

Melcangi, Davide; Sarpietro, Silvia

Working Paper

Nonlinear firm dynamics

Staff Report, No. 1088

Provided in Cooperation with:

Federal Reserve Bank of New York

Suggested Citation: Melcangi, Davide; Sarpietro, Silvia (2024) : Nonlinear firm dynamics, Staff Report, No. 1088, Federal Reserve Bank of New York, New York, NY, <https://doi.org/10.59576/sr.1088>

This Version is available at:

<https://hdl.handle.net/10419/300454>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

NO. 1088
MARCH 2024

Nonlinear Firm Dynamics

Davide Melcangi | Silvia Sarpietro

Nonlinear Firm Dynamics

Davide Melcangi and Silvia Sarpietro

Federal Reserve Bank of New York Staff Reports, no. 1088

March 2024

<https://doi.org/10.59576/sr.1088>

Abstract

This paper presents empirical evidence on the nature of idiosyncratic shocks to firms and discusses its role for firm behavior and aggregate fluctuations. We document that firm-level sales and productivity are hit by heavy-tailed shocks and follow a nonlinear stochastic process, thus departing from the canonical linear. We estimate a state-of-the-art model to flexibly capture the rich dynamics uncovered in the data and characterize the drivers of nonlinear persistence and non-Gaussian shocks. We show that these features are crucial to get empirically plausible volatility and persistence of micro-originated (granular) aggregate fluctuations.

JEL classification: D22, E23, E32

Keywords: firm dynamics, productivity, nonlinearities, non-Gaussian shocks, granular fluctuations

Melcangi: Federal Reserve Bank of New York (email: davide.melcangi@ny.frb.org). Sarpietro: Bologna University (email: silvia.sarpietro@unibo.it). The authors thank David Dam, Meghana Gaur, Matthew Fisher, and Brendan Moore for excellent research assistance, as well as Martin Almuzara, Manuel Arellano, Salome Baslandze, Augustin Belin, Vasco Carvalho, Minsu Chang, Marco Del Negro, Julian di Giovanni, Jonathan Dingel, Giulio Fella, Sebastian Heise, Ben Hyman, Fatih Karahan, Vincent Sterk, and Yulong Wang for their helpful comments.

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in this paper are those of the author(s) and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the author(s).

To view the authors' disclosure statements, visit
https://www.newyorkfed.org/research/staff_reports/sr1088.html.

1 Introduction

Firms are hit by idiosyncratic shocks of various nature and magnitude. The features of this firm-level variation are crucial for many economic questions that relate to firm dynamics.¹ Despite their importance, however, the statistical properties of firms' fundamentals remain elusive, especially when it comes to higher-order moments and rich time series processes.² The vast majority of firm dynamics models assume that firm-level productivity follows an AR(1) process with normal innovations, but potential departures from this benchmark are likely to have substantial consequences. Rare, nonlinear, idiosyncratic shocks can even become a source of aggregate fluctuations.

In this paper, we document that, even among relatively large, publicly listed, US firms, firm-level sales and productivity substantially depart from standard AR(1) models with Gaussian innovations. We find that heavy-tailed, extreme, shocks are a pervasive feature of firm-level growth rates. Changes to revenue-based productivity are even more heavy-tailed than sales, suggesting that firms are able to absorb some of these fundamental fluctuations. In addition, we document substantial nonlinearities in persistence: firms can be hit by large shocks that erase the impact of previous shocks. We estimate a model of productivity that separates persistent and transitory shocks, is consistent with our empirical findings, and helps us characterize their drivers. Finally, we show analytically and quantitatively how non-Gaussianity and nonlinearities affect the importance of idiosyncratic (granular) shocks for aggregate fluctuations. Simultaneously accounting for these features is crucial to obtain empirically plausible volatility and autocorrelation of granular fluctuations.

Our starting point is an agnostic analysis of firm sales, using data from US Compustat. We document that firm-level sales growth is very concentrated – relative to a Gaussian benchmark – but exhibits heavy tails. Sales growth is fairly symmetric, but very large changes of either sign are likely to occur: kurtosis is much higher than what implied by a Gaussian distribution at all quantiles of firm sales, and especially high for small firms.

We then investigate the dynamics of firm sales. Following recent work on household earnings by [Arellano et al. \(2017\)](#), we estimate quantile autoregressions and document that the persistence of lagged firm sales is not constant in the size and sign of sales changes as

¹For example, firm-level stochastic processes matter for the role played by input adjustment costs (e.g., [Bloom \(2009\)](#) and [Roys \(2016\)](#)), misallocation losses from financial frictions (see [Moll \(2014\)](#)), and in general are key inputs to macroeconomic models with firm heterogeneity (e.g., [Khan and Thomas \(2008\)](#)). Moreover, idiosyncratic shocks can even be a source of aggregate fluctuations (see [Gabaix \(2011\)](#), [Di Giovanni et al. \(2014\)](#) [Carvalho and Grassi \(2019\)](#)).

²Complementary to our analysis, [Salgado et al. \(2019\)](#) show that the skewness of sales and productivity growth is procyclical. A burgeoning literature has recently documented non-normality and nonlinearities in household earnings dynamics (see [Arellano et al. \(2017\)](#) and [Güvenen et al. \(2021\)](#)): our analysis is indeed inspired by this influential recent work.

would be implied by a standard AR(1) process. In particular, the approach introduced by [Arellano et al. \(2017\)](#) provides a measure of persistence of past shocks, often referred to as the persistence of history. We find that such persistence is heterogeneous, with small firms being hit by large positive shocks to their sales having the lowest persistence. This means that a large positive shock hitting a small firm can wipe out the history of past negative shocks. However, among large firms, good and bad past shocks exhibit similar persistence.

Firm sales clearly represent the combination of exogenous fundamental shocks and endogenous inputs. We find that our findings on sales are not solely driven by endogenous responses, as revenue-based labor productivity and revenue-based Total Factor Productivity (TFP) do not seem to be characterized by either a AR(1) process or to face Gaussian innovations. In fact, we find that TFP growth is even more concentrated and heavy-tailed than labor productivity and, in turn, sales. Nearly half of TFP growth rates are less than 5% in absolute value, 5 times more likely than if the data was normally distributed and more than twice as likely than for sales growth. Rare, large, TFP changes of either sign occur: kurtosis is larger than for sales. TFP growth exhibits mildly negative skewness among large firms, but these firms seem to absorb left-tailed shocks given what we find for sales. An illustrative exercise suggests that this is consistent with firms adjusting their inputs increasingly less as the productivity shock becomes larger.³

TFP also exhibits nonlinear persistence of history: in fact, one that is more pronounced than for sales, with double the range of estimated persistence values. In particular, we find evidence of microeconomic disasters hitting large firms. Since this feature holds for productivity but not for sales, it further confirms that large firms are equipped to absorb the impact of disastrous negative shocks. In summary, we conclude that firms face fundamental shocks that feature non-Gaussianity and nonlinearities.

In order to investigate further the drivers and the implications of our empirical results, we estimate a rich, state-of-the-art, model of productivity dynamics, following work by [Arellano et al. \(2017\)](#) and [Arellano et al. \(2021\)](#) on household earnings. The model features both transitory and persistent components, and allows for non-Gaussianity and nonlinear persistence of history, hence making it consistent with our empirical findings. The estimated model indicates that both persistent and, especially, transitory shocks are non-Gaussian. While a decomposition of this type is common in the household literature, transitory shocks are often overlooked in firm dynamics. Our results suggest that they should be taken seriously, especially given the estimated degree of heavy-tailedness. Shocks of varying degree

³In complementary recent work, [Fella et al. \(2021\)](#) find that Spanish firm-level productivity exhibits non-linearity and non-normality. They find that accounting for these productivity features alters the estimation of capital adjustment costs.

of persistence have been shown to be important for various aspects of firm dynamics, such as input adjustment costs (Cooper and Haltiwanger (2006), Fella et al. (2021)), financial frictions (Moll (2014)), insurance provision (Guiso et al. (2005)).

We focus on a different aspect: rare, large, idiosyncratic shocks as those we document can be a direct source of aggregate fluctuations. Indeed, Gabaix (2011) supports his granular hypothesis with many historical examples that are consistent with our empirical findings. Among others, the major strike at General Motors in 1970, or natural disasters in the U.S. (Barrot and Sauvagnat (2016)), Japan (Boehm et al. (2019)), and more generally rare economic disasters (Barro (2006)).⁴ Large shocks can also be positive, such as those at Walmart in 2002.⁵

Conceptual frameworks studying the granular hypothesis are typically static and feature little or no discussion on the nature of the shocks. One exception is Carvalho and Grassi (2019), who build a systematic framework linking firm dynamics and micro-originated fluctuations; however, the nature of firm-level stochastic processes is not central to their conclusions. To the best of our knowledge, we are the first to study the role played by non-Gaussian shocks and nonlinear persistence for the granular hypothesis. We proceed in two steps.

First, we show analytically the role played by the heavy-tailed distribution of shocks in a stylized theoretical setting. Intuitively, non-Gaussian shocks can provide a reason for a heavy-tailed distribution of firm sizes, which is the typical cornerstone of the granular hypothesis. Moreover, for a given size distribution, non-Gaussian shocks play an ambiguous role for granular aggregate fluctuations, depending on whether heavier tails are coupled with larger idiosyncratic variance, and thus greater aggregate volatility.

Second, we use our estimated model to quantitatively assess the importance of non-Gaussianities and nonlinearities for granular fluctuations. To do so, we compute the volatility and autocorrelation of the growth rate in an aggregate measure of TFP, generated by simulating our baseline estimated model and four alternative models. Separating a transitory and a persistent component of productivity, even in an AR(1) setup, increases aggregate volatility towards the empirical counterpart, but at the expense of counterfactually negative autocorrelation. In contrast, modeling a single component with nonlinear persistence, using estimates from empirical quantile autoregressions, makes the autocorrelation of aggregate TFP growth rates excessively large. Adding Gaussian transitory shocks to this framework results in excessively high volatility and negative autocorrelation. It is only with our full model, which

⁴Other examples of large negative shocks are scandals that reduce firms' profitability. For example, in January 2016 55 E. Coli bacteria infections were linked to food served at Chipotle restaurants. Q1 2016 Chipotle revenues dropped 23%.

⁵Sterk et al. (2021) study high-growth, "gazelles", startups, who make persistent contributions to aggregate productivity growth.

features nonlinear persistence *and* non-Gaussian transitory and persistent shocks, that we can generate aggregate fluctuations that align with the data.

Our exercise thus suggests that non-Gaussian shocks as typically observed in the data are associated with less volatile micro-originated aggregate fluctuations, compared to a Gaussian counterpart. We therefore propose a new, empirically grounded, dampening channel, complementary to, but conceptually distinct from, the size-variance tradeoff documented by Stanley et al. (1996) and Yeh (2019). Our exercise also shows that nonlinearities and non-Gaussianities need to be simultaneously accounted for; failure to do so generates counterfactual behavior, not just at the firm level, but also in the aggregate.

Our paper is closely related to three recent studies investigating heavy-tailed and nonlinear firm-level shocks. Jaimovich et al. (2023) focus on the implications for the stationary distribution of firms and the effects of firm subsidies. Fella et al. (2021) discuss how non-normalities influence the quantitative importance of capital adjustment costs. Salgado et al. (2019) focus on the cyclicity of firm-level skewness. Complementary to these studies, we explore a novel dimension through which the nature of firm-level shocks matters: micro-originated aggregate fluctuations.

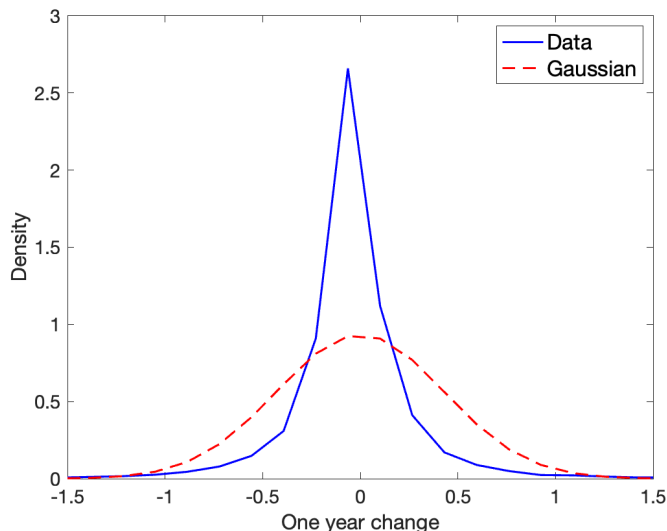
The paper is organized as follows. In Section 2 we present our empirical results on firm sales. Section 3 shows that TFP also follows a non-Gaussian and nonlinear stochastic process. Section 4 estimates a model consistent with the empirical findings. Section 5 shows, analytically and quantitatively, how non-Gaussianity and nonlinearities affect the granular hypothesis. Section 6 concludes.

2 Non-Gaussianity and nonlinearities in firm sales

We start by documenting the empirical features of firm sales, using data from US Compustat over the period between 1987 and 2017. We keep firms that remain in operation for at least 2 consecutive years and apply standard selection criteria. We focus on the manufacturing sector but we show robustness to the whole economy in Appendix B.4. Appendix A provides details on the sample. We regress the log of firm sales on industry dummies and a linear time trend, and retain the residuals as our variable of interest. By doing so, we absorb observable permanent differences and low frequency aggregate fluctuations. As such, our measures can be thought as the combination of idiosyncratic shocks and heterogeneous sensitivities to aggregate shocks, if any.⁶ For the remainder of the paper, when mentioning sales and productivity we always refer to their residualized counterparts.

⁶For some further discussion on this, see Section 5. In Appendix B.6 we show how our results change if we do not detrend our data.

Figure 1: Distribution of one-year log sales changes



Notes: Empirical densities of residualized log sales changes between year t and $t - 1$ depicted in solid blue. The standard deviation is 0.427, Kelley Skewness 0.052 and Crow-Siddiqui Kurtosis 6.457. Gaussian density with the same standard deviation of data shown in dashed red. See main text for details on the data.

Figure 1 shows how the distribution of sales growth, measured as annual changes in residualized log sales, is non-Gaussian. In the data there are far more extreme sales changes than what is implied by a Gaussian distribution with the same standard deviation as in the data. Indeed, sales growth displays high concentration and heavy tails. In the data, 22% of the time sales growth is less than 5% in absolute value, compared to only 9% under the Gaussian distribution. Two thirds of firms experience an annual sales growth rate smaller than 20% in absolute terms; if the data was normally distributed, only one third of the observations would experience these growth rates.

