publishing, Computer Software, and Petroleum play a more important role at the top compared to the market). I construct a benchmark portfolio that weights each industry similar to the industry represented in the top. I find a similar exposure of 1 to this industry weighted portfolio. This suggests that the exact industry composition of individuals at the top does not matter much for the growth of top wealth shares.

I use this factor model to decompose the wealth growth of top households $R_{\text{top households}}$ into a term due to the financial returns of top households (which can itself be decomposed into a term due to the risk-free rate and a term due to the exposure of households wealth to priced factors), a positive term due to labor income, a negative term due to tax paid as a proportion of wealth, and a residual, i.e.:

$$R_{\text{top households}} = R_f + \sum_{1 \leq k \leq K} \beta_k \times (R_k - R_f) + \frac{\sum_{i \in \mathcal{T}} (\text{Labor}_i - \text{Tax}_i)}{S_t} + \epsilon \tag{A46}$$

where Labor_i denotes the labor income, Tax_i denotes the total tax paid by households i, R_f is the risk-free rate and R_k is the return of factor $k \leq K$. I obtain the total tax paid and total wage income received by the top 400 individuals by income from the IRS.[47] The dataset is only available after 1992, so I use the average of this term in 1992-1995 to impute it starting from 1983.

Panel A of Table A3 reports decomposition (A46) using the market return as a factor. Because labor income and taxes are very small as a proportion of total wealth, they play a very small role in the within term. The residual, which can be interpreted as the difference between an eventual alpha of top households minus a consumption rate, appears to be negative, and increases over time.

To understand better what drives the increase of this residual over time, Panel B of Table A3 reports the decomposition of the industry-weighted return as a factor, instead of the market return. The industry-weighted portfolio overweights industries that are over-represented in the top. It has particularly low returns in the 1980s, due to the poor performance of the Real Estate and Petroleum industries during this decade. After using this industry-weighted portfolio, the residual ϵ appears to be constant over time, reflecting that the industry composition of top percentiles plays a substantial role in understanding the fluctuations of the within term.

C.5 Demography Term

I now examine the level and the dynamics of the demography term through the lens of the model presented in Section 2. I first focus on the term due to death r_{death}. According to the theoretical model, this term

[47] https://www.irs.gov/pub/irs-soi/13intop400.pdf.

equals

$$R_{\text{death}} = \frac{\chi_t^\zeta - 1}{\zeta}\delta_t \tag{A47}$$

where ζ is the Pareto tail of the wealth distribution, δ_t is the death rate of households at the top, and χ_t captures the extent to which deceased households are able to bequest their wealth to their offspring. If $\chi_t = 0\%$, deceased households are not able to pass their wealth to their offspring. If $\chi_t = 100\%$, deceased households are perfectly able to pass their wealth to their offspring. $1 - \chi_t$ can be interpreted as the average estate tax paid by deceased households.

Given an estimate from ζ and δ_t that can be obtained from the data, one can always compute the implicit inheritance parameter χ_t that explains the magnitude of the death term R_{death} using (A47). I report the result of this decomposition in Lemma 4. I find that the inheritance parameter χ_t averages to 60% and is pretty stable over time. It corresponds to an average estate tax $1 - \chi_t = 40\%$ which is close to the actual top marginal estate tax rate during the period.

I now focus on the term due to population growth. According to the theoretical framework in Section 2, the term equals

$$R_{\text{demography}} = \left(1 - \frac{1}{\zeta}\right)\eta_t \tag{A48}$$

where ζ is the Pareto tail of the wealth distribution and η_t is the population growth rate. I compare the actual population growth rate to the term predicted by the model, using the estimate for ζ used in Section 4, and the population growth rate η_t. I find that the two terms are very close, with a difference ϵ close to zero. By definition of the population growth term in the accounting decomposition, any difference between the population growth and the model-predicted term purely reflects the fact that the ratio between the wealth at the lower threshold to the average wealth of the population $q_t p/S_t$ may not be exactly equal from $1 - 1/\zeta$, where ζ is an estimate of the Pareto exponent. The fact that the two terms are equal reflects the fact that the right tail of the wealth distribution is very close to Pareto.

