New Evidence on Productivity Dynamics

Alexander Amundsen[1]   Markus Poschke[2]

[1]Department of Finance Canada

[2]McGill University

HEC Montréal, Dec 3, 2021

Motivation

- Firm level productivity varies enormously across firms even within narrowly defined industries.
- Top productivity firms are much more productive than others, and account for a large share of economic activity.
- In recent years, dispersion of productivity has widened further.
- Some have linked this to a global decline in aggregate productivity growth.
- Yet there has been no systematic exploration of the dynamics of top productive firms in the literature, mostly due to data limitations.

We provide new evidence on productivity dynamics

- using data on the universe of Canadian firms.
- The sample size allows us to *focus on top firms*.

Background and Research Questions

- Productivity is usually modelled as a stochastic process common to all firms.
- Generally a first-order Markov process, such as an AR(1).
- Typically also used in structural models of heterogenous firms.

Our aim: go beyond this.

- What do firm productivity dynamics look like, beyond an AR(1) estimate?
- Are they different for top firms?
- How well do standard models fit observed dynamics of productivity?
- Better models?

This Paper

1. Use data on the universe of Canadian firms to establish new facts on productivity dynamics, with a particular focus on top firms.
2. Estimate a rich parametric process that can account for central features of the data.
3. Show how departure from AR(1) matters for the effects of frictions and policies. [in progress]

Key data patterns

Confirm known patterns:

1. Mean reversion: Mean TFP growth declines with TFP level.
2. TFP levels and growth rates are fat-tailed.

Three new patterns:

1. g more dispersed, very left-skewed for top firms.
2. The top is very persistent: Half of all firms that make it to the top 1% TFP are still in the top 10% of TFP a decade later.
3. Mean g and probability of staying in top 1% *increase* with time at the top.

Key data patterns

Confirm known patterns:

1. Mean reversion: Mean TFP growth declines with TFP level.
2. TFP levels and growth rates are fat-tailed.

Three new patterns:

1. g more dispersed, very left-skewed for top firms.
2. The top is very persistent: Half of all firms that make it to the top 1% TFP are still in the top 10% of TFP a decade later.
3. Mean g and probability of staying in top 1% *increase* with time at the top.

We estimate a model that accounts for all these patterns

3-state regime switching model

3 AR(1) regimes for TFP that differ in mean, variance and persistence:

1. "regular" regime
2. "high variance" regime
3. "persistent high growth" regime

Heterogeneity and composition effects explain all three new facts:

1. Most top firms are "high growth", but some are "high variance".
   - $\Rightarrow$ left-skewed g for top firms.
2. Most top firms have highly persistent TFP.
3. Composition effects explain how growth increases with time at the top.

In contrast, AR(1) only allows mean g to vary with TFP.

Implications

- Effects of frictions (adjustment costs, financial frictions) and policies (EPL, ...) depend on the persistence of TFP.
- Top firms are different.

Data

Data

- Dataset from Statistics Canada: T2-LEAP
- Universe of incorporated firms in Canada that file a T2 (Corporation Income Tax Return) and register a payroll with the Canada Revenue Agency.
- Confidential, longitudinal, 2000-2015
- Covers all industries, includes both private and publicly listed firms, and includes information on investment and capital.

⇒ Not limited to manufacturing, unlike in US data

Our sample:

- 2002-2014 (to be sure to capture full years)
- keep firms with ≥ 5 employees, for most industries
- require: sales, net income before taxes and extraordinary items, payroll, capital, investment, amortization

⇒ balanced panel of 70,159 firms over 13 years: 912,067 observations

sample details

Measuring productivity

Production function:

$$VA_{it} = \exp(z_{it}) K_{it}^{\alpha} L_{it}^{\beta}$$

- Today: results using log *Total Factor Productivity (TFP)* z estimated using method by Collard-Wexler & De Loecker (2016)
  - allow α, β to differ by 4-digit industry
  - mean estimates: $\hat{\alpha} = 0.73, \hat{\beta} = 0.17$
  - productivity demeaned at industry-year level
  - productivity percentiles reported below are industry-specific
  - often focus on productivity growth $g_{it} \equiv z_{it} - z_{i,t-1}$
- Productivity dynamics similar when using *Labor Productivity* (VA/L) or TFP estimated using industry cost shares.

