Top productivity dynamics*

Alexander Amundsen[†] Markus Poschke[‡]

April 11, 2026

Abstract

This paper presents new facts on the dynamics of firm level productivity, with a particular focus on the top 1% of firms in the productivity distribution. We do so by applying a variety of non-parametric techniques inspired by the recent literature on earnings dynamics to two datasets each containing 1-2 million observations on around 100,000 firms: one dataset covering the universe of incorporated employer firms in Canada, and another one covering several European countries. We show three new facts: 1) more productive firms face a growth rate distribution with a higher variance and more negative skewness; 2) top productivity status is very persistent: notably, half of the firms that make it to the top 1% productivity are still in the top 10% of productivity a decade later; and 3) the longer a firm remains in the top 1%, the more likely it is to stay there, and the higher its expected productivity growth rate. These patterns are inconsistent with the common assumption that firm level productivity follows an autoregressive process. Instead, we show that a parsimonious hidden type autoregressive model with three types fits the data, including the new facts, well. Top firms are drawn from two types, a "high growth and persistence" one and a "high variance" one. Dynamic selection implies that those persistently at the top are from the high growth and persistence type. We show that these findings matter for the predictions of macroeconomic models with heterogeneous firms.

Keywords: Firm dynamics; Productivity distribution
JEL Code: D24; C14; L10; O30

*First version: Dec 2022. The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management. Masaya Takano and Harpreet Singh provided outstanding research assistance. We thank Vasco Carvalho, Lukas Freund, David Rivers and Kjetil Storesletten for their comments, and Mons Chan for an oustanding discussion. We have also benefited from the input of seminar participants at HEC Montréal, McGill University and the Federal Reserve Bank of St. Louis, at the meetings of the Canadian Economic Association and the Brazilian Econometric Society, the Banff Innovation, Productivity and Growth Conference 2024, the Swiss Macro Workshop 2025 and the Canadian Macroeconomic Study Group.

[†] International Monetary Fund. E-mail: aamundsen@imf.org

[‡] Department of Economics, McGill University. E-mail: markus.poschke@mcgill.ca

1 Introduction

Productivity varies enormously across firms even within narrowly defined industries (Syverson, 2011). Top firms are much more productive than others and account for a large share of economic activity. How do these firms grow to become so large and productive? How does their productivity evolve, on the way to the top and once there? How persistent is the top? What is the chance that a median-productivity firm reaches the top?

Answers to these questions matter for understanding the nature of firm and aggregate growth, as well as for the effects of policies.[1] Yet, existing work has provided only partial anwers to these questions, by suggesting the presence of fixed effects that differ across firms.[2] To our knowledge, there is no work investigating how the *dynamics* of productivity vary across the distribution, especially with a focus on top firms.[3] This state of affairs possibly reflects the data requirements of such an analysis, since few existing data sources allow measuring productivity for detailed groups of firms and following them over time, while also covering the entire economy, and not just selected sectors.

Our first contribution is to document in detail the dynamics of productivity and its variation across the productivity distribution. We do so drawing on two very large data sets, each containing 1-2 million observations on close to 100,000 firms. Our analysis uncovers novel facts on productivity dynamics, showing that they differ enormously between the top of the distribution and the rest. Stochastic processes typically used to model productivity in quantitative work, like a common autoregressive process, cannot capture this. We provide a parsimonious alternative that can. We also show that these new facts have important implications for macroeconomic analyses of firm dynamics and productivity.

Key to our analysis is the wealth of data we use. The first dataset is a large administrative data set, called T2-LEAP, from Statistics Canada. This data set covers essentially the universe of incorporated employer firms in Canada. It has some unique advantages for our analysis. Comparing to US sources for example, it covers the entire economy (unlike the Annual Survey of Manufactures), and allows us to construct a measure of capital for firms in all sectors (unlike the Longitudinal Business Database). Our main sample is a balanced panel of more than seventy thousand firms, which we follow over 13 years, implying close to a million observations. Results are similar for an even larger unbalanced panel. As a result, the data provide good coverage throughout the distribution, even at the top.[4]

Our second data set is Orbis, which is a global dataset containing balance sheet and statement

[1] For example, policies giving investment support to young or small firms are more beneficial if these firms actually have the potential to reach the top. The effect of policies aiming to allocate resources towards more productive firms, like structural reforms, will also depend on the persistence of TFP. Finally, it is well known that the persistence of productivity matters for the effect of frictions and adjustment costs (Moll, 2014).

[2] Luttmer (2011) argues that accounting for the relatively young age of large firms (mere decades) requires persistent heterogeneity in firm growth rates. Sterk et al. (2021) account for patterns in firm employment growth with a model featuring fixed effects in both long-run productivity and initial conditions.

[3] We discuss one exception below.

[4] In ongoing work, we are expanding the time coverage to include more recent years.

of income information of private and listed firms. We focus on a set of European countries with the broadest and most consistent coverage. By combining the observations across countries, we obtain a large dataset of 2 million observations on over 100,000 firms, which implies that we have detailed coverage of the top of the distribution. Strikingly, the facts we document are not only qualitatively but also quantitatively quite similar in the two datasets, suggesting that they are not specific to the Canadian context, and that they are not driven by data issues.

Our analysis proceeds in four steps. We first estimate total factor productivity (TFP) at the firm level. For robustness, we do so in several ways. The main results we present rely on a measure of TFP derived from a new production function estimation technique by Collard-Wexler and De Loecker (2025) (CWDL) that explicitly accounts for measurement error in capital. Results are similar for traditional measures of TFP relying on input cost shares, or for labour productivity, as we show in an Appendix. These data confirm significant concentration of TFP. The distribution of TFP levels has fat tails, and the most productive firms are large. We also find that TFP growth rates are fat tailed.

In the second stage of our analysis, we study the dynamics of productivity across the distribution. The methods we use here are inspired by the literature on heterogenous earnings dynamics (see e.g. Guvenen et al. (2014, 2015)). These techniques allow for a very rich description of productivity dynamics, and have not yet been applied to data from firms.

The results from this analysis can be stated as three new facts. First, the distribution of growth rates varies significantly with a firm's current TFP level. Not only is expected TFP growth lower for more productive firms – productivity is mean reverting – but more productive firms also have a much larger variance of TFP growth, as well as a more left-skewed distribution of growth rates. This is particularly pronounced for the top 1% most productive firms. This is a clear departure from standard autoregressive processes popular in the literature, in particular in work in quantitative macroeconomics, starting with Hopenhayn and Rogerson (1993).

Second, following top firms before and after their time at the top, we find that the top is a highly persistent state. Twelve years before or after a year in the top 1%, firms on average are still in percentile 75. 18% are still in the top 1% twelve years later, and around half are still in the top 10%. Hence, despite strong mean reversion, the top is a very persistent state.

Third, we show that firms that have been in the top percentile longer have higher expected future productivity growth and a higher probability of remaining at the top.

These three new facts indicate clear heterogeneity in productivity dynamics across the distribution. In a third step, we estimate a parsimonious parametric stochastic process that can capture this heterogeneity and that is useful for applied, quantitative work. We show that a hidden-type first order autoregressive process with three types captures the data well. It provides a good fit to all new facts, as well as to the densities of TFP levels and growth. Estimates imply that firms belong to one of three types: a "regular" type with intermediate shock persistence (around 0.76 annually) and variance; a "high variance" type with much lower shock persistence, but much more dispersed innovations; and a "high growth and persistence" type with higher overall productivity and annual persistence above 0.9.

Firms of the latter type are overrepresented at the top, explaining the high persistence of top productivity. Nevertheless, firms of the high variance type also make up a non-negligible share of the top productivity group – but typically not for long, so that dynamic selection results in growth rates that increase with time spent at the top. The highly variable, left-skewed growth rates of top firms are the result of the mixture of the growth rate distributions of high growth and high variance firms. This brief discussion already reveals that all features of the parsimonious process we write down are required to match the data. Somewhat surprisingly, additional firm TFP fixed effects do not improve the ability of the model to match the new facts. We believe that the process we estimate will be very useful for future work, since it is easy to integrate in quantitative models of firm dynamics.

In a fourth and final step, we show the importance of type heterogeneity in a standard model of firm dynamics with financial frictions. [in progress]

Related literature. Our work is part of a primarily empirical literature that seeks to advance our understanding of productivity dynamics. It also speaks to a large literature in quantitative macroeconomics that relies on estimated processes of firm TFP dynamics as an input, and to a further literature that aims to understand the sources of firm-level TFP growth.

The first literature is small, as it is limited by the availability of large, high quality data sources. Closest to our work are Sterk et al. (2021) and Jaimovich et al. (2025). Sterk et al. (2021) estimate auto-covariance functions of firm-level employment in data from the US LBD and show that a model with permanent firm heterogeneity can account for the slow decay of auto-covariances over time. Their work is similar in spirit to ours in highlighting the importance of heterogeneity, but focusses on heterogeneity in levels, whereas our findings highlight heterogeneity in the stochastic process for TFP across firms. Jaimovich et al. (2025) use data on revenue of Spanish firms from Orbis. Like us, they show departures of dynamics from a Gaussian AR(1) process, focussing in particular on fat tails in revenue dynamics. Neither of these papers addresses how dynamics differ across the TFP distribution. Our focus on the top 1% of firms – which account for a disproportionate share of value added and input use – is particularly novel to the literature. In addition, we estimate a particularly parsimonious stochastic process for firm-level productivity.

Such stochastic processes are a key input in quantitative models of firm dynamics. Starting with Hopenhayn and Rogerson (1993) and with few exceptions, this literature – which is too large to cite here – tends to assume that firm productivity follows an AR(1) process. It is well known that details of this process, like productivity persistence, matter crucially for model predictions, for example on the importance of financial frictions (Moll, 2014). Sterk et al. (2021) provide a similar result, and Jaimovich et al. (2025) highlight the importance of accurate estimates of revenue dynamics for understanding firm exit over the business cycle. The parametric stochastic process we estimate might be the smallest necessary departure from the AR(1) process popular in this literature that can generate the new evidence on productivity dynamics, and hence should be a highly useful input here.

Our findings, by highlighting a small number of firm types, also shed light on the nature of firm growth. In this, they connect to a literature on innovation. Here, Luttmer (2011) has highlighted in a theoretical analysis that persistent heterogeneity in firm growth rates can explain how firms can grow to become very large in decades, not centuries. Benhabib et al. (2021) study how innovation versus adoption decisions generate this heterogeneity and shape the productivity distribution. Empirical work in this space has typically focussed on innovation vs adoption, and has thus relied on patent or R&D data (see e.g. König et al., 2022; Akcigit and Ates, 2023), not productivity dynamics.

The remainder of the paper is organized as follows. Section 2 describes the data and descriptive statistics. Section 3 discusses the methods to estimate productivity. Section 4 presents the empirical analysis. In Section 5, we show that a parametric process fits the empirical patterns well. Section 6 explores the implications of our findings in a model of firm dynamics. Section 7 concludes with a discussion.

2 Data

In order to conduct a robust investigation into the properties of the productivity process, we require a dataset that is sufficiently large across industries to allow even the top 1% of the productivity distribution to be well-populated by firms. To do this, we leverage two comprehensive datasets: i) Administrative Canadian data from Statistics Canada covering the universe of firms; and ii) Orbis European data covering a wide set of countries (Belgium, France, Italy, Spain, Sweden, United Kingdom, Finland, and Portugal).

The dataset for Canada is referred to as T2-LEAP, and is the combination of two files, the Corporate Tax Statistical Universe (T2) and the Longitudinal Employment Analysis Program (LEAP). It provides annual longitudinal information on the universe of incorporated firms in Canada that file a T2 and register a payroll with the Canada Revenue Agency. It covers the time period 2000-2015 and includes all industries. There is detailed information on the firm's balance sheet and statement of income. The investment data come from T2 Schedule 6, 8 and 10, and provide information on the acquisition and disposition of capital. This data set is comparable to the U.S. Annual Survey of Manufactures, but with the advantage that it covers all sectors.

The dataset for Europe comes from Moody's (Bureau van Dijk) Orbis, which is a global dataset containing balance sheet and statement of income information of private and listed firms. Since Orbis is not administrative, it can vary over time in terms of coverage and composition. We therefore focus the analysis on countries with the broadest and most consistent coverage. By combining the observations across countries, we obtain a large dataset with sufficiently detailed coverage of the top. The time span covers 2005-2022.

For the purposes of this paper, we impose restrictions on each dataset, but maintain a mindset of striking a balance between excessive data cleaning and excluding spurious firms and/or measurement errors.

For the Canadian data, we begin by dropping the first and last year of each firm, as well as the preceding and subsequent year around gaps in each firm's panel. This is to ensure that the data we observe is for a full operating year. Firms must have a non-missing industry. This permits us to demean each productivity measure by its respective 4-digit NAICS industry-year mean, allowing us to combine the analysis across industries. Firms must have a positive age, where age is defined as the difference between the current year and the year the firm first started paying payroll. This ensures that we exclude the first (partial) year of the firm's life as an employer. Each firm must have at least 5 or more employees in each year. This type of restriction is common in the literature, and in administrative datasets is needed to exclude non-producing firms. For example, there are tax advantages to incorporating oneself and if that firm paid a payroll to its owner, it would then appear in our dataset. Similarly, small family run firms could take advantage of employing family members to distribute profits across lower marginal tax rates. These sort of issues could lead to spurious results on their productivity dynamics. We also require that firms have positive non-missing values of value added, payroll, capital, and intermediate inputs. Finally, we exclude some industries to focus on industries where there are extensive private markets and capital used is mostly tangible, as opposed to financial capital.[5]

Table 1: Variable Definitions in T2LEAP

Variables	Definition
S_{it}	Sales (gross output). Total revenue for the firm.
L_{it}	Payroll. Total payroll for the firm.
K_{it}	Capital. Total capital stock for the firm. Sum of tangible and intangible assets - accumulated amortization of tangible and intangible assets in the first year of each firm. Built up over time by the Perpetual Inventory Method.
I_{it}	Investment. Total investment for the firm. Sum of net investment in tangible and intangible capital.
VA_{it}	Value added. Sum of net income before taxes and extraordinary items + payroll + amortization of tangible capital + amortization of intangible capital.
M_{it}	Intermediate Inputs. Sales - Value added.

For the European data, we impose similar restrictions to mimic the Canadian sample while adjusting for its unique characteristics. We follow the procedures in Díez et al. (2021) and Kalemli-Özcan et al. (2024) for cleaning up duplicate observations. We also focus on twelve-month annual filings, firms with non-missing 3-digit NACE industry (which are used to demean productivity by

[5]We exclude the following industries from the analysis: Utilities (NAICS 22), Finance and Insurance (NAICS 52), Real Estate and Rental and Leasing (NAICS 53), Management of Companies and Enterprises (NAICS 55), Educational Services (NAICS 61), Health Care and Social Assistance (NAICS 62), and Public Administration (NAICS 91).

industry-year across countries), and firms with positive age. For our baseline sample, we further drop micro firms (< 5 employees), observations with negative or missing values of value added, payroll, capital, or intermediate inputs, and a similar set of utilities, public sector and financial industries as in the Canadian data.[6]

In Table 1 we provide the variable definitions used to measure productivity. Payroll is provided in LEAP and represents the total payroll of the firm. Capital is measured in the first sample year of each firm as the sum of tangible and intangible capital minus the accumulated amortization of tangible and intangible capital. It is then built up over time by the Perpetual Inventory Method, using investment in tangible and intangible capital and 2-digit NAICS industry-year specific depreciation rates. Value added is not explicitly reported in T2LEAP, so we compute it as the sum of net income before taxes and extraordinary items + payroll + amortization of tangible capital + amortization of intangible capital. Intermediate inputs are backed out as the difference between sales and value added. All variables are deflated by their respective 2-digit NAICS industry deflators from Statistics Canada.[7]

Most variables in Orbis align with those in T2LEAP. Age is defined as the difference between the current year and the firm's incorporation date. Value-added is defined as Earnings Before Interest, Taxes, Depreciation, and Amortization + Payroll. All variables are deflated using deflators from Eurostat and from the World Bank's World Development Indicators. The main difference with respect to T2LEAP concerns investment, which is not available in Orbis. We thus measure it with the perpetual inventory method, using 2-digit NACE depreciation rates estimated from the data. To do so, we measure capital as the after-depreciated after-amortized book value of tangible and intangible assets.

The T2LEAP final sample is a balanced panel that covers the years 2002 to 2014. It contains data on 70,159 firms over 13 years, or 912,067 firm-year observations. This sample represents 43% of sales and 39% of payroll for the respective industries. The final Orbis sample is a balanced panel covering the year 2005 to 2022. It contains 2,084,058 firm-year observations, covering 115,781 firms.[8]

[6]We drop the following industries: NACE 35: Electricity, gas, steam and air conditioning supply; NACE 36: Water collection, treatment and supply; NACE 37: Sewerage; NACE 64: Financial service activities, except insurance and pension funding; NACE 65: Insurance, reinsurance and pension funding; NACE 66: Activities auxiliary to financial services and insurance: NACE 68 Real estate activities; NACE 70: Activities of head offices; management consultancy activities; NACE 77: Rental and leasing activities; NACE 84: Public administration and defence; compulsory social security; NACE 85: Education; NACE 86: Human health activities; NACE 87: Residential care activities; NACE 88: Social work activities without accommodation.

[7]We use Statistics Canada (7 01a) to derive the depreciation rates and deflate capital and investment, Statistics Canada (LEAP) to deflate payroll, and Statistics Canada (7 01b) to deflate sales and value added. Intermediate inputs is deflated by the double deflator method from sales and value added.