D Appendix for Section 5

I now discuss how the law of motion of top wealth shares can also shed light on the behavior of the tail index of the distribution. This exercise relates my paper to the results to Gabaix et al. (2016), that discusses the convergence of Pareto tails after changes in wealth dynamics.

The ratio between the share of wealth owned by a top percentile compared to another top percentile is a good proxy for the tail index of the distribution. Indeed, for a distribution with a Pareto tail with tail index ζ, the ratio of the wealth shares of two top percentiles p and p' is:

$$\log\left(\frac{S_t(p)}{S_t(p')}\right) = \left(1 - \frac{1}{\zeta}\right)\log\left(\frac{p}{p'}\right) \tag{A49}$$

The dynamics of the ratio between two top wealth shares therefore captures the dynamics of the right tail of the distribution.

I examine the case where the individual volatility depends on the wealth level, i.e.

$$\frac{dw_{it}}{w_{it}} = \mu_t dt + \nu_t(w_{it})dB_{it} \tag{A50}$$

The law of motion of the ratio between two top wealth shares is now:

$$d\log\left(\frac{S_t(p)}{S_t(p')}\right) = \left(\frac{g_t(q_t(p))q_t(p)^2}{2S_t(p)}\nu_t(q_t(p))^2 - \frac{g_t(q_t(p'))q_t(p')^2}{2S_t(p')}\nu_t(q_t(p'))^2\right)dt \tag{A51}$$

At the first order, the growth of this ratio depends on two terms: the difference in the shape of the wealth distribution at percentile p and at percentile p', but also the difference in idiosyncratic volatility for households at the lower threshold of percentile p and those at the lower threshold of the percentile p'.

Section 5 shows that the difference in the shape of the wealth distribution between the top 1% and the top 0.01% can only account for a 0.2% yearly difference in growth rate between the top two percentiles ($(1.6 - 1.5)/2 \times 0.2^2$). In contrast, the difference in the idiosyncratic variance of wealth growth between the top 1% and the top 0.01% can account for a 1.2% yearly difference in growth rates between the top two percentiles $(1.6 - 1)/2 \times (0.25^2 - 0.15^2)$). In conclusion, only differences in the idiosyncratic volatility of households at the top of the wealth distribution 0.01% threshold, compared to households at the 1% threshold, can generate a rapid thickening of the tail of the distribution, as discussed in Gabaix et al. (2016).

E Appendix for Section 6

Proof of Lemma 1. We can express the average time $T_q(w_{it})$ by backward induction.

$$T_q(w_{it}) = \delta\Delta t \times 0 + (1 - \delta\Delta t) \times (\Delta t + \mathrm{E}[T_q(w_{it+\Delta t})]) \tag{A52}$$

Therefore

$$0 = (1 - \delta\Delta t)(\Delta t + \mathrm{E}[T_q(w_{it+\Delta t}) - T_q(w_{it})]) - \delta\Delta t T_q(w_{it}) \tag{A53}$$

Taking $\Delta t \to 0$, we obtain a forward-looking expression for $T_q(w_{it})$:

$$0 = dt + \mathrm{E}[dT_q(w_{it})] - T_q(w_{it})\delta dt \tag{A54}$$

Applying Ito's lemma, we obtain an ODE satisfied by T_q:

$$1 + T_q'(w)\mu w + T_q''(w)\frac{\nu^2 w^2}{2} - \delta T_q(w) = 0 \tag{A55}$$

The solution has the form:

$$T_q(w) = c_1 w^{\zeta_+} + c_2 w^{\zeta_-} + \frac{1}{\delta} \tag{A56}$$

where ζ_+ and ζ_- are respectively the positive and negative zeros of $\zeta \to \mu\zeta + \frac{\zeta(\zeta-1)}{2}\nu^2 - \delta$. Note that this function is convex, converges to infinity as ζ converges to infinity, and equals $-\delta$ in zero, therefore there are exactly two zeros for this function, one negative, and one positive.

