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

Discussion by Jaroslav Borovička (NYU and Federal Reserve Bank of Minneapolis)

November 2018

# OUTLINE

1. Brief summary of the paper
    - role of theory for measurement
    - questions
2. An alternative investigation
    - large deviation theory
    - non-local mobility

# BACKGROUND

Wealth dynamics became an important research area

- inequality, wealth mobility, impact on aggregate growth, business dynamism, monopoly power due to concentration, political clout, etc.
- way too many papers to list here

# BACKGROUND

Wealth dynamics became an important research area

- inequality, wealth mobility, impact on aggregate growth, business dynamism, monopoly power due to concentration, political clout, etc.
- way too many papers to list here

How can we measure contributions to the wealth growth of top wealth percentiles?

- due to incumbents and due to new entrants (displacement)

## BACKGROUND

Wealth dynamics became an important research area

- inequality, wealth mobility, impact on aggregate growth, business dynamism, monopoly power due to concentration, political clout, etc.
- way too many papers to list here

How can we measure contributions to the wealth growth of top wealth percentiles?

- due to incumbents and due to new entrants (displacement)

With enough data $\Longrightarrow$ simply count!

- trivial if we have large panel datasets and study large groups
- not the case of Forbes 400 (top 0.000157% of U.S. adult population)
    - noise, correlations between individuals (Waltons, Page/Brin/Schmidt, Gates/Ballmer/Allen), ...

# IDEA

Impose elementary theoretical restrictions on individual wealth dynamics

- relative wealth follows an Itô process

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

- compute evolution $dS_t$ of wealth share in upper quantile $p$

$$S_t = \int_{q_t(p)}^{\infty} w g_t(w)\,dw$$

# IDEA

Law of motion

$$dS_t = S_t \underbrace{E\left[\mu_t(w) \mid w \geq q_t\right]}_{\text{within}} dt + \underbrace{\frac{1}{2}\left[q_t \nu_t(q_t)\right]^2 g_t(q_t)}_{\text{displacement}} dt$$

- $q_t$ is the relative wealth level at quantile $p$

Where does the displacement term come from?

- probability current

$$J(w,t) = \underbrace{w\mu_t(w)\, g_t(w)}_{\text{deterministic drift}} - \frac{\partial}{\partial w}\underbrace{\left[\frac{1}{2}\left(w\nu_t(w)\right)^2 g_t(w)\right]}_{\text{`churning'}}$$

# PARAMETRIC RESTRICTIONS

Utilize empirical evidence on the Pareto shape of the upper tail of the wealth distribution.

$$P\left(w_{it} \geq w\right) = Cw^{-\zeta}$$

- $\zeta > 1$ is the shape parameter (higher $\zeta$, less inequality)
- e.g., when $\mu_t$ and $\nu_t$ are independent of $w$ and constant over time

# PARAMETRIC RESTRICTIONS

Utilize empirical evidence on the Pareto shape of the upper tail of the wealth distribution.

$$P\left(w_{it} \geq w\right) = Cw^{-\zeta}$$

- $\zeta > 1$ is the shape parameter (higher $\zeta$, less inequality)
- e.g., when $\mu_t$ and $\nu_t$ are independent of $w$ and constant over time

Instead of nonparametric estimation, infer the steady-state shape parameter $\zeta$ and determine the displacement term as

$$\frac{1}{2}\left(\zeta - 1\right)\nu^2$$

- the distribution is not in steady state but the approximation is good and $\zeta$ moves only slowly over time.
- $\zeta$ can be estimated from a richer cross-section

# RESULTS

- displacement accounts for more than half of the top wealth growth
- role of displacement declines over time
    - consistent with recent literature on wealth and business dynamics
- diffusion model predicts the displacement term well
- higher-order (jump) terms have a small effect
    - except during the dot-com boom
- a larger number of robustness checks and alternative specifications

# COMMENT 1: CAVEAT WITH PARAMETRIC RESTRICTIONS

If $\mu_t$ and $\nu_t$ are independent of $w$ and constant over time then the only possible choice is $\mu = 0$.

- $w_{it}$ is wealth relative to aggregate, then aggregate and individual growth rates must be the same $\implies \mu = 0$.
- but then $\zeta = 1 - 2\mu/\nu^2 = 1$ and the distribution does not have a finite mean

How to resolve this?

- Pareto shape only aplies to the tail
- wealth relative to a different benchmark

# COMMENT 2: BETWEEN AND WITHIN INDUSTRIES DECOMPOSITION

The between/within industry decomposition would deserve more explanation.

- It seems that the decomposition uses two terms

$$\frac{1}{2}(\zeta - 1)\,\nu^2_{\text{within}} \quad \text{and} \quad \frac{1}{2}(\zeta - 1)\,\nu^2_{\text{between}}$$

  where $\nu^2_{\text{within}}$ and $\nu^2_{\text{between}}$ are simply the within and between variances according to Fama–French industry portfolios
- Decomposition attributes most of the displacement effect to the within industry component (higher $^2_{\text{within}}$).

But how is it related to the within and between variances in the portfolios in the Forbes 400 list?

- These portfolios are highly selective, is FF representative?
- What about non-traded wealth?

