# Ups and Downs: How Idiosyncratic Volatility Drives Top Wealth Inequality*

Matthieu Gomez †

This version: October 2018
First version: November 2016

**Abstract**

This paper examines the role of idiosyncratic volatility in driving the recent rise in top wealth inequality. Because the composition of households in top percentiles changes over time, the growth of top wealth shares is not simply equal to the average wealth growth of households in top percentiles relative to the economy. It also depends on a displacement term, which is driven by the entry and exit of households in top percentiles. I relate analytically the displacement term to the dispersion of wealth shocks among top households. Using the Forbes 400 list, I document that the displacement term accounts for more than half the rise in top wealth inequality in the United States since 1983. I discuss the implications of this result for wealth mobility, as well as for the relationship between inequality and technological innovation.

*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, 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? A common hypothesis is that this phenomenon is driven by the rapid growth of top households relative to the overall economy, i.e. that "the rich are getting richer" (Piketty (2014) or Hubmer et al. (2016)). This view implicitly assumes that the composition of households in top percentiles remains constant over time. In reality, 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. These large composition changes naturally drive a wedge between the growth of top wealth shares and the wealth growth of households in the top percentiles. This paper examines the importance of these composition effects for the dynamics of top wealth shares.

The paper is organized in three parts. First, I relate in continuous time the dynamics of top wealth shares to the dynamics of individual wealth. The growth of the wealth share of a top percentile is the sum of three terms. The first term (a "within" term) corresponds to the average wealth growth of households in the top percentile relative to the rest of the economy. This is the "rich getting richer" driver of wealth inequality. The second term (a "displacement" term) corresponds to the wealth of households entering the top percentile minus the wealth of households exiting the top. This can be seen as a "rich getting displaced" driver of wealth inequality. The third term (a "demography" term), which I find to be empirically negligible, is driven by the death of households in the top percentile as well as population growth.

Intuitively, the displacement term is driven by the dispersion of wealth shocks for households at the top. To formalize this, I examine displacement when wealth follows a diffusion process (i.e. normal shocks). When the wealth distribution has a Pareto tail, I show that the displacement term equals $1/2(\zeta - 1)\nu^2$, where $\nu$ denotes the idiosyncratic volatility of wealth growth and $\zeta$ denotes the power law exponent of the wealth distribution. The displacement term increases with the variance of wealth shocks $\nu^2$ and decreases with wealth inequality (i.e. increases with $\zeta$). The higher the level of wealth inequality, the bigger the gap between wealthy households and the rest of the population, and the less likely it is for households to enter and exit top percentiles.

The rapid rise of a few entrepreneurs at the top of the distribution indicates that non-normal shocks may play an important role for displacement. I examine this hypothesis by studying the dynamics of top wealth shares when wealth follows a diffusion process with jumps (i.e. non-normal shocks). When the wealth distribution has a Pareto tail with power law exponent $\zeta$, the displacement term equals $\sum_2^{+\infty} 1/j!(\zeta^{j-1} - 1)\kappa_j$ where $\kappa_j$ denotes the $j$-th cumulant of wealth growth. All higher-order cumulants, not just the variance of wealth growth, matter for the growth of top wealth shares. Negative skewness tends to decrease top wealth shares while positive kurtosis tends to increase top wealth shares.

Second, I use this theoretical framework to document the role of displacement in the data. I present an accounting framework that decomposes the growth of top wealth shares into a within term, a displacement term, and a demography term, as suggested by the theory introduced above. Applying this framework on the share of wealth owned by the 400 wealthiest in the U.S., I find that the displacement term accounts for more than half the increase in top wealth inequality since 1983. More precisely, the 3.9% annual growth of top wealth inequality from 1983 to 2017 can be decomposed as follows: the within term accounts for an annual growth of 1.9%, the displacement term accounts for an annual growth of 2.3%, and the demography term is negligible. In other words, a researcher who neglects compositional changes would overestimate by a factor of two the average wealth growth of top households relative to the economy.

I use the theoretical model discussed above to shed light on the displacement term. With a Pareto tail $\zeta \approx 1.5$ and an 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%. Higher-order cumulants do not matter much for displacement. Intuitively, wealth inequality is so high that most of the entry in the top percentile is driven by households already close to the top percentile, rather than entrepreneurs from the bottom of the distribution with extremely high wealth realization. Moreover, most of displacement happens within industries rather than between industries: the disruption of oil fortunes by tech entrepreneurs only accounts for a small share of overall displacement.

I then use the 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 $\nu$ of wealth at these top percentiles by interacting the share of wealth invested in equity with the cross-sectional standard variance of firm-level returns. Overall, displacement follows an inverted U-shape over the 20th century. It first peaked during the Great Depression, remained low during the World Wars and the postwar economic boom, before peaking again during the technological revolutions of the 1980s and 1990s.

Third, I examine the implications of this displacement term along two dimensions. I first study the relation between wealth inequality and technological innovation. The existing literature suggests that innovation has an ambiguous effect on top wealth shares. On the one hand, a rise in technological innovation tends to reduce the market capitalization of incumbent firms, which tends to decrease the average growth of households in top percentiles, and therefore to decrease top wealth shares (Gârleanu et al. (2012), Jones and Kim (2016)). On the other hand, a rise in technological innovation increases the dispersion of wealth shocks among top, which tends to increase displacement in top percentiles, and therefore to increase top wealth shares (Kogan et al. (Forthcoming), Gârleanu and Panageas (2017)). Decomposing the growth of top wealth shares allows me to separate these two opposite effects of innovation. Overall, I find that the second effect dominates: technological innovation tends to increase top wealth shares. 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.

Finally, I examine the importance of this decomposition for wealth mobility. While both the within term and the displacement term increase inequality they have opposite effects on wealth mobility. More precisely, a rise in the average wealth growth of top households decreases mobility, whereas a rise in the dispersion of wealth shocks increases mobility. Using the decomposition above, I estimate that the average time a household in the top 0.01% remains in the top percentile decreased from 25 years in 1983 to 20 years now. The importance of displacement in the recent rise in top wealth shares suggests that wealth mobility will remain high in the 21rst century, even as wealth inequality continues to increase.

**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) and Garbinti 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) define 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.

This work adds to a growing literature that studies the dynamics of inequality through the lens of random growth models (Wold and Whittle (1957), Acemoglu and Robinson (2015), Jones (2015)). Luttmer (2012), Gabaix et al. (2016) and Jones and Kim (2016) develop tools to study the dynamics of wealth inequality 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

[1]To examine the impact of an increase of idiosyncratic volatility on top wealths shares, Gabaix et al. (2016) use

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) 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) calibrate consumption/saving models to match the recent rise in top wealth shares. By disentangle non-parametrically the within and the displacement term in the recent rise in top wealth shares, my paper provides new moments that could be used to further discipline these models.

Recent empirical papers stress the importance of idiosyncratic wealth shocks at the very top. Campbell (2016) proposes a decomposition of the change in the variance of log wealth into a term due to differences in expected wealth growth and a term due to differences in unexpected wealth shocks. Roussanov (2010) argues that rich households may be more likely to own assets with idiosyncratic risk if they care about their ranks in the distribution. Bach et al. (2015) and Fagereng et al. (2016) stress the dispersion of wealth growth across households, using administrative data from Sweden and Norway. My contribution is to quantify theoretically and empirically the role of the dispersion of wealth shocks in the dynamics of top wealth shares. Bach et al. (2017) use the decomposition presented in this paper to decompose the dynamics of top wealth shares in Sweden.

Finally, this paper contributes to a large literature on the importance of innovation for the overall economy. I document a strong relationship between displacement and innovation, which supports similar evidence in Aghion et al. (2015) and Kogan et al. (Forthcoming). Furthermore, I document a decline in displacement in the last two decades. This ties my paper to a growing literature documenting the secular decline in business dynamism (Decker et al. (2016a)), and in particular in the number of young high-growth firms since 2000 (Decker et al. (2016b)).

**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

Kolmogorov Forward 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 derives the dynamics of top wealth shares directly.

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 and jumps. Finally, Section 2.3 derives the dynamics of top wealth shares in presence of death and population growth.

## 2.1 Displacement in the Baseline Model

**Wealth Process.** I first examine the dynamics of top wealth shares in a very simple framework. I assume that wealth follows a geometric Brownian motion. More precisely, the wealth of household $i$ relative to the total wealth in the economy, $w_{it}$, follows the diffusion process:

$$\frac{dw_{it}}{w_{it}} = \mu_t dt + \nu_t dB_{it} \tag{1}$$

where $B_i = \{B_{it}, \mathcal{F}_t, t \geq 0\}$ is an idiosyncratic Brownian motion for household $i$, in a probability space $(\Omega, P, \mathcal{F})$ equipped with a filtration $F = \{\mathcal{F}_t, t \geq\}$ with the usual conditions. The instantaneous drift $\mu_t$ and the instantaneous volatility $\nu_t$ of wealth are allowed to depend on time.

I focus on the dynamics of $S_t$, the share of wealth owned by households in a top percentile $p$. Denote $g_t$ the density of relative wealth in the economy and $q_t$ the relative wealth of household at the percentile threshold (formally, the $1-p$ quantile). The top wealth share $S_t$ is simply given by the total amount of relative wealth owned by households above the threshold $q_t$:[2]

$$S_t = \int_{q_t}^{+\infty} w g_t(w) dw \tag{2}$$

**Proposition 1** (Dynamics of Top Wealth Share)**.** *When wealth follows the law of motion (1), the top wealth share $S_t$ follows the law of motion:*

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

[2]Here and for the rest of the paper, I assume that the top wealth share is finite, i.e., for a distribution with a heavy tail, that its power law exponent is higher than one.

The growth of the top wealth share $S_t$ does not only depend on the average wealth growth of households in top percentiles, $\mu_t$. It also depends on a term due to the variance of wealth shocks, $\nu_t^2$. For the rest of the paper, the first term is referred as the "within" term, and the second term is referred as the "displacement" term.

Proposition 1 can be seen as an integrated version of the Kolmogorov Forward equation.[3] Kolmogorov Forward equation relates the dynamics of the density of wealth to the dynamics of individual wealth:

$$dg_t(w) = -\mu_t dt \partial_w (w g_t(w)) + \frac{\nu_t^2 dt}{2} \partial_{ww}(w^2 g_t(w)) \tag{4}$$

This equation similarly decomposes the change in the wealth density into a term due to the average wealth growth and a term due to idiosyncratic volatility. Compared to Equation (3), this equation cannot be easily brought to the data, since it requires to estimate the first and second derivatives of the wealth density.

I now present a heuristic derivation for the expression of the displacement term. This derivation follows the graphical explanation given in Figure 1. During a short period of time $dt$, the Brownian motion can be approximated by a discrete process: with probability half, the wealth of a household is multiplied by $(1+\nu_t\sqrt{dt})$, otherwise, it is multiplied by $(1-\nu_t\sqrt{dt})$. This dispersion of wealth shocks generates entry and exit in the top percentile. First, households with a wealth between $q_t/(1+\nu_t\sqrt{dt})$ and $q_t$ with a positive 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 $q_t$. The total increase of $S_t$ due to these entries is therefore $q_t$: $\int_{\frac{q_t}{1+\nu_t\sqrt{dt}}}^{q_t} \frac{((1+\nu_t\sqrt{dt})w-q_t)}{S_t}\frac{g_t(w)}{2}dw$. Using the midpoint method, it can be written the product of the mass of households that enter the top percentile during the time period $dt$, $1/2 q_t \nu \sqrt{dt} g_t(q_t)$, and the average increase of top wealth share per entry, $q_t\nu_t\sqrt{dt}/(2S_t)$. Second, households with a wealth between $q_t$ and $q_t/(1-\nu_t\sqrt{dt})$ with a negative shock exit the top percentile. Each exiting household is replaced by a household at the lower percentile threshold, with wealth $q_t$. The total increase of $S_t$ due to these exits is therefore $\int_{\frac{q_t}{1+\nu_t\sqrt{dt}}}^{q_t} \frac{((1+\nu_t\sqrt{dt})w-q_t)}{S_t}\frac{g_t(w)}{2}dw$. Using the midpoint method, it can be written the product of the mass of households that exit the top percentile during the time period $dt$, $1/2 q_t \nu \sqrt{dt} g_t(q_t)$, and the average increase of top wealth share per exit, $q_t\nu_t\sqrt{dt}/(2S_t)$. Summing the term due to entry and the term due to exit, we obtain Proposition 1.

[3] It is actually derived by integrating the Kolmogorov Forward equation, see the proof of Proposition 1 in Appendix A.

**Heavy Tail.** One of the most ubiquitous regularities in economics and finance is that many distributions are well approximated by a power law. In this case, we can simplify Proposition 1.