Using the limit condition:

$$T_q(q) = 0 \tag{A57}$$

$$\lim_{w \to +\infty} T_q(w) = \frac{1}{\delta} \tag{A58}$$

we obtain $c_1 = 0$ and $c_2 = -1/(\delta q^{\zeta_-})$, therefore

$$T_q(w) = \frac{1}{\delta}\left(1 - \left(\frac{w}{q}\right)^{\zeta_-}\right) \tag{A59}$$

□

The average first passage time $T_q(w)$ for the case $\delta = 0$ can be obtained by taking the limit as $\delta \to 0$. Since $\zeta_- \to 0$, both the numerator and the denominator in the expression for $T_q(w)$ tend to zero. We can obtain the limit of $T_q(w)$ when using L'Hôpital's rule:

$$\lim_{\delta \to 0} T_q(w) = -\frac{\partial \zeta_-}{\partial \delta}(\delta = 0) \log \frac{w}{q} \tag{A60}$$

Using the implicit function theorem to compute the derivative of ζ_- with respect to δ, we obtain

$$\lim_{\delta \to 0} T_q(w) = \frac{1}{\nu^2/2 - \mu} \log \frac{w}{q} \tag{A61}$$

I now derive comparative statics of $T_q(w)$ with respect to μ. By the definition of ζ_+ and ζ_- as the implicit function theorem, we have $\frac{\partial \zeta_-}{\partial \mu} \leq 0$ and $\frac{\partial \zeta_-}{\partial \nu^2} \geq 0$. Therefore, given a distance to the quantile w/q, an increase in μ increases the average time before exit $T_q(w)$. Similarly, an increase in ν^2 decreases the average time before exit $T_q(w)$.

The average time before exit $T_q(w)$ decreases in δ. The comparative statics of $T_q(w)$ with respect to δ is a little bit harder to prove, since δ appears directly in the formula, as well as through ζ_-. The derivative of T with respect to δ is:

$$\frac{\partial T_q(w)}{\partial \delta} = -\frac{1}{\delta^2}\left(1 - \left(\frac{w}{q}\right)^{\zeta_-}\left(1 - \delta \frac{\partial \zeta_-}{\partial \delta} \log\left(\frac{w}{q}\right)\right)\right) \tag{A62}$$

This derivative has the same sign as the function

$$f : \delta \to 1 - \left(\frac{w}{q}\right)^{\zeta_-}\left(1 - \delta \frac{\partial \zeta_-}{\partial \delta} \log\left(\frac{w}{q}\right)\right) \tag{A63}$$

The function is nonnegative when δ tends to 0, and tends to $+\infty$ when δ tends to $+\infty$. Its derivative with respect to δ is

$$\frac{\partial f}{\partial \delta} = \left(\frac{w}{q}\right)^{\zeta_-} \log\left(\frac{w}{q}\right) \delta \left(\left(\frac{\partial \zeta_-}{\partial \delta}\right)^2 \log \frac{w}{q} + \frac{\partial^2 \zeta_-}{\partial \delta^2}\right) \tag{A64}$$

which is always positive since ζ_- is a convex function of δ by the implicit function theorem. We conclude that f is nonnegative and therefore $T_q(w)$ is decreasing in δ.

Proof of Proposition 8. When $\delta \neq 0$, the average time is

$$\begin{aligned}
T(p) &= \mathrm{E}^g[T_q(w_{it})|w_{it} \geq q] \\
&= \frac{\int_q^{+\infty} T_q(w)g(w)dw}{\int_q^{+\infty} g(w)dw} \\
&= \frac{1}{\delta}\left(1 - \frac{\int_q^{+\infty}(w/q)^{\zeta_-} w^{-\zeta_+-1}dw}{\int_q^{+\infty} w^{-\zeta_+-1}dw}\right) \\
&= \frac{1}{\delta}\left(1 - \frac{1}{q^{\zeta_-}}\frac{\frac{q^{\zeta_- - \zeta_+}}{\zeta_+ - \zeta_-}}{\frac{q^{-\zeta_+}}{\zeta_+}}\right) \\
&= \frac{1}{\delta}\frac{1}{1-\zeta_+/\zeta_-}
\end{aligned} \tag{A65}$$