What are the features of this non-Gaussianity? To investigate this further, we look at the skewness and kurtosis of sales growth, by quantiles of past sales.⁷ First, shocks are roughly symmetric even for subsets of firms, as shown in Figure 2a. We report Kelley (1947) skewness, defined as $S_K = \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{P_{90} - P_{10}}$. A positive Kelley Skewness implies that the right part of the distribution ($P_{50} - P_{10}$) accounts for a larger share of the overall dispersion ($P_{90} - P_{10}$) than the lower part. This seems to be particularly true for firms with low sales in $t - 1$, although even for this group of firms the departure from symmetry is mild. These relatively

⁷The population skewness and kurtosis might be infinite with heavy-tailed data, see for instance Sarpietro et al. (2022). Thus, we consider quantile-based measures of skewness and kurtosis, which are robust to the existence of moments, and refer to them simply as skewness and kurtosis throughout the rest of the paper.

small deviations from symmetry are in sharp contrast with substantially negative skewness of household earnings growth that has been documented, among others, by [Guvenen et al. \(2021\)](#). Our findings are instead consistent with the symmetric distribution of employment growth rates among mature firms documented by [Decker et al. \(2016\)](#).

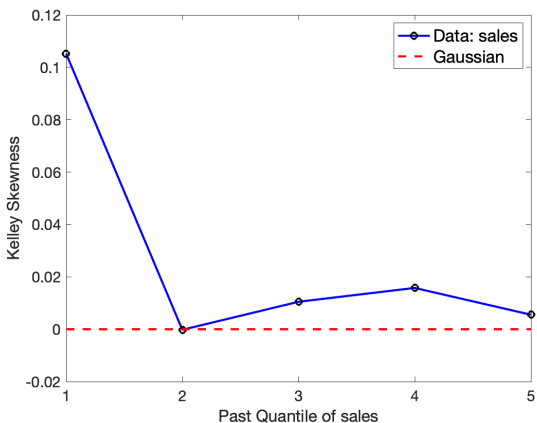
Kurtosis, rather than skewness, is the main driver of non-Gaussianity in firm sales growth. As shown in [Figure 2b](#), firms of different size all experience heavy tailedness of shocks. We measure this with the Crow-Siddiqui measure of kurtosis, defined as $K_K = \frac{P97.5 - P.25}{P75 - P25}$, which purges the measure from outliers. Kurtosis is highest at the bottom of the sales distribution and is always higher than 5 for all other quintiles of past sales. These values are significantly higher than the benchmark kurtosis of 2.906 for the standard Normal distribution. Even the largest firms face sales growth that is twice as heavy-tailed as its Gaussian counterpart. Large values of conditional kurtosis are common to previous findings on household earnings. Our results are also consistent with German data on investment surprises, which display excess kurtosis but no significant skewness, as showed by [Bachmann et al. \(2017\)](#). This study interprets their findings as at least partly driven by fat-tailed underlying firm risk, a conclusion that is consistent with our later findings on productivity.

Moreover, our findings on kurtosis do not hinge on specific aggregate conditions.⁸ Indeed, kurtosis is above 4 for all years in our sample and high across all quintiles of past sales even during recessions. We thus conclude that the heavy tailedness is a key feature of the distribution of sales growth and represents a clear departure from normality.

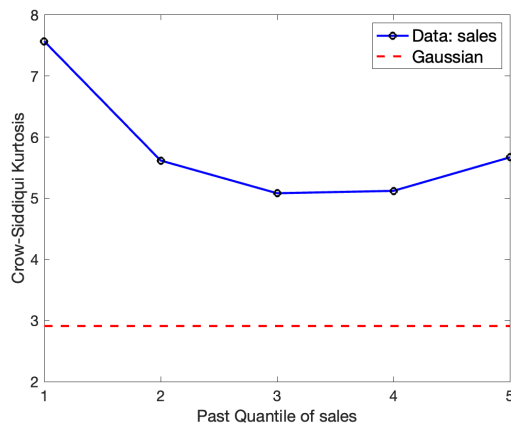
We perform some robustness checks to validate our findings. First, we investigate whether the departures from normality are driven by transitory changes only. Following [Guvenen et al. \(2021\)](#), we consider a measure of persistent changes defined as the difference between two 3-year averages of residualized log sales as follows: $\Delta(\tilde{y}) = \bar{y}_{t-1,t+1} - \bar{y}_{t-2,t-4}$, where $\bar{y}_{t-j,t-k}$ is the average of log sales over the period $[t-j, t-k]$. Through averaging before differencing, this alternative measure allows us to purge the effects of the transitory component of sales changes. We repeat our analysis with this measure of persistent changes and obtain qualitatively similar results, as shown in [Appendix B.2](#). We find that persistent changes are non-Gaussian, although they exhibit lower concentration and relatively less heavy-tailedness compared to the benchmark measure of overall sales changes. As a second robustness check, we repeat the analysis for the whole economy, besides manufacturing. Results are broadly unchanged and shown in [Appendix B.4](#); skewness departs little from Gaussian symmetry and the distribution has heavier tails compared to the reference normal density, with kurtosis ranging from 4.8 to 7.5. Third, we restrict the sample by trimming relatively small

⁸For an analysis on how higher order moments of sales growth evolve over the business cycle see [Higson et al. \(2002\)](#) and [Bloom et al. \(2018\)](#).

Figure 2: Skewness and kurtosis of one-year log sales changes



(a) Kelley Skewness



(b) Crow-Siddiqui Kurtosis

Notes: Observations are ordered by quantiles of the distribution of $t - 1$ residualized log sales (x-axis). For each quantile, we show on the y-axis the Kelley Skewness of residualized log sales changes in Figure 2a and the Crow-Siddiqui Kurtosis in Figure 2b.

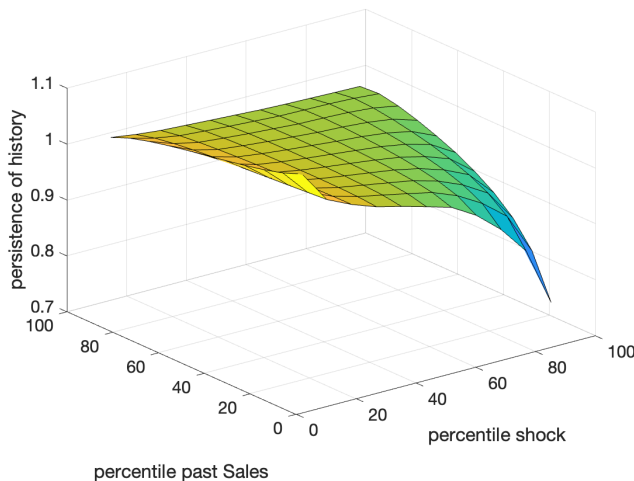
firms.⁹ All the results are basically unaffected, suggesting that the differential behavior at the bottom of the distribution is not driven by outliers. To conclude our robustness analysis, we compute confidence bands via bootstrapping. The tight confidence bands shown in Appendix B.5 confirm the precision and reliability of our findings.

Finally, we focus on the dynamics of firm sales. Following Arellano et al. (2017), we plot a measure of persistence in the data in Figure 3. The figure shows the estimated average derivative of the conditional quantile function of residualized log sales $y_{i,t}$ given $y_{i,t-1}$.¹⁰ Derivatives are taken with respect to $y_{i,t-1}$ and evaluated at a given percentile of $y_{i,t-1}$, and at a given percentile of the shock. This derivative effect can be interpreted as a measure of persistence in an autoregressive process for log sales. As such, it provides a measure of persistence of past shocks, often referred to as the persistence of history. As discussed by Arellano et al. (2017), a standard AR(1) process would deliver the same level of history persistence for all firms, regardless of their past sales and/or shocks: in other words, a flat three-dimensional figure. Figure 3, instead, shows some nonlinearities, with persistence of firm sales history ranging from 0.76 to slightly above 1. History persistence is lowest for low-sales firms hit by a large positive shock. Thus, a large positive shock to firms at the bottom of the sales distribution can wipe out the history of past negative shocks (see the right end of the graph). The opposite does not hold true: for high-sales firms hit by the smallest (most

⁹In particular, we drop firms with less than 10 employees, or either sales and total assets less than 100 thousand dollars. As such, we exclude about 3% of the firms relative to the baseline sample.

¹⁰As done in the literature, quantile functions are set as third-order Hermite polynomials.

Figure 3: Quantile autoregressions of log sales



Notes: Estimated average derivative of the conditional quantile function of residualized log sales $y_{i,t}$ given $y_{i,t-1}$. Derivatives are taken with respect to $y_{i,t-1}$ and evaluated at a given percentile of $y_{i,t-1}$, and at a given percentile of the shock. Details on data and estimation of quantile functions in the text.

negative) shocks persistence is around 1. Hence, negative shocks do not cancel out the good sales history of large firms. All these results are broadly unaffected when looking at the whole economy (beside manufacturing) and at recession years.

The one-sided nonlinearity that we document for firm sales differs from recent findings on household earnings by [Arellano et al. \(2017\)](#), with high-income households facing “micro disasters”. It is possible that these micro disasters also hit large firms and make them exit the market, thus eluding our selected sample. However, we believe that this channel should play a limited role. First, publicly quoted firms are less likely to exit than all firms in the population, thus making Compustat less prone to this selection issue. Second, we are also keeping firms that stay in the sample for only two years, thus limiting selection as much as possible. Third, we find that the probability of firm exit monotonically decreases with past sales. As such, even if present, this selection issue is less likely to be relevant among high-sales firms. Fourth, as we show next, large firms do face micro disasters – in the form of fundamental shocks – but they seem to absorb them via input adjustment. As such, persistence of history is more nonlinear for productivity than for sales.

To conclude, we have shown several statistical properties of firm sales that are inconsistent with a simple random walk process with gaussian shocks, even in a fairly homogeneous sample as Compustat. In the following sections, we investigate whether these departures from linearity and Gaussianity mainly come from shocks or responsiveness. To do so, we investigate whether our findings are solely driven by endogenous responses turning to two

alternative productivity measures: Labor Productivity (LP) and Total Factor Productivity (TFP).

3 Non-Gaussianity and nonlinearities in productivity

In this section we investigate whether the documented non-Gaussianity and nonlinearities in firm sales are features of the distributions of more exogenous measures of firm performance, or they are the result of peculiar ways in which firms’ inputs respond to fundamental shocks.

We look at two measures of productivity: revenue labor productivity and revenue-based total factor productivity. We assume that firms operate with a Cobb-Douglas production, using a composite measure of variable inputs, C , and capital stock, K . Then, for firm $i = 1, \dots, N$ at time $t = 1, \dots, T$, we estimate TFP as a residual of the revenue function as follows:

$$z_{i,t} = y_{i,t} - \beta_c c_{i,t} - \beta_k k_{i,t} \tag{1}$$

where $y_{i,t}$ is the log sales of firm i at time t , $c_{i,t}$ is the log of the variable input and $k_{i,t}$ is the logged capital stock. Following the analysis by [De Loecker et al. \(2020\)](#), also with Compustat data, we use cost of goods sold (COGS) for $c_{i,t}$. This variable combines all expenses directly attributable to the production of the goods sold by the firm; as such, it includes not only labor costs, but also other inputs such as materials, intermediates, and energy. Compustat does not provide a detailed breakdown of these inputs.¹¹ We construct the capital stock following [Ottonello and Winberry \(2020\)](#), as we discuss in Appendix A. Since we do not observe firm-level prices, our TFP measure effectively combines idiosyncratic technical efficiency shocks with demand shocks. With this caveat in mind, this remains the relevant measure of fundamental shocks that affect firms’ revenues, as discussed by [Decker et al. \(2020\)](#) and [Blackwood et al. \(2021\)](#).¹² In order to estimate the revenue elasticities, β_x , with $x = \{c, k\}$, we follow [Decker et al. \(2020\)](#) and note that the first-order condition on a given production input x from static profit maximization implies that β_x equates the share of that factor’s costs in total sales. We construct these ratios using industry-level data from the KLEMS database constructed by the Bureau of Labor Statistics (BLS). Our revenue elasticities are year-specific and industry-specific. The industry granularity we use is 3-digit NAICS for the manufacturing sector and 2-digit for the whole economy. The first order

¹¹[Keller and Yeaple \(2009\)](#) impute a measure of intermediate inputs, but this requires additional assumptions.

¹²See [De Loecker and Goldberg \(2014\)](#) for a discussion of output and input price heterogeneity in the context of TFP estimation. [DeVera \(2023\)](#) use Portuguese data to disentangle productivity from demand shocks: they find nonlinear persistence in either component, consistent with our results.

condition pinning down revenue elasticities holds if there are no factor adjustment costs or wedges; as is typically assumed in the literature (see for instance [Decker et al. \(2020\)](#)), we require this condition to hold on average across firms within an industry. As suggested by [Syverson \(2011\)](#), we further strengthen this argument by averaging the revenue elasticities over 6-year windows. We choose this approach over some alternatives for TFP estimation to avoid imposing restrictive assumptions on the stochastic process. In [Section 4](#), we will then introduce a state-of-the art model that best represents the relevant empirical features of our TFP measure. In the literature, our TFP measure has been called TFPS (for “TFP-shares”), as discussed by [Decker et al. \(2020\)](#).

As for sales, our baseline sample consists of firms operating in the manufacturing sector from US Compustat over the period between 1987 and 2017, but we discuss robustness to different sets of firms in [Appendix B](#). As done for sales in the previous section, we regress the log of firm productivity measures on industry dummies and a linear time trend and retain the residuals as our variable of interest.

We document that growth rates (i.e., one-year log changes) of both productivity measures display clear departures from non-Gaussianity, in the form of high concentration and heavy tails. Indeed, firms typically experience small or no changes in productivity, but they are occasionally hit by extreme, negative and positive, shocks. In fact, [Table 1](#) reports how these features are more markedly pronounced for productivity than for sales. Sales growth less than 5% in absolute value has a 22% probability of occurring in the data, more than twice as likely as if the data was normally distributed, with the same empirical standard deviation. Changes in productivity are even more concentrated: nearly half of TFP growth rates are less than 5% in absolute value, 5 times more likely than if the data was normally distributed. In [Appendix B.1](#) we show this concentration towards negligible growth rates is compensated by the presence of rare and heavy-tailed events.