[8]For robustness, we also analyze unbalanced panels. Th unbalanced T2LEAP panel represents 60% of sales and 56% of payroll for the respective industries.

3 Productivity Estimation

Productivity is a measure of efficiency in production: how much output can be obtained for a given set of inputs. In the literature, a number of methods have been developed to estimate productivity, each with their advantages and disadvantages. Since we are interested in the dynamics of productivity, we need to estimate it for each firm-year observation. To ensure our results are robust to deficiencies in any particular method, multiple measures will be estimated and compared.[9]

3.1 Approaches to productivity measurement

The simplest measure of productivity is labour productivity. Labour productivity has the advantage that it is a single-factor measure, making it easy to understand and does not require any estimation to compute. Its biggest disadvantage is that it is affected by the intensity of use of the excluded inputs. For example, if one firm uses capital more intensively than another firm, their labour productivity may differ substantially when in reality they both have the same level of total factor productivity. Let VA_{it} be value added and L_{it} be payroll. Then log labour productivity for firm i at time t is

$$LP_{it} = log\left(\frac{VA_{it}}{L_{it}}\right). \tag{1}$$

We use its logged form to facilitate comparability to other productivity measures that are typically denoted in logs.

To accommodate capital in the productivity measure, the second measure that is widely used is the factor share (FS) method. This method measures total factor productivity (TFP) and is invariant to the intensity of use of the inputs, reflecting variation in output for a given set of inputs. The FS method relies on a number of assumptions to make it feasible as a productivity measure. It assumes firms are cost-minimizers, where they equate each input's output elasticity to its cost share.[10] If cost shares are not observable, we can make further assumptions of perfect competition and constant returns to scale, which results in the output elasticities equaling the revenue shares (Syverson (2011)).[11]

Suppose firms face a Cobb-Douglas production function with Hicks-neutral TFP:

$$Y_{it} = A_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L} \tag{2}$$

where Y_{it} is output, A_{it} is TFP, K_{it} is capital, L_{it} is labour, α_K is the output elasticity for capital and α_L is the output elasticity for labour. Since the Cobb-Douglas is log-additive, it can be re-expressed

[9] See Van Beveren (2012) for an excellent survey of the productivity estimation literature.

[10] An important caveat are adjustment costs. If they exist, then the input's output elasticity will not be equal to the cost share. This can be mitigated by estimating the cost shares across firms either cross-sectionally or longitudinally, but nevertheless the problem remains (Syverson (2011)).

[11] With constant returns to scale, one simply needs payroll and value added to calculate both the output elasticity of labour and the output elasticity of capital.

in its log-linear form:

$$log(Y_{it}) = log(A_{it}) + \alpha_K log(K_{it}) + \alpha_L log(L_{it}) \tag{3}$$

Under the assumptions stated above, revenue shares can be used to calculate α_K and α_L.[12] Then TFP follows as

$$log(A_{it}^{FS}) = log(Y_{it}) - \alpha_K log(K_{it}) - \alpha_L log(L_{it}) \tag{4}$$

where $log(A_{it}^{FS})$ denotes the TFP measure from the FS method.

An alternative to these techniques is to estimate the coefficients of the Cobb-Douglas production function for each 4-digit NAICS industry. Estimating (3) involves estimating:

$$log(Y_{it}) = \omega_0 + \beta_K log(K_{it}) + \beta_L log(L_{it}) + \omega_{it}, \tag{5}$$

where β_K and β_L are the parameter estimates for capital and labour, respectively, and $log(A_{it}) = \omega_0 + \omega_{it}$ such that ω_0 is the common TFP component across all firms in each 4-digit NAICS industry and ω_{it} is the idiosyncratic TFP component of each firm. The term we are interested in is ω_{it}, which measures the relative position of the firm in the TFP distribution.

As is well known, there is a major concern to estimating (5): endogeneity bias. As mentioned already by Marschak and Andrews (1944), inputs in the production function are not independently chosen and are a function of a firm's productivity. If the firm has a prior belief on their productivity, then there will be correlation between the regressors and the error term. This violates the exogeneity assumption for OLS, and promotes the use of more advanced semi-parametric techniques that the literature has developed.[13] This includes: Olley and Pakes (1996) (OP), who proxy for the unobserved productivity shock with investment; Levinsohn and Petrin (2003) (LP), who proxy for the unobserved productivity shock with intermediate inputs; Ackerberg et al. (2006) (ACF), who demonstrate an identification problem for the labour coefficient in OP and LP and suggest an alternative estimation procedure; and Wooldridge (2009) (WOOL), who implements the OP, LP and ACF estimators using a one-step GMM approach.

One issue common to these estimation procedures is that they typically yield unrealistically low capital coefficients. Collard-Wexler and De Loecker (2025) argue that this reflects the fact that among the variables used in the estimation, capital is particularly susceptible to measurement error. Using Monte Carlo simulations, they find that this leads to a downward bias in the capital coefficient when using standard approaches. They therefore develop a framework that incorporates log-additive classical measurement error in the capital stock. Using the LP framework of intermediate inputs to control for unobserved productivity, they show that lagged investment can be used as an instrument

[12] The revenue shares are calculated in each 4-digit NAICS industry using the ratio of total payroll and total value added, where the total is taken as the sum across firms and years within the industry. For each 4-digit NAICS industry: $\alpha_L = (\sum_{j=1}^{N} L_j)/(\sum_{j=1}^{N} VA_j)$ and $\alpha_K = 1 - \alpha_L$, where j is the firm-year observation.

[13] A second concern for estimating (5) is the selection bias. The firm's choice of inputs is dependant on its survival, and if the firm has a prior belief on their productivity, which affects its exit, then there will exist correlation between the regressors and the error term (Wedervang (1965)). Empirically, using an unbalanced sample seems to mitigate the selection bias to a large degree (Van Beveren (2012)).

to control for measurement error, since it is uncorrelated with the error term but correlated with the capital stock. Testing their measure on data from several countries, they find capital coefficients that are roughly twice as large when compared to standard measures. In our analysis of the data, we confirm much higher capital coefficients using the IV One-Step Control approach from CWDL, suggesting that measurement error may be an issue for capital in the Canadian and Orbis data as well. This motivates our use of the measure for the analysis.

The semi-parametric techniques, including CWDL, assume that unobserved productivity follows a first-order Markov process. This allows for the invertibility condition, where the proxy demand function is inverted to so that productivity is a function of observables (Ackerberg et al. (2007)). The methods can in principle be adapted to use higher order processes and Ackerberg et al. (2007) show this in the OP framework for a second-order Markov process, as well as another case where two productivity processes exist, a controlled process that is a function of the firm's R&D and an uncontrolled process that is exogenous. The majority of the empirical literature, however, employs the standard methods that assume a first-order Markov process. This raises an important question of the validity of their results if productivity deviates substantially from this process.

For our purposes, there is a tradeoff between using different types of productivity measures. We consider labour productivity and TFP computed using the FS and CWDL measures. Each method has its advantages and disadvantages. Labour productivity is easily observed, and for this reason is also widely used. At the same time, it is not TFP, and is affected by the excluded inputs. Moreover, in frictionless settings with competitive factor markets, labour productivity should be equalized even across firms that differ in TFP (Hsieh and Klenow (2009)). TFP does not suffer from these shortcomings, but is harder to measure. FS productivity makes strong assumptions about firm behaviour, market competition and returns to scale. CWDL remedies some of these shortcomings, but at the cost of strong assumptions about the unobserved productivity process.

To address these issues, we conduct our analysis using all three productivity measures. The fact that our results are consistent across these measures implies that our findings on productivity dynamics are robust to the estimation method.

3.2 Production function estimates

Both the FS and CWDL methods require measurement of cost shares, or production function estimation. We provide the production function estimates in Table 2 for both the FS method and IV One-Step Control approach from CWDL. Since we are the first to apply the CWDL approach using Canadian data, we include the FS estimates as a point of reference. We estimate production functions at the 4-digit industry level. The table presents the mean, median and standard deviation of the capital coefficient, labour coefficient and returns to scale. It shows these statistics both across all observations and across 4-digit NAICS industries.[14] For the FS estimates, we find a capital coefficient of 0.33, in line with typical estimates. Since constant returns to scale is assumed, the

[14]In this case, each 4-digit NAICS industry has equal weight.

labour coefficient is its remainder from one. For the CWDL estimates, we find a lower capital coefficient and a labour coefficient close to the typical value around 0.67. While the CWDL capital coefficient is lower than the FS one, it is still significantly greater than what is typically obtained using alternative methods like OP. Finally, means across industries or across observations are similar.

Table 2: Production Function Estimates

	Capital Coefficient	Labour Coefficient	Returns to Scale
Factor Share (FS)			
Across Observations			
Mean	0.29	0.71	1.00
Median	0.26	0.74	1.00
Standard Deviation	0.11	0.11	0.00
Factor Share (FS)			
Across 4-digit Industries			
Mean	0.33	0.67	1.00
Median	0.30	0.70	1.00
Standard Deviation	0.16	0.16	0.00
IV One-Step Control (CWDL)			
Across Observations			
Mean	0.17	0.73	0.90
Median	0.15	0.73	0.89
Standard Deviation	0.09	0.08	0.08
IV One-Step Control (CWDL)			
Across 4-digit Industries			
Mean	0.20	0.67	0.88
Median	0.18	0.69	0.89
Standard Deviation	0.19	0.17	0.26

This table presents the mean, median and standard deviation of the capital coefficient, labour coefficient and returns to scale across all observations and across equally weighted 4-digit industries. The estimates derived from a Cobb-Douglas production function using the factor share (FS) method and the IV One-Step Control approach from Collard-Wexler and De Loecker (2025) (CWDL).

To provide a first simple description of TFP dynamics, we estimate an AR(1) process on the CWDL measure. The estimates, ρ and σ, are found by estimating the following equation by OLS:

$$\omega_{it} = \rho\omega_{it-1} + \epsilon_{it}, \tag{6}$$

where ω_{it} is the idiosyncratic TFP component from the CWDL measure, ρ is the persistence parameter and ϵ_{it} is the error term, which is normally distributed $N(0, \sigma^2)$.[15] We find $\rho = 0.76$ and $\sigma = 0.24$. In the remainder of the paper, when we present results for the AR(1), we compute them using a synthetic dataset that is 5 times larger than our sample (4,560,335 observations) with a burn-in equal to 5 times our sample length (65 periods).[16] This helps to minimize simulation error due to random sampling.

4 A description of top productivity dynamics

4.1 Summary statistics

In this subsection, we explore summary statistics for the main variables of interest. Table 3 presents the mean, median and standard deviation of value added, capital, payroll and employment for the overall sample and for firms in the top 1% of the TFP distribution. Here and in the following, TFP percentile groups are always measured within 4-digit NAICS industry-year groups.

We find in the overall sample that the mean values across value added, capital, payroll and employment are large, and much larger than the median values. This reflects positive skewness, where there is a long tail to the right. The most productive firms are significantly larger, using about 20 times as much capital, and having an almost ten times larger payroll, to produce 17 times as much output as the mean firm. Given skewness in the top itself, the median top TFP firm is much closer to the sample median. This also implies that not all firms in the top TFP group are large. We also find that the standard deviation of the top 1% TFP has a higher value for each variable compared to the overall sample.

Given their size, the most productive firms account for a substantial part of aggregates in our sample: As 21% of aggregate value added, 23% of total capital, and 12% of total payroll. These values are substantial, and imply that shocks to these firms could have consequences for aggregate dynamics.

4.2 Distributions of TFP and TFP Growth

In this subsection, we explore the distribution of TFP and TFP growth with a focus on its tails, and in particular the right tail. The first result we highlight is that both TFP and TFP growth are fat tailed. Figure 1 shows the log density of TFP and TFP growth from the data. It also shows the counterparts that would be implied by a normal distribution with the same mean and standard deviation.[17] Presenting the density in logs helps to better illustrate the thickness of the tails. We define the tails as lying beyond the point at which the empirical log density of TFP (TFP growth)

[15] There is no intercept because this is captured by the industry mean ω_0.

[16] The synthetic dataset is created by randomly drawing from the normal distribution, $N(0, 0.24^2)$, in the initial period for 350,795 $(70,159 \times 5)$ firms and then applying (6) for 78 periods, where the first 65 (13×5) periods are dropped. This ensures that the distribution is ergodic.

[17] In order to maintain confidentiality, the left axis has been removed and statistics concerning the plots have been rounded to the nearest two digits.

Table 3: Summary Statistics

	Value added (In millions)	Capital (In millions)	Payroll (In millions)	Employment (Persons)
Sample				
Mean	4.86	4.23	2.87	66
Median	0.824	0.325	0.620	19
Standard Deviation	90.9	109	34.7	788
Top 1% by TFP				
Mean	83.4	80.9	27.4	715
Median	3.10	0.292	0.964	27
Standard Deviation	514	608	152	4910

This table presents the mean, median and standard deviation of value added, capital, payroll and employment for the sample, firms in the top 5% of the payroll distribution, and firms in the top 1% of the TFP distribution.

crosses the normal density, illustrated by the vertical dashed lines.[18] We find that there are thick tails in both the top and bottom group of firms. For TFP, the fat left tail is made up of 1% of firms, while the fat right tail is made up of 2% of firms. The middle section makes up the remaining 97%. For TFP growth, the values are 1%, 1% and 98%, respectively.

It is also informative to inspect the percentiles across the TFP and TFP growth distributions. We present them in Table 4, with the AR(1) counterparts. We can see that TFP is highly dispersed. Taking for example the $P99/P50$ ratio, we find that top 1% firms are at least 275% times more productive than the median firm. The distribution of log TFP is largely symmetric with the 90% and 10% occurring at roughly equal absolute values. Compared to the AR(1), the TFP distribution is more dispersed at the top but is less dispersed within the 10%-95%. This shows that the AR(1) is crucially missing out on top productivity firms.

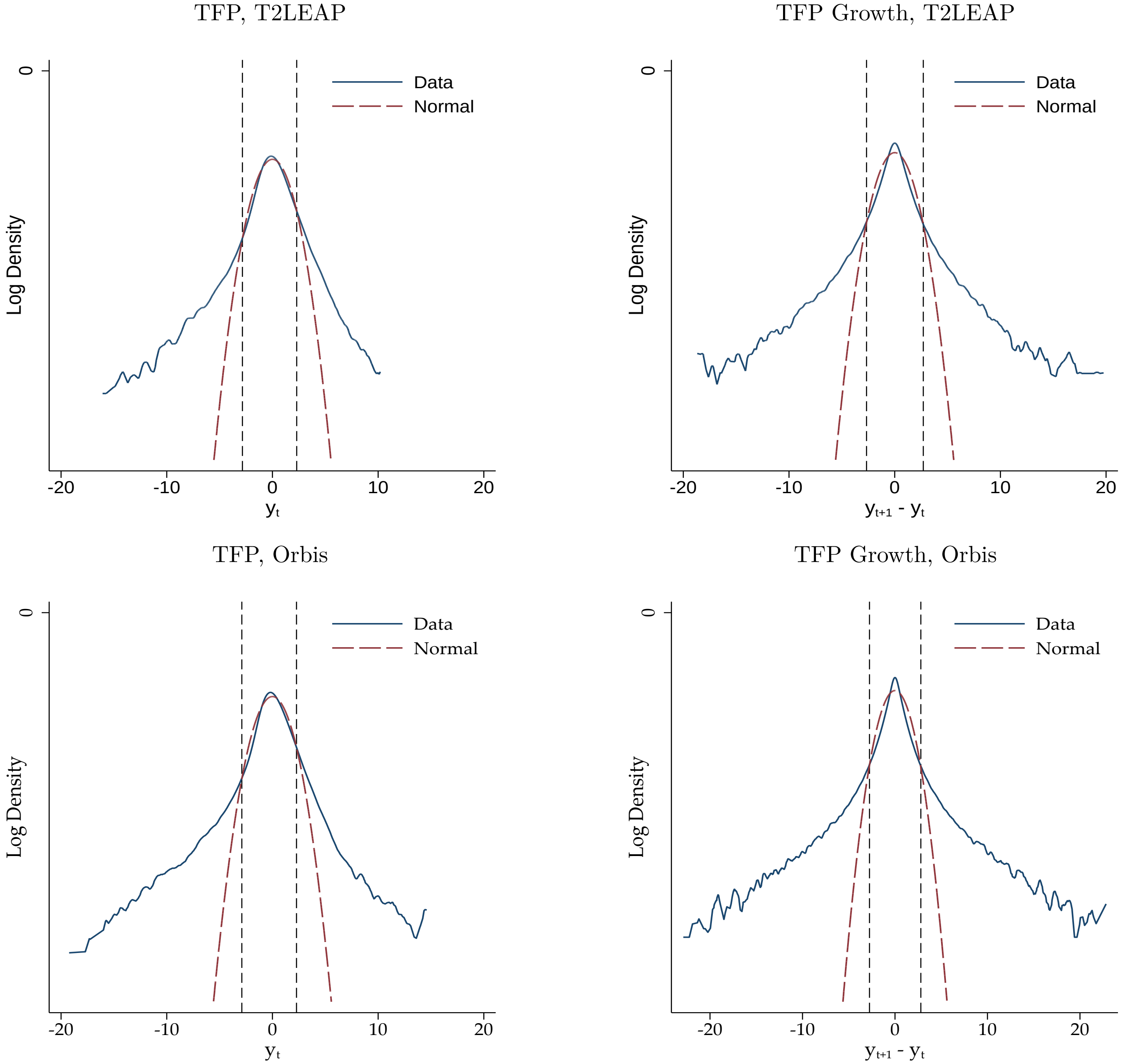
The table also clearly reveals the fat tails of the growth rate distribution. This implies that movements across the distribution can be large. For example, if a firm were to receive a shock equal to the 99^{th} percentile of the growth rate distribution, 0.84, it would move from around the 10^{th} percentile to around the 90^{th} percentile in the TFP distribution. The AR(1) has a much lower value in its growth rate distribution for the 99^{th} percentile.