**Definition 1.** *A distribution is Pareto if the CDF is a power law of wealth, i.e.:*

$$\mathbb{P}(w_{it} \geq w) = Cw^{-\zeta} \tag{5}$$

*$\zeta$ is called the power law exponent of the distribution.*

**Proposition 2** (Dynamics of Top Wealth Share with Pareto Distribution)**.** *Suppose that the wealth distribution at time $t$ is Pareto with power law exponent $\zeta$, and that the instantaneous law of motion of relative wealth $w_{it}$ is given by (1). Then, the top wealth share $S_t$ follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt}_{dr_{within}} + \underbrace{\frac{\zeta - 1}{2}\nu_t^2 dt}_{dr_{displacement}} \tag{6}$$

Proposition 2 gives a strikingly simple formula for the displacement term. It depends on only two parameters: $\zeta$, the power law exponent of the wealth distribution, and $\nu$, the geometric volatility of wealth shocks. For a distribution with a power law exponent $\zeta \approx 1.5$ and idiosyncratic variance equals to $\nu \approx 27\%$, we can expect the displacement term to average 2.0% per year, which is large.

The displacement term does not depend on the top percentile $p$. The role of idiosyncratic volatility for the dynamics of the wealth share is quantitatively the same everywhere in the distribution. Of course, this result relies on the strong assumption that the wealth distribution is Pareto everywhere. Proposition 2 can be adapted to the more general case in which the wealth distribution only has a heavy tail, i.e. $\mathbb{P}(w_{it} \geq w) = L(w)w^{-\zeta}$ where $L(w)$ is a slowly varying function.[4] In this case, Proposition 2 holds true in the limit, i.e.:

$$\frac{dS_t}{S_t} \sim \mu_t dt + \frac{\zeta - 1}{2}\nu_t^2 dt \text{ as } p \to 0 \tag{8}$$

The displacement term decreases in the power law exponent of the wealth distribution $\zeta$. Intuitively, as the distribution becomes more unequal ($\zeta$ decreases), displacement decreases for two reasons. First, there are fewer households near the lower percentile threshold, i.e. the mass of households that enter or exit the top percentile relative to the mass of households in the top percentile $\zeta\nu\sqrt{dt}$

[4] A slowly varying function $L$ is defined as

$$\lim_{w \to +\infty} \frac{L(tw)}{L(w)} \to 1 \tag{7}$$

decreases[5]. Second, the ratio of wealth at the lower percentile threshold to the average wealth above the percentile decreases, i.e. the average growth of top wealth share per entry or exit $(1-\frac{1}{\zeta})\nu\sqrt{dt}$ decreases. In the limit where $\zeta$ converges to one (Zipf's law), the displacement term converges to zero: idiosyncratic shocks have no impact on the growth of the top wealth share $S_t$.

While Proposition 2 assumes that the wealth distribution has a Pareto tail at time $t$, whether the wealth distribution continues to have a Pareto tail going forward depends on the exact specification of the model.[6] In a wide class of models, however, Gabaix et al. (2016) show that the right tail of the distribution moves very slowly.[7] Empirically, we will see that Equation (6) continues to be a good approximation of the displacement term along the transition path of the wealth distribution.

**Long Run.** When $\mu_t$ and $\nu_t$ are constant over time, the wealth distribution converges to a stationary wealth distribution that is Pareto.[8] In this case, Proposition 1 can be used to characterize the power law exponent of the stationary wealth distribution $\zeta$: it is the index such that the (positive) displacement term exactly compensates the (negative) within term:

$$0 = \mu dt + \frac{\zeta - 1}{2}\nu^2 dt \tag{9}$$

that is, $\zeta = 1 - 2\mu/\nu^2$. While this well-known formula for the power law exponent $\zeta$ is usually derived from the law of motion of the wealth density (Kolmogorov-Forward equation)[9], deriving it from the law of motion of top wealth shares makes it clear that it represents a balance equation for top wealth shares.

## 2.2 Displacement in Extended Models

Proposition 1 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, and jumps. I derive analytical expressions for the displacement term in all these cases

[5] See the heuristic derivation of Proposition 1 discussed above.

[6] In particular, whether there is a reflecting barrier at a certain level of wealth.

[7] In the particular setup of Proposition 2, one can show the instantaneous change in $g_t(q_t)q_t^2/S_t$ for a top percentile is zero. To prove it, combine the law of motion for the top wealth share given in Proposition 1, the law of motion for the quantile given in Proposition 7, and the law of motion for the wealth density from Kolmogorov-Forward Equation.

[8] One needs to impose that $\mu \leq 0$ and that there is a reflecting barrier at some low level of wealth, see Gabaix (1999).

[9] See for instance Gabaix (2009).

**Scale Dependence.** Proposition 1 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 drift and volatility of wealth depend on the wealth level $w$, that is

$$\frac{dw_{it}}{w_{it}} = \mu_t(w_{it})dt + \nu_t(w_{it})dB_{it} \tag{10}$$

where $\mu$ and $\nu$ are differentiable functions of wealth $w_{it}$. In this case, I show in the appendix that the top wealth share $S_t$ follows the law of motion:[10]

$$\frac{dS_t}{S_t} = \underbrace{\mathrm{E}^{gw}[\mu_t(w)|w \geq q_t]dt}_{dr_{\text{within}}} + \underbrace{\frac{g_t(q_t)q_t^2}{2S_t}\nu_t(q_t)^2 dt}_{dr_{\text{displacement}}} \tag{11}$$

where $\mathrm{E}^{gw}$ 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.

Remarkably, 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.

This equation reveals that the law of motion of $S_t$ does not depend on the dynamics of wealth of households below the top percentile. Therefore, Proposition 1 still holds true when wealth is a geometric Brownian motion only above the lower threshold $q_t$.

**Heterogeneity.** Proposition 1 was derived under the assumption that all households have the same process for relative wealth. In reality, different households may have different average wealth growth, different exposure to aggregate risks, 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, and that the relative wealth of households in group $n$ evolves according to

[10] See Proposition 8 in Appendix A.

the law of motion:

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

where $Z_t = \{Z_{it} \in \mathbb{R}^d, \mathcal{F}_t, t \geq 0\}$ is a $d-$dimensional aggregate Brownian motion. Each group $n$ has a different exposure to aggregate risk given by $\sigma_{nt}$. 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).

In this case, I show in the appendix that the top wealth share $S_t$ follows the law of motion:[11]

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

where $\mathrm{E}^{gw}$ (resp. $\mathrm{Var}^{gw}$) denotes the wealth-weighted cross-sectional average (resp. variance) along the wealth distribution.

The within term can still be interpreted as the total wealth growth of households in the top percentile. It is the sum of a drift term and a stochastic term, where the stochastic term is driven by the wealth-weighted average exposure to aggregate risk of households in the top.

The displacement term now depends on the average of idiosyncratic variance at the threshold $\mathrm{E}^{gw}[\nu_{nt}^2|w_{it} = q_t]$, but also on the cross-sectional variance in exposure for households at the lower percentile threshold $\mathrm{Var}^{gw}[\sigma_{nt}|w_{it} = q_t]$. Heterogeneous exposures to aggregate risks — not just idiosyncratic risk — may drive the displacement term. The displacement term can still be interpreted as the cross-sectional variance of the wealth growth of individuals around the wealth threshold $q_t$.

For short time periods, fluctuations in the top wealth share $S_t$ are driven by fluctuations in the wealth growth of households in the top, rather than fluctuations in displacement. This is because the within term is exposed to aggregate risk $dZ_t$, while the displacement term is not. Over the long-run, however, these aggregate shocks average out, and the effect of displacement becomes more apparent.

**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

[11] See Proposition 10 in Appendix A.

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).

I derive the dynamics of the top wealth share $S_t$ when wealth follows a jump-diffusion process. I focus here on the case in which the wealth distribution has a Pareto tail with power law exponent $\zeta$ at time $t$. Under certain conditions, I show in the appendix that the growth of the top wealth share $S_t$ is given by:[12,13]

$$\frac{dS_t}{S_t} = \frac{1}{\zeta}\mathrm{E}_t\left[\frac{dw_{it}^{\zeta}}{w_{it}^{\zeta}}\right] \tag{14}$$

Adding and subtracting the instantaneous growth of wealth gives the growth of the top wealth share $S_t$ as the sum of a within term and a displacement term:

$$\frac{dS_t}{S_t} = \underbrace{\mathrm{E}_t\left[\frac{dw_{it}}{w_{it}}\right]}_{dr_{\text{within}}} + \underbrace{\frac{1}{\zeta}\mathrm{E}_t\left[\frac{dw_{it}^{\zeta}}{w_{it}^{\zeta}}\right] - \mathrm{E}_t\left[\frac{dw_{it}}{w_{it}}\right]}_{dr_{\text{displacement}}} \tag{15}$$

The displacement term is positive and increases in $\zeta$, as in the baseline model.[14] By expanding $\mathrm{E}_t[dw_{it}^{\zeta}/w_{it}^{\zeta}]$ as a power series in $\zeta$, the displacement term can be re-written as the sum of the cumulants of wealth growth:[15]

$$\begin{aligned} dr_{\text{displacement}} &= \sum_{j=2}^{+\infty}\frac{\zeta^{j-1}-1}{j!}\kappa_{jt}dt \\ &= \frac{\zeta-1}{2}\nu_t^2 dt + \frac{\zeta^2-1}{6}\nu_t^3\cdot\text{skewness}\cdot dt + \frac{\zeta^3-1}{24}\nu_t^4\cdot\text{excess kurtosis}\cdot dt + \dots \end{aligned} \tag{16}$$

This formula generalizes the displacement term for the case of jumps. The displacement term depends on the variance, but also on all higher-order cumulants of wealth growth. A negative skewness tends to decrease the displacement term while a positive kurtosis tends to increase the displacement term.

As the power law exponent of the wealth distribution $\zeta$ increases, the importance of higher-order cumulants increases relative to the variance component. To take a simple example, going from a

[12] See Proposition 11 in Appendix A.

[13] This expression characterizes the power law exponent $\zeta$ of the eventual stationary distribution, giving a continuous-time version of Champernowne (1953).

[14] See Proposition 12 in Appendix A.

[15] Formally, I define $\kappa_{jt}$ as the $j$-th coefficient in the Taylor expansion of $\zeta \to \mathrm{E}_t\left[dw_{it}^{\zeta}/w_{it}^{\zeta}\right]$, i.e.

$$\mathrm{E}_t\left[\frac{dw_{it}^{\zeta}}{w_{it}^{\zeta}}\right] = \sum_{j=1}^{+\infty}\frac{\zeta^j}{j!}\kappa_{jt}dt$$

distribution with $\zeta = 1.5$ (the power law exponent of the wealth distribution) to a distribution with $\zeta = 2.5$ (the power law exponent of the distribution of labor income), the term due to the variance of wealth shocks is multiplied by 3, while the term due to kurtosis is multiplied by 6. 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 households close to the percentile threshold (i.e. due to the variance of wealth growth).

### 2.3 Demography Term

**Demographic Forces.** For simplicity, the preceding analysis assumed away any change in top wealth shares due to demography. In reality, households at the top may die, which changes the composition of households at the top. 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.

I model death by assuming that households in the top percentile $p$ die with a hazard rate $\delta$. The hazard rate $\delta$ can vary over time, though I omit the time subscript for notational simplicity. 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 $\chi \in [0, 1]$ of their initial wealth. Other newborns are born below the top percentile.

The parameter $\chi$ controls the extent to which top fortunes are able to maintain themselves. When $\chi = 100\%$ (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 $\chi = 0\%$ (i.e. "no inheritance") there is no transmission of wealth across generations. Economically, the fraction of wealth that cannot be passed to offspring, $1 - \chi$, can be interpreted as the average estate tax.

Finally, I assume that population grows with rate $\eta$. Like the death rate $\delta$, the population growth rate $\eta$ can vary over time, though I omit the time subscript for notational simplicity.

**Proposition 3** (Dynamics of Top Wealth Share with Demographic Forces)**.** *When wealth follows the law of motion (1), with death rate $\delta$, inheritance parameter $\chi$, and population growth $\eta$, the top wealth share $S_t$ follows the law of motion:*

$$\frac{dS_t}{S_t} = dr_{within} + dr_{displacement} + dr_{demography} \tag{17}$$

*where the within term $dr_{within}$ and the displacement term $dr_{displacement}$ are defined in Proposition 1,* *and where the demography term $dr_{demography}$ is given by:*

$$dr_{demography} \equiv \underbrace{\left(\frac{\chi S_t(\pi_t p) + (1-\pi_t)q_t p}{S_t} - 1\right)\delta dt}_{dr_{death}} + \underbrace{\frac{q_t p}{S_t}\eta dt}_{dr_{pop.\ growth}} \tag{18}$$

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

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 is always negative. It increases with the degree of inheritance $\chi$. When $\chi = 100\%$ (i.e. perfect inheritance), $\pi_t = 1$ and therefore the term equals 0. In this case, death does not decrease the growth of top wealth share $S_t$. By contrast, the term due to population growth is always positive.[16] As population grows, the top percentile $p$ includes more and more households, which increases the top wealth share $S_t$. Overall, the demography term has an ambiguous sign.