We further explore the features of this non-Gaussianity by investigating moments of firm-level changes. We focus on TFP, but report similar results for labor productivity in [Appendix B.3](#). [Figure 4a](#) plots the standard deviation of firm-level one-year log-changes in sales and TFP, for 5 quintiles of past sales. Both measures decay with size; this behavior could be consistent with the fact that larger firms are more stable and established. Not only they face less volatile sales, but also less disperse fundamental shocks. The standard deviation of sales growth is slightly higher than for productivity, a feature consistent with standard theoretical frameworks of firm decisions.

These results may suggest that sales inherit most of the underlying features of the productivity process. However, we show that the picture is more complicated when we look at higher order conditional moments of productivity changes, which lie at the core of our results

Table 1: Concentration of sales, labor productivity and TFPS growth rates

| S : | [-0.05, 0.05] | | | [-0.10, 0.10] | | | [-0.20, 0.20] | | |
|------------------------------|---------------|--------------------------------|-------|---------------|--------------------------------|-------|---------------|--------------------------------|-------|
| | Data | $\mathcal{N}(\mu_x, \sigma_x)$ | Ratio | Data | $\mathcal{N}(\mu_x, \sigma_x)$ | Ratio | Data | $\mathcal{N}(\mu_x, \sigma_x)$ | Ratio |
| $Prob(\Delta Sales \in S)$ | 21.9 | 9.4 | 2.3 | 41.7 | 18.4 | 2.3 | 66.9 | 35.9 | 1.9 |
| $Prob(\Delta LP \in S)$ | 29.9 | 9.7 | 3.1 | 50.9 | 19.2 | 2.7 | 72.7 | 37.3 | 2.0 |
| $Prob(\Delta TFP \in S)$ | 47.2 | 10.5 | 4.5 | 69.3 | 20.8 | 3.3 | 85.3 | 39.9 | 2.1 |

Note: One year changes in logged variables, using baseline sample as described in the text. Columns for $\mathcal{N}(\mu_x, \sigma_x)$ refer to draws, of size equal to the dataset, from a normal distribution with mean μ_x and standard deviation σ_x , where x denotes the variable considered in each row. Standard deviations are 0.427, 0.413, and 0.384 for changes in log sales, LP, and TFP, respectively.

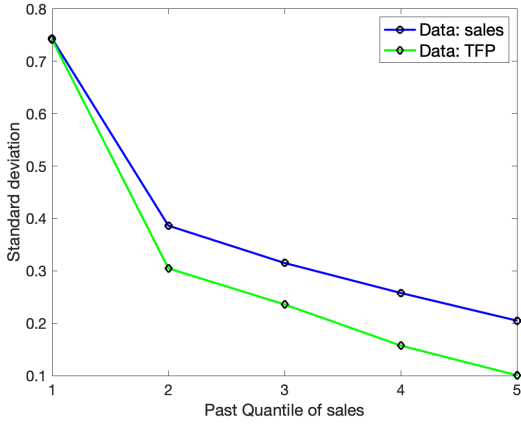
on non-Gaussianity and nonlinearities. Specifically, we focus on quantile-based measures of skewness and kurtosis as done for sales in Section 2. Figure 4b plots Kelley Skewness. Similar to what shown for sales, there is little departure from symmetry. However, conditional skewness of productivity growth falls more clearly with firm size, while the same is not true for sales. Most importantly, large firms display negative tail risk in productivity, but symmetric sales growth. This suggests that they are more likely to be hit by large negative productivity shocks, but are better able to insure against them, thus avoiding large sale losses.

We also find that the heavy tails in sales growth are most likely driven by heavy-tailed underlying exogenous shocks. Figure 4c displays Crow-Siddiqui kurtosis for our different measures, conditional on past quantiles of sales. Departures from gaussianity are even more pronounced for TFP than for sales. Firms at the bottom of the distribution are particularly likely to face very large shocks, of either sign.

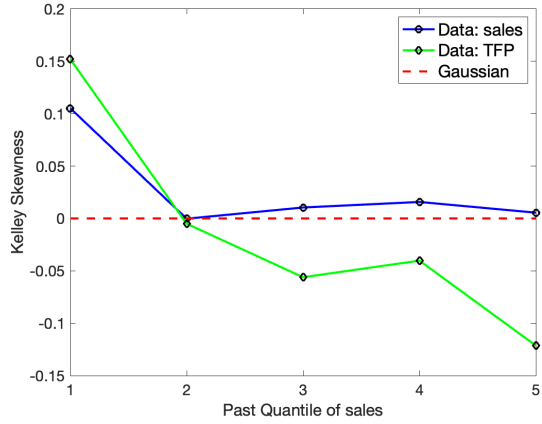
We perform several robustness checks to validate our findings and report them in Appendix B. As for sales, we confirm that non-Gaussianity is not entirely driven by a transitory component of productivity growth. Indeed, we show that there are consistent departures from non-Gaussianity when considering persistent changes in productivity, equivalent to $\Delta(\tilde{y})$ described in the previous section. Our results are also robust to enlarging the sample to the whole economy, besides manufacturing, and to excluding the smallest firms.

In summary, we conclude that firms face shocks that are fundamentally non-Gaussian, and therefore what documented for firm sales is not solely the result of endogenous responses to these shocks. In fact, firms behave in a way that reduces the degree of heavy tailedness and asymmetry of their sales, relative to the shocks that they face. While investigating this is beyond the scope of this paper, in Appendix C we estimate a simple reduced form setup that can shed some light on these findings. We find that firms respond to shocks

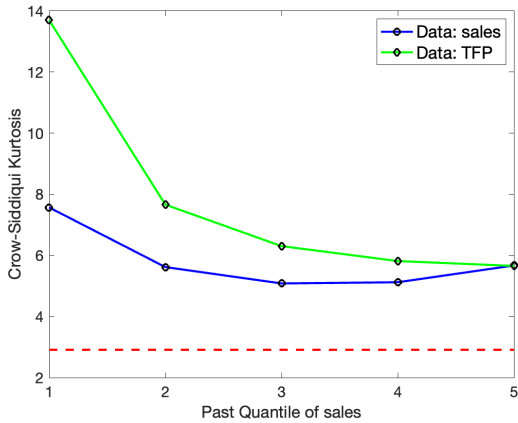
Figure 4: Conditional moments of firm-level one year log-changes



(a) Standard Deviation



(b) Kelley Skewness



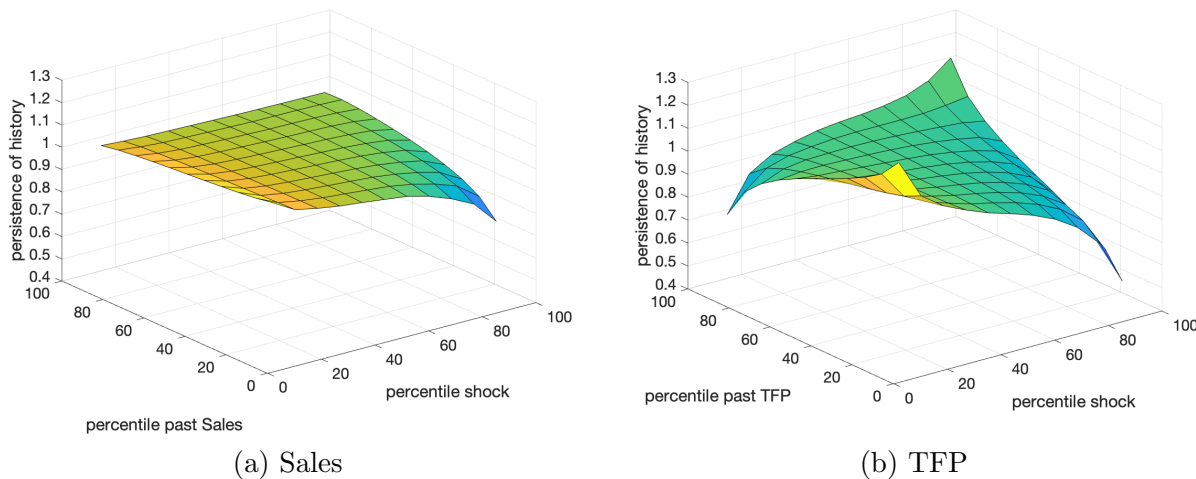
(c) Crow-Siddiqui Kurtosis

Notes: Solid-circled blue lines refer to the log of firm sales, and solid-diamond green lines to log TFP. For each measure, observations are ordered by quantiles of the sales distribution of $t - 1$ (x-axis). For each quantile, we show on the y-axis the Standard Deviation in Figure 4a, the Kelley Skewness in Figure 4b and Crow-Siddiqui Kurtosis in Figure 4c of log-changes.

increasingly less as the productivity change becomes of larger magnitude. This behavior is broadly consistent with various economic models of firm behavior, such as those featuring hiring costs and financial frictions.

Finally, we study the dynamics of productivity. We have previously shown how firm sales depart from a standard AR(1) process and display lower persistence of history among small firms being hit by large positive shocks. In other words, large positive shocks erase part of the history of previous bad realizations for firms with low sales. We confirm that this behavior seems to be inherent to fundamental shocks, rather than firm behavior, as shown in the right panel of Figure 5. Moreover, we highlight the appearance of a nonlinear left tail. Such behavior implies that very negative productivity shocks cancel the history of past

Figure 5: Quantile autoregressions



Notes: Estimated average derivative of the conditional quantile function of residualized values $y_{i,t}$ given $y_{i,t-1}$. Derivatives are taken with respect to $y_{i,t-1}$ and evaluated at a given percentile of $y_{i,t-1}$, and at a given percentile of the shock. Details on data and estimation of quantile functions in the text.

positive productivity shocks when hitting firms with highest productivity. An equivalent mechanism has been highlighted for household earnings by [Arellano et al. \(2017\)](#); we show it is also present among firms, albeit quantitatively more limited.¹³ Persistence of history is more heterogeneous for TFP than for sales, ranging between 0.5 and 1.3.

We have shown that firm-level productivity features non-Gaussian shocks and nonlinear persistence. What statistical process is consistent with our findings and how can we use an estimated model to learn about micro and macro fluctuations? This is what we investigate in the next section.

4 A nonlinear model of productivity

The empirical evidence presented so far suggests that the dynamics of firm productivity are characterized by a rich process. In this last part of the paper, we investigate further its drivers and implications. In this section, we estimate a model of productivity that can rationalize the empirically observed nonlinearities and non-Gaussian shocks. We use a state-of-the-art framework and techniques proposed by [Arellano et al. \(2017\)](#) (thereafter, ABB) and [Arellano et al. \(2021\)](#) (thereafter, ABBL) and leverage two main features. First, we assume that the stochastic process of productivity is a combination of a persistent and a transitory component, allowing us to disentangle the drivers of our empirical findings showed

¹³All the results presented in this section are qualitatively unchanged when considering all industries.

so far. Second, we introduce potential nonlinearities in the persistence of shocks using a quantile-based panel data framework.

Allowing for shocks of different degrees of persistence is not common in the study of firm dynamics, where an AR(1) process for the log of productivity is typically assumed.¹⁴ This stands in sharp contrast with the literature on household earnings dynamics, in which transitory and persistent shocks are typically separated, even in a standard framework without nonlinearities.¹⁵ Persistence of idiosyncratic productivity is, however, of first-order importance for firm dynamics too. Among others, [Moll \(2014\)](#) shows how the persistence of productivity shocks interacts with financial frictions and matters for the size and the speed of aggregate transitions. Moreover, transitory and persistent shocks are likely to affect firms' inputs differently, depending on the nature of their adjustment costs, and whether they face a static or forward-looking input demand (see seminal work by [Cooper and Haltiwanger \(2006\)](#) and a specific discussion by [Roys \(2016\)](#)). A further example of the importance of this distinction can be found in the corporate finance literature: [Décamps et al. \(2016\)](#) show how permanent and transitory cash-flow shocks are associated with very different corporate policies. For instance, their relative importance substantially alters the cash-flow sensitivity of cash.

Motivated by these considerations, we estimate the model proposed by ABB. Let $o_{i,t}$ denote the residualized log of productivity, net of a linear time trend and industry effects, of firm i at time t , where $t = 1, \dots, T$ and $i = 1, \dots, N$. We model $o_{i,t}$ as the sum of a general Markovian persistent component and a transitory innovation, as follows:

$$o_{it} = \eta_{it} + \epsilon_{it} \tag{2}$$

We allow the persistent component η_{it} to follow a first-order Markov process:

$$\eta_{it} = Q(\eta_{it-1}, u_{it}) \quad (u_{it} | \eta_{it-1}, \eta_{it-2}, \dots) \sim \text{Uniform}(0, 1), \quad t = 2, \dots, T \tag{3}$$

where $Q(\eta_{it-1}, u_{it})$ is the u_{it} -th conditional quantile of η_{it} given η_{it-1} . The transitory component, ϵ_{it} , has zero mean and is independent over time and of η_{is} for all $s, s = 1, \dots, T$.¹⁶

We take from ABB the argument for identification, which relies on the literature on

¹⁴[Guiso et al. \(2005\)](#) and [Roys \(2016\)](#) are some notable exceptions, although both of these works do not allow for any nonlinearities.

¹⁵In the earnings literature, the transitory-persistent time series structure can be dated back to [Friedman \(1957\)](#) and [MaCurdy \(1982\)](#).

¹⁶Note that the process can be extended to allow for a higher order Markov process for η , a moving average for ϵ , or to allow for unobserved time-invariant firm-specific effects. As noted in the literature, it is not possible to disentangle measurement error from the transitory innovation without additional information. We believe, however, that this should be less of a concern for our dataset – audited balance sheets – than for survey data commonly used in the earnings literature.

nonlinear models with latent variables (see [Hu and Schennach \(2008\)](#), [Wilhelm \(2015\)](#)). This quantile-based model allows us to capture nonlinear persistence and conditional heteroskedasticity. In particular, it highlights a role for the persistence of histories. Consider a shock u_{it} to η_{it-1} ; its impact on TFP depends on the shock occurring at time $t + 1$, u_{it+1} , whose impact on TFP will in turn depend on the shock occurring at time $t + 2$, and so on. More formally, nonlinear persistence is:

$$\rho(\eta_{it-1}, \tau) = \frac{\delta Q(\eta_{it-1}, \tau)}{\delta \eta} \quad (4)$$

which is a measure of the impact of η_{it-1} on η_{it} and, as a result, on o_{it} , when firm i is hit by a shock u_{it} that has rank τ .