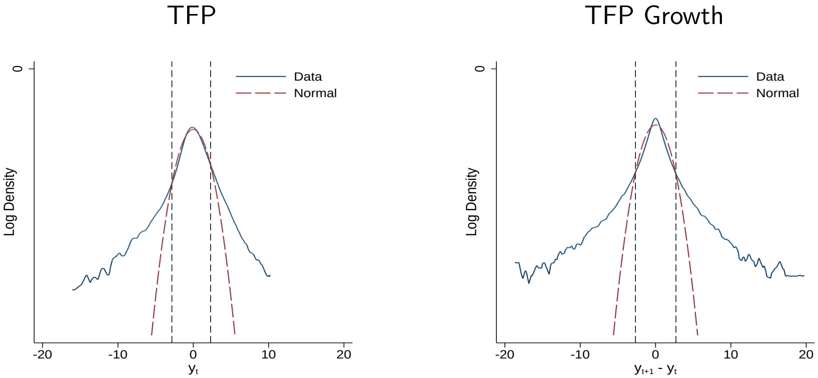
CWDL details

Data patterns

Summary Statistics

	Value added (million $)	Capital (million $)	Payroll (millions $)	Employment (persons)
Full Sample				
Mean	4.86	4.23	2.87	66
Standard Deviation	90.9	109	34.7	788
Top 1% by TFP				
Mean	83.4	80.9	27.4	715
Standard Deviation	514	608	152	4910
Share of total	0.21	0.23	0.12	

TFP and TFP growth have fat tails



This figure shows the density in logs of TFP and TFP growth compared to their normal distribution counterparts. TFP and TFP growth have been standardized. The normal distribution is derived using the estimated mean and standard deviation of each variable in its probability density function. The proportion of firms in the left tail past the vertical dashed line is equal to 1% and 1% for TFP and TFP growth, respectively. The proportion of firms in the right tail past the vertical dashed line is equal to 2% and 1% for TFP and TFP growth, respectively. Growth is defined as the *ln* difference between time $t+1$ and t. TFP is derived from the IV One-Step Control approach from CWDL.

log log plots

New Fact 1: g of top firms more dispersed, left-skewed

TFP percentile	mean	SD	Kelly skewness
0%-10%	0.16	0.41	0.28
10%-50%	0.02	0.20	0.05
50%-90%	-0.03	0.21	-0.08
90%-95%	-0.08	0.25	-0.20
95%-99%	-0.12	0.30	-0.29
99%-100%	-0.23	0.48	-0.40

Kelly skewness: $((p_{90} - p_{50}) - (p_{50} - p_{10}))/(p_{90} - p_{10})$.

The distribution of growth rates differs with TFP level:

Known: mean reversion

New: g of top firms more dispersed, very left-skewed

more

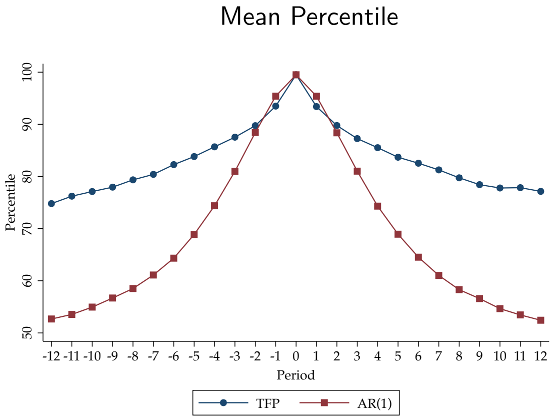
Some firms spend many years at the top

years in top 1%	fraction of all firms	fraction of those in top 1%
0	93.63	
1	3.35	52.55
2	1.15	18.05
3	0.57	8.96
4	0.36	5.71
5	0.29	4.52
6	0.17	2.60
7	0.12	1.95
8	0.10	1.52
9	0.07	1.10
10	0.05	0.83
11	0.05	0.72
12	0.04	0.63
13	0.06	0.87

For reference: AR(1) implies 0.01% spend 9 years in top 1%, 0.001% 12 years.