The derivative with respect to idiosyncratic variance ν^2 is

$$\begin{aligned}
\frac{\partial T}{\partial \nu^2} &= \frac{1}{\delta}\frac{1}{(1-\zeta_+/\zeta_-)^2}\frac{\partial(\zeta_+/\zeta_-)}{\partial \nu^2} \\
&= \frac{1}{\delta}\frac{1}{(1-\zeta_+/\zeta_-)^2}\frac{\zeta_+}{-\zeta_-}\left(\frac{1}{\zeta_-}\frac{\partial \zeta_-}{\partial \nu^2} - \frac{1}{\zeta_+}\frac{\partial \zeta_+}{\partial \nu^2}\right)
\end{aligned} \tag{A66}$$

As ν^2 increases, $T(p)$ decreases only if the percentage decrease of ζ_- is higher than the percentage decrease of ζ_+. □

Table A1: Trends in the Decomposition of the Growth of Top Wealth Share

	Dependent Variable			
	R_{total} (%)	R_{within} (%)	$R_{\text{displacement}}$ (%)	$R_{\text{demography}}$ (%)
	(1)	(2)	(3)	(4)
Panel A: Period Dummies				
Constant	4.23**	1.44	2.94***	−0.07
	(2.04)	(1.75)	(0.51)	(0.23)
$\text{Dummy}_{1994 \leq \text{year} \leq 2004}$	−0.55	0.15	−0.47	−0.25
	(5.10)	(4.81)	(0.65)	(0.31)
$\text{Dummy}_{2005 \leq \text{year} \leq 2016}$	−0.61	1.24	−1.51***	−0.40
	(2.53)	(2.24)	(0.55)	(0.25)
R^2	0.00	0.00	0.21	0.07
Period	1983-2016	1983-2016	1983-2016	1983-2016
N	34	34	34	34
Panel B: Linear Year Trend				
Constant	3.83**	1.93	2.26***	−0.29***
	(1.68)	(1.64)	(0.21)	(0.10)
Year†	−0.13	−0.04	−0.07***	−0.02***
	(0.11)	(0.10)	(0.02)	(0.01)
R^2	0.02	0.00	0.25	0.13
Period	1983-2016	1983-2016	1983-2016	1983-2016
N	34	34	34	34

† The variable Year is demeaned so that the intercept of the regression corresponds to the average of the dependent variable.

Notes. Panel A regresses the terms in the accounting decomposition on period dummies. Panel B regresses the terms in the accounting decomposition in (20) on year trends.

Estimation via OLS. Standard errors in parentheses and estimated using Newey-West with 3 lags. *,**,*** indicate significance at the 0.1, 0.05, 0.01 levels, respectively. Data from Forbes 400.

Table A2: Factor Model

	$R_{\text{top households}} - R_f$				
	Market	FF 3-factors	FF 5-factors	Bond Factors	Industry
	(1)	(2)	(3)	(4)	(5)
market	0.97***	1.08***	0.98***	0.98***	
	(0.26)	(0.27)	(0.29)	(0.26)	
smb		−0.88*	−0.84*		
		(0.51)	(0.50)		
hml		0.32	0.61		
		(0.36)	(0.54)		
cma			−0.76		
			(0.68)		
rmw			−0.05		
			(0.46)		
ltg				0.44	
				(0.55)	
crd				−0.03	
				(0.61)	
industry					1.01***
					(0.26)
Constant	−0.03	−0.05	−0.02	−0.06	−0.03*
	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)
R^2	0.31	0.39	0.42	0.33	0.34
Period	1983-2016	1983-2016	1983-2016	1983-2016	1983-2016
N	32	32	32	32	32

Notes. The table reports the results of regressing the wealth growth of top households on excess stock market returns, and a set of other factors. The left hand side is the three year excess wealth growth of top households, to allow for circumventing the stale pricing model of holdings outside private equity. More precisely, denote $R_{\text{top, t}}$ the average wealth growth of top at time t. I report the coefficients β_i obtained when estimating the linear model

$$(1 + R_{\text{top,t}})(1 + R_{\text{top,t+1}})(1 + R_{\text{top,t+2}}) - (1 + R_{ft})^3 = \alpha + \sum_{1 \leq i \leq f} \beta_i (R_{it} - R_{ft}) + \epsilon$$

Portfolio returns of Fama-French factor models, as well as industry portfolios, are from the Fama-French Data Library. Corporate bond returns are obtained from Ibbotson's Stocks, Bonds, Bills and Inflation Yearbook.
Estimation via OLS. Standard errors in parentheses and estimated using Newey-West with 3 lags. *,**,*** indicate significance at the 0.1, 0.05, 0.01 levels, respectively.