4.3 Mean Reversion and the Distribution of TFP Growth Rates across the TFP Distribution

In this subsection, we explore the mean reversion properties of TFP. We begin by providing in Table 5 the summary statistics of TFP growth conditional on a firm's TFP level. We report the mean,

[18] The plots are created by first standardizing TFP and TFP growth and then binning the observations into 1000 equally spaced bins from which the log density is calculated. The normal distribution is derived by using the mean and standard deviation of each variable in its probability density function.

Figure 1: Log Density Plot



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. In T2LEAP, 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 Collard-Wexler and De Loecker (2025).

standard deviation, skewness, Kelly skewness and kurtosis of a firm's TFP growth rate between

Table 4: Percentiles

	10%	25%	50%	75%	90%	95%	99%
T2LEAP							
TFP	-0.390	-0.210	-0.020	0.200	0.430	0.590	0.990
TFP Growth	-0.22	-0.09	0.00	0.10	0.22	0.34	0.84
Orbis							
TFP	-0.380	-0.210	-0.030	0.200	0.440	0.610	1.070
AR(1)							
TFP	-0.47	-0.25	-0.00	0.25	0.47	0.61	0.89
TFP Growth	-0.33	-0.17	0.00	0.17	0.33	0.42	0.62

This table presents the percentiles of TFP, TFP growth, the AR(1) and AR(1) growth. Growth is defined as the ln difference between time $t+1$ and t. The AR(1) is simulated with parameters $\rho = 0.76$ and $\sigma = 0.24$. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

period t and $t+1$, conditional on its TFP in period t.[19]

It is clear from Table 5 that TFP is mean reverting. For firms in the lowest productivity decile, the mean growth rate is 0.16, which implies a significant move towards the mean. Conversely, the mean realized growth rate among the top 1% most productive firms is -23%, also implying a significant move *down* towards the mean. The mean growth rate in each TFP group declines smootly as we move up the TFP distribution. Average growth rates of firms closer to the middle of the distribution are smaller in absolute value.

The standard deviation of growth rates is just below that of the AR(1) close to the middle of the distribution, grows as we move up the TFP distribution, and is much higher at the top.[20]

The skewness of growth rates similarly varies strongly across the TFP distribution. Both central and Kelly skewness show that growth rates are right skewed at the bottom of the TFP distribution, and very left skewed at the top. That is, for top firms, the TFP growth distribution features a much longer left tail – of negative growth rates – compared to its right tail.

Finally, the kurtosis of TFP growth is very high, reflecting the fat tails of the TFP growth distribution. In the top TFP group that is our focus, kurtosis is 12, much higher than the value of 3 associated with normal TFP innovations. These patterns are very similar in the unbalanced panel, as shown in **??**. They are also similar when limiting the analysis to the largest five percent of firms, as shown in Table 16. Growth rates are higher and slightly less left skewed in this group. We summarize these findings as

Fact 1. *The distribution of growth rates varies with the level of TFP. For the most productive firms, it has a lower mean, a greater standard deviation, and is left skewed.*

[19] The Kelly skewness is given by $((90^{th} - 50^{th}) - (50^{th} - 10^{th}))/(90^{th} - 10^{th})$ and is a skewness measure that is robust to outliers.

[20] It is also higher at the bottom.

The table also illustrates that an AR(1) process can only approximate empirical patterns of mean reversion, but not of higher moments, since it misses the higher dispersion of growth rates of top firms, and also misses the skewness and kurtosis so prominent in observed growth rates.

Table 5: TFP Growth by Level

T2LEAP

Percentile	Mean	SD	Kelly Skew	Skewness	Kurtosis
0-10%	0.160	0.410	0.280	2.030	38.320
10-50%	0.020	0.200	0.050	-1.240	47.260
50-90%	-0.030	0.210	-0.080	-2.450	70.790
90-95%	-0.080	0.250	-0.200	-1.520	21.470
95-99%	-0.120	0.300	-0.290	-1.540	16.710
99-100%	-0.230	0.480	-0.400	-1.830	12.080

Orbis

Percentile	Mean	SD	Kelly Skew	Skewness	Kurtosis
0-10%	0.170	0.450	0.360	1.720	27.510
10-50%	0.010	0.200	0.050	-3.180	63.850
50-90%	-0.030	0.200	-0.090	-2.380	43.750
90-95%	-0.060	0.230	-0.190	-2.030	24.050
95-99%	-0.090	0.270	-0.250	-2.490	29.850
99-100%	-0.160	0.410	-0.390	-2.740	23.950

AR(1)

Percentile	Mean	SD	Skewness	Kelly Skew	Kurtosis
0%-10%	0.16	0.24	0.01	-0.00	3.00
10%-50%	0.05	0.24	0.00	0.00	3.00
50%-90%	-0.05	0.24	0.00	-0.00	3.00
90%-95%	-0.13	0.24	-0.01	-0.00	2.98
95%-99%	-0.17	0.24	-0.01	-0.00	3.02
99%-100%	-0.23	0.24	-0.02	-0.00	2.99

This table presents summary statistics of growth by different percentile groups in levels. The table presents the mean, standard deviation (SD), skewness (Skew.), Kelly skewness (Kelly skew.) and kurtosis (Kurt.). The Kelly skewness is given by $((90^{th} - 50^{th}) - (50^{th} - 10^{th}))/(90^{th} - 10^{th})$. Growth is defined as the ln difference between time $t+1$ and t. The AR(1) is simulated with parameters $\rho = 0.76$ and $\sigma = 0.24$. Results in this panel do not always exactly equal their theoretical counterparts due to simulation error in the finite sample of 100,000 firms. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

4.4 How long do firms stay at the top?

Thanks to the size of our dataset, we can further deepen the analysis of productivity dynamics by exploring the amount of time firms spend at the top of the distribution.

We begin by providing the distribution of time firms spend in the top 1% in Table 6. This table presents the distribution of the number of years in the top 1% for T2LEAP and Orbis.

We find that in Canada, 6.37% of firms are present in the top group at some point in time. Of

Years	T2LEAP	Orbis	AR(1)
0	93.630	94.470	92.06
1	3.350	2.400	4.94
2	1.150	0.900	1.79
3	0.570	0.530	0.71
4	0.360	0.350	0.29
5	0.290	0.240	0.13
6	0.170	0.210	0.05
7	0.120	0.160	0.02
8	0.100	0.120	0.01
9	0.070	0.100	0.00
10	0.050	0.070	0.00
11	0.050	0.070	0.00
12	0.040	0.080	0.00
13	0.060	0.050	0.00
14		0.040	0.00
15		0.040	0.00
16		0.040	0.00
17		0.050	0.00
18		0.070	0.00

Table 6: Distribution of Time in the Top 1%

these firms, half spent only 1 year in the top, with the remaining half spending 2 to 13 years at the top. A small fraction of firms spend a large number of years in the top. An interesting pattern is that the tail of the distribution of years spent at the top is fairly flat from 9 years on, and does not thin out.

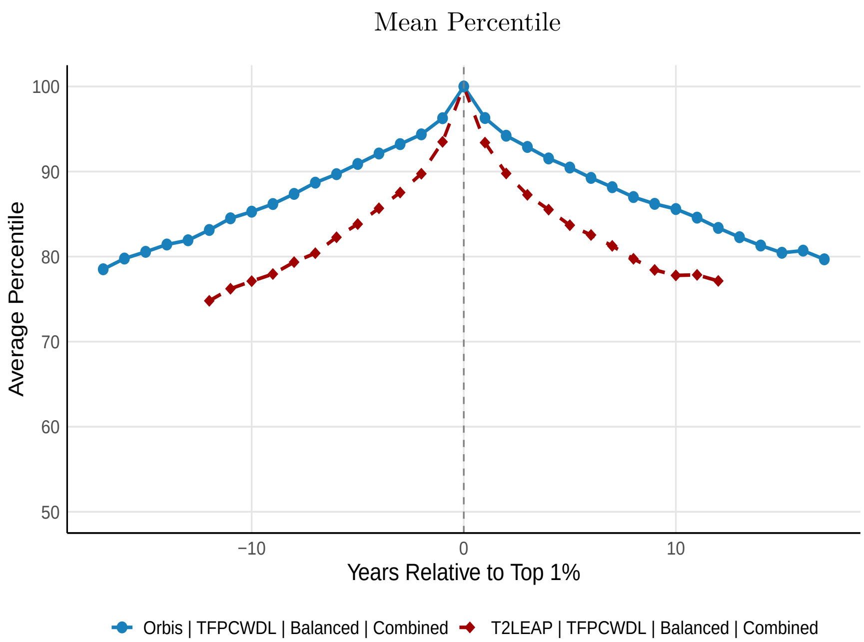
These patterns are similar in Orbis. Here, 5.53% of firms ever reach the top. Just under half only spend a single year in the top 1%. The tail of the duration distribution is extermely similar, with around 0.1% of firms spending 8 years in the top in both data sets, and 0.05% 13 years. The longer sample in Orbis allows us to see that the share of firms spending a long time in the top is strictly positive. Close to 0.1% of firms spend the entire sample period of 18 years in the top 1%.

For comparison, with an AR(1) process, more firms reach the top (8%), but they spend less time there. More than half of them spend only a single year at the top, and virtually no firm spends a long time at the top. For example, only 0.01% of firms spend 8 or more years at the top, compared to 0.10% of firms in the data. This shows that top productivity is a highly persistent process in the data.

4.5 TFP Trajectories of Top Firms

Again thanks to the size of the data set, we can describe the trajectories of top 1% TFP firms before and after their spell at the top. To do so, we track a firm's percentile in the TFP distribution in the years before and after a year spent in the top group. Figure 2 presents the mean TFP percentile across firms for each year before (-) and after (+) a year in the top 1%. Since we use a balanced sample that is 13 years in length, we can trace firms up to 12 years before or after being in the top group in T2LEAP. In Orbis, this extends to 17 years.[21]

Figure 2: Tracking the Percentile of Top 1% Firms



This figure provides the mean percentile across firms for each year before (-) and after (+) being in the top 1%. The AR(1) is simulated with parameters $\rho = 0.76$ and $\sigma = 0.24$. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

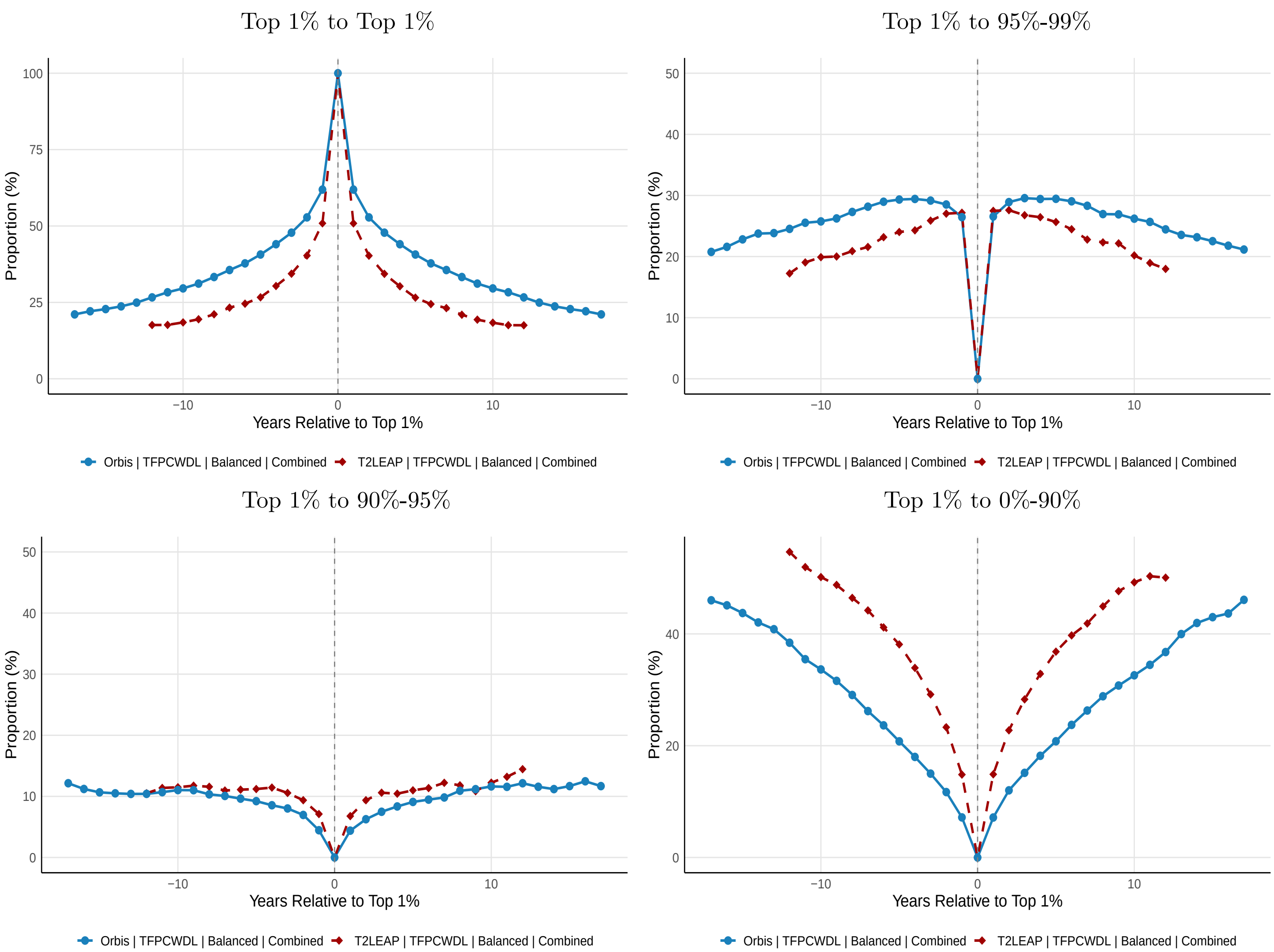
We find that TFP of top firms is highly persistent. In Canada, firms who spent a year at the top on average are in the 77^{th} percentile of the distribution even a decade before or after spending a year in the top. This number is even higher in Orbis, where even fifteen years before or after a stint in the top, firms on average are still around the 80th percentile of the productivity distribution. This is clearly higher than the patterns implied by an AR(1) process (not graphed). In that case, there are no persistent differences across firms and so after 12 years all firms, including those who entered the top, on average are at the 50^{th} percentile of the distribution as one moves backward/forward in time. We summarize these patterns as

Fact 2. *Top productivity status is highly persistent.*

To delve deeper into where in the distribution the top 1% firms come from and go to, we present in Figure 3 the mean proportion of firms in different percentile groups for each year before (-) and after (+) being in the top 1%. The percentile groups are: 0%-90%, 90%-95%, 95%-99% and top 1%. The information displayed is the same as in Figure 2. We again see that TFP of firms that reach the top is very persistent. For instance, in Canada, 12 years before or after a year in the top

[21]Years before and after a year at the top can also be at the top, or elsewere in the distribution.

Figure 3: Tracking the Trajectories of Top 1% Firms



These plots provide the mean proportion of firms in each percentile group for each year before (-) and after (+) being in the top 1%. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

1%, 18% of firms are already/still in the top 1%, 40% in the top 5%, and half in the top 10%. In Europe, top status is even more persistent, with almost a quarter of firms still in the top 1% even fifteen years before/after a spell at the top. Fewer than half of top firms leave the top 10% over 17 years. It is very clear that firms that make it to the top tend to remain close to the top. With an AR(1) process, in contrast, essentially no firms remain in the top 1% after a decade, and only around 10% remain in the top 10% (not graphed).

4.6 Persistence of Top Productivity

In this final subsection, we further investigate the persistence of TFP highlighted in the previous results. Table 7 shows the one year transition matrix of TFP, as observed in T2LEAP and Orbis (top two panels) and as implied by an AR(1) (bottom panel). The transition matrix represents the frequency of transitions observed in the data between any two consecutive periods, and highlights how firms move around in the distribution conditional on their previous level. The transition matrix illustrates that TFP is persistent, as diagonal elements are quite large. This is particularly strong at the top 1%, at almost 51%, despite the relatively small size of the origin group. TFP in Orbis is even more persistent, with almost 62% of top 1% firms remaining in that percentile year to year.

Most firms that do not stay in the top move to the next lower group, percentiles 95 to 99. The probability of moving far down the distribution is small but not zero – as also implied by the negative skewness of growth rates of the most productive firms shown above. An autoregressive process with a persistence parameter of 0.76, as fits the sample best, only features a probability of 33% of remaining in the top, with transition probabilities that drop off to zero quickly for large moves.