I now present an heuristic derivation for the term due to demographic forces. Start with the term due to death. Between $t$ and $dt$, a mass $\delta p dt$ of households in the top die, which decreases total wealth in the top percentile by $S_t \delta dt$. A proportion $\pi_t$ of these households has a wealth high enough that their newborn offspring enters the top percentile. Since the average wealth of these offsprings 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 $(\chi S_t(\pi_t p) + (1-\pi_t)q_t p - S_t)\delta dt$. I now turn to the term due to population growth. Between $t$ and $dt$, a new mass $\eta p dt$ 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 $\eta p q_t dt$. The total increase of $S_t$ due to population growth is $\eta p q_t dt$.

**Heavy Tail.** Similarly to the preceding analysis in Section 2.1, the demography term dramatically simplifies when the wealth distribution is Pareto.

**Proposition 4** (Dynamics of Top Wealth Share with Demographic Forces and Pareto Distribution)**.** *Suppose that the wealth distribution at time $t$ is Pareto with power law exponent $\zeta$, and that the*

[16] 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 $\theta$ the ratio between the wealth of a newborn household relative to the average wealth a household in the economy, the within term would equal $\mu dt + \theta\eta dt$ while the population growth term would equal $\left(\frac{q_t p}{S_t} - \theta\right)\eta dt$. In any case, the displacement term remains unchanged.

*instantaneous law of motion of relative wealth $w_{it}$ is given by (1), with death rate $\delta$, inheritance parameter $\chi$, and population growth $\eta$. Then the top wealth share $S_t$ follows the law of motion:*

$$\frac{dS_t}{S_t} = dr_{within} + dr_{displacement} + dr_{demography} \tag{19}$$

*where the within term $dr_{within}$ and the displacement term $dr_{displacement}$ are given in Proposition 2, while the demography term $dr_{demography}$ is given by:*

$$dr_{demography} \equiv \underbrace{\frac{\chi^{\zeta} - 1}{\zeta}\delta dt}_{dr_{death}} + \underbrace{\left(1 - \frac{1}{\zeta}\right)\eta dt}_{dr_{pop.\ growth}} \tag{20}$$

The term due to demography is greatly simplified when the wealth distribution has a Pareto tail. The term due to death now depends on only three parameters: $\zeta$, the power law exponent of the wealth distribution, $\delta$, the death rate of top households, and $\chi$, the fraction of wealth that can be passed to offspring. The term due to population growth depends on only two parameters: $\zeta$, the power law exponent of the wealth distribution, and $\eta$, the population growth rate.

Like the displacement term, the demography term increases in $\zeta$. As $\zeta$ 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.

Quantitatively, we can expect the demography term to be small. To take realistic parameters, for a distribution with a power law exponent $\zeta \approx 1.5$, a death rate $\delta \approx 1.5\%$, a fraction of wealth passed to offspring $\chi \approx 60\%$,[17] and a population growth rate $\eta \approx 1\%$, we obtain the demography increases the growth rate of top wealth shares by $dr_{\text{demography}} \approx -0.2\%$ per year, which is small.

## 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 share into a within term, a displacement term, and a demography term using panel data.

**Case without Demographic Forces.** I first present the accounting decomposition in the case without demographic forces, i.e. without death or population growth. This makes it easier to

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

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:

$$\frac{S_{t+\tau}-S_t}{S_t}=\underbrace{\frac{\sum_{i\in\mathcal{T}} w_{it+\tau}}{\sum_{i\in\mathcal{T}} w_{it}}-1}_{R_{within}}+\underbrace{\frac{\sum_{i\in\mathcal{E}} w_{it+\tau}-\sum_{i\in\mathcal{X}} w_{it+\tau}}{\sum_{i\in\mathcal{T}} w_{it}}}_{R_{displacement}} \tag{21}$$

The within term $R_{\text{within}}$ is the wealth change for households in the top percentile at time $t$, whether or not these 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:

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

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 Demographic Forces.** 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 $\mathcal{T}' = (\mathcal{T} \cup \mathcal{E} \cup \mathcal{E}_\mathcal{D}) \setminus (\mathcal{X} \cup \mathcal{X}_\mathcal{D})$.[18]

**Proposition 5** (Accounting Decomposition)**.** *The growth of the top wealth share $S_t$ between $t$ and $t+\tau$ can be decomposed as follows:*

$$\frac{S_{t+\tau} - S_t}{S_t} = R_{within} + R_{displacement} + R_{demography} \tag{23}$$

*where the within term $R_{within}$ is defined as*

$$R_{within} \equiv \frac{\sum_{i \in \mathcal{T} \setminus \mathcal{X}_\mathcal{D}} w_{it+\tau}}{\sum_{i \in \mathcal{T} \setminus \mathcal{X}_\mathcal{D}} w_{it}} - 1 \tag{24}$$

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

$$R_{displacement} \equiv \frac{\sum_{i \in \mathcal{E}} (w_{it+\tau} - q_{t+\tau})}{\sum_{i \in \mathcal{T}} w_{it}} + \frac{\sum_{i \in \mathcal{X}} (q_{t+\tau} - w_{it+\tau})}{\sum_{i \in \mathcal{T}} w_{it}} \tag{25}$$

*and the demography term $R_{demography}$ is defined as*[19]

$$R_{demography} \equiv \underbrace{\frac{\sum_{i \in \mathcal{E}_\mathcal{D}} w_{it+\tau} + (|\mathcal{X}_\mathcal{D}| - |\mathcal{E}_\mathcal{D}|)\, q_{t+\tau} - \sum_{i \in \mathcal{X}_\mathcal{D}} (1 + R_{within}) w_{it}}{\sum_{i \in \mathcal{T}} w_{it}}}_{R_{death}} + \underbrace{\frac{(|\mathcal{T}'| - |\mathcal{T}|) q_{t+\tau}}{\sum_{i \in \mathcal{T}} w_{it}}}_{R_{pop.\ growth}} \tag{26}$$

The new demography term $R_{\text{demography}}$ is the sum of a term due to death $R_{\text{death}}$ 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.[20] 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,[21] when the wealth of households in the top percentile follows the law of motion (1), the decomposition converges to the theoretical decomposition presented in Proposition 3 as the time period $\tau$ tends to zero. Therefore, this accounting framework can be used to disentangle the different drivers of top wealth shares.

[18] 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.

[19] $|.|$ denotes the number of elements in the set .

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

[21] See the proof of Proposition 5.

# 4 Empirical Analysis

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.[22] Because Forbes nominatively identifies the 400 wealthiest individuals in the U.S, one can track the wealth of the same individuals over time,[23] 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.[24] 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.[25] To obtain the wealth share of this percentile by dividing 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

[22] 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."

[23] 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.

[24] 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.

[25] 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.

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 the period.

### 4.2 Results

**Description.** 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%.

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 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.

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.[26]

**Displacement Through the Lens of the Model.** I examine the displacement term through the lens of the theoretical framework laid out in Section 2.[27] 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 $\zeta$ is the power law exponent of the wealth distribution and $\nu^2$ is the idiosyncratic variance of wealth growth. To compare the prediction of this model with the actual displacement term, I estimate the power-law exponent of the wealth distribution $\zeta$ as well as the standard deviation of

[26]See Equation (13).

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

wealth growth $\nu$ in Table 2. I obtain a model-predicted displacement term equal to 2.0%, which comes from an average $\zeta$ equal to 1.5 and an average $\nu$ equal to 27%. This 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 $\kappa_{jt}$ denotes the $j-$th cumulant of wealth growth.[28] 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 a power law exponent 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. Intuitively, wealth inequality is so high that entry in the top percentile is mostly driven by households already close to the percentile threshold, rather than entrepreneurs from the bottom of the distribution with extremely high wealth realization (similarly, exit is mostly driven by households already close to the percentile threshold rather than households from the top of the distribution).

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.

**What Drives Idiosyncratic Volatility?** I now examine the role of firm-level returns in driving the dispersion of wealth shocks for households in the top percentile. To test this hypothesis, 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 $\nu_{\text{stocks}}$ 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

[28]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.

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.

**Explaining the Decline in Displacement.** 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 $\nu$ and a decline in the shape of the wealth distribution $\zeta$ 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 power law exponent 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.[29]

**Within and Between Industries.** 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

[29]Formally, the power law exponent of the stationary wealth distribution decreases with the idiosyncratic volatility of wealth, as shown in Equation (9). Appendix D explores the dynamics of the power law exponent after changes in idiosyncratic volatility.

[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.

time: the only time when the displacement between industries is quantitatively important is the height of the dot-com bubble.[32]

### 4.3 Robustness

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 5.

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 over-estimate 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 6 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 $-2\rho$, where $\rho$ 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

[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.

[33] This includes in particular Donald Trump.

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 (Piketty (2017)). 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 displacement term can be approximated by the term predicted by the diffusion model $1/2(\zeta - 1)\nu^2$.

**Methodology.** In this section, I proxy for the displacement term for the top $1\%, 0.1\%$, and $0.01\%$ from 1916 to 2012 using the term predicted by the diffusion model $1/2(\zeta-1)\nu^2$. In the simplest model presented in Section 2, the displacement term $1/2(\zeta - 1)\nu^2$ does not depend on the top percentile $p$. This is driven by two key assumptions: the wealth distribution is assumed to be exactly Pareto and the idiosyncratic volatility is assumed to be the same for all households. In a more realistic setting, as shown in Proposition 8, the displacement term is $\frac{g_t(q_t)q_t^2}{2S_t}\nu_t^2(q_t)$, which varies along the wealth distribution if $\frac{g_t(q_t)q_t^2}{2S_t}$ varies along the wealth distribution, or if the idiosyncratic volatility of wealth $\nu_t^2(q_t)$ depends on the wealth level.

I first estimate the shape of the wealth distribution $\frac{g_t(q_t)q_t^2}{2S_t}$ 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 7 reports the estimated $\zeta$ for top percentiles. Over the time period, $\zeta$ equals 1.5 for the Top 1% and 1.7 for the Top 0.01%. The

[34] See Footnote 24.

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

estimate $\zeta$ does not move much in the right tail of the distribution, 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 7 reports the estimated $\nu$ for top percentiles. Over the time period, $\nu$ equals 14% for the Top 1%, and 21% for the top 0.01%. The fact that $\nu$ increases in the right tail of the distribution reflects the fact that top percentiles tend to invest more in equity. .

**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 in 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 $\nu^2$ for the top 0.01%. Most of the fluctuations in the model-predicted displacement term arises from fluctuation in the idiosyncratic variance of wealth rather than fluctuations in the power-law exponent 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 7 suggest that the differences in displacement between the two percentiles can explain almost half of this gap.[36]

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

# 6 Implications

## 6.1 Inequality and Technological Innovation

I now examine the link between the growth of top wealth shares and technological innovation in light of the accounting decomposition done in the previous sections.

One strand of the literature argues that a rise in technological innovation reduces inequality. For instance, in Jones and Kim (2016), in periods of high innovation, existing businesses are disrupted, which decreases the average growth of households at the top. Similarly, in Gârleanu et al. (2012), positive shocks to the productivity of new capital depresses the returns of existing firms, which decreases the average wealth growth of existing households. On the other hand, another strand of the literature argues that a rise in technological innovation may increase inequality. In Kogan et al. (Forthcoming) and Gârleanu and Panageas (2017), when technological innovation is high, innovative entrepreneurs steal market share from less innovative entrepreneurs, which increases the dispersion of wealth shocks among top households.

My empirical decomposition allows me to tease out these opposite effects by examining separately the effect of innovation on the within term and on the displacement term. I proxy for technological innovation using the economic value of patents issued during the year. This measure is constructed by Kogan et al. (2017) as follows: first, the value of each patent issued from public firms is estimated using the stock market's response to news about the patent, second, the measure of economy-wide innovation is defined by aggregating the value of all patents every year, normalized by the total market capitalization in the economy.

I examine the relation between innovation and the dispersion of wealth growth in Table 8. In Column (1) I regress the variance of the log wealth growth of households in the top on aggregate patent activity. The estimate for the slope is strongly significant, with a $R^2$ equal to 36%: patent innovation correlates strongly with the variance of log wealth growth. In Column (2) of Table 8, 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 inequality. 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 $\zeta$.[37] 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. In conclusion, the accounting decomposition allows me to measure more precisely the role of innovation for top wealth inequality.

