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

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.
1. Brief summary of the paper
   • role of theory for measurement
   • questions

2. An alternative investigation
   • large deviation theory
   • non-local mobility
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
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)
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 \(\Rightarrow\) 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), ...
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 \( p \)

\[
S_t = \int_{q_t(p)}^{\infty} w g_t(w) \, dw
\]
Law of motion

\[ dS_t = S_tE[\mu_t(w) | w \geq q_t]dt + \frac{1}{2} [q_t\nu_t(q_t)]^2 g_t(q_t)dt \]

- \( q_t \) is the relative wealth level at quantile \( p \)

Where does the displacement term come from?

- probability current

\[ J(w, t) = w\mu_t(w)g_t(w) - \frac{\partial}{\partial w} \left[ \frac{1}{2} (w\nu_t(w))^2 g_t(w) \right] \]

- deterministic drift

- ‘churning’
Utilize empirical evidence on the Pareto shape of the upper tail of the wealth distribution.

\[ P(w_{it} \geq w) = Cw^{-\zeta} \]

- \(\zeta > 1\) is the shape parameter (higher \(\zeta\), less inequality)
- e.g., when \(\mu_t\) and \(\nu_t\) are independent of \(w\) and constant over time
 Utilize empirical evidence on the *Pareto shape of the upper tail* of the wealth distribution.

\[
P(w_{it} \geq w) = Cw^{-\zeta}
\]

- \( \zeta > 1 \) 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

\[
\frac{1}{2} (\zeta - 1) \nu^2
\]

- 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
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 $\Rightarrow \mu = 0$.
- 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 applies to the tail
- wealth relative to a different benchmark
The between/within industry decomposition would deserve more explanation.

- It seems that the decomposition uses two terms
  \[
  \frac{1}{2} (\xi - 1) \nu_{\text{within}}^2 \quad \text{and} \quad \frac{1}{2} (\xi - 1) \nu_{\text{between}}^2
  \]

  where \( \nu_{\text{within}}^2 \) and \( \nu_{\text{between}}^2 \) 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 \( \nu_{\text{within}}^2 \)).

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?
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 $\varepsilon$

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

We want to study

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

- other state variables possible (suppressed here)
Consider a class of wealth processes indexed by \( \varepsilon \)

\[
dw^\varepsilon_t = \mu_t (w^\varepsilon_t) \, dt + \sqrt{\varepsilon} \sigma_t (w^\varepsilon_t) \, dB_t.
\]

We want to study

\[
P (w^\varepsilon_T \geq r) \quad \text{given } w^\varepsilon_0 = \bar{w}.
\]

\[\cdot \text{ other state variables possible (suppressed here)}\]

Construct the function \( h_r (w) = 1 \{ w \geq r \} \). Then

\[
P (w^\varepsilon_T \geq r) = E_0 [ h_r (w^\varepsilon_T) ].
\]
VARADHAN’S LEMMA

We are interested in the limit

$$\lim_{\varepsilon \searrow 0} \varepsilon \log E_0 \left[ h_r \left( w_T^\varepsilon \right) \right] \overset{!}{=} -I (\bar{w}, r, T)$$

• as $\varepsilon \searrow 0$, the threshold $r$ is more relatively more distant, given the underlying uncertainty.
We are interested in the limit

\[
\lim_{\varepsilon \searrow 0} \varepsilon \log E_0 [h_r (W_T^\varepsilon)] = -I(\bar{w}, r, T)
\]

• 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:

\[
I(\bar{w}, r, T) = \inf_u \int_0^T \frac{1}{2} |u_t|^2 \, dt
\]

subject to

\[
\dot{w}_t = \mu_t (w_t) + \sigma_t (w_t) u_t, \quad w_0 = \bar{w}, w_T \geq r.
\]

• choosing a particular path of shock realizations leading to \( w_T \geq r \).
The associated Hamilton–Jacobi–Bellman equation is

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

- optimal control (limiting most likely path)

\[ u_t^* = -\sigma_t(w) l_w(w, t) \]
The associated Hamilton–Jacobi–Bellman equation is

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

- optimal control (**limiting most likely path**)

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

Hence we obtain a **Riccati equation**

\[ 0 = -\frac{1}{2} \sigma_t^2 (w) l_w (w, t)^2 + \mu_t (w) l_w (w, t) + l_t (w, t) \]

with boundary condition \( l (w, T) = \infty \) if \( w < r \) and \( l (w, T) = 0 \) otherwise.
\[ u_t^* = -\sigma_t(w) l_w(w, t) \]

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, \( \sigma_t(w) = w\nu_t(w) \))

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

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

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