Table 7: One Year Transition Matrix (%)

T2LEAP

From	0-10%	10-50%	50-90%	90-95%	95-99%	Top 1%
0-10%	51.710	40.680	6.740	0.420	0.330	0.130
10-50%	9.970	68.480	20.650	0.540	0.280	0.080
50-90%	1.660	20.670	69.800	5.500	2.030	0.340
90-95%	0.900	4.450	43.750	32.110	17.070	1.720
95-99%	0.850	3.050	20.260	21.460	46.090	8.280
Top 1%	1.140	2.870	10.880	6.770	27.480	50.870

Orbis

From	0-10%	10-50%	50-90%	90-95%	95-99%	Top 1%
0-10%	55.410	37.760	6.130	0.370	0.250	0.080
10-50%	9.360	72.130	17.970	0.340	0.170	0.030
50-90%	1.520	17.970	73.940	5.010	1.420	0.140
90-95%	0.700	2.910	39.680	39.060	16.670	0.980
95-99%	0.690	1.790	14.370	20.720	55.160	7.280
Top 1%	0.650	1.330	5.160	4.390	26.550	61.920

AR(1)

From	0-10%	10-50%	50-90%	90-95%	95-99%	99-100%
0%-10%	52.17	44.91	2.91	0.01	0.00	0.00
10%-50%	11.22	61.36	26.69	0.58	0.14	0.00
50%-90%	0.72	26.70	61.34	7.05	3.85	0.35
90%-95%	0.01	4.67	56.37	18.79	17.18	2.98
95%-99%	0.00	1.45	38.43	21.57	29.09	9.46
99%-100%	0.00	0.14	13.70	15.13	37.87	33.16

This table presents in percentages the one year transition matrix between percentile groups of 0%-10%, 0%-50%, 50%-90%, 90%-95%, 95%-99% and 99%-100%. The AR(1) is simulated with parameters $\rho = 0.76$ and $\sigma = 0.24$. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

Table 8 presents the five year transition matrix. This matrix shows where firms end up five periods in the future. We again see significant persistence across the distribution, and in particular at the top. The top 1% of TFP have a probability of almost 27% of remaining in the top 1% five years later in Canada, and over 40% in Orbis. This probability is much lower in the AR(1) case, in line with the histories shown in the previous subsection.

Note also that the probability of being in the top five years later is much greater than what would be implied by the one-year transition matrix. These patterns are in line with Fact 2.

Appendix Tables 14 and 15 show that these patterns are very similar for the unbalanced panel. The main difference is that in the unbalanced panel, the probability of dropping to very low levels of productivity is somewhat larger, at the expense of the productivity of staying at the top or very close to it. Nevertheless, these transition matrices still exhibit very strong persistence at the top – much stronger than implied by the AR(1), for example – with a probability of being in the top 1% of 49% after one year and 24% after five years. When analyzing the five percent largest firms only, productivity is even more persistent, as shown in Tables 17 and 18.

Finally, Table **??** shows how the distribution of growth rate changes as a function of the time a firm has spent in the top TFP group. (Due to shrinking sample size, we stop at 4+ years.) Subsequent growth clearly increases in time at the top. (Although mean reversion remains, even for firms that have spent a lot of time at the top.) The table shows that this also implies a greater probability of staying at the top. We summarize this as These patterns are very pronounced in both T2LEAP and Orbis.[22]

Fact 3. *Mean growth and the probability of remaining in the top TFP group increase with time spent in the top TFP group.*

5 Heterogeneous TFP dynamics

The previous section has documented three new facts: (1) TFP growth is more dispersed and very left skewed for top firms, (2) being at the the top of the TFP distribution is a very persistent state, and (3) mean growth and the probability of staying at the top increase with time at the top. All of these facts are novel to the literature, and none can easily be explained by stochastic processes commonly used to describe TFP dynamics, in particular an AR(1).

In this section, we estimate a simple parametric process that can account for all these patterns. Besides fitting the new facts, our criterion in choosing a specific type of process is simplicity or parsimony, as well as ease of use in applied work, in particular in quantitative macroeconomics.

A very parsimonious and easy to deal with process that, as we show below, fits the data well, is a hidden type autoregressive process. We thus specify that TFP follows the process

$$z^*_{it} = \alpha^k + \rho^k z^*_{it-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim \sigma^k t(df), \tag{7}$$

[22] By construction, this pattern does not appear in an AR(1) process.

Table 8: Five Year Transition Matrix (%)

T2LEAP

From	0-10%	10-50%	50-90%	90-95%	95-99%	Top 1%
0-10%	25.620	50.070	21.090	1.730	1.160	0.330
10-50%	12.130	53.500	30.850	1.980	1.270	0.260
50-90%	5.350	31.090	52.850	6.140	3.750	0.810
90-95%	3.390	14.940	51.070	15.270	12.630	2.710
95-99%	2.860	11.090	38.680	17.080	22.980	7.310
Top 1%	2.790	9.120	24.910	10.990	25.650	26.540

Orbis

From	0-10%	10-50%	50-90%	90-95%	95-99%	Top 1%
0-10%	36.800	47.960	13.570	0.890	0.620	0.160
10-50%	11.470	60.110	26.560	1.140	0.600	0.110
50-90%	3.650	26.140	60.640	6.140	3.010	0.420
90-95%	2.310	9.200	48.320	22.410	15.740	2.020
95-99%	2.110	6.480	29.350	19.660	34.320	8.070
Top 1%	1.880	4.270	14.640	9.090	29.450	40.670

AR(1)

From	0-10%	10-50%	50-90%	90-95%	95-99%	99-100%
0%-10%	19.36	48.40	28.48	2.20	1.35	0.21
10%-50%	12.02	43.79	37.06	3.86	2.72	0.55
50%-90%	7.17	37.00	43.80	5.93	4.86	1.23
90%-95%	4.40	30.82	47.52	8.06	7.14	2.06
95%-99%	3.28	27.53	48.66	8.97	8.70	2.86
99%-100%	2.03	22.14	49.11	10.72	11.35	4.64

This table presents in percentages the five year transition matrix between percentile groups of 0%-10%, 0%-50%, 50%-90%, 90%-95%, 95%-99% and 99%-100%. The AR(1) is simulated with parameters $\rho = 0.76$ and $\sigma = 0.24$. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

where k indexes the hidden type. With this process, each firm's TFP follows an autoregressive process, but the population is heterogeneous in terms of type, and the parameters of the autoregressive process vary by type.

In addition, we allow for the presence of classical measurement error. That is, we assume that we do not directly observe "true" TFP z^*_{it}, but instead observe

$$z_{it} = z^*_{it} + u_{it}, \quad u_{it} \sim N(0, \sigma_u^2). \tag{8}$$

Concretely, the intercept, persistence, and variance of the process differ across types.[23] The resulting model is a mixture of autoregressive processes. Different parts of the productivity distribution contain different mixtures of firm types. This permits the model to fit the moments

[23] One intercept can be normalized to zero. We also experimented with letting measurement error and the degrees of freedom of the t distribution vary by type. This did not meaningfully improve the fit of the model. We thus restrict them to take on a common value.

Years in top 1%	T2LEAP: Mean growth	T2LEAP: Pr(stay)	Orbis: Mean growth	Orbis: Pr(stay)
any	-0.230	50.870	-0.160	61.920
1	-0.310	37.790	-0.240	44.380
2	-0.190	55.780	-0.140	61.930
3	-0.150	64.240	-0.120	70.690
4+	-0.100	77.710	-0.070	82.520

Table 9: Growth and Persistence by Number of Consecurive Years in the Top 1%

described above. For instance, consider the higher persistence of TFP at the top. The model can match this if a type with higher persistence ρ^k is more prevalent at the top of the TFP distribution.

5.1 Estimation

Given the confidential nature of the T2LEAP data, we estimate the model by simulated method of moments (SMM). Letting K be the number of types, the model has $3K+1$ parameters: $\alpha^k, \rho^k, \sigma^k, df$ and σ_u (normalizing $\alpha^1 = 0$). We estimate them using 68 data moments, belonging to 10 sets of 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)

We thus require the model not only to fit the degree of mean reversion (mean growth by productivity level) and the dispersion of growth rates in the data, but the variation in higher moments of the growth rate distribution with productivity (fact 1), variation in persistence (fact 2), as well as variation of growth by time at the top (fact 3).

We search for the parameter vector θ that minimizes the sum of squared differences between model and data moments. Given the high degree of precision in measuring moments in the administrative data we use, we weight moments equally. Because moments differ in scale, we rescale them to a similar order of magnitude.[24]

Next, we report estimation results for a model with three hidden types. This model provides a noticeably better fit (reported below) than a model with two types, which does not capture persistence at the top as well. Adding a fourth types, in contrast, does not lead to a marked improvement in fit.

The estimation implies that latent TFP z_t^* follows the following process:

$$z_{it}^* = \begin{cases} 0.7888 z_{it-1}^* + \varepsilon_{it} & \text{if } k = 1: \text{ “regular” type} \\ 0.0048 + 0.3029 z_{it-1}^* + \varepsilon_{it} & \text{if } k = 2: \text{ “high variance” type} \\ 0.0152 + 0.9277 z_{it-1}^* + \varepsilon_{it} & \text{if } k = 3: \text{ “high growth \& persistence” type} \end{cases}$$

where $\varepsilon_{i,t} \sim \sigma_i \, t(5.1271)$, $\sigma_1 = 0.1142$, $\sigma_2 = 0.3175$, $\sigma_3 = 0.1157$ and $\sigma_u = 0.0247$. 59.8% of firms belong to type 1, 14.3% to type 2, and the remaining 25.9% to type 3.[25]

These estimates imply clear differences across types. The first type is characterized by parameters quite close to those of the AR(1) process that best fits the population overall, with intermediate persistence and shock size. The second type has much lower persistence, slightly higher mean growth, and much higher variance. The third type has much higher mean growth, extremely persistent TFP, and a variance close to the “regular” type. Estimates are qualitatively and even quantitatively similar in both the Canadian T2LEAP and the Orbis data.

5.2 Model fit

Figure 4 shows the model fit in terms of the distribution of TFP growth rates across the TFP distribution.

Like an AR(1) process, the hidden type model fits the level of mean reversion in the data closely. In addition, the heteregeneous type model also fits the fact that the most productive firms experience more dispersed, left-skewed, and fat-tailed TFP growth rates. While the model does not exactly fit non-targeted skewness and kurtosis moments for TFP groups below the top, it generates qualitative patterns clearly in line with Fact 1.

Figure 5 shows model-predicted histories of TFP percentiles in the 12 years before and after a year at the top. The model slightly overstates the degree of mean-reversion right before the top spell, but nonetheless generates a remarkably close fit. For example, the mean productivity percentile 12

[24] Some technical details: we use a burn-in period of 186 periods, and then measure model moments on the following 13 simulated periods. Following GKOS, Guvenen et al. (2015), we measure the distance between data and model moments x and y using the arc-percent distance $\frac{x-y}{(|x|+|y|)/2+\psi}$, where ψ is the tenth percentile of the distribution of moments, or the second smallest in a set. This provides a systematic way of comparing deviations for moments of very different scale. We have also bootstrapped standard errors. These turn out to be small so, for brevity, we do not report or discuss them.

[25] In the Orbis data, the corresponding estimates are $\rho^1 = 0.7955$, $\rho^2 = 0.3156$, $\rho^3 = 0.9438$, $\alpha^2 = 0.0056$, $\alpha^3 = 0.0151$, $\sigma^1 = 0.1200$, $\sigma^2 = 0.3721$, $\sigma^3 = 0.1168$, $df = 4.8318$, $\sigma_u = 0.0204$. 65.0% of firms belong to type 1, 7.2% to type 2, and the remaining 27.8% to type 3.

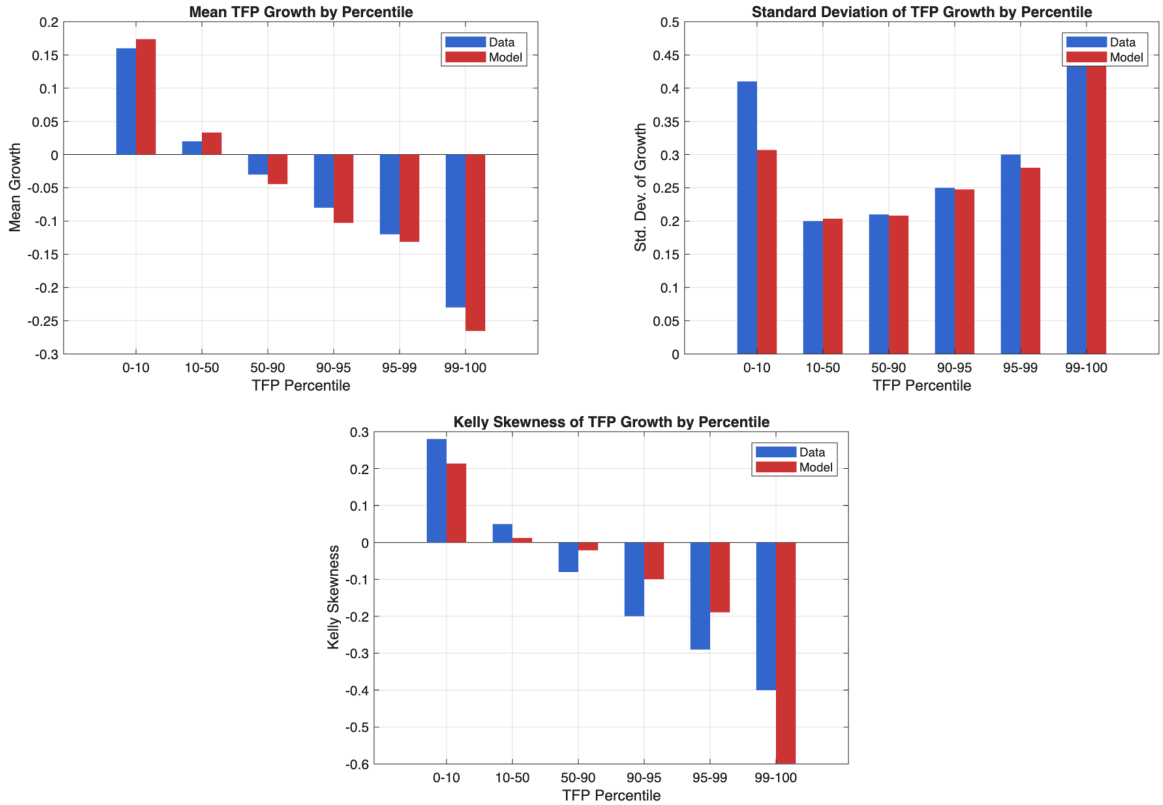


Figure 4: Model fit: the distribution of growth rates by TFP level, T2LEAP

years before (after) a year at a top in model data is 75 (76), exactly as in the data (77 in the data). At a horizon of 12 years, the equivalent figure for an AR(1) process with coefficient 0.76 is close to the median. Appendix Figure 9 shows transition probabilities between the stint in the top 1% and four percentile groups for 12 years before and after. The model also fits these closely.

Figure 6 graphically represents the last row of the TFP transition matrices shown in Tables 7 and 8 above. At both the one- and five-year horizons and with very few exceptions, the model (red) generates transition probabilities very close to the data (blue). The model thus fits Fact 2.

Finally, Figure **??** productivity dynamics at the top of the distribution as a function of time in the top for model and data. Again, model moments match target moments closely, for both mean productivity growth and the probability of remaining in the top 1%. The model thus also fits Fact 3.

Overall, model moments closely match target moments. The model closely reproduces all three new facts. The fit is similarly good for the Orbis data, as shown in Appendix A.1.

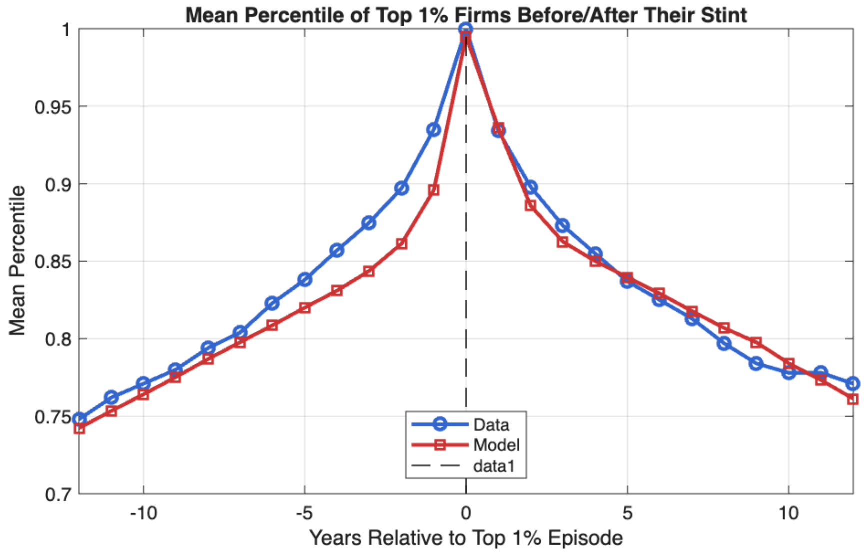


Figure 5: Model fit: top TFP trajectories, T2LEAP

5.3 Inspecting the mechanism

The fit to the new facts is achieved through the mixture of types in the hidden type model. In this section, we illustrate this for each fact.

Figure 8 shows the model-implied probability density of growth rates for firms in the top percentile of productivity. The left graph shows this for all firms. The graph clearly shows the left skewness and the fat tail. The right graph shows how the overall density is composed of densities for the three types. This reveals that the left skewness reflects both the lower mean growth of types 1 and particularly 2, as well as the higher variance of type 2.

The high persistence of top productivity results from the fact that most firms in the top percentile are of type 3. Since productivity of this type is more persistent, firms in the top percentile have more persistent productivity than firms in lower percentiles of the distribution.

Finally, the third fact arises from the fact that conditional on being in the top percentile, type 3 firms stay there longer than type 1 or 2 firms, due to their higher persitence. That is, dynamic selection implies that over time, the composition of a cohort of firms entering the top will shift towards firms of type 3. Since these have higher mean growth and higher persistence, firms that have been at the top longer have higher mean growth, and a greater probability of remaining at the top.

5.4 Robustness

In searching for a model to fit the new facts on top productivity dynamics, we have chosen a balance between fit and parsimony. How does the fit change as this is modified? We have explored a large battery of robustness checks.