In conclusion, in periods of high innovation, the dispersion of wealth shocks rises which tends to increase top wealth shares through displacement. This time series evidence relates this paper to the cross sectional evidence of Aghion et al. (2015), that document a positive relationship between innovation and top income inequality across U.S. states.

The reason innovation increases the dispersion of wealth shocks is that, when innovation is high, households owning the innovative firms enter the top percentile, displacing the households that own the non innovative firms. To test this channel directly, I regress a measure of firm innovation on a dummy that is equal to one if the household enters the top in the year, and zero if the household is already at the top in Table 9. As a proxy for the innovation of each firm in a given year, Column (1) uses the number of patents 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). To compare household within the same year and industry, regressions are done with year and industry fixed effects. 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.

### 6.2 Inequality and 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

[37] Since innovation is serially correlated, high innovation today is correlated with high innovation in previous years, and therefore with a low power law exponent of the wealth distribution $\zeta$. See Appendix D.

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 this notion of "downward" mobility is that it only depends on the wealth dynamics of individuals in the right tail of the distribution.[38] 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.

$$T_q(w) \equiv \mathrm{E}[\inf\{\tau \text{ s.t. } w_{it+\tau} \leq q \text{ or } i \text{ dies }\}|w_{it} = w] \tag{27}$$

In the remaining of the section, I assume that the law of motion of wealth is given by

$$\frac{dw_{it}}{w_{it}} = \mu dt + \nu dB_{it} \tag{28}$$

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* (28)*, the average first passage time for $w \geq q$ is:*[39]

$$T_q(w) = \frac{1}{\delta}\left(1 - \left(\frac{w}{q}\right)^{\zeta_-}\right) \tag{29}$$

*where $\zeta_-$ 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 $\zeta_-$ 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 $\mu$ but decreases in the idiosyncratic volatility $\nu$.[40] Intuitively, the higher the dispersion of wealth

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

[39] 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.

[40] See the proof in Appendix E.

shocks, the more likely it is to have a negative wealth shock, 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,

$$T(p) \equiv \mathrm{E}^g[T_q(w_{it})|w_{it} \geq q] \tag{30}$$

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

**Proposition 6** (Average First Passage Time for an Average Household with Inheritance and Population Growth)**.** *Consider the stationary distribution in an economy where household wealth follows the process given in (A28) with death rate $\delta$, inheritance parameter $\chi$, and population growth $\eta$. Then, the average time someone in the top percentile $p$ remains in the top percentile is:*

$$T(p) = \frac{1}{\delta(1 - \zeta_+/\zeta_-)} \tag{31}$$

*where $\zeta_-$ is defined in Lemma 1, i.e. the negative zero of $\zeta \to \mu\zeta + \frac{\zeta(\zeta-1)}{2}\nu^2 - \delta$ and $\zeta_+$ denotes the Pareto tail of the stationary wealth distribution, i.e. the positive zero of $\zeta \to \mu\zeta + \frac{\zeta(\zeta-1)}{2}\nu^2 + (\chi^\zeta - 1)\delta + (\zeta - 1)\eta$*

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 $\zeta_+$ and $\zeta_-$. Intuitively, $-\zeta_-$ controls the average first passage time from a given distance to the threshold, while $\zeta_+$ corresponds to the power law exponent 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 $\mu$ 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 ($\zeta_+$ decreases). These two forces combine to decrease mobility.

In contrast, as the idiosyncratic volatility of wealth $\nu$ increases, $T$ tends to decrease (i.e. mobility increases). On the one hand, as $\nu$ increases, the average first passage time before at a given wealth level decreases, which tends to increase mobility ($-\zeta_-$ decreases). On the other, 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. $\zeta_+$ decreases). For realistic parameters, this long-run effect on the wealth distribution is not strong enough to compensate the first force. Overall, mobility increases.

To take a simple example, suppose that, in the initial pre-1980 economy, the relative wealth growth of top households is $\mu = 0\%$, the idiosyncratic volatility is $\nu = 10\%$, the death rate is $\delta = 2\%$ with inheritance parameter $\chi \approx 50\%$, and the population rate is 1.5%.[41] Proposition 6 says that, in this economy, the average time a top household remains in a top percentile is $T \approx 25$ years. Now, suppose that the relative wealth growth of top households increases permanently to $\mu = 2\%$ and that the idiosyncratic volatility of wealth growth increases to $\nu = 27\%$.[42] Applying Proposition 6, I obtain that the average time a top household remains at the top becomes $T \approx 20$ years. Even though wealth inequality increase between these two states, wealth mobility slightly increases. This supports 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. My theoretical framework suggests that wealth mobility will remain higher even as wealth inequality continues to increase.

# 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 $\zeta$ denotes the power law exponent of the wealth distribution and $\nu$ denotes the idiosyncratic volatility of wealth. This formula is useful to

[41] 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. I choose the drift $\mu$ so that the Pareto tail of the stationary wealth distribution is 1.8.

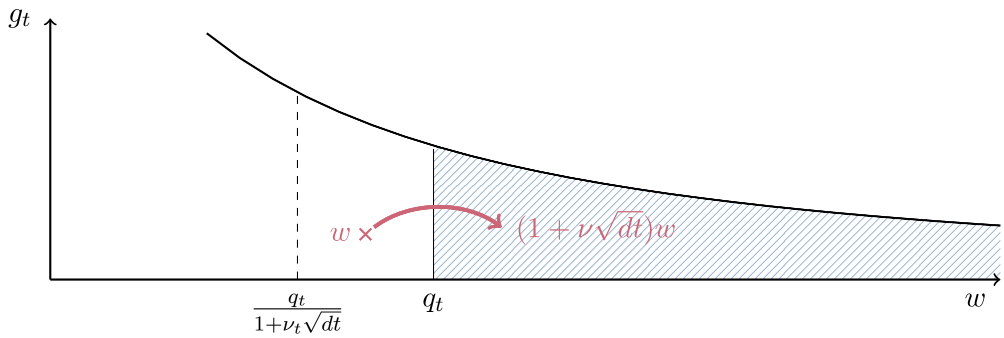
[42] The value of $\mu$ and $\nu$ are taken from the average accounting decomposition (see Table 1).

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 connexion is an important direction for future research.

Figure 1: Growth of Top Wealth Share $S_t$ due to Idiosyncratic Volatility

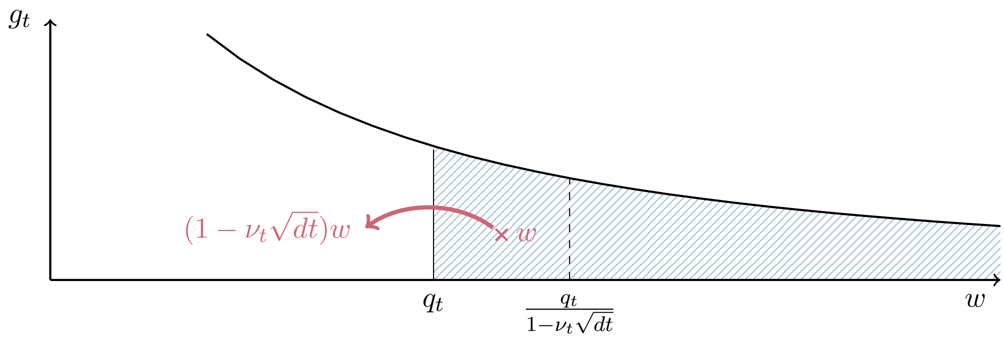
(a) Growth of $S_t$ due to Entry




During a short period of time $dt$, individuals below the top percentile $p$ with a positive wealth shock may enter the top percentile. The entry of an individual with initial wealth $w$ increases wealth in the top by the difference between their new wealth and the wealth of the last individual in the top, i.e. $(1+\nu_t\sqrt{dt})w - q_t$. Summing this quantity over the mass of individuals with initial wealth between $q_t/(1+\nu_t\sqrt{dt}$ and $q_t$, the growth of the top wealth share due to entry can be written as:

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

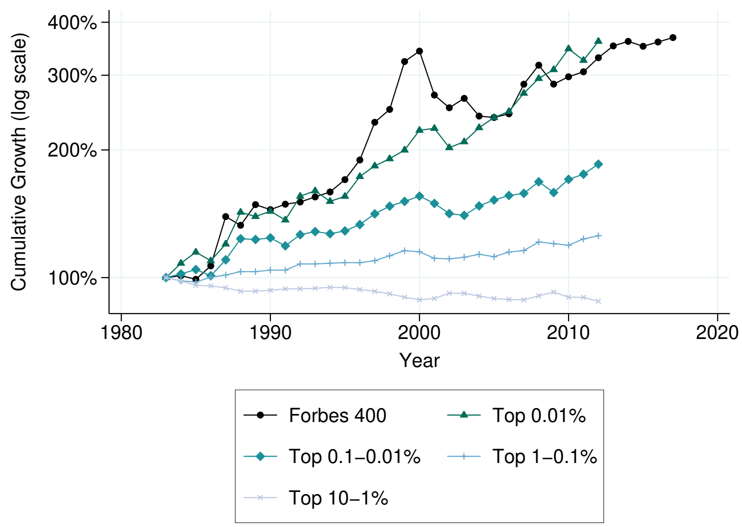
(b) Growth of $S_t$ due to Exit




During a short period of time $dt$, individuals in the top percentile $p$ with a negative wealth shock may exit the top percentile. The exit of an individual with initial wealth $w$ increases wealth in the top by the difference between the wealth of the last individual at the top that replaces them and their new wealth, i.e. $q_t - (1-\nu_t\sqrt{dt})w$. Summing this quantity over the mass of individuals with initial wealth between $q_t$ and $q_t/(1-\nu_t\sqrt{dt}$, the growth of the top wealth share due to exit can be written as:

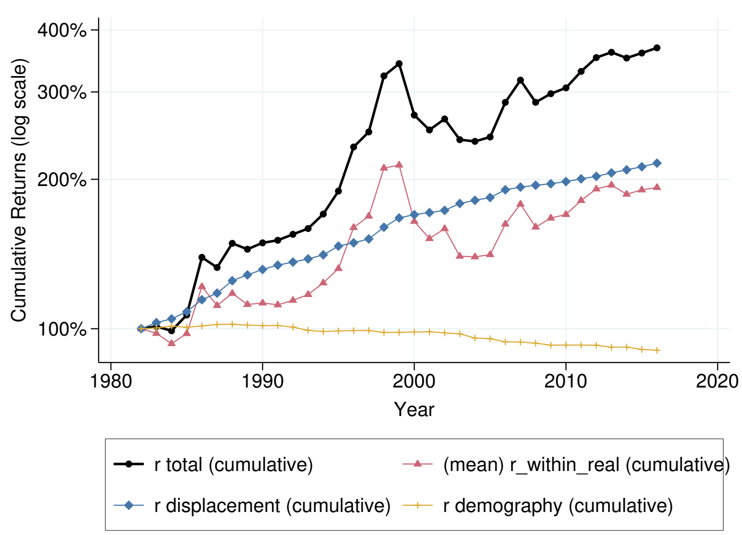
$$dr_{\text{exit}} \approx \int_{q_t}^{q_t/(1-\nu_t\sqrt{dt})} \frac{q_t - (1-\nu_t\sqrt{dt})w}{2S_t} g_t(w)dw \approx \frac{g_t(q_t)q_t^2}{4S_t}\nu_t^2 dt$$

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).

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 (21). It plots the cumulative log terms, i.e. the sum of log terms from 1983 to $t$. The plot 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.

Figure 4: Displacement Term: Data vs Theory

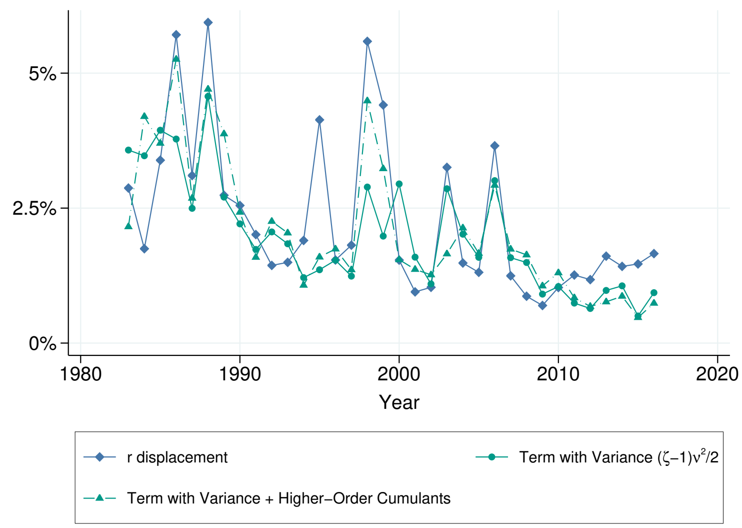



Figure 4 plots the displacement term (defined in Proposition 5) as well as the term predicted by the random-growth model for a diffusion model (normal shocks) and a jump diffusion (non-normal shocks). The power law exponent $\zeta$ is estimated yearly using $\zeta - 1 = g_t(q_t)q_t^2/S_t$, where the density $g_t(q_t)$ is estimated from the mass of households with a wealth 30% higher or lower than $q_t$. The variance 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.