Table A3: Decomposing the Within Term

Year	R_{within}						
	Total (%)	$R_{\text{top households}}$ (%)					$-R_{\text{U.S.}}$(%)
		Total	R_f	$\beta_M(R_M - R_f)$	Labor-Tax	ϵ	
Panel A: R_M is Market Return							
All Years	1.9	5.8	1.6	6.5	–0.8	–1.6	–4.0
1983-1993	1.5	5.0	3.7	5.5	–0.8	–3.4	–3.5
1994-2004	1.6	7.4	2.0	6.8	–0.7	–0.7	–5.9
2005-2016	2.7	5.1	–0.6	7.3	–0.8	–0.7	–2.8
Panel B: R_M is Industry-Weighted Return							
All Years	1.8	5.6	1.6	6.5	–0.8	–3.0	–4.0
1983-1993	1.5	5.0	3.7	3.9	–0.8	–1.8	–3.5
1994-2004	1.6	7.4	2.0	7.3	–0.7	–1.1	–5.9
2005-2016	2.7	5.1	–0.6	7.5	–0.8	–0.9	–2.8

Notes. The table reports the decomposition of the within term R_{within} according to the theoretical model (A45) and (A46), using R_M as a benchmark return. In Panel A, the benchmark return is the (value-weighted) market return. In Panel B, the benchmark return is the industry-weighted return, using the industry composition of households in the top percentile. All returns in real terms. Data for the risk-free rate R_f and market returns come from Fama-French Data Library. Industries are defined using the Fama-French 49 industry classification. Data from Forbes 400.

Table A4: Decomposing the Demography Term

Year	$R_{\text{demography}}$								
	Total (%)	$R_{\text{death}} = \frac{\chi_t^{\zeta}-1}{\zeta}\delta_t$				$R_{\text{pop. growth}} = \left(1 - \frac{1}{\zeta}\right)\eta_t + \epsilon$			
		Total (%)	ζ	δ_t (%)	χ_t (%)	**Total** (%)	ζ	η_t (%)	ϵ (%)
All Years	–0.3	–0.7	1.5	1.8	54	0.4	1.5	1.2	0.0
1983-1993	–0.1	–0.7	1.8	1.9	60	0.6	1.8	1.3	0.0
1994-2004	–0.3	–0.8	1.4	1.9	56	0.4	1.4	1.4	0.0
2005-2016	–0.5	–0.7	1.4	1.6	46	0.3	1.4	0.9	0.0

Notes. The table reports the decomposition of the demography term $R_{\text{demography}}$ according to the theoretical model (A24). The death rate δ corresponds to the yearly death rate of households in the top percentile. The population growth rate η corresponds to the yearly growth of the U.S. population. The tail index is estimated using $\zeta - 1 = g_t(q_t)q_t^2/S_t$, where the density $g_t(q_t)$ is estimated from the mass of households with a wealth between q_t and $1.3q_t$.

Table A5: Comparison Method using Imputed Wealth of Drop-offs vs Reported Wealth

Year	$E\left[\frac{w_{t+1}}{w_t} - 1 \mid \text{Drop-off}\right]$ (%)		Wealth Share Drop-offs (%)	$E\left[\frac{w_{t+1}}{w_t} - 1\right]$ (%)	
	Imputed	Actual		Imputed	Actual
2011	−23.2	−37.1	1.6	7.7	7.5
2012	−28.0	−30.6	1.8	7.0	6.9
2013	−39.7	−19.9	1.7	1.8	2.1
2014	−32.9	−31.0	2.2	−3.2	−3.2
2015	−42.3	−50.8	3.1	2.6	2.4
2016	−41.7	−19.4	2.2	1.1	1.6
2011-2016	−34.6	−31.5	2.1	2.8	2.9

Notes. The table compares the estimate for the within term R_{within} obtained using imputed data compared to the wealth of drop-offs reported after 2011. The difference in R_{within} between the two methods is reported in Column (7). It can be obtained as the product of the difference in the estimate of the return of drop-offs in Column (3) times the share of wealth represented by drop-offs in Column (4).