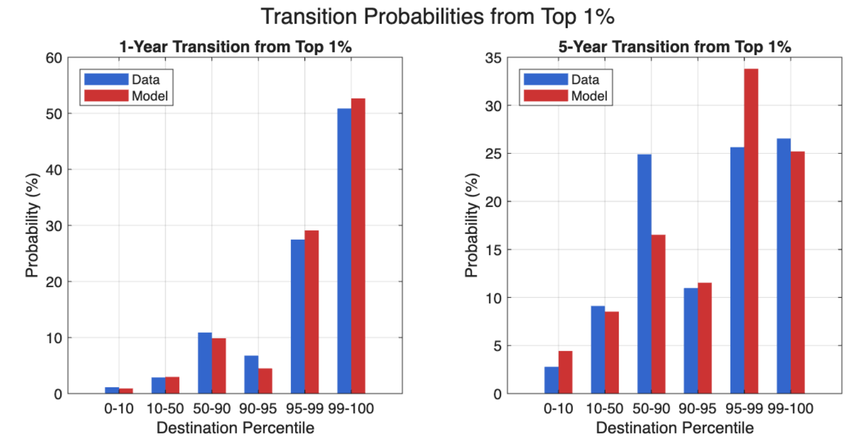


Figure 6: Model fit: TFP transition matrices, T2LEAP

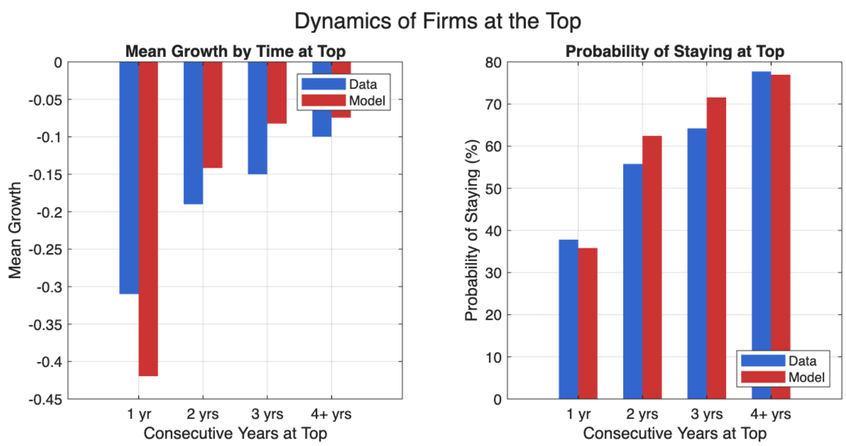


Figure 7: Model fit: TFP dynamics at the top, T2LEAP

If types do not differ in mean growth or in persistence, top firms return to the median more quickly, and expected growth falls with time at the top. If types do not differ in variance, the model cannot generate highly dispersed, left-skewed growth rates at the top observed in the data. As mentioned above, allowing for more types necessarily improves the fit, but not much. Nor does it generate additional insights.

Finally, measurement error latters little.

Adding firm-fixed effects does little to improve the fit in the dimensions of focus here, since they affect TFP for all firms in similar ways, and do not help to explain differences in dynamics across the distribution. An AR(1) process with fat tails, as in the hidden type process, cannot generate differences in growth rates by productivity, nor with time spent at the top.

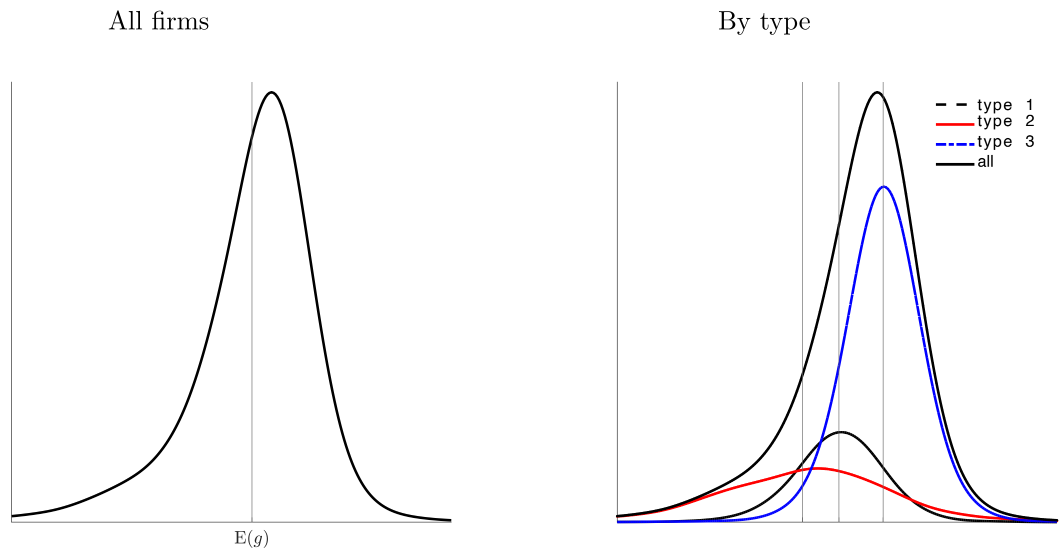


Figure 8: The distribution of growth rates at the top, by type

6 Macroeconomic implications

[in progress]

7 Discussion & Conclusion

This paper has investigated the dynamics of firm level productivity, with a particular focus on the most productive firms. We use new non-parametric techniques inspired by the recent literature on earnings dynamics to exploit a comprehensive dataset that covers the universe of incorporated employer firms in Canada. This allows us to characterize the TFP process in a way not done before in the productivity literature.

Our analysis has yielded new insights into several aspects of productivity dynamics, and highlighted that dynamics at the top of the distribution differ from those elsewhere. First, firms with higher TFP face a growth rate distribution with higher variance and more negative skewness. Second, the top is a highly persistent state. For example, half of firms that spend a year in the top productivity percentile are part of the top decile even twelve years before or later. Third, persistence at the top increases with time spent there, as firms that have been at the top longer on average experience greater subsequent growth.

These findings naturally suggest heterogeneous in productivity dynamics. Indeed, we show that a parsimonious hidden-type autoregressive process with three types can fit the data well, and can be a useful input for quantitative work on firm dynamics. In ongoing work, we explore the importance of type heterogeneity for the macroeconomic analysis of firm dynamics.

Our work also connects to recent work that has highlighted widening dispersion of the productivity distribution in recent years, driven by growth in the relative productivity of the most productive firms (Andrews et al. (2015); Gu (2019)). This divergence has been linked to a global decline in aggregate productivity growth rates (Andrews et al. (2016a); Andrews et al. (2016b)). Connecting this to the type heterogeneity we document, and exploring whether widening dispersion reflects changes in productivity processes by type or in the proportion of types is an exciting agenda for future research.

References

Ackerberg, D., Benkard, C. L., Berry, S., and Pakes, A. (2007). Econometric tools for analyzing market outcomes. *Handbook of Econometrics*, 6:4171–4276.

Ackerberg, D., Caves, K., and Frazer, G. (2006). Structural identification of production functions. MPRA Paper 38349, University Library of Munich, Germany.

Akcigit, U. and Ates, S. T. (2023). What happened to us business dynamism? *Journal of Political Economy*, 131(8):2059–2124.

Andrews, D., Criscuolo, C., and Gal, P. (2016a). The global productivity slowdown, technology divergence and public policy: A firm level perspective. OECD Productivity Working Papers 5.

Andrews, D., Criscuolo, C., and Gal, P. N. (2015). Frontier firms, technology diffusion and public policy: Micro evidence from OECD countries. OECD Productivity Working Papers 2.

Andrews, D., Criscuolo, C., and Gal, P. N. (2016b). The best versus the rest: The global productivity slowdown, divergence across firms and the role of public policy. OECD Productivity Working Papers 5.

Benhabib, J., Perla, J., and Tonetti, C. (2021). Reconciling models of diffusion and innovation: A theory of the productivity distribution and technology frontier. *Econometrica*, 89(5):2261–2301.

Collard-Wexler, A. and De Loecker, J. (2025). Productivity and capital measurement error. mimeo, duke university.

Díez, F. J., Fan, J., and Villegas-Sanchez, C. (2021). Global declining competition? *Journal of International Economics*, 132:103492.

Gu, W. (2019). Frontier firms, productivity dispersion and aggregate productivity growth in Canada. *International Productivity Monitor*, (37):96–119.

Guvenen, F., Kaplan, G., and Song, J. (2014). The glass ceiling and the paper floor: Gender differences among top earners, 1981-2012. Working Paper 20560, National Bureau of Economic Research.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2015). What do data on millions of U.S. workers reveal about life-cycle earnings risk? Working Paper 20913, National Bureau of Economic Research.

Hopenhayn, H. and Rogerson, R. (1993). Job turnover and policy evaluation: A general equilibrium analysis. *Journal of Political Economy*, 101(5):915–938.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Jaimovich, N., Terry, S., and Vincent, N. (2025). The empirical distribution of firm dynamics and its macro implications. mimeo, hec montréal.

Kalemli-Özcan, Ş., Sørensen, B. E., Villegas-Sanchez, C., Volosovych, V., and Yeşiltaş, S. (2024). How to construct nationally representative firm-level data from the orbis global database: New facts on smes and aggregate implications for industry concentration. *American Economic Journal: Macroeconomics*, 16(2):353–374.

König, M., Storesletten, K., Song, Z., and Zilibotti, F. (2022). From imitation to innovation: Where is all that chinese r&d going? *Econometrica*, 90(4):1615–1654.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341.

Luttmer, E. G. (2011). On the mechanics of firm growth. *The Review of Economic Studies*, 78(3):1042–1068.

Marschak, J. and Andrews, W. H. (1944). Random simultaneous equations and the theory of production. *Econometrica*, 12(3/4):143–205.

Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, 104(10):3186–3221.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

Statistics Canada (LEAP). Longitudinal employment analysis program. Government of Canada.

Statistics Canada (Table: 36-10-0097-01a). Flows and stocks of fixed non-residential capital, by sector of industry and type of asset, canada (x 1,000,000). Government of Canada.

Statistics Canada (Table: 36-10-0217-01b). Multifactor productivity, gross output, value-added, capital, labour and intermediate inputs at a detailed industry level. Government of Canada.

Sterk, V., Sedláček, P., and Pugsley, B. (2021). The nature of firm growth. *American Economic Review*, 111(2):547–579.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–65.

Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of Economic Surveys*, 26(1):98–128.

Wedervang, F. (1965). Development of a population of industrial firms. *Oslo: Scandinavian University Books*.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114.