Figure 5: Contribution of $\zeta$ and $\nu$ to Displacement

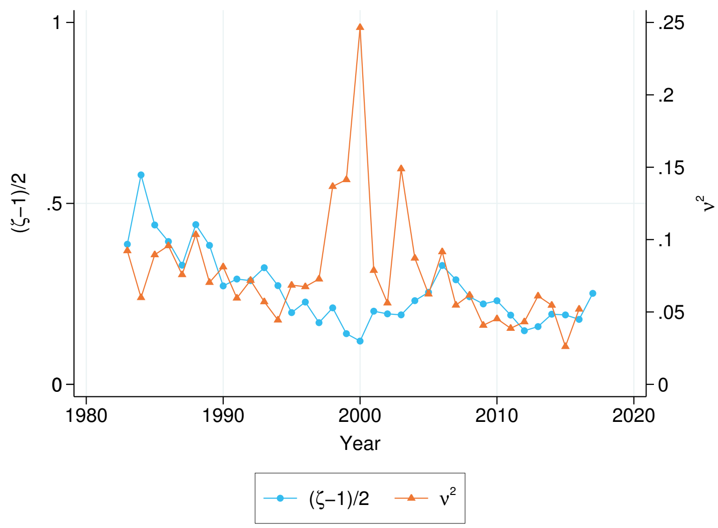
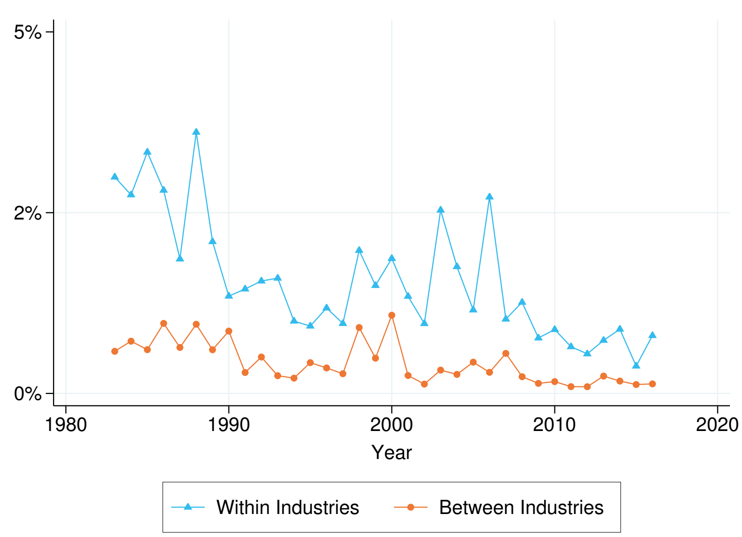



Figure 5 decomposes the model-predicted term into its two components, $(\zeta - 1)/2$ and $\nu^2$. The power law exponent of the wealth distribution $\zeta$ is estimated yearly using $\zeta - 1 = g_t(q_t)q_t^2/S_t$, where the density $g_t(q_t)$ is estimated from the mass of households with a wealth 30% higher or lower than $q_t$. The variance of wealth growth $\nu^2$ 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.

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 $\nu^2$ is the sum of the average variance within groups $\nu^2_{\text{within}}$ and the variance between groups $\nu^2_{\text{between}}$. Industries are defined using the Fama-French 49 industry classification. Data from Forbes 400.

Figure 7: Model-Predicted Displacement Term

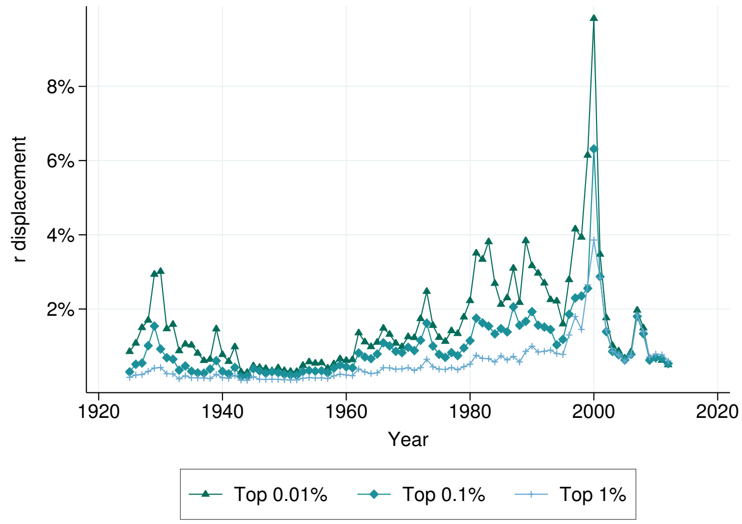



Figure 7 plots the model-predicted displacement term $(\zeta - 1)/2\nu^2$ for the Top 0.01%, 0.1%, and 1%. The power law exponent $\zeta$ is estimated as $1/(1 - \frac{q_t}{S_t/p})$. The idiosyncratic volatility of wealth $\nu$ is estimated by interacting the share of wealth invested in equity at each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

Figure 8: Contribution of $\zeta$ and $\nu$ to Model-Predicted Displacement

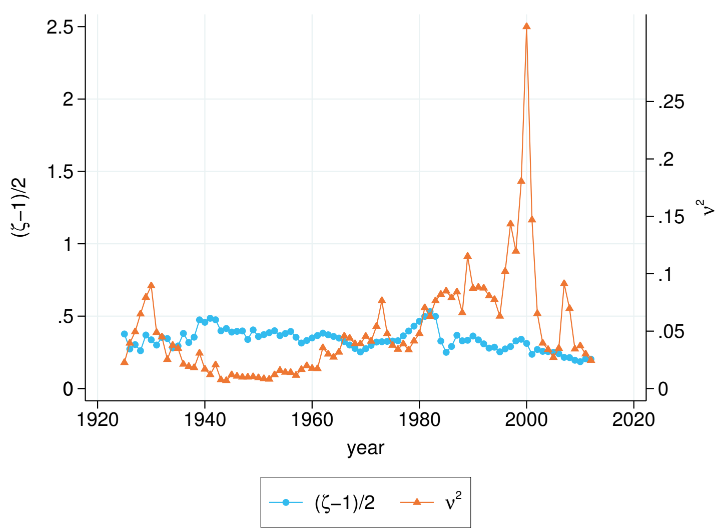



Figure 8 decomposes the model-predicted displacement term into its two components, $(\zeta-1)/2$ and $\nu^2$. The power law exponent of the wealth distribution $\zeta$ is estimated yearly using $\zeta - 1 = g_t(q_t)q_t^2/S_t$. The idiosyncratic volatility at each percentile $\nu$ is estimated by interacting 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).

Table 1: Decomposing the Growth of Top Wealth Share

| Period | **Total** (%) | $R_{\text{within}}$ (%) | $R_{\text{displacement}}$ (%) | $R_{\text{demography}}$ (%) |
|---|---|---|---|---|
| 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% $R_{\text{total}}$, 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 5. All terms in percentage. Data from Forbes 400.

Table 2: Power-law Exponent and Cumulants of Wealth Growth

| Period | Power-law ($\zeta$) | Volatility ($\nu$) | 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 power-law exponent of the wealth distribution and higher-order cumulants of log wealth growth. The power law exponent of the wealth distribution $\zeta$ 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 30% higher or lower than $q_t$. 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 $\sum_2^{+\infty} \frac{\zeta^j-1}{j!}\kappa_j$ (%) | | | | | $\epsilon$ (%) |
| | | **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 $\epsilon$. The term predicted by diffusion-jump is split into different cumulants. The power law exponent of the wealth distribution$\zeta$ is estimated yearly 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 30% higher or lower than $q_t$. 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 $1/2(\zeta-1)\nu^2$ (%) | | | $\epsilon$ (%) |
| | | **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 $\nu^2$ is the sum of the average variance within groups $\nu^2_{\text{within}}$ and the variance between groups $\nu^2_{\text{between}}$. 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

| | $\nu^2$ |
|---|---|
| | (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 $\nu$ 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: 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 7: Displacement Along the Wealth Distribution

| | Top 1% | Top 0.1% | Top 0.01% | Top 400 |
|---|---|---|---|---|
| *Panel A: 1926-2012* | | | | |
| Power-law $\zeta$ | 1.5 | 1.6 | 1.7 | |
| Idiosyncratic Volatility $\nu$ (%) | 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 $\zeta$ | 1.5 | 1.5 | 1.6 | 1.5 |
| Idiosyncratic Volatility $\nu$ (%) | 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 power law exponent $\zeta$ is estimated as $1/(1-\frac{q_t}{S_t/p})$. The idiosyncratic volatility of wealth $\nu$ 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).

Table 8: Regressing Measures of Wealth Inequality on Aggregate Patent Activity

| | $\nu^2$ | $R_{\text{displacement}}/\left(1/2(\zeta-1)\right)$ | $R_{\text{displacement}}$ | $R_{\text{total}}$ |
|---|---|---|---|---|
| | (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 9: Patent Activity of New Entry

| | 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 a 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.

# A Appendix for Section 2

*Proof of Proposition 1.* 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 $q_t$

$$0 = -g_t(q_t)\frac{dq_t}{dt} + \int_{q_t}^{+\infty} \frac{dg_t(w)}{dt}dw \tag{A1}$$

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

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

Injecting the law of motion for $q_t$, we obtain the law of motion for $S_t$:

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

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.

The Kolmogorov Forward equation corresponding to the wealth process is

$$dg_t = -\partial_w(\mu_t dt w g_t(w)) + \partial_w^2(\nu_t^2 dt w^2 g_t(w)/2) \tag{A4}$$

Substituting the law of motion for $dg_t$ (A4) into (A3), and integrating by parts:

$$\begin{aligned} dS_t &= \int_{q_t}^{\infty} (w - q_t)(-\partial_w(\mu_t dt w g_t(w)) + \partial_w^2(\nu_t^2 dt w^2 g_t(w)/2))dw \\ &= -\int_{q_t}^{+\infty} (-\mu_t dt w g_t(w) + \partial_w(\nu_t^2 dt w^2 g_t(w)/2)dw \\ &= \mu_t S_t dt + \frac{g_t(q_t)q_t^2}{2}\nu_t^2 dt \end{aligned} \tag{A5}$$

□

For the sake of completeness, I also give the law of motion of top quantiles in the baseline model, which is originally proven in Steinbrecher and Shaw (2008).

**Proposition 7** (Dynamics of Quantile)**.** *Assume that the law of motion for wealth is given by* (1)*. Then the top quantile* $q_t$ *follows the law of motion:*

$$\frac{dq_t}{q_t} = \mu_t dt - \frac{1}{2}\frac{\partial_w(w^2 g_t(w))}{g_t(q_t)q_t}\nu_t^2 dt \tag{A6}$$

*Proof of Proposition 7.* Combining the definition of the quantile (A1) with the Kolmogorov Forward equation (A4), we obtain

$$
\begin{aligned}
dq_t &= \frac{1}{g_t(q_t)} \int_{q_t}^{+\infty} dg_t dw \\
&= \frac{1}{g_t(q_t)} \int_{q_t}^{+\infty} (-\partial_w(\mu_t dt w g_t(w)) + \frac{1}{2}\partial_w^2(\nu_t^2 dt w^2 g_t(w)) dw \\
&= \mu_t q_t dt - \frac{1}{2}\frac{\partial_w(w^2 g_t(w))}{g_t(q_t)} \nu_t^2 dt
\end{aligned} \tag{A7}
$$

□

*Proof of Proposition 2.* The density function $g_t$ and the quantile function $q_t(g)$ can be expressed as the derivatives of the top wealth share function: $\partial_p S_t = q_t$ and $\partial_{pp} S_t = -1/g_t(q_t)$.