Average nonlinear persistence across η is:

$$\rho(\tau) = E \left[\frac{\delta Q(\eta_{it-1}, \tau)}{\delta \eta} \right] \quad (5)$$

The ρ 's are measures of persistence of productivity histories. Here one can see how this flexible model can allow the persistence of the shocks to the persistent component, η_{it-1} , to depend on the size and sign of current and future shocks $u_{it}, u_{it+1} \dots$. Within this framework it is possible to analyze how unusual shocks to firm productivity can significantly reduce or increase the persistence of past shocks. For instance, micro-disasters that wipe out the memory of all past productivity shocks.¹⁷ We therefore study a framework in which unusual shocks can occur, which wipe out the productivity effect of a series of past shocks. Similarly, the persistence of current shocks is subject to the realization of future shocks.

We closely follow the estimation strategy proposed by ABBL. We specify the conditional quantile function describing the process for the persistent component η_{it} using Hermite polynomials. As a result, we write the expression in equation 3 as:

$$\eta_{it} = Q(\eta_{it-1}, \tau) = \sum_{\kappa=0}^K a_{\kappa}^Q(\tau) \varphi_{\kappa}(\eta_{i,t-1}) \quad (6)$$

where φ_{κ} denotes the basis of the Hermite polynomials and a_{κ}^Q the corresponding coefficients. In implementation, we follow ABBL and set $K = 3$. We define $\tau_l = l/(L + 1)$,

¹⁷Note that this nonlinear model nests the canonical (linear) model, typically used in the literature, as a special case. The canonical model is obtained by restricting η to follow a random walk, i.e. $\eta_{i,t} = \eta_{i,t-1} + \nu_{i,t}$. In the canonical model, all shocks are associated with the same persistence, irrespective of the history of a firm's productivity. A shock ν_{it} enters η_{it} linearly with persistence $\rho = 1$, so that the shock transmits entirely to o_{it} , regardless of a firm's past sequence of shocks to TFP.

with $L = 11$, and specify the functions a_{κ}^Q to be piecewise-linear in the interval $[\tau_1, \tau_L]$. In addition, we follow ABBL and treat a_0^Q as quantiles of an exponential distribution in the ranges $(0, \tau_1]$ and $[\tau_L, 1)$. This introduces two tail parameters, those defining the exponential distributions; we estimate each not only for $\eta_{i,t}$, but also for the transitory shock and for the persistent component in the first period, $\eta_{i,0}$. Computationally, estimation is implemented using the Sequential Monte Carlo methods proposed by ABBL. We defer to their paper for an extensive discussion of the computational details of the algorithm.¹⁸

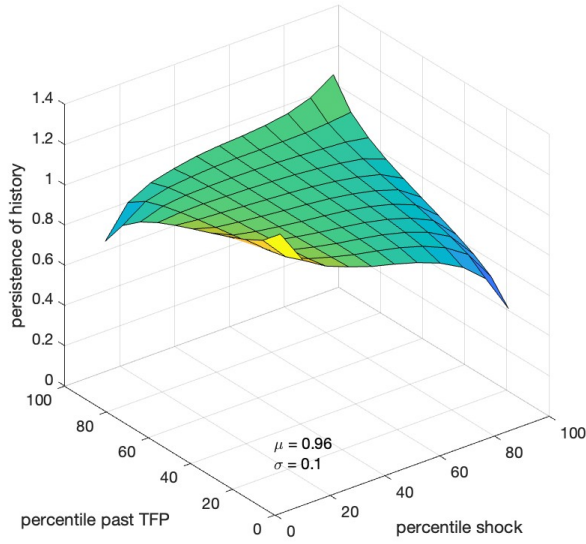
In the remainder of this section, we analyze the drivers of the empirically documented nonlinearities in productivity through the lens of the estimated model. Figure 6 displays estimates of the average derivative of the conditional quantile function of residualized log-TFP. The top-left panel 6a reproduces the empirical persistence for the data, also showed previously in Figure 5b. Persistence of productivity is far from homogeneous, as it depends both on the initial level of productivity and the size of the productivity shock, and it ranges between 0.5 and 1.3. The nonlinear model does a good job at replicating the empirical patterns, as shown in Figure 6b. In Figure 6c, we show the derivatives of the conditional quantile function for the persistent component of the estimated productivity process. The nonlinear patterns of persistence are similar to those estimated for the entire productivity process, although slightly less pronounced, suggesting a marginal role of transitory shocks for nonlinear dynamics. Our estimates also shed light on the behavior of specific firms. In Appendix E we show how large positive and negative shocks to the persistent component of productivity asymmetrically affected a large firm in the sample (Apple). This case study provides an example of how micro disasters can erase the persistence of history.

Another important feature of the nonlinear model is that it allows for non-Gaussianity of both components of the productivity process. As shown in Appendix E, we find that both shocks are non-Gaussian, even when sorting firms by their productivity.¹⁹ Indeed, even the most productive firms face persistent and transitory shocks that feature significant departures from Gaussianity. We further explore this finding in the next section, which investigates the aggregate importance of idiosyncratic productivity shocks.

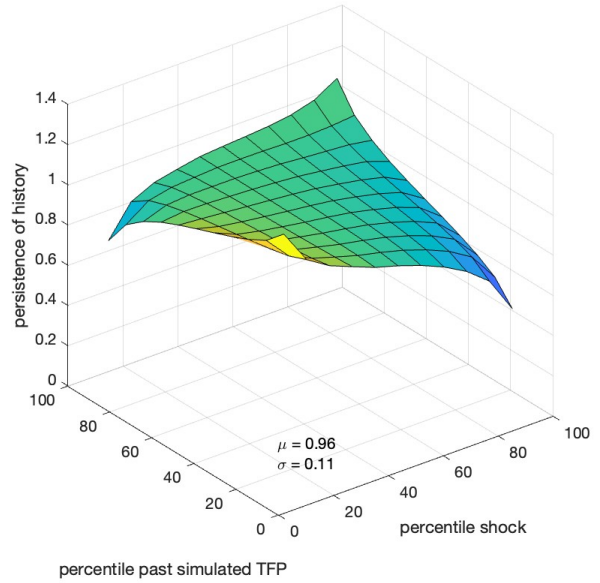
¹⁸The Laplace tails assumption is only imposed in the implementation and can be relaxed and replaced by a more heavy-type tails assumption. However, we do not pursue this idea here and follow instead the same implementation strategy as suggested by ABB and ABBL.

¹⁹Reassuringly, the model is able to match well the conditional moments of productivity growth presented in Section 3, as we show in Appendix E.

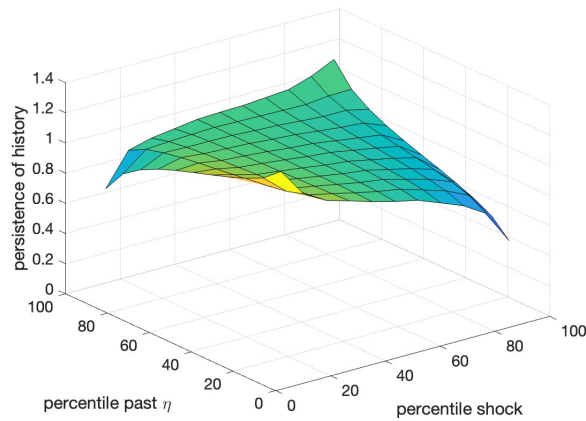
Figure 6: TFP persistence in the data and in the model



(a) Data



(b) Nonlinear model



(c) η_{it} , nonlinear model

Notes: Panels (a) and (b) plot the estimated average derivative of the conditional quantile function of $o_{i,t}$ given $o_{i,t-1}$. Derivatives are taken with respect to $o_{i,t-1}$ and evaluated at a given percentile of $o_{i,t-1}$, and at a given percentile of the shock. Panel (a) is based on the Compustat data, while panel (b) is based on data simulated according to the nonlinear model, with parameters set to their estimated values. Panel (c) shows the same object for the persistence component of the nonlinear model, η . μ and σ are the average and standard deviation of the persistence of history in each panel.

5 An analysis of micro-originated fluctuations

In this last section we revisit one reason why firm-level shocks may be very important for macroeconomic fluctuations: the granular hypothesis first proposed by [Gabaix \(2011\)](#).²⁰ In a nutshell, if the firm size distribution is heavy-tailed, idiosyncratic firm-level shocks do not average out, and shocks to large firms can explain a significant part of aggregate movements. For instance, when firm sizes are distributed according to the Zipf-law, the central limit theorem breaks down and diversification becomes of order $1/(\ln N)$, with N being the number of firms, rather than of $1/\sqrt{N}$ as in the case where firms sizes are normally distributed. As a result, shocks to big firms matter and transmit to the aggregate.

While the literature has typically focused on the firm size distribution, we investigate what role the features of the *shocks* play for the granular hypothesis. First, we highlight some key mechanisms in a stylized theoretical setting. Second, we discuss a quantitative exercise that showcases the role of non-Gaussianity and nonlinearities for the volatility and persistence of micro-originated aggregate fluctuations.

Our approach and interpretation of firm-level shocks are similar in spirit to [Salgado et al. \(2019\)](#), who argue that business cycles can originate from fluctuating skewness of idiosyncratic shocks. As in their framework, one might wonder whether the features of the distribution of the idiosyncratic shocks derive endogenously from macro shocks rather than being the sources of aggregate fluctuations. This decomposition is beyond the scope of our paper. Conceptually, we effectively consider a world without aggregate shocks, where aggregate fluctuations in productivity are solely generated by idiosyncratic shocks. Within this setting, we study the role that the features of the distribution of firm-level shocks and their dynamics play for aggregate fluctuations.

5.1 Non-Gaussian granular shocks

In this section, we investigate analytically the role played by the distribution of shocks for the granular hypothesis. We consider a stylized setting that builds on [Gabaix \(2011\)](#). We assume that there are N firms and each firm i produces a quantity S_{it} of consumption good at time t . Then the firm growth rate is:

$$\frac{\Delta S_{it+1}}{S_{it}} = \frac{S_{it+1} - S_{it}}{S_{it}} = U_{it+1} \tag{7}$$

²⁰There have been several recent empirical applications of this idea. Among others, see [Amiti and Weinstein \(2018\)](#) on granular bank supply shocks and aggregate investment.

where U_{it+1} are uncorrelated random variables with mean 0 and variance σ (i.e., all firms are assumed to have the same volatility). To see a parallel with the model introduced in Section 4, note that the difference $o_{it+1} - o_{it}$, where o_{it} is defined in Equation (2), approximates $\frac{\Delta S_{it+1}}{S_{it}}$ in Equation (7) for small $\frac{\Delta S_{it+1}}{S_{it}}$. Hence, U_{it+1} approximates innovations to the components $\eta_{it} + \epsilon_{it}$ in Equation (7).²¹

Aggregate GDP is $Y_t = \sum_{i=1}^N S_{it}$ and GDP growth is:

$$\frac{\Delta Y_{t+1}}{Y_t} = \sum_{i=1}^N U_{it+1} \frac{S_{it}}{Y_t}$$

Under the assumption that the shocks U_{it+1} are uncorrelated random variables, then the standard deviation of GDP growth is:

$$\sigma_{GDP} = \left(\sum_{i=1}^N \sigma^2 \left(\frac{S_{it}}{Y_t} \right)^2 \right)^{1/2} \quad (8)$$

Gabaix (2011) proposes the granular hypothesis by deriving the asymptotic behavior of σ_{GDP} as $N \rightarrow \infty$. He shows that, as $N \rightarrow \infty$, if the size of firms, S_{it} , has a distribution with finite variance, then the CLT applies, and the decay rate of the shocks is \sqrt{N} . Otherwise, if the distribution of firms sizes has an infinite variance and potentially even infinite mean, then the rate of decay is $N^{1-1/\zeta}$ or $\ln N$, respectively (a Pareto law distribution where the Pareto exponent ζ is ≤ 2 (≤ 1) has infinite variance (mean)). Hence, granular shocks matter for the aggregate when the distribution of sales has a power law behavior.

Is the granular hypothesis affected by non-Gaussianity in U_{it+1} and how? In order to answer this question, we analyze how the distribution of U_{it+1} may alter the distribution of S_{it+1} . From Equation (7), we can rewrite next-period sales as:

$$S_{it+1} = S_{it}(1 + U_{it+1}) \quad (9)$$

The tail behavior of next-period sales depends on the tail behavior of the two factors on the right hand side of Equation (9). Let us assume that $S_{i,t}$, which has positive support, and U_{it+1} are independent. To fix ideas, we consider different cases. If both S_{it} and U_{it+1} are Pareto distributed, then S_{it+1} will have a power tail behavior, i.e., S_{it+1} will have a Pareto-type tail, being regularly varying (RV) at ∞ in the upper tail.²² If, instead, $S_{i,t}$ has

²¹To make this parallel we have obviously assumed that S is a measure of productivity. For simplicity, consider $o_{it} = \epsilon_{it}$. Then, $U_{it+1} = \epsilon_{it+1} - \epsilon_{it}$. If ϵ_{it} and ϵ_{it+1} are uncorrelated and feature heavy tails, then U_{it+1} will also feature heavy tails by Remark 1.3.5 in **Mikosch (1999)**.

²²A random variable X is said to be regularly varying at ∞ with Pareto exponent $\alpha(X^+)$ if, for any $x > 0$

an exponential upper tail, then the power law-behavior of the tail of U_{it+1} is enough to make S_{it+1} have a power tail behavior. As such, non-Gaussian shocks can provide a reason for heavy-tailed size distributions. Finally, if $S_{i,t}$ is Pareto, then $S_{i,t+1}$ will have a power law behavior regardless of whether the distribution of U_{it+1} has an upper heavy tail or not. To establish these results, see, for instance, Proposition 1.3.9 in Mikosch (1999).

While non-Gaussianity of U_{it+1} maintains that S_{it+1} will have a power law behavior, it is not obvious that S_{it} will dominate the power law behavior of S_{it+1} . The heaviness of the tails and variance of U_{it+1} might alter the features of the distribution of S_{it+1} . Thus, the distribution of the error term might be relevant in shaping the distribution of S_{it+1} .