New Fact 2: Top status is very persistent

Where do top 1% firms come from/go to?



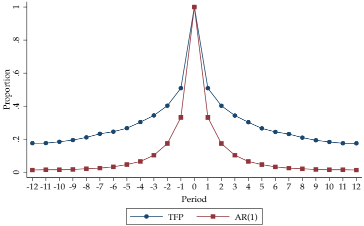
- Mean reversion from top slow.
- 12 years before/after a year in the top 1%, firms on average still at *p75*.

The AR(1) is simulated with parameters $\rho = 0.76$ and $\sigma = 0.24$.

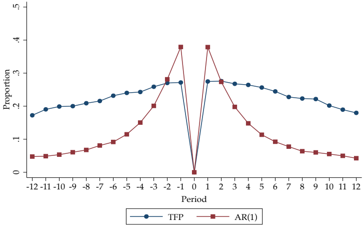
details

New Fact 2: Top status is very persistent

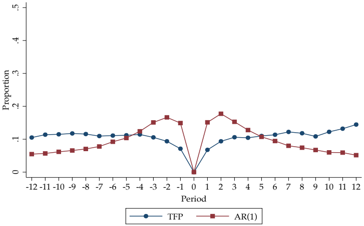
Top 1% to Top 1%



Top 1% to 95%-99%



Top 1% to 90%-95%



12 years before/after a year in the top 1%:

20% still in top 1%

40% still in top 5%

50% still in top 10%

back

These plots provide the mean proportion of firms in each percentile group for each year before (-) and after (+) being in the top

Transitions: Top TFP persists, but large shocks possible

	0%-10%	10%-50%	50%-90%	90%-95%	95%-99%	99%-100%
One-Year Transition Matrix (%)						
0%-10%	51.71	40.68	6.74	0.42	0.33	0.13
10%-50%	9.97	68.48	20.65	0.54	0.28	0.08
50%-90%	1.66	20.67	69.80	5.50	2.03	0.34
90%-95%	0.90	4.45	43.75	32.11	17.07	1.72
95%-99%	0.85	3.05	20.26	21.46	46.09	8.28
99%-100%	1.14	2.87	10.88	6.77	27.48	50.87
Five-Year Transition Matrix (%)						
0%-10%	25.62	50.07	21.09	1.73	1.16	0.33
10%-50%	12.13	53.50	30.85	1.98	1.27	0.26
50%-90%	5.35	31.09	52.85	6.14	3.75	0.81
90%-95%	3.39	14.94	51.07	15.27	12.63	2.71
95%-99%	2.86	11.09	38.68	17.08	22.98	7.31
99%-100%	2.79	9.12	24.91	10.99	25.65	26.54

For reference: with AR(1), bottom right elements are 32% and 4%.

New Fact 3: Those who persist at the top do better

years in top 1%	mean growth	Pr(stay in top 1%)
any	-0.23	50.9
1	-0.31	37.8
2	-0.19	55.8
3	-0.15	64.2
4+	-0.10	77.7

details

Data patterns: Summary

Confirm known patterns:

1. TFP levels and growth rates fat-tailed
2. Mean reversion: Mean TFP growth declines with TFP level

Three new patterns:

1. g more dispersed, very left-skewed for top firms
2. The top is very persistent: Half of all firms that make it to the top 1% TFP are still in the top 10% of TFP a decade later
3. g and probability of staying in top 1% *increase* with time at the top.

$\Rightarrow$ Want a model that matches all these features.

(AR(1) and similar processes miss all three new facts.)

Data patterns: Summary

Confirm known patterns:

1. TFP levels and growth rates fat-tailed
2. Mean reversion: Mean TFP growth declines with TFP level

Three new patterns:

1. g more dispersed, very left-skewed for top firms
2. The top is very persistent: Half of all firms that make it to the top 1% TFP are still in the top 10% of TFP a decade later
3. g and probability of staying in top 1% *increase* with time at the top.