Appendix

A Additional tables and figures

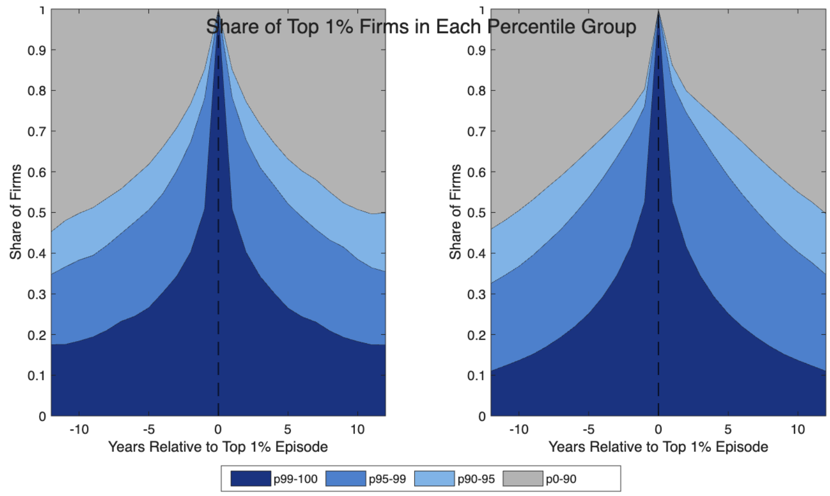


Figure 9: Model fit: top TFP trajectories

A.1 Model fit for Orbis data

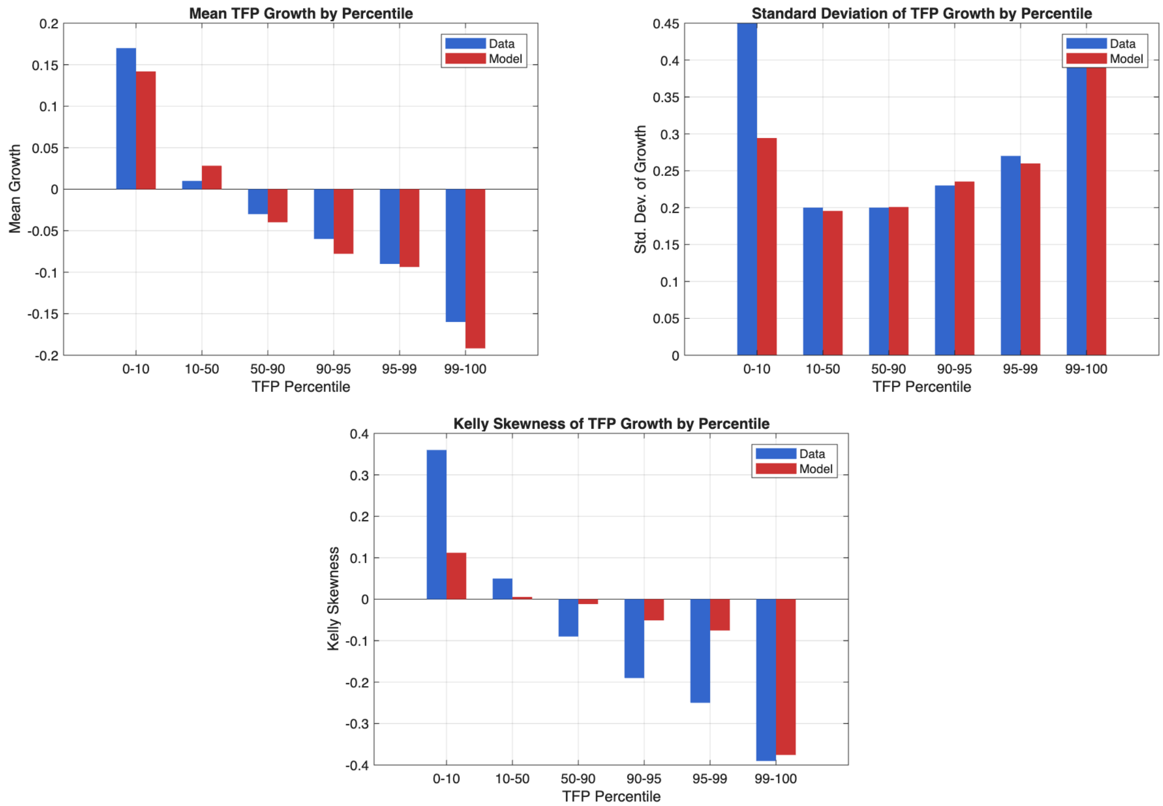


Figure 10: Model fit: the distribution of growth rates by TFP level, Orbis

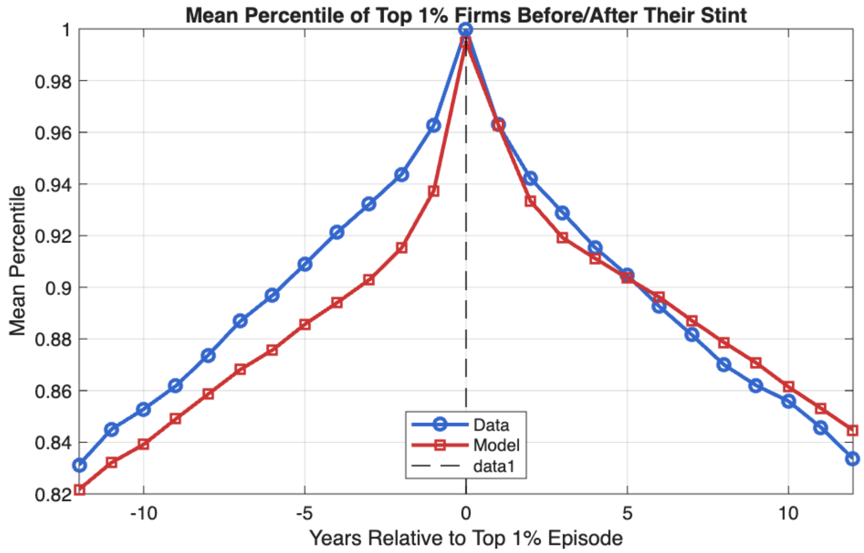


Figure 11: Model fit: top TFP trajectories, Orbis

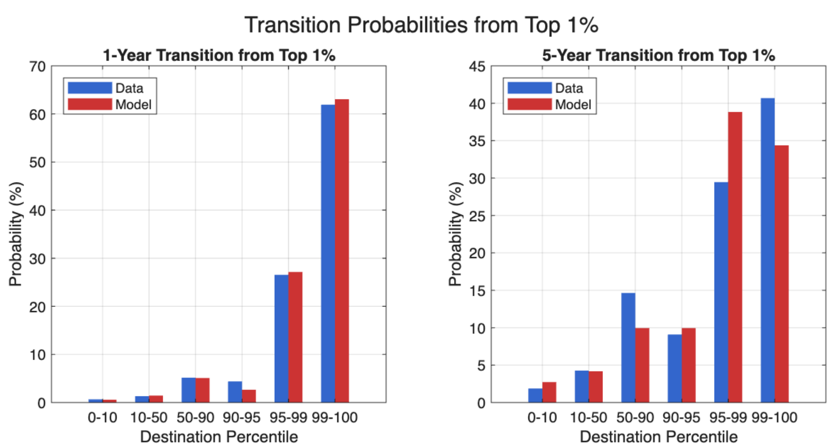


Figure 12: Model fit: TFP transition matrices, Orbis

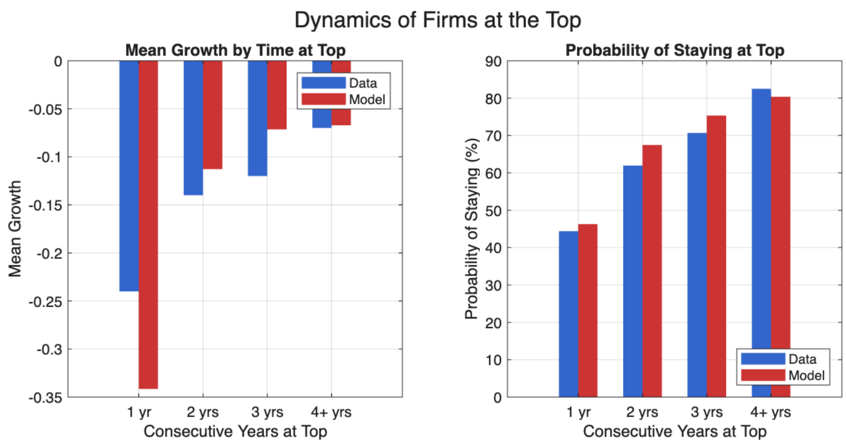


Figure 13: Model fit: TFP dynamics at the top, Orbis

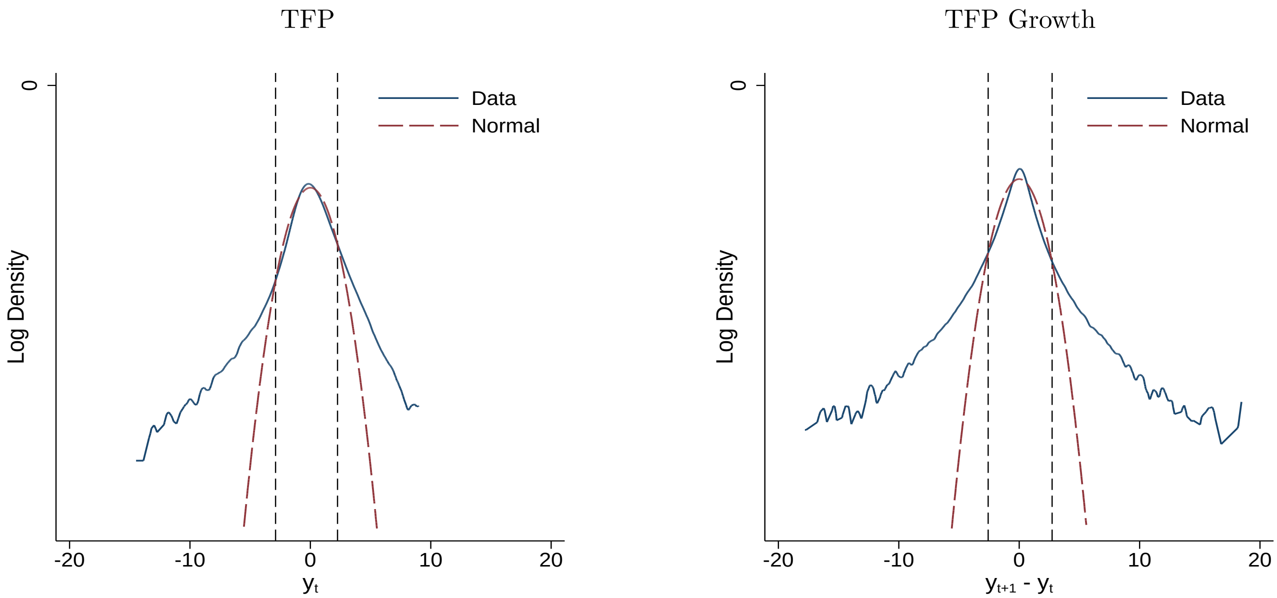
B Empirical patterns for different measures of productivity

B.1 Factor share TFP

Table 10: Growth by Level, Factor Share TFP

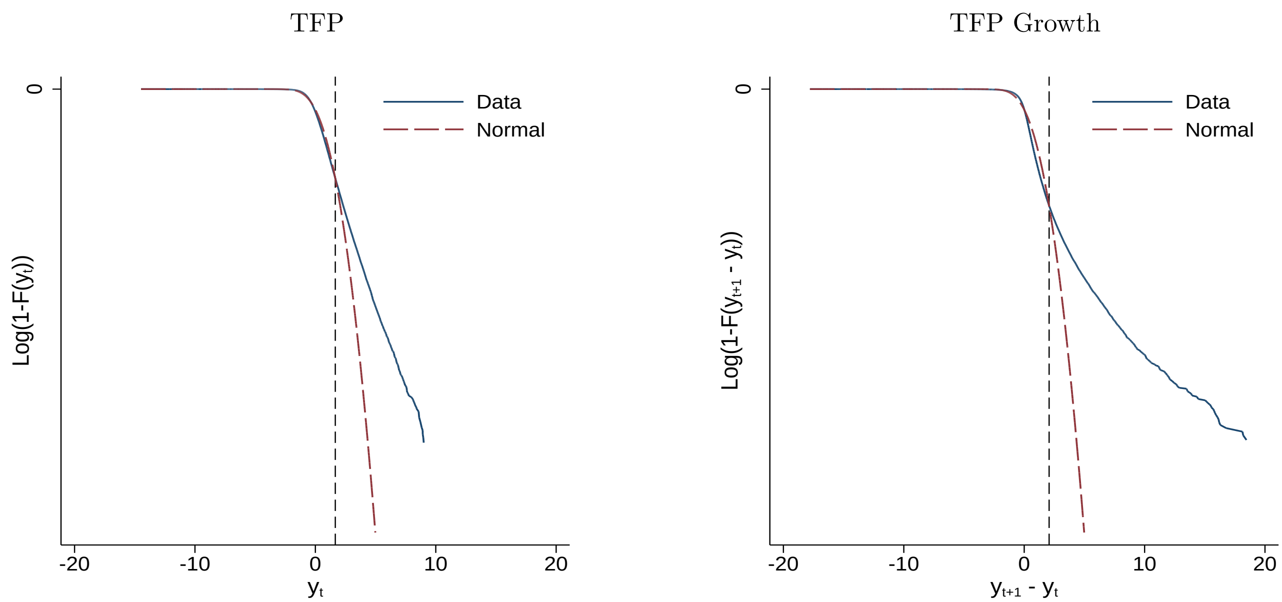
	Mean	SD	Skew.	Kelly Skew.	Kurt.
TFP					
0%-10%	0.16	0.41	1.94	36.31	0.27
10%-50%	0.02	0.21	-1.08	28.25	0.02
50%-90%	-0.03	0.23	-1.72	38.47	-0.11
90%-95%	-0.09	0.28	-2.80	60.38	-0.21
95%-99%	-0.13	0.34	-1.56	12.81	-0.30
99%-100%	-0.22	0.49	-1.94	12.04	-0.41

Figure 14: Log Density Plot, Factor Share TFP



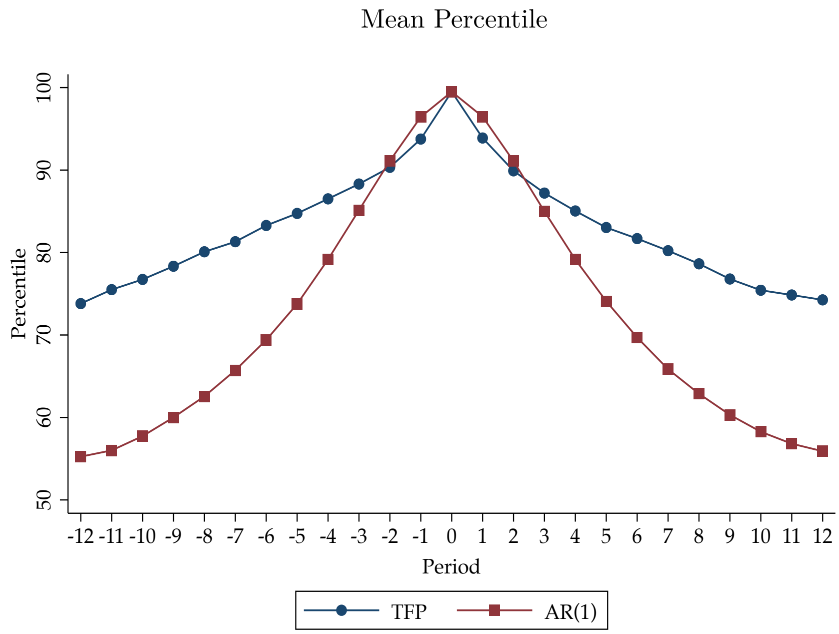
This figure shows the density in logs of TFP and TFP growth compared to their normal distribution counterparts. TFP and TFP growth have been standardized. The normal distribution is derived using the estimated mean and standard deviation of each variable in its probability density function. The proportion 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 Collard-Wexler and De Loecker (2025).

Figure 15: Log-Log Plot, Factor Share TFP



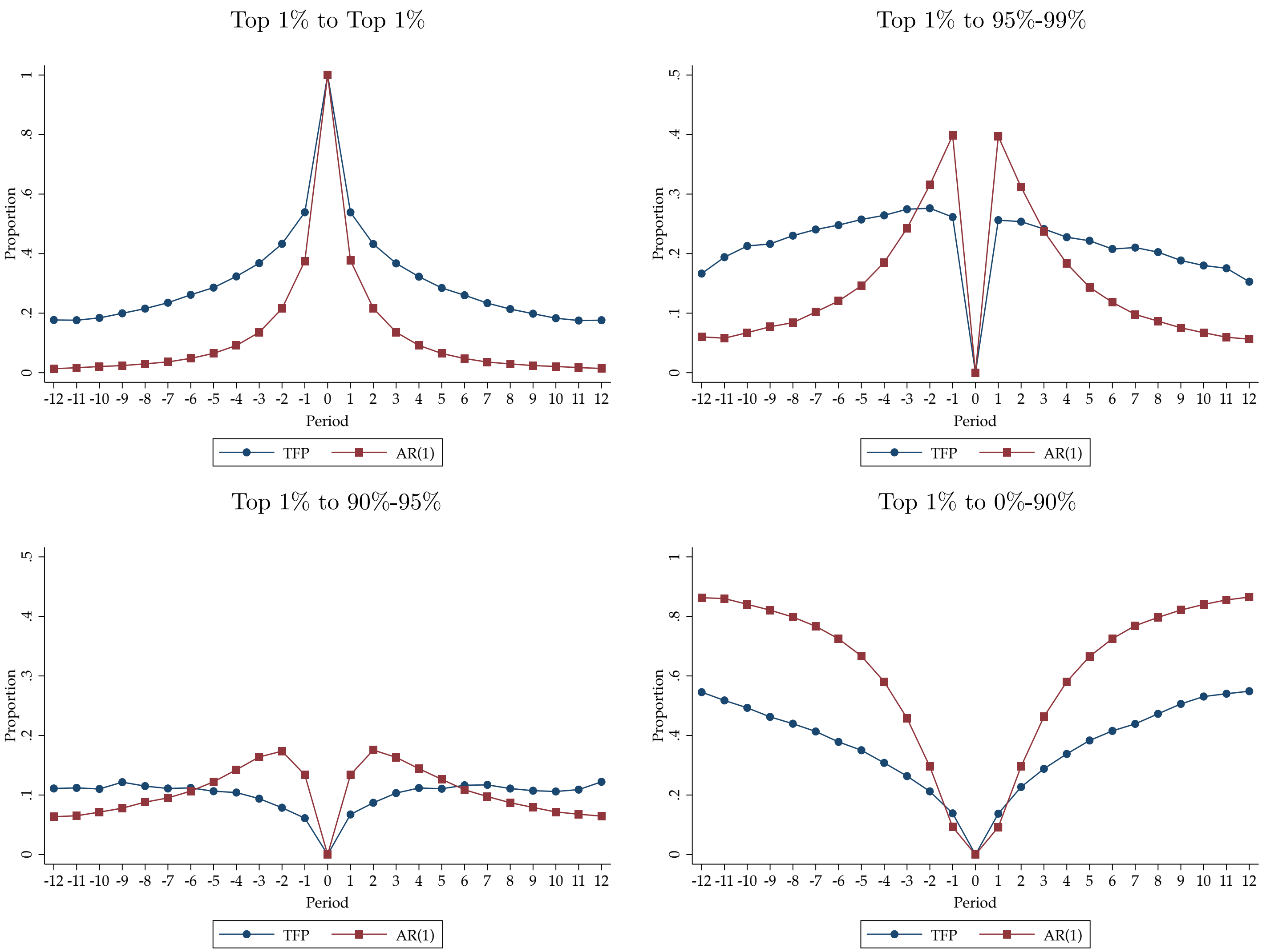
This figure provides the log-log plot of TFP and TFP growth compared to their normal distribution counterparts. TFP and TFP growth have been standardized. The normal distribution is derived using the estimated mean and standard deviation of each variable in its probability density function. The slope for the right tail past the vertical dashed line for TFP and TFP growth are equal to -1.04 and -0.44, 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 Collard-Wexler and De Loecker (2025).

Figure 16: Tracking the Percentile of Top 1% Firms, Factor Share TFP



This figure provides the mean percentile across firms for each year before (-) and after (+) being in the top 1%. The AR(1) is simulated with parameters $\rho = 0.76$ and $\sigma = 0.24$. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

Figure 17: Tracking the Proportion of Top 1% Firms, Factor Share TFP



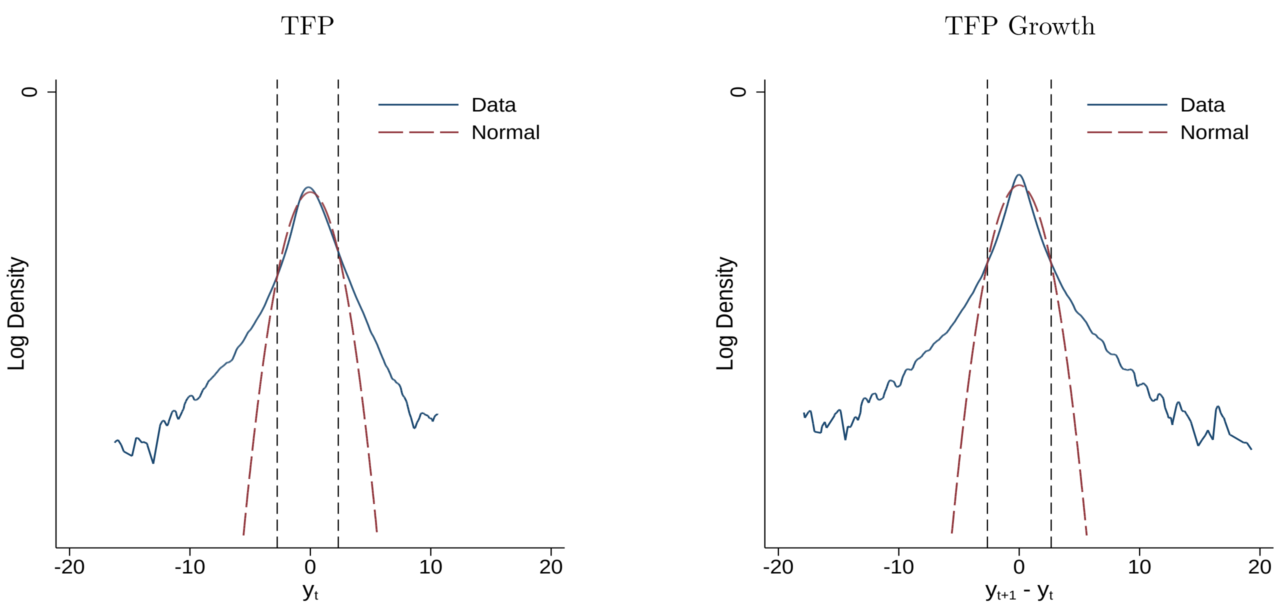
These plots provide the mean proportion of firms in each percentile group for each year before (-) and after (+) being in the top 1%. The AR(1) is simulated with parameters $\rho = 0.76$ and $\sigma = 0.24$. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

B.2 Labor productivity

Table 11: Growth by Level, Labor Productivity

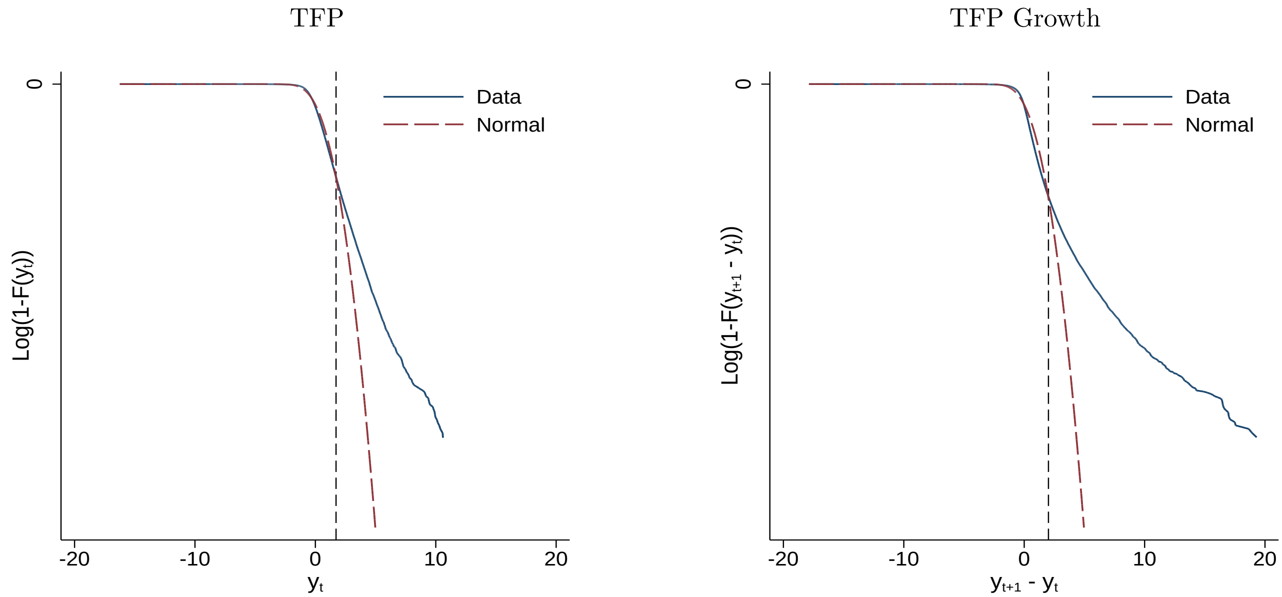
	Mean	SD	Skew.	Kelly Skew.	Kurt.
TFP					
0%-10%	0.18	0.41	1.89	34.76	0.28
10%-50%	0.02	0.20	-1.48	37.77	0.07
50%-90%	-0.03	0.21	-1.93	58.42	-0.05
90%-95%	-0.08	0.26	-1.51	22.68	-0.14
95%-99%	-0.13	0.31	-1.30	14.85	-0.24
99%-100%	-0.26	0.48	-1.62	9.86	-0.37

Figure 18: Log Density Plot, Labor Productivity



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 Collard-Wexler and De Loecker (2025).