Table A6: Wealth Reallocation Within Families

Year	$R_{\text{displacement}}$				
	Total (%)	Displacement Predicted by Variance $1/2(\zeta-1)\nu^2$ (%)			ϵ (%)
		Total	Within Families	Between Families	
All Years	2.3	2.0	0.1	1.9	0.3
1983-1993	3.0	2.9	0.2	2.8	0.0
1994-2004	2.5	1.9	0.1	1.8	0.6
2005-2016	1.4	1.2	0.0	1.2	0.2

Notes. The table decomposes the model-predicted displacement term $1/2(\zeta-1)\nu^2$ into a displacement "within" families $1/2(\zeta-1)\nu^2_{\text{within}}$ and a displacement "between" families $1/2(\zeta-1)\nu^2_{\text{between}}$. The decomposition follows from the law of total variance: the variance of wealth growth ν^2 is the sum of the average variance within groups ν^2_{within} and the variance between groups ν^2_{between}. Data from Forbes 400.

References

Acemoglu, Daron and James A Robinson, "The Rise and Decline of General Laws of Capitalism," *Journal of Economic Perspectives*, 2015, *29* (1), 3–28.

Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hémous, "Innovation and top income inequality," *The Review of Economic Studies*, 2015.

Aït-Sahalia, Yacine, Julio Cacho-Diaz, Tom R Hurd et al., "Portfolio choice with jumps: A closed-form solution," *The Annals of Applied Probability*, 2009, *19* (2), 556–584.

Atkinson, Anthony, "Concentration among the Rich," *Personal Wealth from a Global Perspective*, 2008, pp. 64–90.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, John Van Reenen et al., "The fall of the labor share and the rise of superstar firms," 2017.

Bach, Laurent, Laurent Calvet, and Paolo Sodini, "From Saving Comes Having? Disentangling the Impact of Savings on Inequality," 2017.

_ , Laurent E Calvet, and Paolo Sodini, "Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy," 2015. Working Paper.

Benhabib, Jess, Alberto Bisin, and Mi Luo, "Wealth distribution and social mobility in the US: A quantitative approach," Technical Report, National Bureau of Economic Research 2015.

_ , _ , and Shenghao Zhu, "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents," *Econometrica*, 2011, *79* (1), 123–157.

_ , _ , and _ , "The wealth distribution in Bewley economies with capital income risk," *Journal of Economic Theory*, 2015, *159*, 489–515.

Bessembinder, Hendrik, "Do stocks outperform treasury bills?," *Journal of financial economics*, 2018, *129* (3), 440–457.

Campbell, John Y, "Restoring rational choice: The challenge of consumer financial regulation," *American Economic Review*, 2016, *106* (5), 1–30.

_ , Martin Lettau, Burton G Malkiel, and Yexiao Xu, "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk," *The Journal of Finance*, 2001, *56* (1), 1–43.

Capehart, Kevin W., "Essays on the Wealthiest Americans." PhD dissertation, American University 2014.

Carroll, Christopher D and Miles S Kimball, "On the concavity of the consumption function," *Econometrica: Journal of the Econometric Society*, 1996, pp. 981–992.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda, "Declining business dynamism: What we know and the way forward," *American Economic Review*, 2016, *106* (5), 203–07.

_ , _ , _ , and _ , "Where has all the skewness gone? The decline in high-growth (young) firms in the US," *European Economic Review*, 2016, *86*, 4–23.

Emery, Kenneth, "Private equity risk and reward: Assessing the stale pricing problem," *The Journal of Private Equity*, 2003, pp. 43–50.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity in returns to wealth and the measurement of wealth inequality," *American Economic Review*, 2016, *106* (5), 651–55.

Gabaix, Xavier, "Power laws in economics and finance," *Annu. Rev. Econ.*, 2009, *1* (1), 255–294.

_ , Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The Dynamics of Inequality," *Econometrica*, 2016. Forthcoming.

Garbinti, Bertrand, Jonathan Goupille-Lebret, Thomas Piketty et al., "Accounting for Wealth Inequality Dynamics: Methods, Estimates and Simulations for France (1800-2014)," 2017. Working Paper.

Gârleanu, Nicolae and Stavros Panageas, "Finance in a Time of Disruptive Growth," 2017.