$\Rightarrow$ Want a model that matches all these features.
(AR(1) and similar processes miss all three new facts.)

Data patterns: Summary

Confirm known patterns:

1. TFP levels and growth rates fat-tailed
2. Mean reversion: Mean TFP growth declines with TFP level

Three new patterns:

1. g more dispersed, very left-skewed for top firms
2. The top is very persistent: Half of all firms that make it to the top 1% TFP are still in the top 10% of TFP a decade later
3. g and probability of staying in top 1% *increase* with time at the top.

$\Rightarrow$ Want a model that matches all these features.

(AR(1) and similar processes miss all three new facts.)

Model

Typical modeling strategies

1. Model TFP as an AR(1) process
   - Captures mean reversion in TFP,

   but:
   - implies only mean g varies with z, not higher moments
   - not enough persistence at the top
2. Natural extension: AR(1) with fat-tailed innovations
   - Captures fat-tailed growth rates, and thus levels.
   - A fatter tail may increase measured persistence in top percentiles.

   but:
   - higher persistence at top, but not enough
   - implies only mean g varies with z, not higher moments

New evidence, new models

New evidence:

1. g more dispersed, very left-skewed for top firms
2. The top is very persistent.
3. g and probability of staying in top 1% increase with time at the top.

Our modeling approach: Markov regime-switching model

With enough regimes, can capture all new facts.

Markov Regime-switching model

Latent TFP z_t^* follows:

$$z_t^* = \begin{cases} \beta_1 z_{t-1}^* + \varepsilon_{1,t} & \text{if } S_t = 1 \\ \alpha_2 + \beta_2 z_{t-1}^* + \varepsilon_{2,t} & \text{if } S_t = 2 \\ \alpha_3 + \beta_3 z_{t-1}^* + \varepsilon_{3,t} & \text{if } S_t = 3 \end{cases}$$

where $\varepsilon_{i,t} \sim \sigma_i t(df), i = 1, 2, 3$.
Transition matrix for S_t:

$$P = \begin{bmatrix} 1 - p_1 & p_1 & 0 \\ \gamma p_2 & 1 - p_2 & (1-\gamma) p_2 \\ 0 & p_3 & 1 - p_3 \end{bmatrix}$$

Observed TFP z_t follows

$$z_t = z_t^* + u_t, \quad u_t \sim N(0, \sigma_u^2).$$

Remarks on model

- 3 regimes, which can differ in
  - persistence (β),
  - mean (α), and
  - variance (σ).
  - After experimentation, common *df* of innovations.
- To aid identification, rule out extreme regime switches.
- Common measurement error

Estimation: Simulated method of moments (SMM)

– confidential data

⇒ compute and export target moments

– hard to estimate process directly on Statscan server

– to estimate 14 parameters, use 68 moments on

- distribution of g by productivity level
- persistence of top
- growth by time at the top
- TFP levels

⇒ target old and new facts about productivity growth

Estimation: Simulated method of moments (SMM)

10 sets of target moments (68 moments)

1. Mean growth by productivity group (S_{μ}, 3 groups)
2. Standard deviation of growth by productivity group (S_{σ}, 3 groups)
3. Skewness of growth in top 1% (S_{skew})
4. Kurtosis of growth in top 1% (S_{kurtosis})
5. Distribution of years spent in top 1% (S_{yrs}, 0-13 years)
6. Average percentile before/after a year in top 1% (S_{pct}, years -12 to 12)
7. 1-year transition probabilities for top 1% firms (S_1, to 6 TFP groups)
8. 5-year transition probabilities for top 1% firms (S_5, to 6 TFP groups)
9. Mean growth by time spent in top 1% ($S_{\mu \text{ by } t}$, 1 to 4 years)
10. TFP level by productivity group (S_z, 6 groups)

Estimation: Simulated method of moments (SMM)