Figure 19: Log-Log Plot, Labor Productivity



This figure provides the log-log plot 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 slope for the right tail past the vertical dashed line for TFP and TFP growth are equal to -1.04 and -0.44, 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 Collard-Wexler and De Loecker (2025).

C Empirical patterns for the unbalanced panel

Table 12: Growth by Level, Unbalanced Panel

	Mean	SD	Skew.	Kelly Skew.	Kurt.
TFP					
0%-10%	0.20	0.55	1.33	0.26	24.35
10%-50%	0.01	0.25	-2.76	0.03	65.74
50%-90%	-0.04	0.24	-2.55	-0.10	51.33
90%-95%	-0.10	0.30	-2.56	-0.21	37.23
95%-99%	-0.14	0.36	-2.19	-0.29	25.59
99%-100%	-0.29	0.57	-2.22	-0.41	14.62

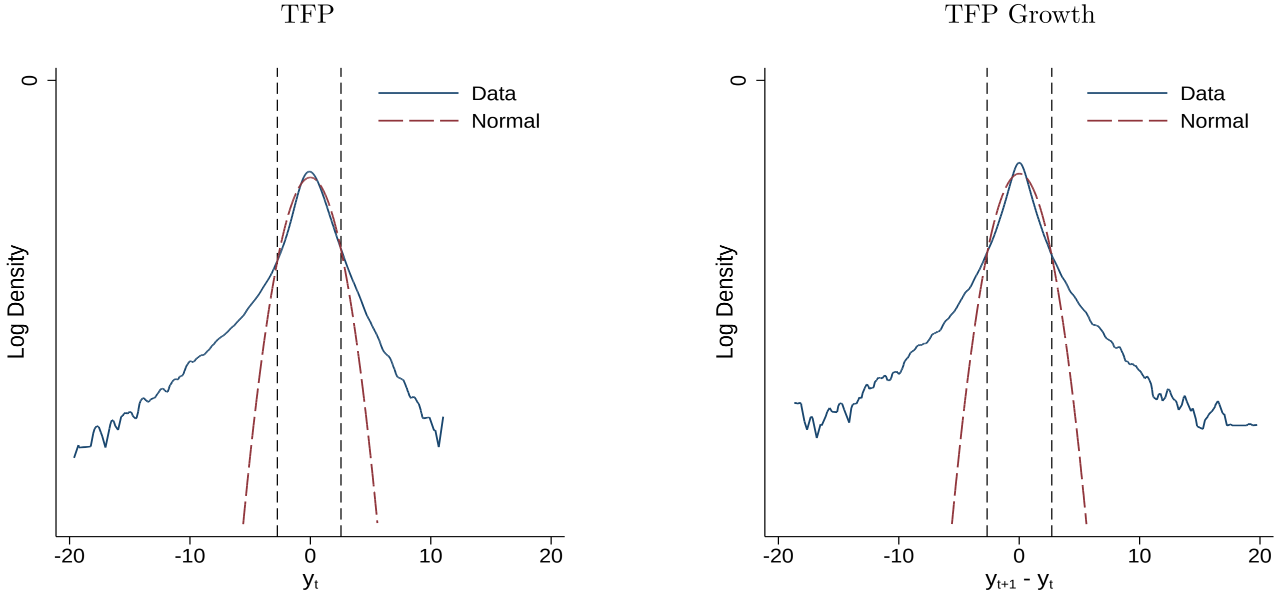
Table 13: Growth by Level, Unbalanced Panel, for the 5% largest firms

	Mean	SD	Skew.	Kelly Skew.	Kurt.
TFP					
0%-10%	0.26	0.63	1.56	19.68	0.37
10%-50%	0.05	0.30	-1.48	37.13	0.08
50%-90%	-0.01	0.24	-2.34	55.16	-0.04
90%-95%	-0.05	0.27	-1.42	28.86	-0.15
95%-99%	-0.09	0.33	-2.82	44.95	-0.25
99%-100%	-0.21	0.56	-2.77	21.03	-0.38

Table 14: One Year Transition Matrix, Unbalanced Panel (%)

	0%-10%	10%-50%	50%-90%	90%-95%	95%-99%	99%-100%
TFP						
0%-10%	51.03	39.87	7.81	0.60	0.50	0.18
10%-50%	10.09	67.92	20.98	0.57	0.34	0.09
50%-90%	2.01	22.04	68.35	5.34	1.95	0.30
90%-95%	1.18	4.93	44.78	31.26	16.41	1.44
95%-99%	1.16	3.52	21.14	21.23	45.38	7.57
99%-100%	1.77	3.08	10.38	6.78	28.92	49.08

Figure 20: Log Density Plot, Unbalanced Panel

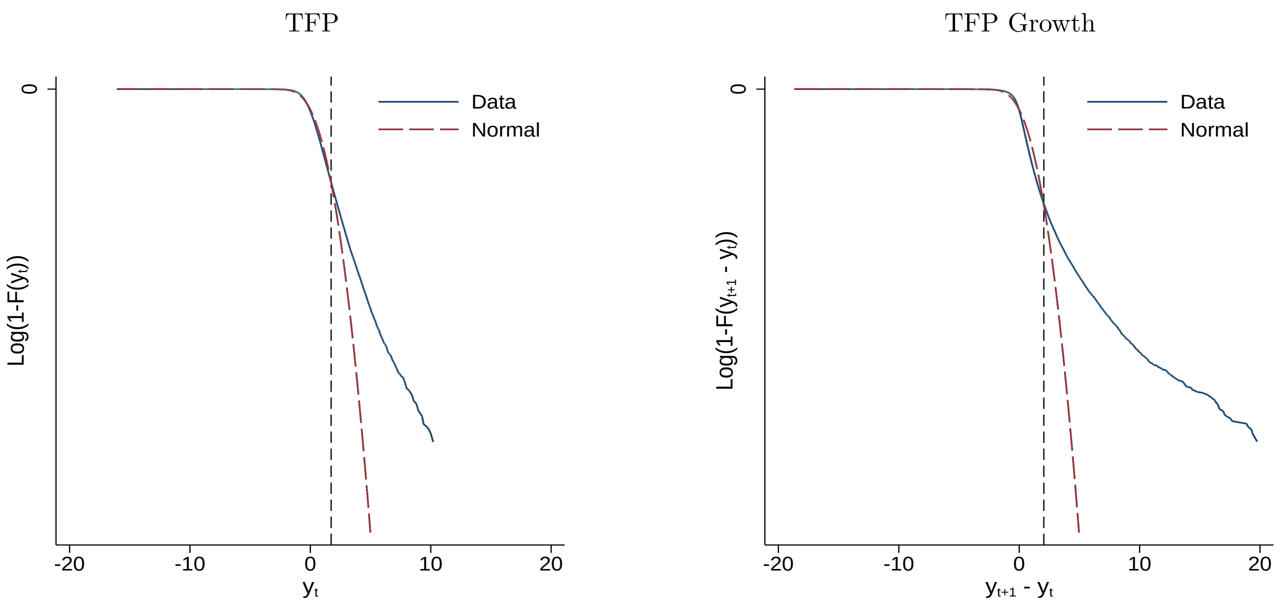


This figure shows the density in logs of TFP and TFP growth compared to their normal distribution counterparts. TFP and TFP growth have been standardized. The normal distribution is derived using the estimated mean and standard deviation of each variable in its probability density function. The proportion 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 Collard-Wexler and De Loecker (2025).

Table 15: Five Year Transition Matrix, Unbalanced Panel (%)

	0%-10%	10%-50%	50%-90%	90%-95%	95%-99%	99%-100%
TFP						
0%-10%	24.61	47.99	23.27	2.14	1.55	0.44
10%-50%	11.63	53.76	31.04	2.01	1.27	0.28
50%-90%	5.80	32.70	51.53	5.78	3.50	0.69
90%-95%	3.96	16.77	50.58	14.44	11.94	2.31
95%-99%	3.77	12.48	39.41	16.19	21.96	6.18
99%-100%	3.90	10.35	25.01	11.70	25.19	23.85

Figure 21: Log-Log Plot, Unbalanced Panel



This figure provides the log-log plot 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 slope for the right tail past the vertical dashed line for TFP and TFP growth are equal to -1.04 and -0.44, 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 Collard-Wexler and De Loecker (2025).

D Empirical patterns by productivity and size

Table 16: Growth by Level, Balanced Panel, for the 5% largest firms

	Mean	SD	Skew.	Kelly Skew.	Kurt.
TFP					
0%-10%	0.27	0.59	2.02	0.40	24.47
10%-50%	0.06	0.28	-0.76	0.12	31.75
50%-90%	0.00	0.22	-0.86	-0.01	26.18
90%-95%	-0.03	0.22	-1.22	-0.11	21.45
95%-99%	-0.06	0.28	-2.00	-0.20	29.12
99%-100%	-0.14	0.44	-2.22	-0.36	16.35

Table 17: One Year Transition Matrix, for the 5% largest firms (%)

	0%-10%	10%-50%	50%-90%	90%-95%	95%-99%	99%-100%
TFP						
0%-10%	51.55	32.71	12.32	1.35	1.26	0.82
10%-50%	9.02	57.37	30.31	1.77	1.22	0.31
50%-90%	1.16	10.80	73.94	9.78	3.61	0.72
90%-95%	0.44	1.86	35.35	40.43	19.99	1.94
95%-99%	0.48	1.54	11.88	21.53	55.27	9.31
99%-100%	0.72	1.04	6.87	4.79	24.65	61.92

Table 18: Five Year Transition Matrix, for the 5% largest firms (%)

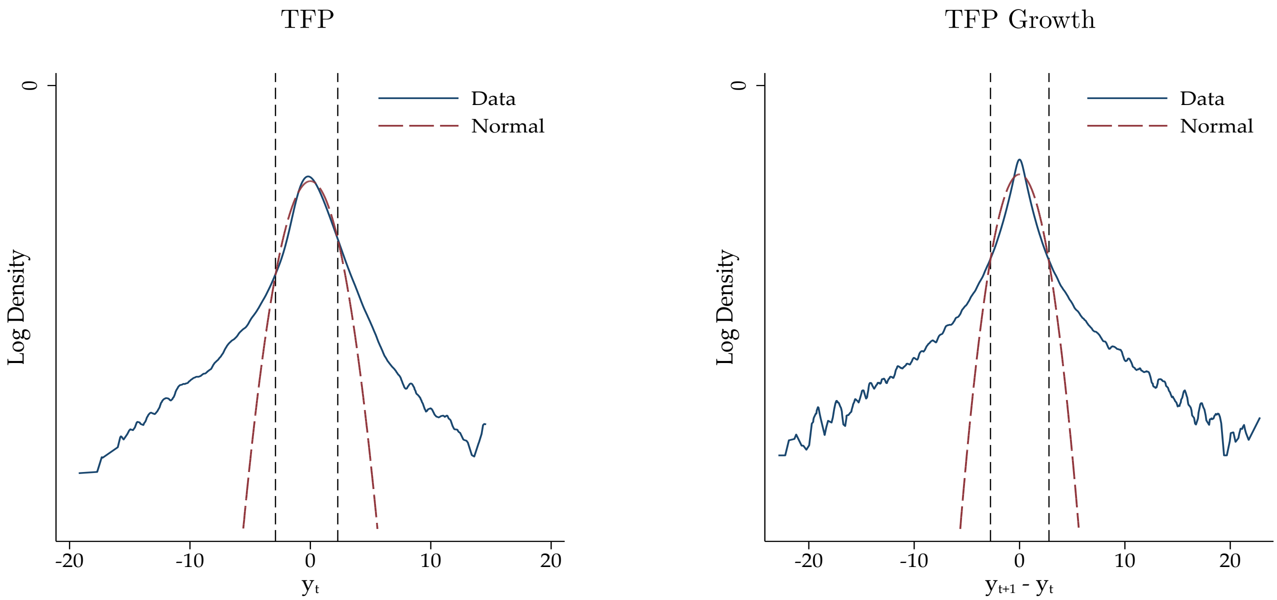
	0%-10%	10%-50%	50%-90%	90%-95%	95%-99%	99%-100%
TFP						
0%-10%	22.03	34.78	32.06	4.64	5.04	1.45
10%-50%	7.82	36.85	44.60	5.38	4.17	1.19
50%-90%	2.83	15.39	61.10	11.82	7.32	1.54
90%-95%	1.62	6.75	48.56	21.77	18.02	3.27
95%-99%	1.03	4.71	28.87	22.17	33.70	9.52
99%-100%	1.70	3.20	15.82	11.05	28.17	40.07

E Evidence from Orbis

Data Description The Orbis database provides comprehensive information on firms worldwide, including financial statements, ownership structures, and key performance indicators. For our analysis, we focus on a balanced panel of firms with consistent data over the study period. This allows us to use data from Belgium, France, Italy, Spain, Sweden, United Kingdom, Finland, and Portugal, for the years 2005-2022. Data cleaning follows Díez et al. (2021) and Kalemli-Özcan et al. (2024)

E.1 TFP measured using CWDL method

Figure 22: Log Density Plot, Orbis CWDL TFP



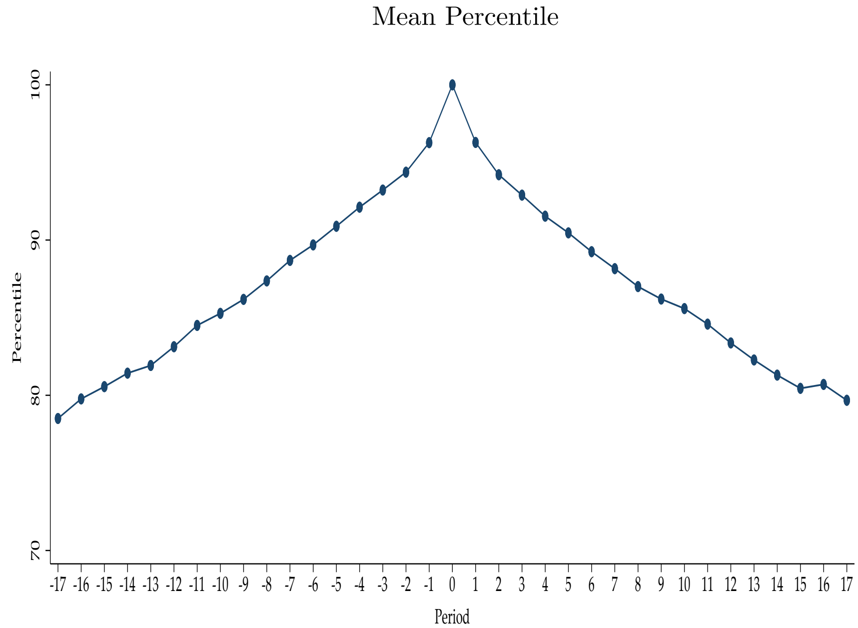
This figure shows the density in logs of TFP and TFP growth compared to their normal distribution counterparts. TFP and TFP growth have been standardized. The normal distribution is derived using the estimated mean and standard deviation of each variable in its probability density function. Growth is defined as the ln difference between time $t+1$ and t. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

Table 19: Growth by Level

Percentile	Mean	SD	Kelly Skew	Skewness	Kurtosis
0-10%	0.170	0.450	0.360	1.720	27.510
10-50%	0.010	0.200	0.050	-3.180	63.850
50-90%	-0.030	0.200	-0.090	-2.380	43.750
90-95%	-0.060	0.230	-0.190	-2.030	24.050
95-99%	-0.090	0.270	-0.250	-2.490	29.850
99-100%	-0.160	0.410	-0.390	-2.740	23.950

This table presents summary statistics of growth by different percentile groups in levels. The table presents the mean, standard deviation (SD), skewness (Skew.), Kelly skewness (Kelly skew.) and kurtosis (Kurt.). The Kelly skewness is given by $((90^{th} - 50^{th}) - (50^{th} - 10^{th}))/(90^{th} - 10^{th})$. Growth is defined as the ln difference between time $t+1$ and t. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