At this point, we may be tempted to conclude that non-Gaussianity is associated with larger granular aggregate fluctuations. In addition to the effect of the tail features of U_{it+1} on S_{it+1} , non-Gaussian shocks may increase the volatility of GDP, all other things being equal. As an extreme case, if the distribution of the error term is such that the variance does not exist (infinite σ), then also aggregate GDP growth volatility in Equation (8) will be infinite. However, larger kurtosis is not necessarily coupled with larger variance. Heavy tailed-ness of a distribution is not informative of its variance, and the latter is the crucial statistic in Equation (8). Aggregate volatility could also be lower when modelling errors as heavy- rather than thin-tailed.

This section illustrated how the features of the cross-sectional distribution of idiosyncratic shocks affect the granular hypothesis. Even in this stylized setting, non-Gaussian shocks play an ambiguous role for granular aggregate fluctuations. Several components have been assumed away but may be important: among others, the correlation between firms' size and the error term, its variance (Yeh (2019)), and the dynamics of the process of firms' productivity. Given the complexity of these layers, in the next section we use numerical simulations to quantitatively assess the role of non-Gaussianities and nonlinearities for aggregate volatility.

5.2 Quantitative exploration of granular shocks

In this section we use the estimates from our model – and other model alternatives – to quantitatively highlight how the different nature of firm-level shocks matters for micro-

as $t \rightarrow \infty$:

$$\frac{\text{Prob}(X > tx)}{\text{Prob}(X > t)} \rightarrow x^{-\alpha(X^+)}$$

originated aggregate fluctuations.²³ In particular, we construct the growth rate of aggregate TFP as $G_t = \log \frac{1}{N_t} \sum_i \exp\{o_{i,t}\} - \log \frac{1}{N_{t-1}} \sum_i \exp\{o_{i,t-1}\}$, where $o_{i,t}$ is the simulated log of TFP, for a firm i at time t , using model-estimated parameters. By dividing by the number of firms N , we abstract from fluctuations along the extensive margin, which are not relevant for our analysis. We consider five different models. First, a plain vanilla AR(1) such that $o_{i,t+1} = \rho o_{i,t} + \nu_{i,t+1}$, where innovations ν are normally distributed with variance σ_ν^2 . We estimate ρ and σ_ν^2 with a standard minimum distance estimation technique proposed by Chamberlain (1984).²⁴ We obtain the following estimates: $\hat{\rho} = 0.931$ and $\hat{\sigma}_\nu^2 = 0.076$. Second, we add Gaussian transitory shocks ε to the AR(1). This slightly increases the estimated persistence to 0.937, and lowers $\hat{\sigma}_\nu^2$ to 0.069, with an estimated variance of transitory innovations $\hat{\sigma}_\varepsilon^2 = 0.107$. Third, we estimate a model that only features a persistent component with nonlinear persistence. To do so, we estimate quantile autoregressions as done for Figure 5, and then use the estimated coefficients to simulate the sample of firms. Fourth, we estimate the flexible model of Section 4, but assume Gaussian transitory shocks. In this setup, the variance of transitory shocks is estimated to be much higher than for the AR(1), at 0.514. Finally, we consider our full model.

For each of these models, we simulate a sample of firms starting from the empirical distribution, and maintaining the structure of the data panel (i.e.: firm-specific spells). We provide details of estimation and simulation in Appendix D. For each model, we compute two objects of interest, which we report in Table 2. First, the volatility of G_t . This is the typical measure in most research on the granular hypothesis, as we have discussed in the previous section. Second, we also compute the one year autocorrelation of G_t . In recent work, Carvalho and Grassi (2019) show that if there is a finite number of firms, rather than a continuum, then the distribution of firms across productivity levels varies stochastically over time. As a result, aggregate productivity also follows a stochastic process, even when all shocks are idiosyncratic.

The standard AR(1) process generates non-zero volatility of aggregate TFP growth and positive persistence, due to the finite number of firms and their fat-tailed productivity distribution. However, the volatility is lower than in the data. Results for this model are shown

²³Strictly speaking, our TFP measures are not solely idiosyncratic shocks but also combine heterogeneous sensitivities to aggregate shocks, as previously discussed. Regardless of their origin, however, this simulation exercise treats these fluctuations as happening at the level of individual firms and then we aggregate them up. We show that if the stochastic process of these firm-level fluctuations features nonlinearities and non-Gaussianities, the implications for the resulting aggregate fluctuations are substantially different.

²⁴We choose the estimator that minimizes the distance between the theoretical autocovariances generated by the model and the empirical autocovariances of productivity. We follow the recommendation of Altonji and Segal (1996) of using the identity matrix to reduce the substantial small sample bias that can result from choosing the optimal weighting matrix suggested by Chamberlain (1984).

Table 2: Micro-originated fluctuations with different idiosyncratic processes

| | Data | (I) | (II) | (III) | (IV) | (V) |
|-----------------------------|-------|-------|--------|-------|--------|-------|
| $\sigma(G_t)*100$ | 0.073 | 0.013 | 0.053 | 0.012 | 0.218 | 0.016 |
| $\text{CORR}(G_t, G_{t-1})$ | 0.162 | 0.045 | -0.161 | 0.448 | -0.445 | 0.248 |

Note: We estimate each model on the baseline sample and simulate out of sample, maintaining the structure of the data panel. We simulate each model 100 times, computing the variance and the autocorrelation of aggregate TFP growth rate for each simulation, and then reporting medians across simulations. Details can be found in Appendix D.

in column (I). Adding a transitory, normally distributed, component, as done in column (II), raises the volatility substantially, but introduces negative autocorrelation as this component dominates the properties of the overall stochastic process.²⁵ A similar pattern is observed in column (IV), which embeds Gaussian transitory shocks in the fully flexible model: volatility overshoots the data counterpart and the autocorrelation becomes even more negative. In column (III), instead, we extend the standard AR(1) with nonlinear persistence. Volatility is little affected. Aggregate growth rates are very persistent in this setup, as witnessed by the autocorrelation. Note that nonlinearities do not necessarily imply higher persistence, especially for the growth rate of aggregate TFP. Indeed, heterogeneity and aggregation can, in principle, work in both directions.

Our full model, showed in column (V), delivers plausible aggregate moments. Aggregate volatility undershoots what observed in the data but is closest to all other models except (II). Differently from the latter, model (V) generates a positive autocorrelation that is close to what seen in the data. It should be stressed that these aggregate moments are not explicitly targeted; as such, we see the good performance of model (V) as an additional success of this modeling framework, previously overlooked in the literature.

By comparing different models, our analysis shows how the nature of firm-level shocks affects micro-originated aggregate fluctuations. Our findings underscore that *both* nonlinearities in persistence and non-Gaussian shocks should be accounted for. In addition, transitory and persistent components should be modeled separately. Although the dispersion and kurtosis of transitory shocks are substantially smaller than those of changes in the persistent component, the results of this section still suggest that it is crucial to account for non-gaussianities in both components. Failure to do so generates counterfactual behavior, not just at the firm level, but also in the aggregate.²⁶ We find that non-Gaussian shocks as

²⁵In Appendix F we show these results analytically, in a simplified setting.

²⁶The model of column (IV) tries to fit rare shocks by estimating an excessively large standard deviation of the Gaussian distribution. This ends up dwarfing the persistent component, even when the latter is allowed

typically observed in the data are associated with less volatile micro-originated aggregate fluctuations, relative to a Gaussian counterpart. Importantly, this dampening channel is separate from, but complementary to, the size-variance tradeoff recently put forward by Yeh (2019).²⁷ Moreover, nonlinear persistence, as estimated by our model, raises the persistence of micro-originated aggregate fluctuations.

6 Conclusion

We have documented that sales and productivity of U.S. publicly listed firms follow a stochastic process likely different from the canonical linear AR(1). Firms face heavy-tailed fundamental shocks and they partly absorb them through input adjustment. In addition, they face nonlinear and heterogeneous persistence. We estimate a quantile-based panel data framework borrowed from the literature on household earnings and show that it is a good representation of firm-level dynamics.

Nonlinear dynamics can affect firm behavior in multiple ways. We have focused on one: micro-originated (granular) aggregate fluctuations. While the literature on the granular hypothesis has typically focused on the firm size distribution, we have investigated the role played by the features of the shocks. Our analysis suggests that transitory and persistent productivity shocks need to be modeled separately, and nonlinearities and non-Gaussianities simultaneously accounted for; failure to do so generates counterfactual behavior, not just at the firm level, but also in the aggregate. We hope that the findings in this paper spur further research in the literature on semi-structural estimation approaches applied to firms.

to be non-gaussian and with nonlinear persistence.

²⁷Note that in our data the variance of TFP growth rates also falls with firm size, as shown previously. As such, this channel is incorporated in our simulations.

References

- Altonji, Joseph G and Lewis M Segal**, “Small-sample bias in GMM estimation of covariance structures,” *Journal of Business & Economic Statistics*, 1996, *14* (3), 353–366.
- Amiti, Mary and David E Weinstein**, “How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data,” *Journal of Political Economy*, 2018, *126* (2), 525–587.
- Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme**, “Earnings and consumption dynamics: a nonlinear panel data framework,” *Econometrica*, 2017, *85* (3), 693–734.
- , – , – , and **Jack Light**, “Heterogeneity of Consumption Responses to Income Shocks in the Presence of Nonlinear Persistence,” 2021.
- Bachmann, Rüdiger, Steffen Elstner, and Atanas Hristov**, “Surprise, surprise—Measuring firm-level investment innovations,” *Journal of Economic Dynamics and Control*, 2017, *83*, 107–148.
- Barro, Robert J**, “Rare disasters and asset markets in the twentieth century,” *The Quarterly Journal of Economics*, 2006, *121* (3), 823–866.
- Barrot, Jean-Noël and Julien Sauvagnat**, “Input specificity and the propagation of idiosyncratic shocks in production networks,” *The Quarterly Journal of Economics*, 2016, *131* (3), 1543–1592.
- Blackwood, G Jacob, Lucia S Foster, Cheryl A Grim, John Haltiwanger, and Zoltan Wolf**, “Macro and Micro Dynamics of Productivity: From Devilish Details to Insights,” *American Economic Journal: Macroeconomics*, 2021, *13* (3), 142–72.
- Bloom, Nicholas**, “The impact of uncertainty shocks,” *Econometrica*, 2009, *77* (3), 623–685.
- , **Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J Terry**, “Really uncertain business cycles,” *Econometrica*, 2018, *86* (3), 1031–1065.
- Boehm, Christoph E, Aaron Flaaen, and Nitya Pandalai-Nayar**, “Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tōhoku earthquake,” *Review of Economics and Statistics*, 2019, *101* (1), 60–75.

- Carroll, Christopher D and Miles S Kimball**, “On the concavity of the consumption function,” *Econometrica: Journal of the Econometric Society*, 1996, pp. 981–992.
- Carvalho, Vasco M and Basile Grassi**, “Large firm dynamics and the business cycle,” *American Economic Review*, 2019, *109* (4), 1375–1425.
- Chamberlain, Gary**, “Panel data’, in (Z. Griliches and M. Intriligator, eds.) *Handbook of Econometrics*,” 1984.
- Cooper, Russell W and John C Haltiwanger**, “On the nature of capital adjustment costs,” *The Review of Economic Studies*, 2006, *73* (3), 611–633.
- Décamps, Jean-Paul, Sebastian Gryglewicz, Erwan Morellec, and Stephane Villeneuve**, “Corporate policies with permanent and transitory shocks,” *The Review of Financial Studies*, 2016, p. hhw078.
- Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda**, “Where has all the skewness gone? The decline in high-growth (young) firms in the US,” *European Economic Review*, 2016, *86*, 4–23.
- , – , – , and – , “Changing business dynamism and productivity: Shocks versus responsiveness,” *American Economic Review*, 2020, *110* (12), 3952–90.
- DeVera, Micole**, “Firm-Level Productivity and Demand Shocks in Imperfectly Competitive Labor Markets: Implications for Wage Dynamics,” 2023.
- Fella, Giulio, Julio Galvez, Beatriz González, Juan Carlos Ruiz-Garcia, and Tatsuro Senga**, “Non-linear productivity and investment dynamics,” Technical Report, Working Paper.[Cited 5 February 2022]. Available from: [https://giuliofella ...](https://giuliofella...) 2021.
- Friedman, Milton**, “Introduction to” A Theory of the Consumption Function”, in “A theory of the consumption function,” Princeton university press, 1957, pp. 1–6.
- Gabaix, Xavier**, “The granular origins of aggregate fluctuations,” *Econometrica*, 2011, *79* (3), 733–772.
- Giovanni, Julian Di, Andrei A Levchenko, and Isabelle Mejean**, “Firms, destinations, and aggregate fluctuations,” *Econometrica*, 2014, *82* (4), 1303–1340.
- Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi**, “Insurance within the firm,” *Journal of Political Economy*, 2005, *113* (5), 1054–1087.

- Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song**, “What Do Data on Millions of US Workers Reveal About Lifecycle Earnings Dynamics?,” *Econometrica*, 2021, *89* (5), 2303–2339.
- Higson, Chris, Sean Holly, and Paul Kattuman**, “The cross-sectional dynamics of the US business cycle: 1950–1999,” *Journal of Economic Dynamics and Control*, 2002, *26* (9-10), 1539–1555.
- Hu, Yingyao and Susanne M Schennach**, “Instrumental variable treatment of nonclassical measurement error models,” *Econometrica*, 2008, *76* (1), 195–216.
- Jaimovich, Nir, Stephen J Terry, and Nicolas Vincent**, “The Empirical Distribution of Firm Dynamics and Its Macro Implications,” Technical Report, National Bureau of Economic Research 2023.
- Keller, Wolfgang and Stephen R Yeaple**, “Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States,” *The Review of Economics and Statistics*, 2009, *91* (4), 821–831.
- Kelley, Truman Lee**, *Fundamentals of statistics*, Harvard University Press, 1947.
- Khan, Aubhik and Julia K Thomas**, “Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics,” *Econometrica*, 2008, *76* (2), 395–436.
- Loecker, Jan De and Pinelopi Koujianou Goldberg**, “Firm performance in a global market,” *Annu. Rev. Econ.*, 2014, *6* (1), 201–227.
- , **Jan Eeckhout, and Gabriel Unger**, “The rise of market power and the macroeconomic implications,” *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.
- MaCurdy, Thomas E**, “The use of time series processes to model the error structure of earnings in a longitudinal data analysis,” *Journal of econometrics*, 1982, *18* (1), 83–114.
- Mikosch, Thomas**, “Regular variation, subexponentiality and their applications in probability theory,” 1999.
- Moll, Benjamin**, “Productivity losses from financial frictions: Can self-financing undo capital misallocation?,” *American Economic Review*, 2014, *104* (10), 3186–3221.
- Ottonello, Pablo and Thomas Winberry**, “Financial heterogeneity and the investment channel of monetary policy,” *Econometrica*, 2020, *88* (6), 2473–2502.