Therefore, Proposition 1 can be rewritten as a PDE on the top wealth share function $p \to S_t(p)$:[43]

$$
\frac{\partial_t S_t}{S_t} = \mu_t dt - \frac{\partial_p S_t^2}{2 S_t \partial_{pp} S_t} \nu_t^2 dt \tag{A8}
$$

For a distribution with Pareto tail $\zeta$, we have $S(p) = Cp^{1-\frac{1}{\zeta}}$, therefore

$$
\frac{\partial_t S_t}{S_t} = \mu_t dt + \frac{\zeta - 1}{2} \nu_t^2 dt \tag{A9}
$$

I now examine the case in which the distribution only has a heavy tail, i.e. $\mathbb{P}(w_{it} \geq w) = L(w) w^{-\zeta}$ where $L(w)$ is a slowly varying function, with a density function $g_t$ that is ultimately monotone. In this case, the Karamata theorem gives, as $w \to +\infty$:[44]

$$
g_t(w) \sim \zeta L(x) w^{-\zeta - 1} \tag{A10}
$$

$$
S_t(w) \sim \frac{\zeta}{\zeta - 1} L(x) w^{-\zeta + 1} \tag{A11}
$$

Therefore, as $w \to +\infty$:

$$
\frac{g_t(w) w^2}{S_t(w)} \sim \zeta - 1 \tag{A12}
$$

□

**Proposition 8** (Dynamics of Top Wealth Share with Scale Dependence)**.** *Assume that the law of motion for wealth is given by* (10)*. Then the top wealth share $S_t$ follows the law of motion:*

$$
\frac{dS_t}{S_t} = \underbrace{\mathrm{E}^{gw}[\mu_t(w) | w \geq q_t] dt}_{dr_{within}} + \underbrace{\frac{g_t(q_t) q_t^2}{2 S_t} \nu_t^2(q_t) dt}_{dr_{displacement}} \tag{A13}
$$

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

[43] This PDE can be used to generate the evolution of the wealth distribution for a given path of $\mu_t, \nu_t$ over time, although the PDE is not linear in $S_t$, contrary to Kolmogorov Forward.

[44] For instance, see Mikosch (1999).

*Proof of Proposition 8.* The Kolmogorov Forward equation corresponding to the wealth process is

$$dg_t = -\partial_w(\mu_t(w)dtwg_t) + \partial_w^2(\nu_t^2(w)dtw^2 g_t/2) \tag{A14}$$

Substitute this equation into (A2) and integrate by part to obtain that

$$\frac{dS_t}{S_t} = \underbrace{\frac{\int_{q_t}^{\infty}\mu_t(w)wdtg_t(w)dw}{S_t}}_{dr_{\text{within}}} + \underbrace{\frac{g_t(q_t)q_t^2}{2S_t}\nu_t^2(q_t)dt}_{dr_{\text{displacement}}} \tag{A15}$$

□

**Proposition 9** (Dynamics of Top Wealth Share with Aggregate Risk)**.** *Assume that the law of motion for wealth is*

$$\frac{dw_{it}}{w_{it}} = \mu_t dt + \sigma_t dZ_t + \nu_t dB_{it} \tag{A16}$$

*where $Z_t = \{Z_{it} \in \mathbb{R}, \mathcal{F}_t, t \geq 0\}$ is an aggregate Brownian motion. Then the top wealth share $S_t$ follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt + \sigma_t dZ_t}_{dr_{within}} + \underbrace{\frac{g_t(q_t)q_t^2}{2S_t}\nu_t^2 dt}_{dr_{displacement}} \tag{A17}$$

*Proof of Proposition 9.* 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 $q_t$

$$0 = -g_t(q_t)\frac{dq_t}{dt} + \int_{q_t}^{+\infty}\frac{dg_t(w)}{dt}dw - \sigma_t[dg_t(q_t)]\sigma_t[dq_t] \tag{A18}$$

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 $S_t$:

$$dS_t = -q_t g_t(q_t)dq_t + \int_{q_t}^{\infty} wdg_t(w)dw - q_t\sigma_t[dg_t(q_t)]\sigma_t[dq_t]dt - \frac{1}{2}g_t(q_t)\sigma_t[dq_t]^2 dt \tag{A19}$$

Injecting (A18) into (A19), we obtain:

$$\begin{aligned} dS_t &= \int_{q_t}^{\infty}(w-q_t)dg_t(w)dw - \frac{1}{2}g_t(q_t)\sigma_t[dq_t]^2 dt \\ &= \int_{q_t}^{\infty}(w-q_t)dg_t(w)dw - \frac{1}{2}\frac{1}{g_t(q_t)}(\int_{q_t}^{\infty}\sigma_t[dg_t(w)]dw)^2 dt \end{aligned} \tag{A20}$$

The Kolmogorov Forward equation corresponding to the wealth process with aggregate risk is:[45]

$$dg_t = -\partial_w(\mu_t dtwg_t + \sigma_t wdZ_t) + \partial_w^2((\sigma_t^2 + \nu_t^2)dtw^2 g_t/2) \tag{A21}$$

[45] See Gomez (2016) for an heuristic derivation.

Substituting the law of motion for $dg_t$ from the Kolmogorov Forward equation with aggregate risk and integrating by parts:

$$
\begin{aligned}
dS_t = &\int_{q_t}^{\infty} (w - q_t)(-\partial_w((\mu_t dt + \sigma_t dZ_t) w g_t(w)) + \partial_w^2((\sigma_t^2 + \nu_t^2) dt w^2 g_t(w)/2) \\
&- \frac{1}{2}\frac{1}{g_t(q_t)}(\int_{q_t}^{+\infty} \partial_w(\sigma_t w g_t(w) dw))^2 dt \\
= &- \int_{q_t}^{+\infty} (-(\mu_t dt + \sigma_t dZ_t) w g_t(w) + \partial_w((\sigma_t^2 + \nu_t^2) dt w^2 g_t(w)/2)) dw \\
&- \frac{1}{2}\frac{1}{g_t(q_t)}(\int_{q_t}^{+\infty} \partial_w(\sigma_t w g_t(w) dw))^2 dt \\
&= \mu_t S_t dt + \sigma_t S_t dZ_t + \frac{g_t(q_t) q_t^2}{2} \nu_t^2 dt
\end{aligned}
\tag{A22}
$$

□

**Proposition 10** (Dynamics of Top Wealth Share with Heterogeneity)**.** *Assume that the law of motion for wealth is given by* (12)*. Then the top wealth share $S_t$ follows the law of motion:*

$$
\begin{aligned}
\frac{dS_t}{S_t} = &\underbrace{\mathrm{E}^{gw}[\mu_{nt} | w_{it} \geq q_t] dt + \mathrm{E}^{gw}[\sigma_{nt} | w_{it} \geq q_t] dZ_t}_{dr_{within}} \\
&+ \underbrace{\frac{g_t(q_t) q_t^2}{2 S_t}(\mathrm{E}^{gw}[\nu_{nt}^2 | w_{it} = q_t] + \mathrm{Var}^{wg}[\sigma_{nt} | w_{it} = q_t]) dt}_{dr_{displacement}}
\end{aligned}
\tag{A23}
$$

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

*Proof of Proposition 10.* Denote $\pi_n$ 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:

$$
g_t = \sum_{1 \leq n \leq N} \pi_n g_{nt} \tag{A24}
$$

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

$$
dg_t = \sum_{1 \leq n \leq N} \pi_n dg_{nt} \tag{A25}
$$

Applying the law of motion for $S_t$ in terms of $g_t$ (A20) and the Kolmogorov Forward Equation for $g_{nt}$ for $1 \leq n \leq N$ (A21), we obtain:

$$
\begin{aligned}
dS_t = &\int_{q_t}^{\infty} (w - q_t) \sum_{1 \leq n \leq N} \pi_n \left(-\partial_w((\mu_{nt} dt + \sigma_{nt} dZ_t) w g_{nt}(w)) + \partial_w^2((\nu_{nt}^2 + \sigma_{nt}^2) dt w^2 g_{nt}(w)/2)\right) dw \\
&- \frac{1}{2}\frac{1}{g_t(q_t)}(\int_{q_t}^{+\infty} \sum_{1 \leq n \leq N} \pi_n \partial_w(\pi_n \sigma_{nt} w g_{nt}(w) dw))^2 dt
\end{aligned}
\tag{A26}
$$

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

$$\frac{dS_t}{S_t} = \frac{\int_{q_t}^{+\infty} \sum_{1\le n\le N} (\mu_{nt}dt + \sigma_{nt}dZ_t) w\pi_n g_{nt}(w)dw}{S_t}$$
$$+ \frac{g_t(q_t)q_t^2}{2S_t} \left( \sum_{1\le n\le N} \frac{\pi_n g_{nt}(q_t)}{g_t(q_t)} \nu_{nt}^2 + \sum_{1\le n\le N} \frac{\pi_n g_{nt}(q_t)}{g_t(q_t)} \sigma_{nt}^2 - \left( \sum_{1\le n\le N} \frac{\pi_n g_{nt}(q_t)}{g_t(q_t)} \sigma_{nt} \right)^2 \right) dt \quad \text{(A27)}$$

□

I now examine the case in which wealth follows a jump-diffusion process, i.e.

$$\frac{dw_{jt}}{w_{jt-}} = \mu_t dt + \nu_t dB_{it} + (e^{J_{it}} - 1)dN_{it} \quad \text{(A28)}$$

where $N_{it}$ is an idiosyncratic jump process with intensity $\lambda_t$. The innovations $J_{it}$ are drawn from an exogenous distribution such that $\mathrm{E}^f[e^{J_{it}}] = 1$, where $\mathrm{E}^f$ denotes the expectation with respect to the jump density $f_t$. In other words, jumps do not change the average wealth growth of households.

**Proposition 11** (Dynamics of Top Wealth Share with Jumps)**.** *Assume that the law of motion for wealth is given by* (A28)*. Then the top wealth share $S_t$ follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt}_{dr_{within}} + \underbrace{\frac{g_t(q_t)q_t^2}{2S_t}\nu_t^2 dt + \frac{\lambda_t dt}{S_t}\mathrm{E}^f[\int_{q_t e^{-J}}^{q_t} (e^J w - q_t) g_t(w)dw]}_{dr_{displacement}} \quad \text{(A29)}$$

*Proof of Proposition 11.* I first present an heuristic derivation, and then a formal derivation.

*Heuristic Derivation* Idiosyncratic jumps increase top shares for two reasons. First, some lucky households *below* the top percentile with *positive* shock enter the top percentile. Because population size in the top percentile is held constant, this displaces marginal households at the threshold with wealth $q_t$. Second, some unlucky households *inside* the top with *negative* jumps exit the top percentile. They are replaced by marginal households at the threshold with wealth $q_t$. The total growth of top wealth share due to entry and exit is given by:

$$dr_{\text{jump}} = \frac{\lambda_t dt}{S_t}\mathrm{E}^f[\int_0^{q_t} (e^J w - q_t)^+ g_t(w)dw] + \frac{\lambda_t dt}{S_t}\mathrm{E}^f[\int_{q_t}^{+\infty} (q_t - e^J w)^+ g_t(w)dw] \quad \text{(A30)}$$

*Formal Derivation* The Kolmogorov Forward equation corresponding to the wealth process is

$$dg_t = -\partial_w(\mu_t dt w g_t) + \partial_w^2(\nu_t^2 dt w^2 g_t/2) + \lambda_t dt \mathrm{E}^f[e^{-J} g_t(we^{-J}) - g_t(w)]. \quad \text{(A31)}$$

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

$$\begin{aligned} dS_t = &\int_{q_t}^{\infty} (w - q_t)(-\partial_w(\mu_t dt w g_t(w)) + \partial_w^2(\nu_t^2 dt w^2 g_t(w)/2) \\ &+ \lambda \int_{q_t}^{+\infty} (w - q_t)(\mathrm{E}^f[e^{-J} g_t(we^{-J})] - g_t(w))dw \\ = &- \int_{q_t}^{+\infty} (-\mu_t dt w g_t(w) + \partial_w(\nu_t^2 dt w^2 g_t(w)/2))dw \\ &+ \lambda \int_{q_t}^{+\infty} (w - q_t)(\mathrm{E}^f[e^{-J} g_t(we^{-J})] - g_t(w))dw \\ = &\ \mu_t S_t dt + \frac{g_t(q_t)q_t^2}{2}\nu_t^2 dt + \lambda_t dt \int_{q_t}^{+\infty} (w - q_t)(\mathrm{E}^f[e^{-J} g_t(we^{-J})] - g_t(w))dw \end{aligned} \tag{A32}$$

The term due to jump 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:

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

This concludes the proof. □

**Proposition 12** (Dynamics of Top Wealth Share with Jumps and Pareto Distribution)**.** *Suppose that the law of motion of relative wealth $w_{it}$ is (1), with a maximum jump size $\overline{J}$ and that the wealth distribution at time $t$ is Pareto with power law exponent $\zeta$ for a wealth level higher than $q_t e^{-\overline{J}}$. Then the law of motion of the top wealth share $S_t$ follows the law of motion:*

$$\frac{dS_t}{S_t} = \underbrace{\mu_t dt}_{dr_{within}} + \underbrace{\frac{\zeta - 1}{2}\nu_t^2 dt + \lambda_t dt \frac{\mathrm{E}^f[e^{J_{it}\zeta}] - 1}{\zeta}}_{dr_{displacement}} \tag{A33}$$

*Alternatively, it can be written as*

$$\frac{dS_t}{S_t} = \frac{1}{\zeta} E_t\left[\frac{dw_{it}^{\zeta}}{w_{it}^{\zeta}}\right] \tag{A34}$$

$$= \underbrace{\mu_t dt}_{dr_{within}} + \underbrace{\sum_{2}^{+\infty} \frac{\zeta^{j-1} - 1}{j!}\kappa_{jt}}_{dr_{displacement}} \tag{A35}$$

*Moreover, the term due to jumps is positive and increases in $\zeta$.*

*Proof of Proposition 12.* When the distribution is Pareto, the growth of top wealth shares due to jumps can be written

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

I now prove that the jump term increases in $\zeta$. The derivative of the jump term with respect to $\zeta$ is

$$\zeta \to \lambda dt \frac{m'(\zeta)\zeta - (m(\zeta)-1)}{\zeta^2} \tag{A36}$$

where $m(\zeta) = \mathrm{E}^f[e^{J\zeta}]$. The function $m$ is convex and equals to zero in zero and 1, therefore $m'(1) \geq 0$. Since the derivative of the numerator is nonnegative, the numerator is always nonnegative, and therefore the jump term increases in $\zeta$.