# COMMENT 3: PERSISTENCE OF WEALTH DYNAMICS

In this model, everybody is ex ante identical

- some people get rich because they are lucky
- aligns with literature that stresses the role of idiosyncratic returns

Alternative: heterogeneity

- entrepreneurial skills, other forms of human capital

The two stories have different predictions for survival patterns in top quantiles

- paper computes expected survival times predicted by the model
- is the data informative to produce reliable hazard rates for survival?

The displacement term is a local concept

- rate of crossing the top $p$-th quantile of the wealth distribution

What about non-local counterparts?

- chance of getting into the top $p$-th quantile, starting from a given level of wealth $\bar{w}$.
- characterize the 'typical' paths to reach the quantile.

Discuss concepts related to the theory of large deviations.

Consider a class of wealth processes indexed by $\varepsilon$

$$dw_t^\varepsilon = \mu_t\left(w_t^\varepsilon\right) dt + \sqrt{\varepsilon}\sigma_t\left(\omega_t^\varepsilon\right) dB_t.$$

We want to study

$$P\left(w_T^\varepsilon \geq \mathrm{r}\right) \qquad \text{given } w_0^\varepsilon = \bar{w}.$$

- other state variables possible (suppressed here)

# LARGE DEVIATION THEORY

Consider a class of wealth processes indexed by $\varepsilon$

$$dw_t^\varepsilon = \mu_t\left(w_t^\varepsilon\right)dt + \sqrt{\varepsilon}\sigma_t\left(\omega_t^\varepsilon\right)dB_t.$$

We want to study

$$P\left(w_T^\varepsilon \geq \mathrm{r}\right) \qquad \text{given } w_0^\varepsilon = \bar{w}.$$

- other state variables possible (suppressed here)

Construct the function $h_{\mathrm{r}}(w) = \mathbf{1}\{w \geq \mathrm{r}\}$. Then

$$P\left(w_T^\varepsilon \geq \mathrm{r}\right) = E_0\left[h_{\mathrm{r}}\left(w_T^\varepsilon\right)\right].$$

We are interested in the limit

$$\lim_{\varepsilon \searrow 0} \varepsilon \log E_0 \left[h_{\mathrm{r}}\left(w_T^{\varepsilon}\right)\right] \doteq -I\left(\bar{w}, \mathrm{r}, T\right)$$

- as $\varepsilon \searrow 0$, the threshold r is more relatively more distant, given the underlying uncertainty.

# VARADHAN'S LEMMA

We are interested in the limit

$$\lim_{\varepsilon \searrow 0} \varepsilon \log E_0 \left[ h_{\mathrm{r}} \left( w_T^{\varepsilon} \right) \right] \doteq -I \left( \bar{w}, \mathrm{r}, T \right)$$

- as $\varepsilon \searrow 0$, the threshold r is more relatively more distant, given the underlying uncertainty.

Solution can be characterized by the following deterministic problem:

$$I \left( \bar{w}, \mathrm{r}, T \right) = \inf_{u} \int_0^T \frac{1}{2} \left| u_t \right|^2 dt$$

subject to

$$\dot{w}_t = \mu_t \left( w_t \right) + \sigma_t \left( w_t \right) u_t, \qquad w_0 = \bar{w}, \, w_T \geq \mathrm{r}.$$

- choosing a particular path of shock realizations leading to $w_T \geq \mathrm{r}$.

# HAMILTON–JACOBI–BELLMAN EQUATION

The associated Hamilton–Jacobi–Bellman equation is

$$0 = \inf_{u} \frac{1}{2} \left| u \right|^2 + \left[ \mu_t \left( w \right) + \sigma_t \left( w \right) u \right] I_w \left( w, t \right) + I_t \left( w, t \right)$$

- optimal control (limiting most likely path)

$$u_t^* = -\sigma_t \left( w \right) I_w \left( w, t \right)$$

# HAMILTON–JACOBI–BELLMAN EQUATION

The associated Hamilton–Jacobi–Bellman equation is

$$0 = \inf_u \frac{1}{2} |u|^2 + [\mu_t(w) + \sigma_t(w) u] I_w(w,t) + I_t(w,t)$$

- optimal control (limiting most likely path)

$$u_t^* = -\sigma_t(w) I_w(w,t)$$

Hence we obtain a Riccati equation

$$0 = -\frac{1}{2}\sigma_t^2(w) I_w(w,t)^2 + \mu_t(w) I_w(w,t) + I_t(w,t)$$

with boundary condition $I(w,T) = \infty$ if $w < \mathrm{r}$ and $I(w,T) = 0$ otherwise.

# MOST LIKELY PATH

$$u_t^* = -\sigma_t(w)\, l_w(w,t)$$

Which shocks get you closer to the top quantile?

- shocks that occur when volatility $\sigma(w)$ is high (static effect)
- shocks that increase the probability of crossing the threshold quickly ($-l_w$ high, dynamic effect)

Compare this to the local displacement (here, $\sigma_t(w) = w\nu_t(w)$)

$$\frac{1}{2}\left[w\nu_t(w)\right]^2 g_t(w)$$

- again, static ($[w\nu_t(w)]^2$) and dynamic ($g_t(w)$) effect

Simple (but elegant) theory to aid measurement.

- leverages evidence on the approximate Pareto shape of the wealth distribution
- turns a non-parametric accounting exercise into a parametric estimation problem
- even without an explicit model of investor optimization etc.
- lots of robustness checks