- Roys, Nicolas**, “Persistence of shocks and the reallocation of labor,” *Review of Economic Dynamics*, 2016, *22*, 109–130.
- Salgado, Sergio, Fatih Guvenen, and Nicholas Bloom**, “Skewed business cycles,” Technical Report, National Bureau of Economic Research 2019.
- Sarpietro, Silvia, Yuya Sasaki, and Yulong Wang**, “Non-Existent Moments of Earnings Growth,” *arXiv preprint arXiv:2203.08014*, 2022.
- Stanley, Michael HR, Luis AN Amaral, Sergey V Buldyrev, Shlomo Havlin, Heiko Leschhorn, Philipp Maass, Michael A Salinger, and H Eugene Stanley**, “Scaling behaviour in the growth of companies,” *Nature*, 1996, *379* (6568), 804–806.
- Sterk, Vincent, Petr Sedláček, and Benjamin Pugsley**, “The nature of firm growth,” *American Economic Review*, 2021, *111* (2), 547–579.
- Syverson, Chad**, “What determines productivity?,” *Journal of Economic literature*, 2011, *49* (2), 326–65.
- Wilhelm, Daniel**, “Identification and estimation of nonparametric panel data regressions with measurement error,” Technical Report, cemmap working paper 2015.
- Yeh, Chen**, “Revisiting the origins of business cycles with the size-variance relationship,” *Federal Reserve Bank of Richmond December*, 2019.

Supplemental Appendix to “Nonlinear Firm Dynamics”

by DAVIDE MELCANGI[†] and SILVIA SARPIETRO[‡]

A Data Appendix

Our data consists of yearly firm-level information from US Compustat, over the period between 1987 and 2017. We exclude firms that are not incorporated in the US. Moreover, we keep firms with positive assets, sales, and cost of goods sold. We exclude firms without an industry classification (“unclassifiable firms”), as well firms whose Standard Industrial Classification (SIC) is financial and real estate activities (6021 – 6799) or utilities (4900 – 4991). This does not affect our baseline results which use only the manufacturing sector (i.e., NAICS code starting with 3), but only additional results for the whole economy. Finally, we keep firms that remain in operation for at least 2 consecutive years.¹ This leaves us with an unbalanced panel of 4360 firms and 51803 firm-year observations. We deflate sales, cost of goods sold, and the capital stock by the GDP deflator. The capital stock is constructed following [Ottonello and Winberry \(2020\)](#). For each firm, the first value of end of period capital is set to the level of gross plant, property, and equipment (ppegst in Compustat). We then compute the evolution of capital, from this period onwards, using changes of net plant, property, and equipment (ppent). If a firm has a missing observation for ppent, we impute its value with linear interpolation only if the missing observation has nonmissing values of ppent in the previous and following year. Capital stock is the lag of the end-of-period capital constructed as just described.

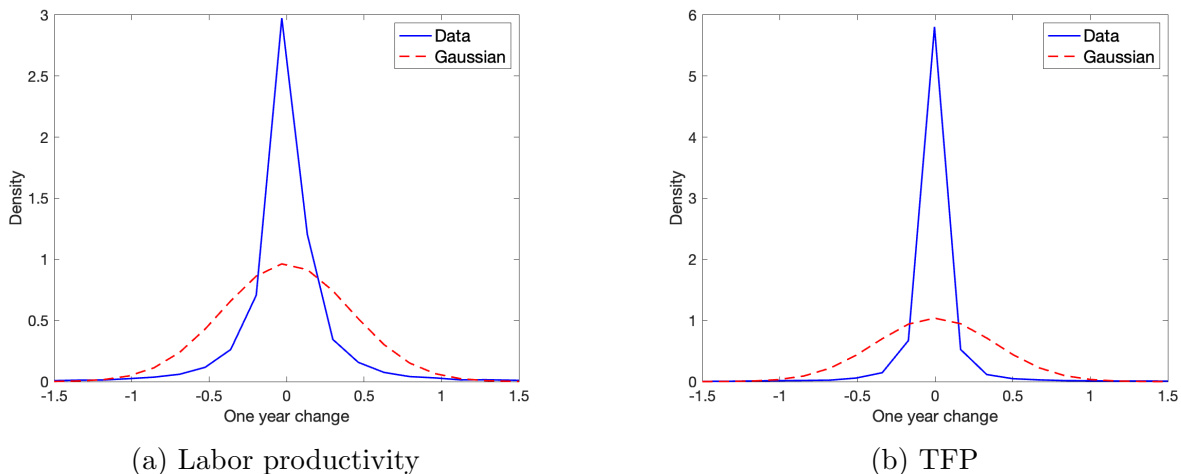
Revenue elasticities used in the calculation of TFPS are constructed using industry-level data from the KLEMS database constructed by the BLS. The revenue elasticity for capital is the ratio of capital cost and value of production, while for the variable input (cost of goods sold) we divide the sum of costs of materials, labor, and energy, by the value of production. Revenue elasticities are year-specific and industry-specific – 3-digit NAICS level for the manufacturing sector, 2-digit when looking at the whole economy. The first order

[†]Federal Reserve Bank of New York. Email: davide.melcangi@ny.frb.org. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

[‡]Bologna University. Email: silvia.sarpietro@unibo.it.

¹In unreported results we find that our results are little sensitive to the exclusion of firms with sales or total assets less \$100,000, or with less than 10 employees.

Figure A.1: Distribution of one-year log changes



Notes: Empirical densities of log changes in labor productivity (left panel (a)) and TFP (right panel (b)) between year t and $t - 1$ depicted in solid blue. In panel (a), the standard deviation is 0.413, Kelley Skewness 0.031 and Crow-Siddiqui Kurtosis 7.293. In panel (b), the standard deviation is 0.384, Kelley Skewness 0.017 and Crow-Siddiqui Kurtosis 9.236. Gaussian density with the same standard deviation of data shown in dashed red. Baseline sample of manufacturing firms.

condition pinning down revenue elasticities holds if there are no factor adjustment costs or wedges; as is typically assumed in the literature (see for instance [Decker, Haltiwanger, Jarmin and Miranda \(2020\)](#)), we require this condition to hold on average across firms within an industry. As suggested by [Syverson \(2011\)](#), we further strengthen this argument by averaging the revenue elasticities over 6-year windows.

After constructing the log of TFPS as the revenue function residual – using sales, cost of goods sold, the capital stock, and the revenue elasticities – we residualize it by regressing TFPS on industry dummies and a linear time trend.

B Additional empirical results

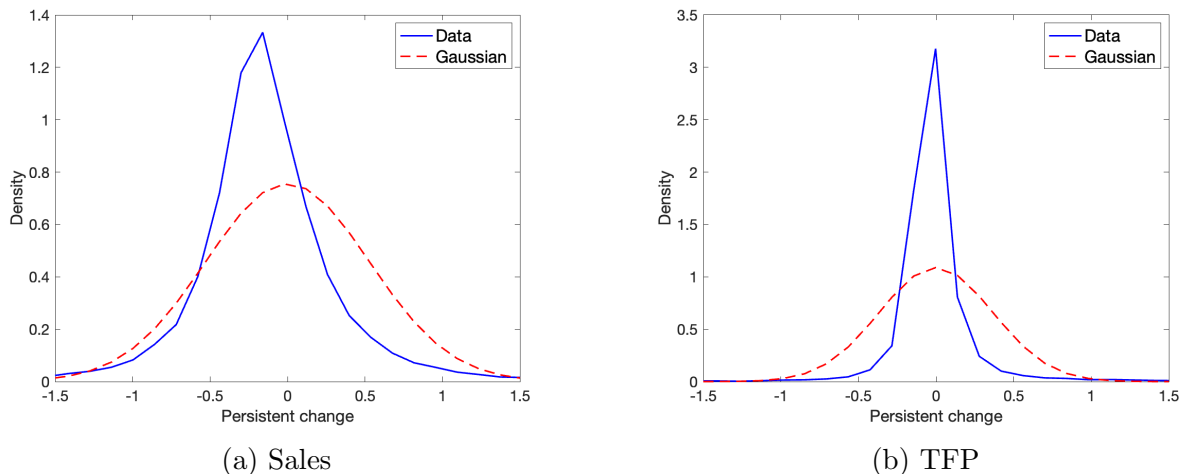
B.1 Productivity distributions

Figure [A.1](#) shows the empirical densities of one-year log changes in labor productivity and TFP. We show how both measures are clearly not distributed in a gaussian way.

B.2 Persistent changes

In this section we report the results for persistent changes, as described in [Section 2](#), In particular, we consider the difference between two 3-year averages of our logged variables.

Figure A.2: Distribution of persistent changes



Notes: Empirical densities of persistent changes ($\Delta(\tilde{y}) = \bar{y}_{t-1,t+1} - \bar{y}_{t-2,t-4}$) in sales (left panel (a)) and TFP (right panel (b)) depicted in solid blue. In panel (a), the standard deviation is 0.528, Kelley Skewness 0.147 and Crow-Siddiqui Kurtosis 4.856. In panel (b), the standard deviation is 0.367, Kelley Skewness 0.223 and Crow-Siddiqui Kurtosis 7.284. Gaussian density with the same standard deviation of data shown in dashed red. Baseline sample of manufacturing firms.

First, we show the distributions. In Figure A.2a we show how the empirical density of these persistent changes in sales is more concentrated than its gaussian counterpart, constructed with the same standard deviation. Compared to the distribution of one-year log changes, shown in Figure 1, the departures from non-Gaussianity are less pronounced, suggesting that transitory shocks are also non-Gaussian.

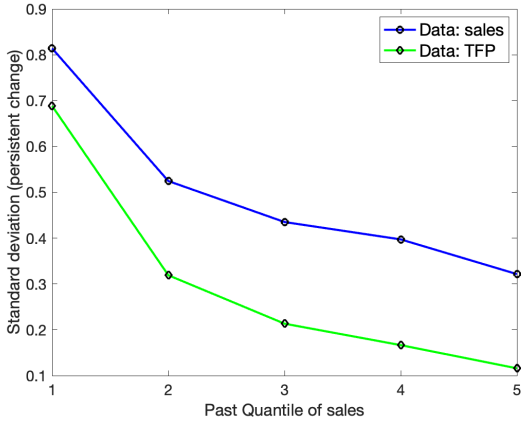
A similar story applies to TFP, as shown in Figure A.2b. Figure A.3 confirms that the main source of non-Gaussianity is heavy tailedness (kurtosis). In addition, small firms display a much larger Crow-Siddiqui kurtosis of persistent changes in TFP, compared to persistent changes in sales, in line with what shown for one-year changes.

B.3 Labor productivity

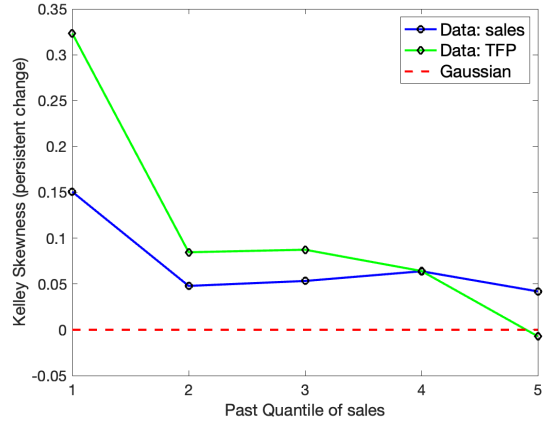
In this section we report higher order moments of log changes in labor productivity, as well as quantile autoregressions, for the baseline sample discussed in Section 2 and 3. In Figure A.4 we show that the dispersion of changes in labor productivity falls as firms sales increase. There is little asymmetry of labor productivity changes, except for a mild positive skewness among smallest firms. Finally, labor productivity shocks are very heavy-tailed, with marked departures from gaussianity.

In Figure A.5 we show that quantile autoregressions of log labor productivity suggest that also this alternative measure of firm profitability has a markedly nonlinear persistence

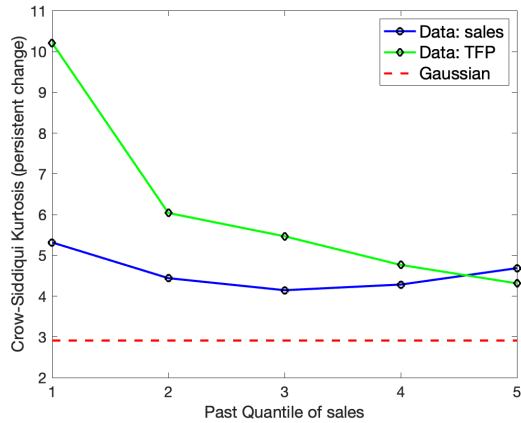
Figure A.3: Conditional moments of firm-level persistent changes



(a) Standard Deviation



(b) Kelley Skewness

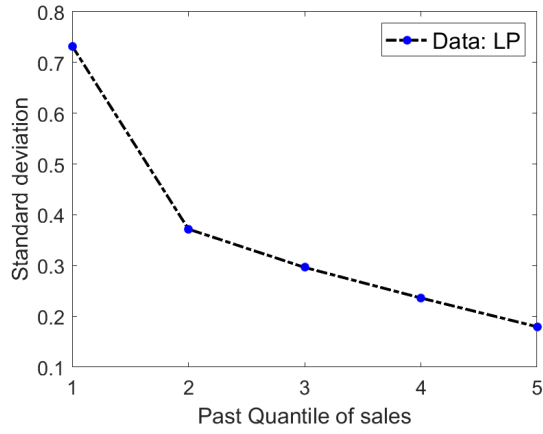


(c) Crow-Siddiqui Kurtosis

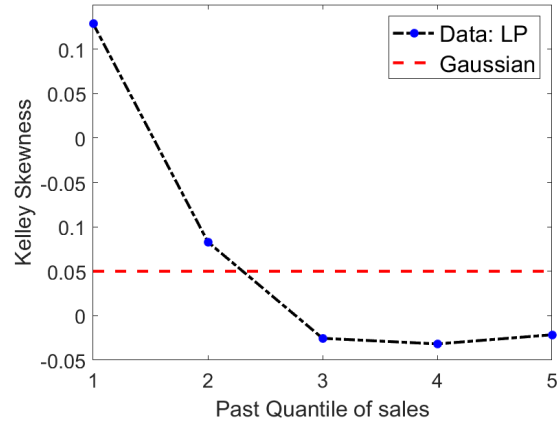
Notes: Solid-circled blue lines refer to the log of firm sales, while solid-diamond green line to log TFP. For each measure, observations ordered by quantiles of the distribution of $\bar{y}_{t-2,t-4}$ as defined in Figure A.2 (x-axis), where y is sales. For each quantile, we show on the y-axis the Standard deviation of persistent changes ($\Delta(\bar{y}) = \bar{y}_{t-1,t+1} - \bar{y}_{t-2,t-4}$) in Figure A.3a, the Kelley Skewness in Figure A.3b and Crow-Siddiqui Kurtosis in Figure A.3c.

of history.