Objective function: find parameter vector θ that minimizes

$$[\hat{m}(\theta) - m]' \, W \, [\hat{m}(\theta) - m]$$

$\theta = \{\beta_1, \beta_2, \beta_3, \sigma_1, \sigma_2, \sigma_3, \alpha_2, \alpha_3, df, p_1, p_2, p_3, \gamma, \sigma_u\}$

Some issues:

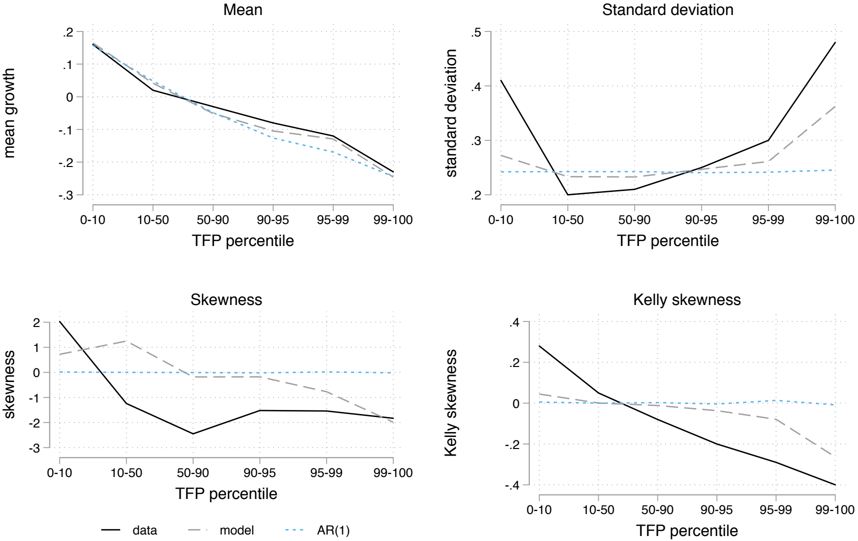
- weights
    - currently have no measures of standard errors of target moments
    - very large dataset: many moments very precise
    - currently, $W = I$
- scaling
    - moments differ in scale
    - some large, others very close to zero
    - currently, rescale moments to similar order of magnitude

Parameters and target moments

parameters	informative moments	(number)
$\beta_1, \beta_2, \beta_3$	$E(g\|z)$	3
$\sigma_1, \sigma_2, \sigma_3$	$var(g\|z)$	3
α_2, α_3	percentiles of z	6
df	kurtosis of g, top 1%	1
p_1, p_2, p_3, γ	skewness of g, top 1%	1
	distr. years at top	14
	avg. percentile 12 years before/after top	24
	tr. matrix, pr.(stay close to top)	6
	$E(g\|z, \text{time at top})$	4
σ_u	tr. matrix, large downward jumps	4
14		68