Guvenen, Fatih, Serdar Ozkan, and Jae Song, "The nature of countercyclical income risk," *Journal of Political Economy*, 2014, *122* (3), 621–660.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh, "The common factor in idiosyncratic volatility: Quantitative asset pricing implications," *Journal of Financial Economics*, 2016, *119* (2), 249–283.

Hubmer, Joachim, Per Krusell, and Anthony A Smith Jr, "The historical evolution of the wealth distribution: A quantitative-theoretic investigation," Technical Report, National Bureau of Economic Research 2016.

Jones, Charles I, "Pareto and Piketty: The macroeconomics of top income and wealth inequality," *The Journal of Economic Perspectives*, 2015, *29* (1), 29–46.

_ and Jihee Kim, "A Schumpeterian Model of Top Income Inequality," 2016. Working Paper.

Kaplan, Edward L and Paul Meier, "Nonparametric estimation from incomplete observations," *Journal of the American statistical association*, 1958, *53* (282), 457–481.

Karlin, Samuel and Howard E Taylor, *A second course in stochastic processes*, Elsevier, 1981.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological innovation, resource allocation, and growth," *The Quarterly Journal of Economics*, 2017, *132* (2), 665–712.

— , — , **and Noah Stoffman**, "Winners and Losers: Creative Destruction and the Stock Market," Forthcoming.

Kopczuk, Wojciech and Emmanuel Saez, "Top Wealth Shares in the United States: 1916-2000: Evidence from Estate Tax Returns," 2004. Working Paper.

— , — , **and Jae Song**, "Earnings inequality and mobility in the United States: evidence from social security data since 1937," *The Quarterly Journal of Economics*, 2010, *125* (1), 91–128.

Kuhn, Moritz, Moritz Schularick, and Ulrike I Steins, "Income and Wealth Inequality in America, 1949-2013," 2017. Working Paper.

Luttmer, Erzo, "Slow convergence in economies with firm heterogeneity," 2012.

Luttmer, Erzo FP, "Measuring economic mobility and inequality: Disentangling real events from noisy data," *Unpublished paper, University of Chicago, Chicago*, 2002.

Martin, Ian WR, "Consumption-based asset pricing with higher cumulants," *The Review of Economic Studies*, 2013, *80* (2), 745–773.

Melitz, Marc J and Sašo Polanec, "Dynamic Olley-Pakes productivity decomposition with entry and exit," *The Rand journal of economics*, 2015, *46* (2), 362–375.

Oh, Simon and Jessica A Wachter, "Cross-sectional skewness," Technical Report, National Bureau of Economic Research 2018.

Papanikolaou, Dimitris et al., "Technological innovation and the distribution of labor income growth," in "2018 Meeting Papers" number 242 Society for Economic Dynamics 2018.

Piketty, Thomas, *Capital in the twenty-first century*, Harvard University Press, 2014.

— **and Emmanuel Saez**, "Income inequality in the United States, 1913–1998," *The Quarterly journal of economics*, 2003, *118* (1), 1–41.

— **and Gabriel Zucman**, "Wealth and Inheritance in the Long Run," *Handbook of Income Distribution*, 2015, *2*, 1303–1368.

Raub, Brian, Barry Johnson, and Joseph Newcomb, "A Comparison of Wealth Estimates for America's Wealthiest Decedents Using Tax Data and Data from the Forbes 400," in "Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association," Vol. 103 JSTOR 2010, pp. 128–135.

Roussanov, Nikolai, "Diversification and its Discontents: Idiosyncratic and Entrepreneurial Risk in the Quest for Social Status," *Journal of Finance*, 2010, *65* (5), 1755–1788.

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

Wachter, Jessica A and Motohiro Yogo, "Why do Household Portfolio Shares Rise in Wealth?," *Review of Financial Studies*, 2010, *23* (11), 3929–3965.

Wang, Chong, Neng Wang, and Jinqiang Yang, "Optimal consumption and savings with stochastic income and recursive utility," *Journal of Economic Theory*, 2016, *165*, 292–331.

Wold, Herman OA and Peter Whittle, "A model explaining the Pareto distribution of wealth," *Econometrica, Journal of the Econometric Society*, 1957, pp. 591–595.