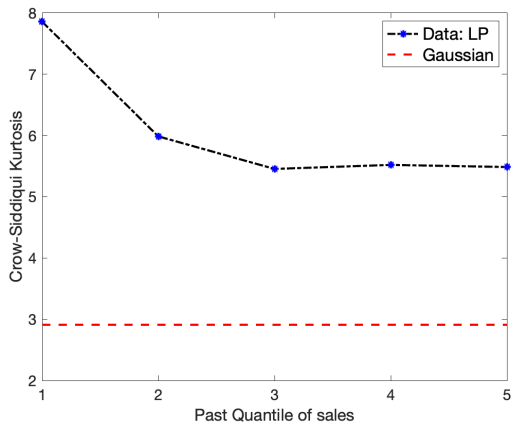
Figure A.4: Conditional moments of firm-level one year log-changes: labor productivity



(a) Standard Deviation



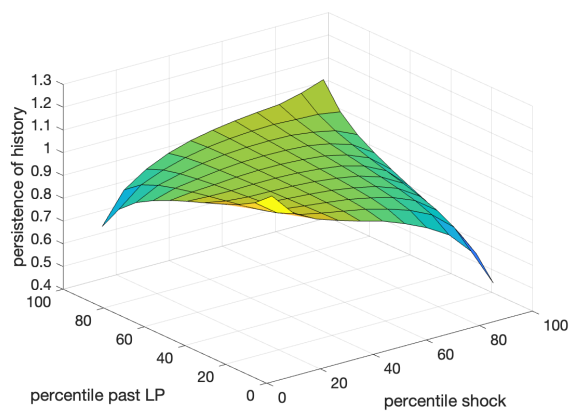
(b) Kelley Skewness



(c) Crow-Siddiqui Kurtosis

Notes: Observations are ordered by quantiles of the sales distribution in $t - 1$ (x-axis). For each quantile, we show on the y-axis the Standard Deviation in Figure A.4a, the Kelley Skewness in Figure A.4b and Crow-Siddiqui Kurtosis in Figure A.4c of log-changes in labor productivity.

Figure A.5: Quantile autoregressions of log labor productivity



Notes: Estimated average derivative of the conditional quantile function of residualized log labor productivity $y_{i,t}$ given $y_{i,t-1}$. Derivatives are taken with respect to $y_{i,t-1}$ and evaluated at a given percentile of $y_{i,t-1}$, and at a given percentile of the shock, labelled τ_{shock} . Details on data and estimation of quantile functions in the text.

Table A.1: Concentration of sales, labor productivity and TFPS growth rates: all industries

| S : | [-0.05, 0.05] | | | [-0.10, 0.10] | | | [-0.20, 0.20] | | |
|------------------------------|---------------|--------------------------------|-------|---------------|--------------------------------|-------|---------------|--------------------------------|-------|
| | Data | $\mathcal{N}(\mu_x, \sigma_x)$ | Ratio | Data | $\mathcal{N}(\mu_x, \sigma_x)$ | Ratio | Data | $\mathcal{N}(\mu_x, \sigma_x)$ | Ratio |
| $Prob(\Delta Sales \in S)$ | 21.6 | 9.3 | 2.3 | 41.9 | 18.5 | 2.3 | 67.5 | 35.7 | 1.9 |
| $Prob(\Delta LP \in S)$ | 30.9 | 9.7 | 3.2 | 51.6 | 19.4 | 2.7 | 72.7 | 37.7 | 1.9 |
| $Prob(\Delta TFP \in S)$ | 44.3 | 11.0 | 4.0 | 65.9 | 21.8 | 3.0 | 83.2 | 42.2 | 2.0 |

Note: One year changes in logged variables, sample of firms using all industries. Columns for $\mathcal{N}(\mu_x, \sigma_x)$ refer to draws, of size equal to the dataset, from a normal distribution with mean μ_x and standard deviation σ_x , where x denotes the variable considered in each row. Standard deviations are 0.430, 0.411, and 0.357 for changes in log sales, LP, and TFP, respectively.

B.4 All industries

In this section we repeat the analysis of Section 3 for the whole economy. Table A.1 confirms that the distributions of sales, labor productivity, and TFP growth rates are much more concentrated than a Normal with the same empirical variance. In addition, concentration is typically highest for TFP, which is arguably the more exogenous measure of fundamental shocks.

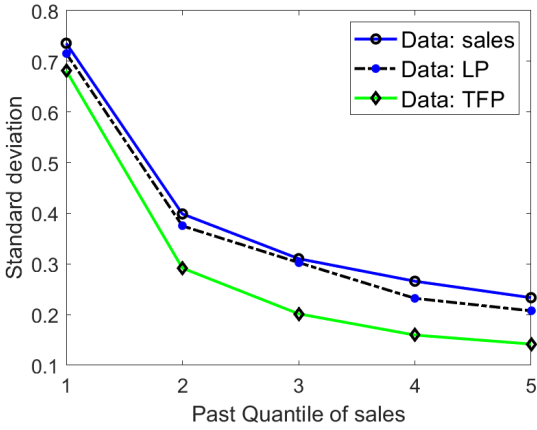
We then turn, in Figure A.6, to conditional moments of those distributions. Our results are similar to what shown for the manufacturing sector in Figure 4. Smaller firms face more dispersion in TFP growth rates, and even more so in sales growth rates. Small firms face shocks that are slightly positively skewed; in contrast, large firms display negative skewness in TFP log changes, but this does not translate into similar asymmetries for sales. We confirm that heavy tails in sales shocks are most likely driven by heavy-tailed underlying exogenous shocks. This is particularly true for firms at the bottom of the sales distribution.

Finally, we compute quantile autoregressions for this larger sample and confirm our main results, as shown in Figure A.7. Firm sales depart from a standard AR(1) process and display lower persistence of history among small firms hit by large positive shocks. This nonlinear persistence is even more pronounced when looking at productivity, and a nonlinear left tail appears. Quantitatively, nonlinearities are almost identical in this broader sample to what shown for the manufacturing sector.

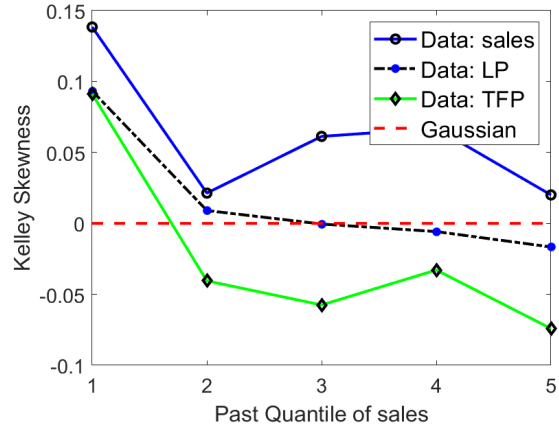
B.5 Confidence intervals

We bootstrap the baseline empirical sample, at the firm level, with replacement, and compute for each bootstrapped sample the quantile autoregressions, conditional Kelley skewness and conditional Crow-Siddiqui Kurtosis. The results showed in Figures A.8 and A.9 point to

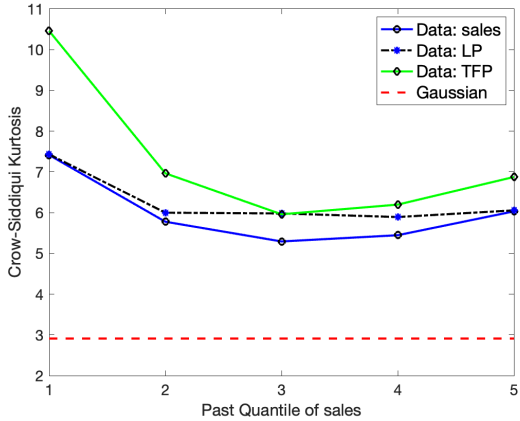
Figure A.6: Conditional moments of firm-level one year log-changes: all industries



(a) Standard Deviation



(b) Kelley Skewness



(c) Crow-Siddiqui Kurtosis

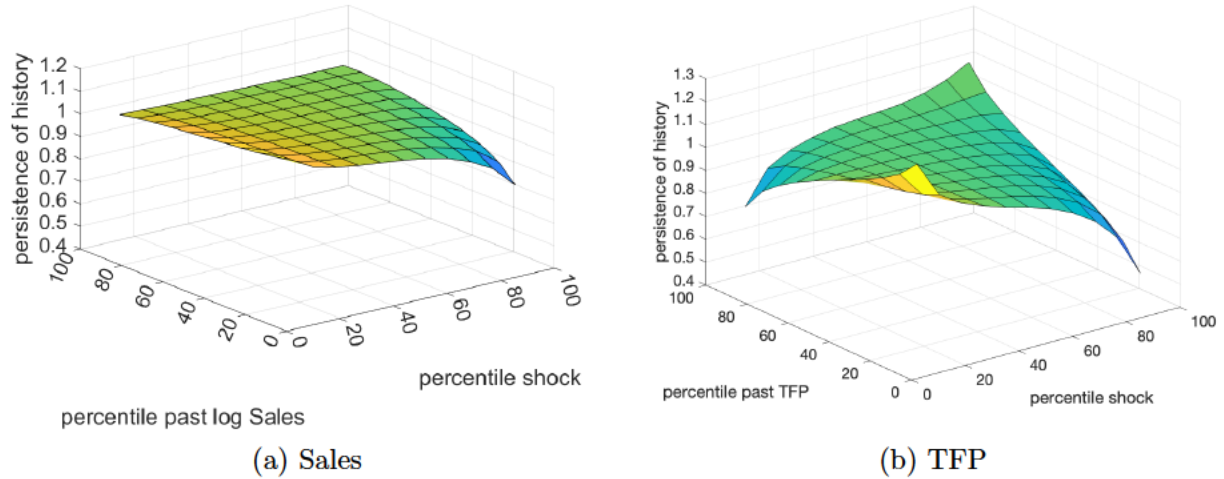
Notes: Solid-circled blue lines refer to the log of firm sales, dash-dotted black line to the log of labor productivity, solid-diamond green line to log TFP. For each measure, observations are ordered by quantiles of the sales distribution in $t - 1$ (x-axis). For each quantile, we show on the y-axis the Standard deviation in Figure A.6a Kelley Skewness in Figure A.6b and Crow-Siddiqui Kurtosis in Figure A.6c of log-changes. Firm-level data on all industries in the sample.

tight 95% pointwise confidence bands.

B.6 No Detrending

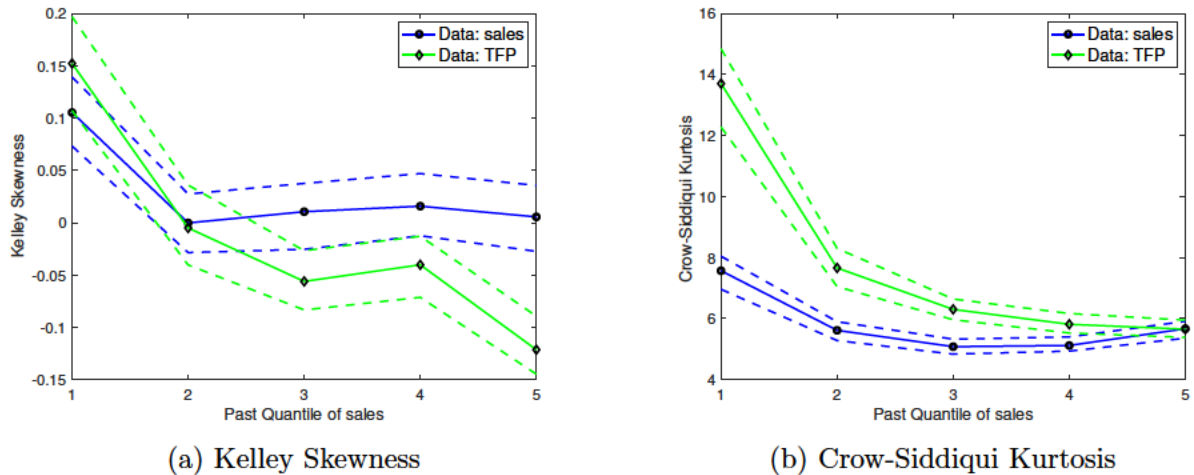
In section we repeat the analysis of Section 3 for data that has not been detrended – only industry fixed effects are subtracted from the variable of interest. As shown in Figure A.10 and A.11, the patterns are virtually identical to our baseline results.

Figure A.7: Quantile autoregressions: all industries



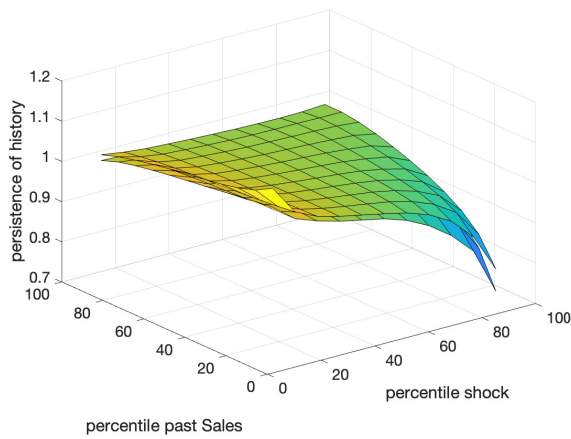
Notes: Estimated average derivative of the conditional quantile function of residualized values $y_{i,t}$ given $y_{i,t-1}$. Derivatives are taken with respect to $y_{i,t-1}$ and evaluated at a given percentile of $y_{i,t-1}$, and at a given percentile of the shock. Details on data and estimation of quantile functions in the text.