Using Ito's lemma, one obtains:

$$\frac{1}{\zeta}E_t\left[\frac{dw_{it}^\zeta}{w_{it}^\zeta}\right] = \mu_t dt + \frac{\zeta-1}{2}\nu_t^2 dt + \lambda_t dt\frac{\mathrm{E}^f[e^{J_{it}\zeta}]-1}{\zeta} \tag{A37}$$

which gives Equation (A34).

Finally, define $\kappa_{jt}$ as the coefficients in the power expansion of $E_t[dw_{it}^\zeta/w_{it}^\zeta]$, i.e.

$$\mathrm{E}_t\left[\frac{dw_{it}^\zeta}{w_{it}^\zeta}\right] = \sum_{j=1}^{+\infty}\frac{\zeta^j}{j!}\kappa_{jt}dt \tag{A38}$$

Plugging this into (A34) gives Equation (A35).

$\kappa_{jt}$ can be called the "instantaneous" cumulant because it can also be defined as the derivative of the $j-$th cumulant of wealth growth between $t$ and $t+\tau$, at $\tau = 0$. To see this, note that:

$$\mathrm{E}_t\left[\frac{dw_{it}^\zeta}{w_{it}^\zeta}\right] = \left(\lim_{\tau\to 0}\frac{1}{\tau}\log E_t\frac{w_{it+\tau}^\zeta}{w_{it}^\zeta}\right)dt = \sum_{j=1}^{+\infty}\frac{\zeta^j}{j!}\left(\lim_{\tau\to 0}\frac{\kappa_{jt}(\tau)}{\tau}\right)dt \tag{A39}$$

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

*Proof of Proposition 3.* The Kolmogorov Forward Equation is

$$dg_t = -\partial_w(\mu_t dt w g_t) + \partial_w^2(\nu_t^2 dt w^2 g_t/2) - \eta dt g_t(w) + \delta dt\left(\frac{g(w/\chi)}{\chi} - g(w)\right) \tag{A40}$$

We obtain the dynamics of the top wealth share $S_t$ by following the same steps as the proof of Proposition 1.

□

*Proof of Proposition 4.* When the wealth distribution is power law with Pareto tail $\zeta$, we have $\pi_t = \chi^\zeta$, and therefore $S_t(\pi p) = \chi^{\zeta-1} S_t(p)$. Therefore, the term due to death can be written as

$$\begin{aligned} dr_{\text{death}} &= \left( \frac{\chi S_t(\pi_t p) + (1-\pi_t) q_t p}{S_t} - 1 \right) \delta dt + \frac{q_t p}{S_t} \eta dt \\ &- ((\chi^\zeta + (1-\chi^\zeta)(1-\frac{1}{\zeta}) - 1)\delta dt \\ &= \frac{\chi^\zeta - 1}{\zeta} \delta dt \end{aligned} \tag{A41}$$

Note that the term is similar to the term that would be generated by a jump process of intensity $\delta$ and size $-\log(\chi)$. As shown in the proof of Proposition 11, the term increases in $\zeta$. □

# B Appendix for Section 3

*Proof of Proposition 5.* 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:

$$\sum_{i \in \mathcal{T}'} w_{it+\tau} = \sum_{i \in \mathcal{T} \setminus \mathcal{X}_\mathcal{D}} w_{it+\tau} + \sum_{i \in \mathcal{E} \cup \mathcal{E}_\mathcal{D}} w_{it+\tau} - \sum_{i \in \mathcal{X}} w_{it+\tau} \tag{A42}$$

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

$$R_{\text{within}} = \frac{\sum_{i \in \mathcal{T} \setminus \mathcal{X}_\mathcal{D}} w_{it+\tau}}{\sum_{i \in \mathcal{T} \setminus \mathcal{X}_\mathcal{D}} w_{it}} - 1 \tag{A43}$$

Equation (A42) giving the total wealth of households in the top at time $t+\tau$ can be rewritten using $r_{t+\tau}$

$$\sum_{i \in \mathcal{T}'} w_{it+\tau} = (1+R_{\text{within}})(\sum_{i \in \mathcal{T}} w_{it} - \sum_{i \in \mathcal{X}_\mathcal{D}} w_{it}) + \sum_{i \in \mathcal{E} \cup \mathcal{E}_\mathcal{D}} w_{it+\tau} - \sum_{i \in \mathcal{X}} w_{it+\tau} \tag{A44}$$

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

$$\begin{aligned} \sum_{i \in \mathcal{T}'} w_{it+\tau} &= (1+R_{\text{within}}) \sum_{i \in \mathcal{T}} w_{it} + \sum_{i \in \mathcal{X}_\mathcal{D}} (q_{t+\tau} - (1+R_{\text{within}}) w_{it}) + \sum_{i \mathcal{E}_\mathcal{D}} (w_{it+\tau} - q_{t+\tau}) \\ &+ \sum_{i \in \mathcal{E}} (w_{it+\tau} - q_{t+\tau}) + \sum_{i \in \mathcal{X}} (q_{t+\tau} - w_{it+\tau}) + (|\mathcal{T}'| - |\mathcal{T}|) q_{t+\tau} \end{aligned} \tag{A45}$$

Dividing by $S_t$ and rearranging, one obtains the accounting decomposition (23). □

I now relate this accounting decomposition to the theoretical decomposition in Proposition 3. 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 (1) for household wealth $w_{it}$, we get for $\tau > 0$

$$\mathrm{E}[R_{\text{within}}] = e^{\int_t^{t+\tau} \mu_s ds} - 1 \tag{A46}$$

$$\sim \mu_t \tau \text{ as } \tau \to 0 \tag{A47}$$

Integrating the law of motion for the top wealth share (17), we get for $\tau > 0$

$$\mathrm{E}[\frac{S_{t+\tau}}{S_t}] - 1 = e^{\int_t^{t+\tau}\left(\mu_s + \frac{g_s(q_s)q_s^2}{2S_s}\nu_s^2 + (\frac{q_s p}{S_s} - 1)\delta_s + \frac{q_s p}{S_t}\eta_s\right)ds} - 1 \tag{A48}$$

$$\sim \mu_t \tau + \frac{g_t(q_q)q_t^2}{2S_t}\nu_t^2 \tau + \left(\frac{q_t p}{S_t} - 1\right)\delta\tau + \frac{q_t p}{S_t}\eta_t \tau \text{ as } \tau \to 0 \tag{A49}$$

The demography term is

$$\mathrm{E}[R_{\text{demography}}] = \left(1 - e^{-\int_t^{t+\tau}\delta_s ds}\right)\left(\frac{q_{t+\tau}p}{S_t} - e^{\int_t^{t+\tau}\mu_s ds}\right) + e^{\int_t^{t+\tau}\eta_s ds} p \frac{q_{t+\tau}}{S_t} \tag{A50}$$

$$\sim (\frac{q_t p}{S_t} - 1)\delta_t \tau + \frac{q_t p}{S_t}\eta_t \tau \text{ as } \tau \to 0 \tag{A51}$$

Therefore,

$$\mathrm{E}[R_{\text{displacement}}] = \mathrm{E}[\frac{S_{t+\tau}}{S_t}] - \mathrm{E}[R_{\text{within}}] - \mathrm{E}[R_{\text{demography}}]$$

$$= e^{\int_t^{t+\tau}\left(\mu_s + \frac{g_s(q_s)q_s^2}{2S_s}\nu_s^2 + (\frac{q_s p}{S_s} - 1)\delta_s + \frac{q_s p}{S_t}\eta_s\right)ds} - e^{\int_t^{t+\tau}\mu_s ds} \tag{A52}$$

$$- \left(1 - e^{-\int_t^{t+\tau}\delta_s ds}\right)\left(\frac{q_{t+\tau}p}{S_t} - e^{\int_t^{t+\tau}\mu_s ds}\right) - e^{\int_t^{t+\tau}\eta_s ds} p \frac{q_{t+\tau}}{S_t} \tag{A53}$$

$$\sim \frac{g_t(q_t)q_t^2}{2S_t}\nu_t^2 \tau \text{ as } \tau \to 0 \tag{A54}$$

Therefore, the expectation of the terms in the accounting decomposition are asymptotically equivalent to the terms in Proposition 3 as the time period $\tau$ tends to zero. In this sense, the accounting decomposition converges to the theoretical decomposition as the time period $\tau$ tends to zero.

# C Appendix for Section 4

## C.1 Left Censoring

The decomposition Section 3 requires 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 drop 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 2% in average. 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 less than 0.1%. 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 $\eta_{it}$ is an i.i.d. process independent of wealth. Moreover, suppose the observed wealth $\tilde{w}_{it} = w_{it} e^{\epsilon_{it}}$ where $\epsilon_{it}$ is an i.i.d process independent of wealth capturing measurement error and $\mathrm{E}[e^{\epsilon_{it}}] = 1$. Denote $\xi$ the ratio between the variance of measurement error and the variance of wealth growth, i.e.

$$\xi \equiv \frac{var(\epsilon_{it})}{var(r_{it})} \tag{A55}$$

The log change in wealth can be written

$$\log\left(\frac{\tilde{w}_{it+1}}{\tilde{w}_{it}}\right) = r_{it+1} + \epsilon_{it+1} - \epsilon_{it} \tag{A56}$$

A regression of wealth growth on past wealth growth estimates the slope coefficient $\rho$:

$$\begin{aligned}\rho &= \frac{cov(\log(\tilde{w}_{it+1}/\tilde{w}_{it}), \log(\tilde{w}_{it+1}/\tilde{w}_{it}))}{var(\log(\tilde{w}_{it+1}/\tilde{w}_{it}))} \\ &= \frac{cov(r_{it+1}+\epsilon_{it+1}-\epsilon_{it}, r_{it}+\epsilon_{it}-\epsilon_{it-1})}{var(r_{it}+\epsilon_{it}-\epsilon_{it-1})} \\ &= -\frac{var(\epsilon_{it})}{var(r_{it})+2var(\epsilon_{it})} \\ &= -\frac{\xi}{1+2\xi} \end{aligned} \tag{A57}$$

Testing whether $\rho$ is equal to zero is a test on whether $\xi$ is different from zero, i.e. that there is measurement error.

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

I now examine how the derivative of the displacement term depends exactly on $\rho$. Even though we cannot reject that $\rho$ is statistically different from zero, it is important to check small values of $\rho$ do not have a disproportionate effect on the displacement term. I examine the bias in the displacement term in a simple setting. I assume that $\epsilon_{it}$ and $r_{it}$ are normal variables and that the wealth distribution has a Pareto tail with power law exponent $\zeta$.

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

$$r_{\text{displacement}} = \frac{\zeta-1}{2} Var(\eta_{it}) \tag{A58}$$

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

$$\tilde{r}_{\text{displacement}} = \frac{\zeta-1}{2}\left(Var(\eta_{it}) + 2Var(\epsilon_{it})\right) \tag{A59}$$

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

$$\begin{aligned}\frac{\tilde{r}_{\text{displacement}} - r_{\text{displacement}}}{r_{\text{displacement}}} &= -\frac{2\rho}{1+2\rho} \\ &\approx -2\rho \end{aligned} \tag{A60}$$

In particular, when $\rho$ is close to zero, the relative bias well approximated by a simple linear function of $\rho$. Since $\rho$ 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 $R_{\text{within}}$ 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 $R_{\text{U.S.}}$:

$$R_{\text{within}} \approx R_{\text{top households}} - R_{\text{U.S.}} \tag{A61}$$

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 comes 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 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.[46] 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.[47].

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 adds 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 adds 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 excess returns corporate bonds. Similarly, only the exposure to the market is significant. Overall, 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.