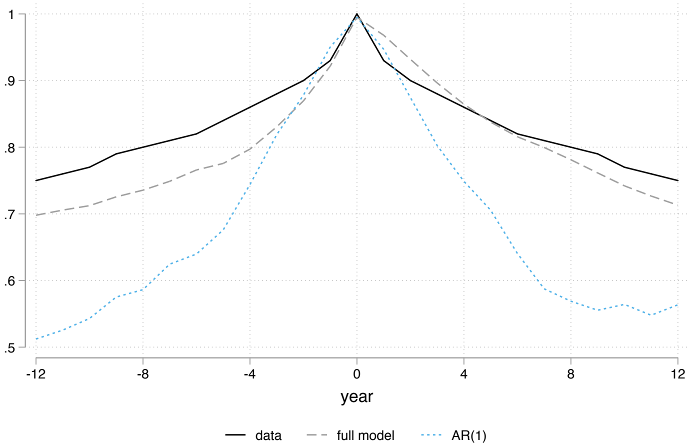
Model fit and estimates

Model fit: growth by TFP level

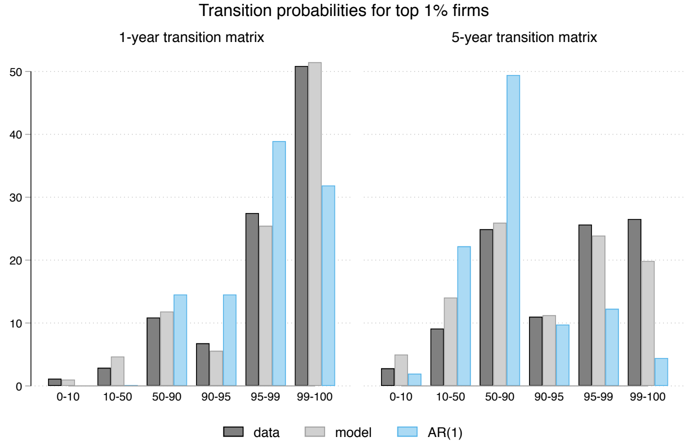


▸ Table

Model fit: Where do top 1% firms come from/go to?



Model fit: Top firms growth and transitions



Mean growth by time in top 1%

years in top 1%	1	2	3	4
data	-0.31	-0.19	-0.15	-0.1
model	-0.29	-0.25	-0.19	-0.13

Parameter estimates

Latent TFP z_t^* follows:

$$z_t^* = \begin{cases} 0.76 z_{t-1}^* + \varepsilon_{1,t} & \text{if } S_t = 1: \quad \text{“regular” regime} \\ 0.018 + 0.61 z_{t-1}^* + \varepsilon_{2,t} & \text{if } S_t = 2: \quad \text{“high variance” regime} \\ 0.033 + 0.91 z_{t-1}^* + \varepsilon_{3,t} & \text{if } S_t = 3: \quad \text{“high growth \& persistence”} \end{cases}$$

where $\varepsilon_{i,t} \sim \sigma_i\, t(4.15)$, $\sigma_1 = 0.14$, $\sigma_2 = 0.37$, $\sigma_3 = 0.12$.

Fraction in S_1, S_2, S_3: 0.79, 0.04, 0.17
Prob. leave: $p_1 : 0.4\%, p_2 : 5.2\%,\ p_3 : 2.6\%$

Measurement error: Observed TFP z_t follows

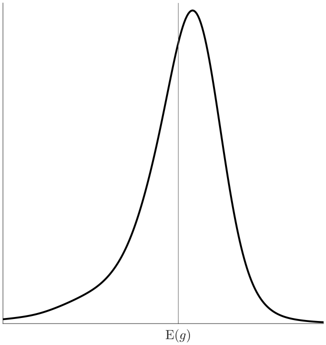
$$z_t = z_t^* + u_t, \quad u_t \sim N(0, 0.069^2).$$

For comparison: AR(1) has $\beta = 0.76, \sigma = 0.24$.

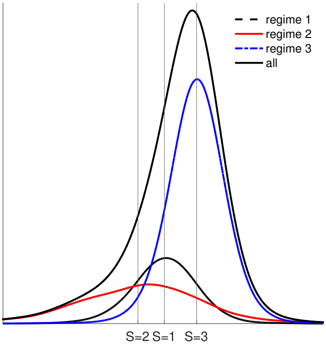
New Fact 1: Growth of the top 1%

High variance & left-skewed because of high-variance regime 2 firms, which are over-represented at the top.