Figure 23: Tracking the Percentile of Top 1% Firms



This figure provides the mean percentile across firms for each year before (-) and after (+) being in the top 1%. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

Table 20: One Year Transition Matrix (%)

From	0-10%	10-50%	50-90%	90-95%	95-99%	Top 1%
0-10%	55.120	37.960	6.250	0.370	0.240	0.060
10-50%	9.410	71.760	18.260	0.350	0.180	0.030
50-90%	1.540	18.260	73.470	5.070	1.500	0.140
90-95%	0.690	3.070	40.280	38.030	16.920	1.010
95-99%	0.660	1.850	15.070	21.010	53.920	7.490
Top 1%	0.550	1.260	5.180	4.560	27.410	61.040

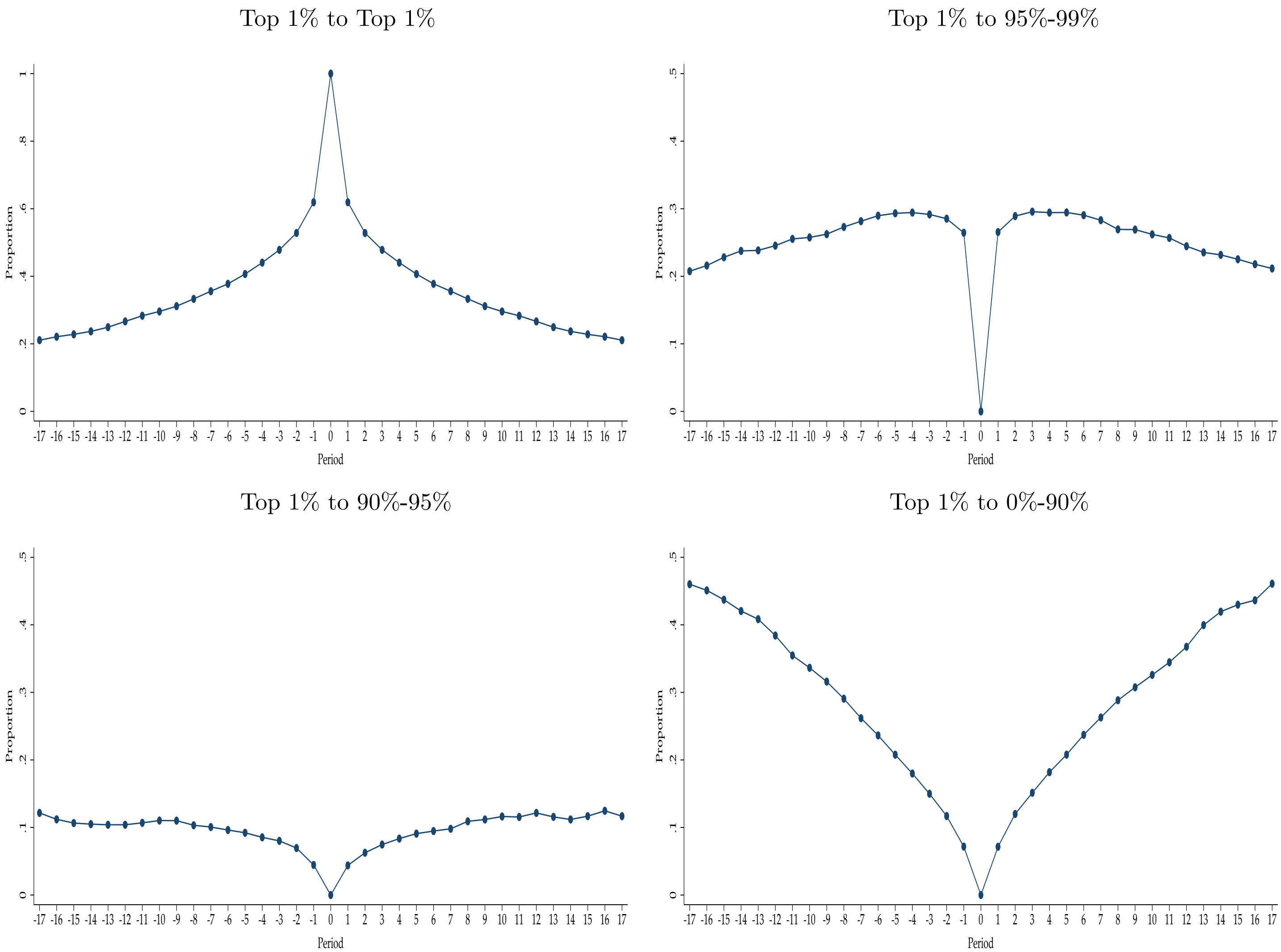
This table presents in percentages the one year transition matrix between percentile groups of 0%-10%, 0%-50%, 50%-90%, 90%-95%, 95%-99% and 99%-100%. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

Table 21: Five Year Transition Matrix (%)

From	0-10%	10-50%	50-90%	90-95%	95-99%	Top 1%
0-10%	36.300	48.200	13.820	0.940	0.600	0.150
10-50%	11.490	59.630	26.930	1.190	0.640	0.110
50-90%	3.740	26.490	60.030	6.190	3.120	0.440
90-95%	2.420	9.480	48.860	21.350	15.780	2.100
95-99%	2.110	6.870	30.440	19.670	32.940	7.960
Top 1%	2.000	4.330	14.900	9.190	29.980	39.600

This table presents in percentages the five year transition matrix between percentile groups of 0%-10%, 0%-50%, 50%-90%, 90%-95%, 95%-99% and 99%-100%. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

Figure 24: Tracking the Trajectories of Top 1% Firms



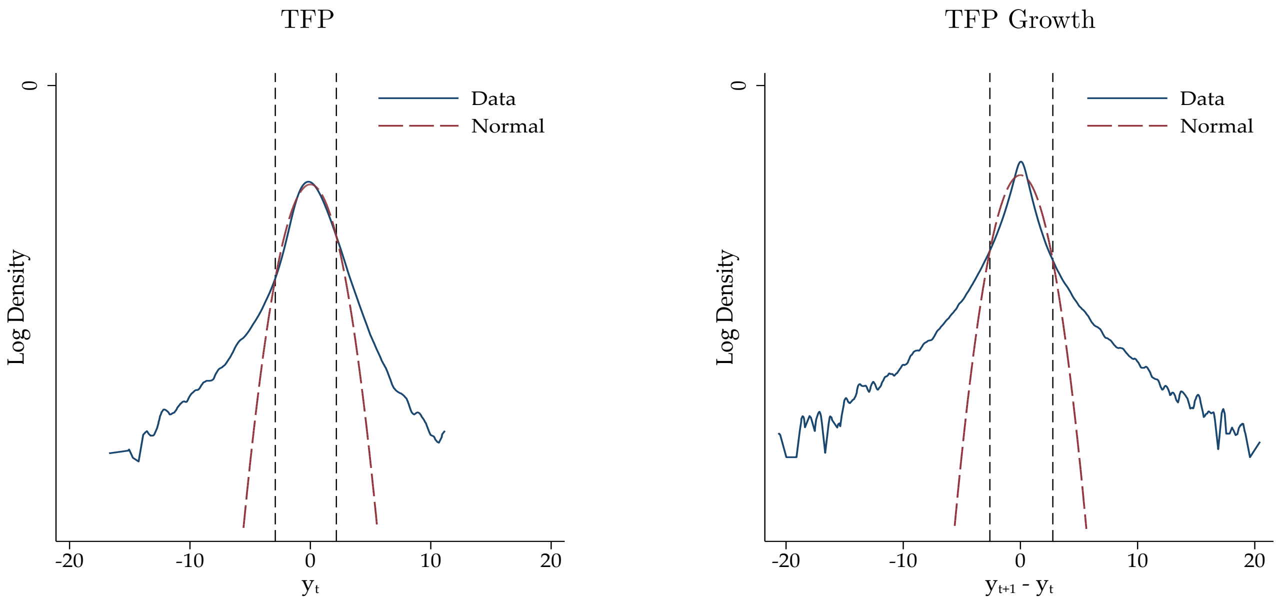
These plots provide the mean proportion of firms in each percentile group for each year before (-) and after (+) being in the top 1%. TFP is derived from the IV One-Step Control approach from Collard-Wexler and De Loecker (2025).

Table 22: Growth and Probability of Staying in Top 1% by Time Spent in the Top 1%

Years in top 1%	Mean growth	Pr(stay in top 1%)
any	-0.150	61.040
1	-0.220	44.210
2	-0.140	61.140
3	-0.120	69.980
4+	-0.080	81.840

E.2 TFP measured using factor shares

Figure 25: Log Density Plot, Orbis Factor Share TFP



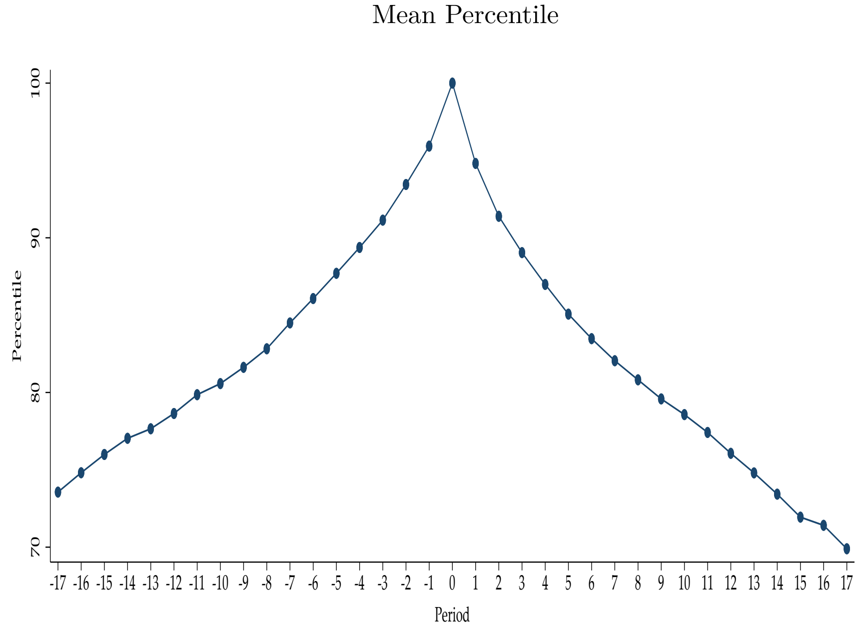
This figure shows the density in logs of TFP and TFP growth compared to their normal distribution counterparts. TFP and TFP growth have been standardized. The normal distribution is derived using the estimated mean and standard deviation of each variable in its probability density function. Growth is defined as the ln difference between time $t+1$ and t. TFP is measured using the factor share method..

Table 23: Growth by Level

Percentile	Mean	SD	Kelly Skew	Skewness	Kurtosis
0-10%	0.150	0.440	0.350	1.810	27.290
10-50%	0.020	0.220	0.040	-1.960	43.360
50-90%	-0.030	0.230	-0.120	-1.720	26.730
90-95%	-0.080	0.290	-0.200	-1.550	15.870
95-99%	-0.120	0.340	-0.250	-1.720	15.930
99-100%	-0.230	0.480	-0.360	-1.730	10.600

This table presents summary statistics of growth by different percentile groups in levels. The table presents the mean, standard deviation (SD), skewness (Skew.), Kelly skewness (Kelly skew.) and kurtosis (Kurt.). The Kelly skewness is given by $((90^{th} - 50^{th}) - (50^{th} - 10^{th}))/(90^{th} - 10^{th})$. Growth is defined as the ln difference between time $t+1$ and t. TFP is measured using the factor share method..

Figure 26: Tracking the Percentile of Top 1% Firms



This figure provides the mean percentile across firms for each year before (-) and after (+) being in the top 1%. TFP is measured using the factor share method..

Table 24: One Year Transition Matrix (%)

From	0-10%	10-50%	50-90%	90-95%	95-99%	Top 1%
0-10%	61.670	33.740	4.110	0.250	0.190	0.050
10-50%	8.130	74.280	17.080	0.310	0.160	0.040
50-90%	1.200	16.680	74.780	5.560	1.600	0.180
90-95%	0.680	3.550	39.930	34.920	19.660	1.270
95-99%	0.540	2.360	18.490	20.140	49.450	9.020
Top 1%	0.510	2.120	8.750	6.560	27.400	54.650

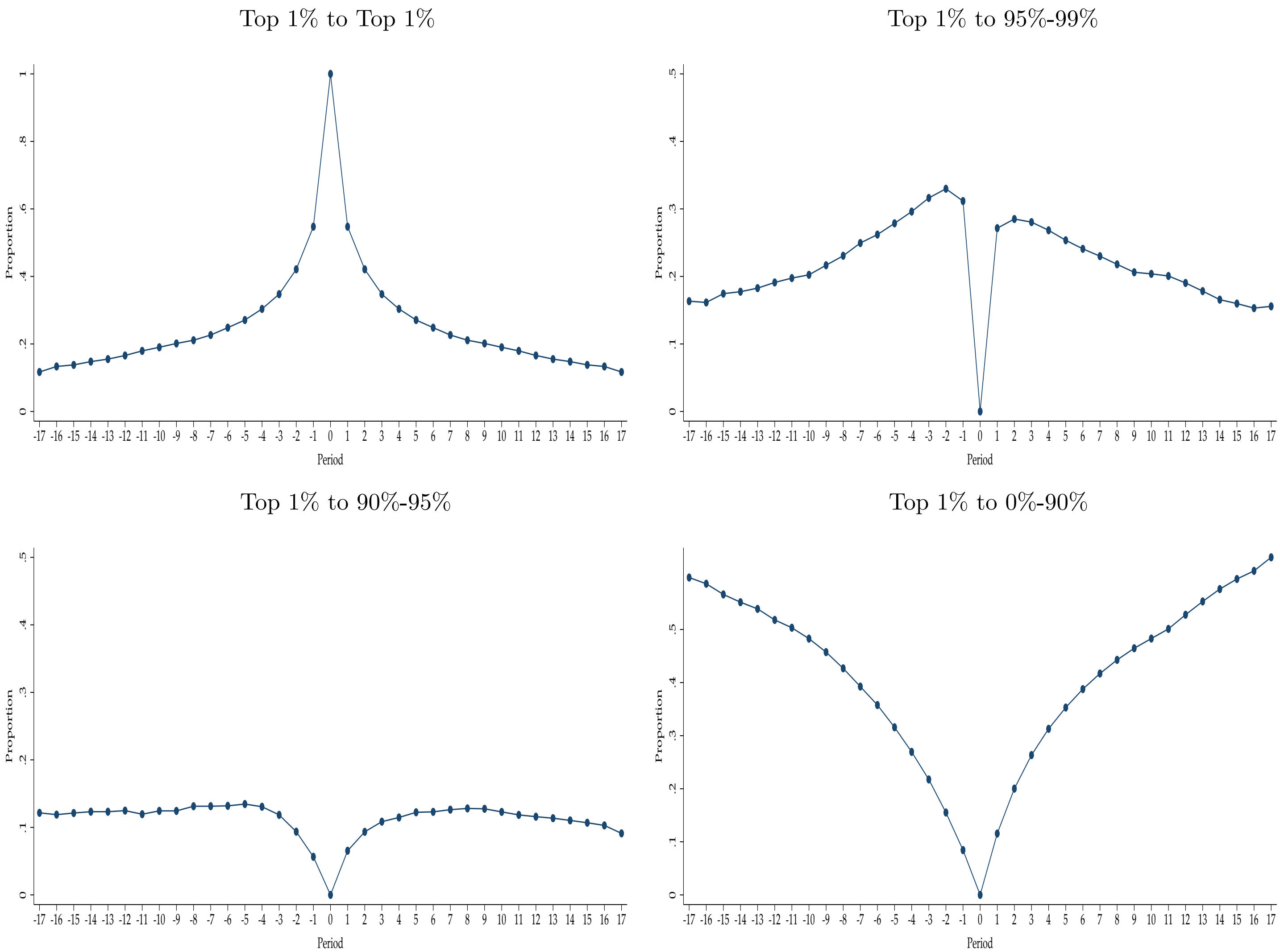
This table presents in percentages the one year transition matrix between percentile groups of 0%-10%, 0%-50%, 50%-90%, 90%-95%, 95%-99% and 99%-100%. TFP is measured using the factor share method..

Table 25: Five Year Transition Matrix (%)

From	0-10%	10-50%	50-90%	90-95%	95-99%	Top 1%
0-10%	41.630	46.130	10.730	0.810	0.550	0.150
10-50%	10.580	59.790	27.510	1.240	0.730	0.160
50-90%	3.300	26.090	59.050	6.820	4.030	0.710
90-95%	2.370	12.510	49.990	17.010	15.100	3.020
95-99%	2.160	9.850	37.620	17.560	24.780	8.020
Top 1%	1.860	8.050	25.550	12.130	25.110	27.300

This table presents in percentages the five year transition matrix between percentile groups of 0%-10%, 0%-50%, 50%-90%, 90%-95%, 95%-99% and 99%-100%. TFP is measured using the factor share method..

Figure 27: Tracking the Trajectories of Top 1% Firms



These plots provide the mean proportion of firms in each percentile group for each year before (-) and after (+) being in the top 1%. TFP is measured using the factor share method..

Table 26: Growth and Probability of Staying in Top 1% by Time Spent in the Top 1%

Years in top 1%	Mean growth	Pr(stay in top 1%)
any	-0.230	54.650
1	-0.270	42.940
2	-0.230	56.360
3	-0.210	63.070
4+	-0.150	75.710