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

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

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 industry of households in top percentiles is 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 is 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.:

$$R_{\text{top households}} = R_f + \sum_{1 \leq k \leq K} \beta_k \times (R_k - R_f) + \frac{\sum_{i \in \mathcal{T}} (\text{Labor}_i - \text{Tax}_i)}{S_t} + \epsilon \tag{A62}$$

where $\text{Labor}_i$ denotes the labor income and $\text{Tax}_i$ 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.[48] The dataset is only available after 1992, so I use the average of this term in 1992-1995 to input it starting from 1983.

Panel A of Table A3 reports decomposition (A62) 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 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 $\epsilon$ 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 $r_{\text{death}}$. According to the theoretical model, this term equals

$$R_{\text{death}} = \frac{\chi^{\zeta} - 1}{\zeta} \delta \tag{A63}$$

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

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

Given an estimate from $\zeta$ and $\delta$ that can be obtained from the data, one can always compute the implicit inheritance parameter $\chi$ that explains the magnitude of the death term $R_{\text{death}}$ using (A63). I report the result of this decomposition in Proposition 3. I find that the inheritance parameter $\chi$ averages to 60% and is pretty stable over time. It corresponds to an average estate tax $1 - \chi = 40\%$ 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

$$R_{\text{demography}} = \left(1 - \frac{1}{\zeta}\right)\eta \tag{A64}$$

where $\zeta$ is the Pareto tail of the wealth distribution and $\eta$ is the population growth rate. I compare the actual population growth rate to the term predicted by the model, using the estimate for $\zeta$ used in Section 4, and the population growth rate $\eta$. I find that the two terms are very close, with a difference $\epsilon$ 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 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 $\zeta$ 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 power law exponent 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 power law exponent of the distribution. Indeed, for a distribution with a Pareto tail with power law exponent $\zeta$, the ratio of the wealth shares of two top percentiles $p$ and $p'$ is:

$$\log\left(\frac{S_t(p)}{S_t(p')}\right) = \left(1 - \frac{1}{\zeta}\right)\log\left(\frac{p}{p'}\right) \tag{A65}$$

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.

$$\frac{dw_{it}}{w_{it}} = \mu_t dt + \nu_t(w_{it})dB_{it} \tag{A66}$$

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

$$d \log \left( \frac{S_t(p)}{S_t(p')} \right) = \left( \frac{g_t(q_t(p))q_t(p)^2}{2S_t(p)} \nu_t(q_t(p))^2 - \frac{g_t(q_t(p'))q_t(p')^2}{2S_t(p')} \nu_t(q_t(p'))^2 \right) dt \tag{A67}$$

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.

$$T_q(w_{it}) = \delta \Delta t \times 0 + (1 - \delta \Delta t) \times (\Delta t + \mathrm{E}[T_q(w_{it+\Delta t})]) \tag{A68}$$

Therefore

$$0 = (1 - \delta \Delta t)(\Delta t + \mathrm{E}[T_q(w_{it+\Delta t}) - T_q(w_{it})]) - \delta \Delta t T_q(w_{it}) \tag{A69}$$

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

$$0 = dt + \mathrm{E}[dT_q(w_{it})] - T_q(w_{it}) \delta dt \tag{A70}$$

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

$$1 + T_q'(w) \mu w + T_q''(w) \frac{\nu^2 w^2}{2} - \delta T_q(w) = 0 \tag{A71}$$

The solution has the form:

$$T_q(w) = c_1 w^{\zeta_+} + c_2 w^{\zeta_-} + \frac{1}{\delta} \tag{A72}$$

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

Using the limit condition:

$$T_q(q) = 0 \tag{A73}$$

$$\lim_{w \to +\infty} T_q(w) = \frac{1}{\delta} \tag{A74}$$

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

$$T_q(w) = \frac{1}{\delta}\left(1 - \left(\frac{w}{q}\right)^{\zeta_-}\right) \tag{A75}$$

□

The average first passage time $T_q(w)$ for the case $\delta = 0$ can be obtained by taking the limit as $\delta \to 0$.[49]

$$T_q(w) = \frac{1}{\nu^2/2 - \mu} \log \frac{w}{q} \tag{A76}$$

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 rule:

$$\lim_{\delta \to 0} T_q(w) = -\frac{\partial \zeta_-}{\partial \delta}(\delta = 0) \log \frac{w}{q} \tag{A77}$$

Using the implicit function theorem to compute the derivative of $\zeta_-$ with respect to $\delta$, we obtain

$$\lim_{\delta \to 0} T_q(w) = \frac{1}{\nu^2/2 - \mu} \log \frac{w}{q} \tag{A78}$$

The average time before exit $T_q(w)$ increases in $\mu$ and decrease in $\nu^2$. By the definition of $\zeta_+$ and $\zeta_-$ 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 $\mu$ increases the average time before exit $T_q(w)$. Similarly, an increase in $\nu^2$ decreases the average time before exit $T_q(w)$.

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

$$\frac{\partial T_q(w)}{\partial \delta} = -\frac{1}{\delta^2}\left(1 - \left(\frac{w}{q}\right)^{\zeta_-}\left(1 - \delta \frac{\partial \zeta_-}{\partial \delta} \log\left(\frac{w}{q}\right)\right)\right) \tag{A79}$$

This derivative has the same sign as the function

$$f : \delta \to 1 - \left(\frac{w}{q}\right)^{\zeta_-}\left(1 - \delta \frac{\partial \zeta_-}{\partial \delta} \log\left(\frac{w}{q}\right)\right) \tag{A80}$$

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

$$\frac{\partial f}{\partial \delta} = \left(\frac{w}{q}\right)^{\zeta_-} \log\left(\frac{w}{q}\right) \delta \left(\left(\frac{\partial \zeta_-}{\partial \delta}\right)^2 \log \frac{w}{q} + \frac{\partial^2 \zeta_-}{\partial \delta^2}\right) \tag{A81}$$

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

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

*Proof of Proposition 6.* When $\delta \neq 0$, the average time is

$$\begin{aligned} T(p) &= \mathrm{E}^g[T_q(w_{it})|w_{it} \geq q] \\ &= \frac{\int_q^{+\infty} T_q(w)g(w)dw}{\int_q^{+\infty} g(w)dw} \\ &= \frac{1}{\delta}\left(1 - \frac{\int_q^{+\infty}(w/q)^{\zeta_-} w^{-\zeta_+-1}dw}{\int_q^{+\infty} w^{-\zeta_+-1}dw}\right) \\ &= \frac{1}{\delta}\left(1 - \frac{1}{q^{\zeta_-}}\frac{\frac{q^{\zeta_- -\zeta_+}}{\zeta_+ - \zeta_-}}{\frac{q^{-\zeta_+}}{\zeta_+}}\right) \\ &= \frac{1}{\delta}\frac{1}{1-\zeta_+/\zeta_-} \end{aligned} \tag{A82}$$

The derivative with respect to idiosyncratic variance $\nu^2$ is

$$\begin{aligned} \frac{\partial T}{\partial \nu^2} &= \frac{1}{\delta}\frac{1}{(1-\zeta_+/\zeta_-)^2}\frac{\partial(\zeta_+/\zeta_-)}{\partial \nu^2} \\ &= \frac{1}{\delta}\frac{1}{(1-\zeta_+/\zeta_-)^2}\frac{\zeta_+}{-\zeta_-}\left(\frac{1}{\zeta_-}\frac{\partial \zeta_-}{\partial \nu^2} - \frac{1}{\zeta_+}\frac{\partial \zeta_+}{\partial \nu^2}\right) \end{aligned} \tag{A83}$$

As $\nu^2$ increases, $T(p)$ decreases only if the percentage decrease of $\zeta_-$ is higher than the percentage decrease of $\zeta_+$.

In the case $\chi = 0$ and $n = 0$, using the implicit function theorem, we have

$$\frac{1}{\zeta_-}\frac{\partial \zeta_-}{\partial \nu^2} = -\frac{(\zeta_- - 1)/2}{\mu + (\zeta_- - \frac{1}{2})\nu^2} \tag{A84}$$

$$\frac{1}{\zeta_+}\frac{\partial \zeta_+}{\partial \nu^2} = -\frac{(\zeta_+ - 1)/2}{\mu + (\zeta_+ - \frac{1}{2})\nu^2} \tag{A85}$$

Therefore

$$\frac{1}{\zeta_-}\frac{\partial \zeta_-}{\partial \nu^2} - \frac{1}{\zeta_+}\frac{\partial \zeta_+}{\partial \nu^2} = \frac{(\zeta_+ - \zeta_-)(\mu + \frac{1}{2}\nu^2)}{2(\mu + (\zeta^- - \frac{1}{2})\nu^2)(\mu + (\zeta_+ - \frac{1}{2})\nu^2)} \tag{A86}$$

The denominator is always negative. The numerator is positive if and only if $\mu + \frac{1}{2}\nu^2 \geq 0$. In this case, $\frac{1}{\zeta_-}\frac{\partial \zeta_-}{\partial \nu^2} - \frac{1}{\zeta_+}\frac{\partial \zeta_+}{\partial \nu^2} \leq 0$ and therefore $T$ decreases in $\nu^2$. Conversely, if $\mu \leq -\nu^2/2$, $T$ increases in $\nu^2$.

□

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

| | Dependent Variable | | | |
|---|---|---|---|---|
| | $R_{\text{total}}$ (%) | $R_{\text{within}}$ (%) | $R_{\text{displacement}}$ (%) | $R_{\text{demography}}$ (%) |
| | (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) |
| $\text{Dummy}_{1994\leq\text{year}\leq2004}$ | −0.55 | 0.15 | −0.47 | −0.25 |
| | (5.10) | (4.81) | (0.65) | (0.31) |
| $\text{Dummy}_{2005\leq\text{year}\leq2016}$ | −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 | 0.04** | 0.02 | 0.02*** | 0.00*** |
| | (0.02) | (0.02) | (0.00) | (0.00) |
| Year† | 0.00 | 0.00 | 0.00*** | 0.00*** |
| | (0.00) | (0.00) | (0.00) | (0.00) |
| $R^2$ | 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 (23) 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 of 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 circumvent 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 $\beta_i$ obtained when estimating the linear model

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

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 | $R_{\text{within}}$ | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Total** (%) | $R_{\text{top households}}$ (%) | | | | | $-R_{\text{U.S.}}$(%) |
| | | **Total** | $R_f$ | $\beta_M(R_M - R_f)$ | Labor-Tax | $\epsilon$ | |
| *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 $R_{\text{within}}$ according to the theoretical model (A61) and (A62), 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** (%) | $R_{\text{death}} = \frac{\chi^{\zeta}-1}{\zeta}\delta$ | | | | $R_{\text{pop. growth}} = \left(1 - \frac{1}{\zeta}\right)\eta + \epsilon$ | | | |
| | | **Total** (%) | $\zeta$ | $\delta$ (%) | $\chi$ (%) | **Total** (%) | $\zeta$ | $\eta$ (%) | $\epsilon$ (%) |
| 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 (18). The death rate $\delta$ corresponds to the yearly death rate of households in the top percentile. The population growth rate $\eta$ corresponds to the yearly growth of the U.S. population. The power law exponent 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 30% higher or lower than $q_t$.

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 $R_{\text{within}}$ obtained using imputed data compared to the wealth of drop-offs reported after 2011. The difference in $R_{\text{within}}$ 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 $1/2(\zeta-1)\nu^2$ (%) | | | $\epsilon$ (%) |
| | | **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 $\nu^2$ is the sum of the average variance within groups $\nu^2_{\text{within}}$ and the variance between groups $\nu^2_{\text{between}}$. 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.

**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.

**Champernowne, David G**, “A model of income distribution,” *The Economic Journal*, 1953, *63* (250), 318–351.

**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 and persistence in returns to wealth,” Technical Report, National Bureau of Economic Research 2016.

**Gabaix, Xavier**, “Zipf’s Law for Cities: an Explanation,” *Quarterly Journal of Economics*, 1999, pp. 739–767.

_ , “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.

**_ , Leonid Kogan, and Stavros Panageas**, “Displacement Risk and Asset Returns,” *Journal of Financial Economics*, 2012, *105* (3), 491–510.

**Gomez, Matthieu**, “Asset Prices and Wealth Inequality,” 2016. Working Paper.

**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,” *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.

**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.

**Mikosch, Thomas**, *Regular variation, subexponentiality and their applications in probability theory*, Eindhoven University of Technology, 1999.

**Piketty, Thomas**, "Capital in the 21st Century," *Harvard University Pressed*, 2014.

_ , *Capital in the twenty-first century*, Harvard University Press, 2017.

_ **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.

**Steinbrecher, György and William T Shaw**, "Quantile mechanics," *European journal of applied mathematics*, 2008, *19* (02), 87–112.

**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.