All firms

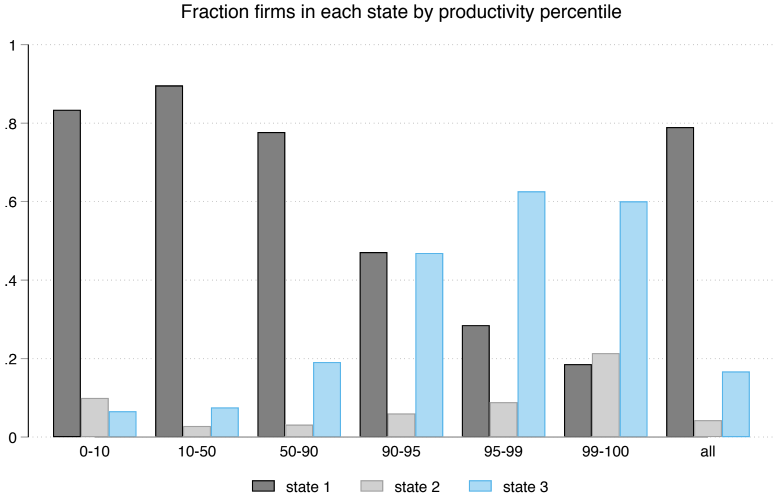


By regime



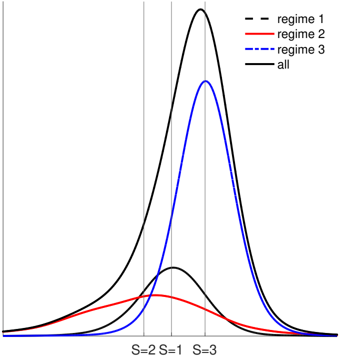
New Fact 2: Top productivity is more persistent

Arises because most top firms are in regime 3 ⇒ high persistence



New Fact 3: g and persistence rise with time at top

Arises because high-growth regime 3 firms more likely to stay at top.

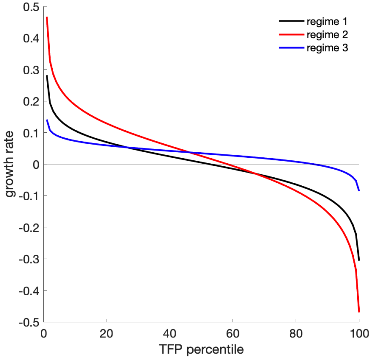


Transition matrix for S_t:

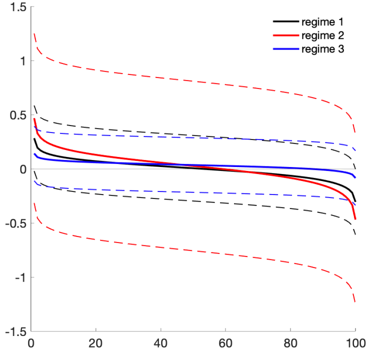
$$P = \begin{bmatrix} 0.9967 & 0.0033 & 0 \\ 0.061 & 0.925 & 0.014 \\ 0 & 0.0035 & 0.9965 \end{bmatrix}$$

Growth rates by regime

Mean growth rate by z and S



Mean and percentiles 5, 95



How the model accounts for key data features

New facts:		Mechanism:
1. differences in higher moments of g by z	$\Leftarrow$	Some top firms have high, persistent growth, others high variance.
2. high persistence of top productivity	$\Leftarrow$	More persistent types at the top.
3. higher growth when at top longer	$\Leftarrow$	Those who stay at top more likely to be high-growth types.

New data patterns explained through heterogeneity and composition effects.

▸ details

All model elements matter

- Top regime features faster growth, is very persistent.
  - If either of these is not in place:
    Top firms return to median more quickly, growth *falls* with time at top.
    figure
- Middle regime has higher variance.
  - Crucial for higher standard deviation and left skewness of growth of top firms. figure
- Measurement error matters little.
- Standard AR(1) misses both key features of top 1% TFP dynamics:
  - high variance and skewness of growth rates
  - high persistence of top status
- AR(1) with fat tails:
  - Generates somewhat higher top persistence (transition matrices, time after top), but figure
  - common higher moments of g by z, common mean g by time in top.

Conclusion

Conclusion & Discussion

We have shown three new data patterns for TFP dynamics:

1. g more dispersed, very left-skewed for top firms.
2. The top is very persistent.
3. g and probability of staying in top 1% increase with time at the top.

A Markov switching model with 3 regimes can account for these patterns.

Our findings have further implications for work on

– productivity estimation

– firm dynamics

Conclusion & Discussion

We have shown three new data patterns for TFP dynamics:

1. g more dispersed, very left-skewed for top firms.
2. The top is very persistent.
3. g and probability of staying in top 1% increase with time at the top.

A Markov switching model with 3 regimes can account for these patterns.

Our findings have further implications for work on

- productivity estimation
- firm dynamics