Figure A.8: Conditional moments of firm-level one year log-changes, 95% pointwise confidence bands

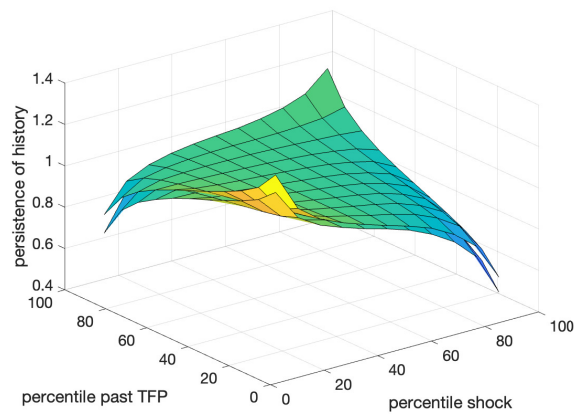


Notes: See notes to Figure 4. Pointwise 95% confidence bands computed over 500 bootstrap replications are showed with dashed lines. Nonparametric bootstrap with replacement, clustered at the firm level. Solid lines report results for the empirical sample, as in Figure 4.

Figure A.9: Quantile autoregressions, 95% pointwise confidence bands



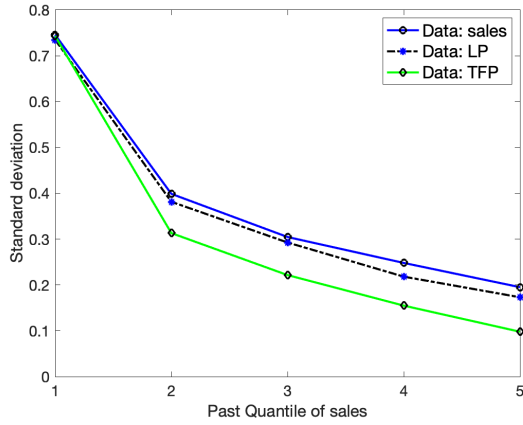
(a) Sales



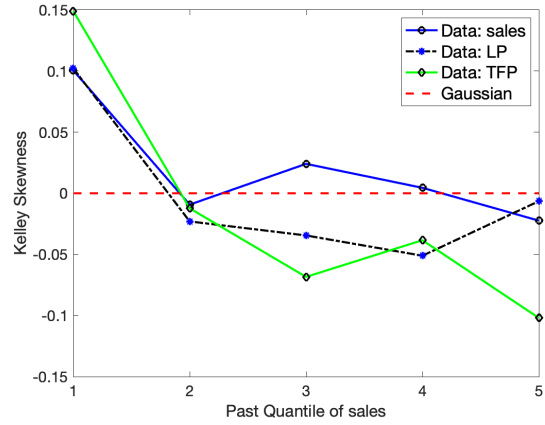
(b) TFP

Notes: See notes to Figure 5. Pointwise 95% confidence bands computed over 500 bootstrap replications. Nonparametric bootstrap with replacement, clustered at the firm level.

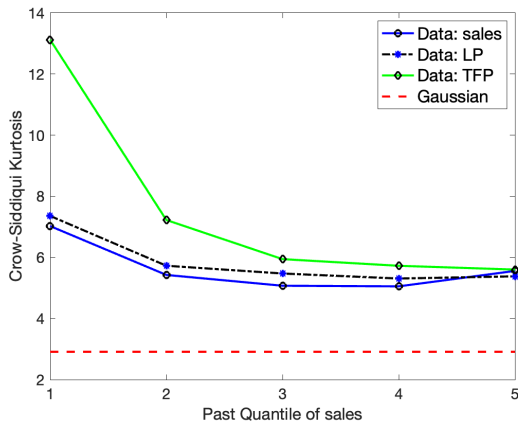
Figure A.10: Conditional moments of firm-level one year log-changes: no detrending



(a) Standard Deviation



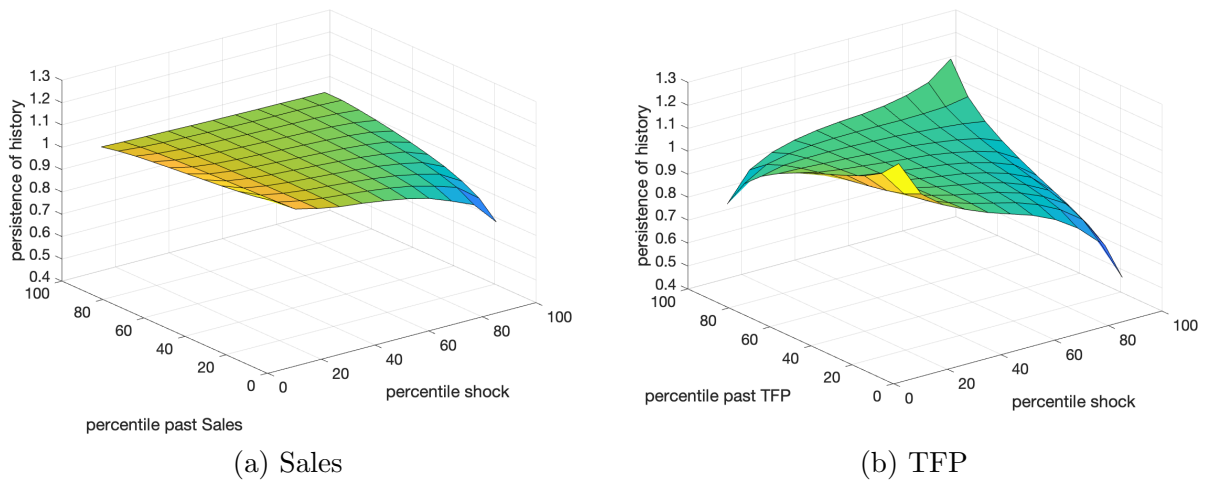
(b) Kelley Skewness



(c) Crow-Siddiqui Kurtosis

Notes: Solid-circled blue lines refer to the log of firm sales, dash-dotted black line to the log of labor productivity, solid-diamond green line to log TFP. For each measure, observations are ordered by quantiles of the sales distribution in $t - 1$ (x-axis). For each quantile, we show on the y-axis the Standard deviation in Figure A.6a Kelley Skewness in Figure A.6b and Crow-Siddiqui Kurtosis in Figure A.6c of log-changes.

Figure A.11: Quantile autoregressions: no detrending



Notes: Estimated average derivative of the conditional quantile function of residualized values $y_{i,t}$ given $y_{i,t-1}$. Derivatives are taken with respect to $y_{i,t-1}$ and evaluated at a given percentile of $y_{i,t-1}$, and at a given percentile of the shock. Details on data and estimation of quantile functions in the text.

C Decaying responsiveness to productivity shocks

What type of firm responsiveness is consistent with the findings shown in Section 3? To shed light on this, we come back to our simple production function in logs presented in Equation (1). Then we take first differences, and assume that changes in the two inputs of production respond to changes in TFP, such that $\Delta c_{i,t} = \phi_c \Delta z_{i,t}$ and $\Delta k_{i,t} = \phi_k \Delta z_{i,t}$. Then:

$$\Delta y_{i,t} = \Delta z_{i,t} + \beta_c \phi_c \Delta z_{i,t} + \beta_k \phi_k \Delta z_{i,t} \quad (10)$$

where y and z denote the log of sales and TFPS, respectively. For simplicity, we set $\phi_k = 0$. We take $\Delta z_{i,t}$ from the data and β_c as the average estimate from our baseline sample. With this we can back out $\Delta y_{i,t}$ from Equation (10) for three examples for ϕ_c and report the relevant moments, with their empirical counterparts, in Table A.2.²

In the first model, ϕ_c is the same for all firms and equal to 1, regardless of the sign and the size of Δz . This framework is consistent with the fact that the dispersion of sales growth is higher than for productivity growth, as long as $\beta_c \phi_c > 0$. With constant responsiveness, however, the Crow-Siddiqui measure of TFP growth is exactly the same as the one for sales growth. This is at odds with our empirical findings, not just in the entire sample, but also conditional on each quintile of sales.

In order to qualitatively match the fact that productivity growth displays heavy-tailed events more often than sales growth, we consider two alternative structures for ϕ_c . We assume a flexible dependence of this sensitivity to productivity shocks, such that $\phi_c = \gamma_1 e^{\gamma_2 |\Delta z|}$. Since the conditional skewness of sales changes is close to zero for nearly all quantiles of firm sales, we maintain a symmetric structure of ϕ_c , but our reasoning can easily be extended to loadings γ_2 that differ by the sign of productivity shocks. When $\gamma_2 < 0$, firms adjust their variable input c increasingly less as the productivity change becomes of larger magnitude; we label this case decaying response. As shown in Table A.2, choosing $\gamma_1 = 1$ and $\gamma_2 = -1$ allows us to qualitatively match what we observe in the data. Not only sales growth remains more dispersed than productivity growth, but it also has a smaller Crow-Siddiqui kurtosis. Not surprisingly, the increasing response features the opposite, and counterfactual, pattern.

This simple exercise suggests that not only firms face non-Gaussian and heavy-tailed fundamental shocks, but they also most likely respond to them by adjusting inputs with double exponential responsiveness. Our suggestive evidence is inconsistent with a simple, frictionless, model in which firms choose labor by setting the marginal product equal to an aggregate wage. That framework would be akin the "constant response" model presented

²In this example we consider all firms, but the intuition can be tailored to each of the quintile-specific result previously shown.

Table A.2: Shocks vs responsiveness

| | Standard Deviation | | | Crow-Siddiqui | | |
|---------------------|--------------------|-------|---------|---------------|-------|-------|
| | Sales | TFP | ratio | Sales | TFP | ratio |
| Data: all firms | 0.427 | 0.384 | 1.11 | 6.457 | 9.236 | 0.699 |
| Constant response | 0.644 | 0.384 | 1.677 | 9.236 | 9.236 | 1 |
| Decaying response | 0.436 | 0.384 | 1.135 | 7.956 | 9.236 | 0.861 |
| Increasing response | 295.033 | 0.384 | 768.315 | 11.359 | 9.236 | 1.230 |

Note: Baseline sample of manufacturing firms as described in Section 2. In the first row, Sales and TFP refer to log changes in residualized sales and TFP, as described in the main body. The last three rows use empirical TFP growth – which is hence the same across models – and obtains sales growth by fitting Equation 10 with different parameter values as discussed in the text.

in Table A.2. Moreover, that model predicts that employment is a convex function of TFP. In contrast, our findings suggest that not only employment is likely a concave function of fundamental shocks, but also that the log of employment is concave in the log of TFP.³ The investigation of the exact source of these relationships is beyond the scope of this paper and the findings of Table A.2 should be treated as suggestive. However, we speculate that this behavior can be seen as broadly consistent with various economic models of firm behavior. For example, hiring costs may limit the extent to which firms are able to reap the benefits of large productivity shocks, reducing their ability to hire more workers. Financial frictions, especially of the type directly constraining labor inputs, could also generate a similar pattern. These frictions are likely to be relevant for small firms: in line with this, these firms display the largest kurtosis gap. On the way down, firing costs would also be consistent with this behavior. Looking back at earlier results, this would rationalize why large firms are more likely to face negative tail risk in productivity, but are able to absorb those and thus face symmetric changes in sales.

D Computational Appendix

We follow Arellano et al. (2017) and Arellano et al. (2021) to estimate the nonlinear hidden quantile model. To do so, we use a stochastic EM algorithm that consists of two steps: 1) drawing the latent components of the model (η_{it} and ϵ_{it}), 2) updating the model parameters given the latent draws. To perform step 1), we follow Arellano et al. (2021) and use a

³In many models of household behavior, such as Carroll and Kimball (1996), consumption is a concave function of wealth, implying that dollar-for-dollar MPCs fall with the size of a transitory income shock. Our findings suggest that a similar pattern may apply to firms too, and even to growth rates of employment.

Sequential Monte-Carlo (SMC) sampling method.⁴ The simulation step is run in parallel across firms to simplify the estimation of the model with unbalanced data. For step 2), as in [Arellano et al. \(2017\)](#), we update the parameters using quantile regressions, after modelling all conditional quantile functions as linear quantile specifications at a grid of 11 equidistant percentiles. For additional details, see [Arellano et al. \(2017\)](#) and [Arellano et al. \(2021\)](#).

We estimate the 5 models described in Section 5.2 on the baseline sample of manufacturing firms, using residualized and detrended log TFP as discussed in the main body. The parameters for the AR(1) models – with and without transitory shocks – are estimated by minimizing the distance between the theoretical autocovariances generated by the model and the empirical autocovariances of productivity. We use an identity matrix. We repeat the estimation for 100 initial conditions and keep the estimated parameters associated with the best goodness of fit (the sum of squared deviations). For model (III), the one that only features nonlinear persistence, we follow [Arellano et al. \(2017\)](#) and estimate quantile autoregressions of residualized log TFP, specifying quantile functions as third-order Hermite polynomials. Model (V) is estimated as described above; model (IV) follows the same structure, adjusting the log-likelihood to assume that transitory shocks follow a Gaussian distribution.

The simulations in Section 5.2 work as follows, using the estimated parameters. We start from the firm’s first empirical observation. We also follow the data in the firm-specific spells – i.e., how long a firm remains in operation. Each model is simulated 100 times. Out-of-sample simulations of models III–V can, rarely, result in explosive behavior of a small number of firms.⁵ To deal with this, we restrict the support of model-simulated log TFP to be the same as in the data: any firm that hits the bounds in the simulation is dropped. This amounts to between 0.02% (model V) and 1.63% (model IV) of all firms in the sample and less than 0.1% of firm–year observations. An alternative approach, winsorizing productivity at the bounds at each time period in the simulation, delivers qualitatively similar results.⁶ We compute the volatility over time and the one-year autocorrelation of aggregate TFP growth rates in each of the simulated panel dataset, for each model, and report the medians across simulations in Table 2 of Section 5.2.

⁴The use of this sampler instead of the Metropolis-Hastings employed by [Arellano et al. \(2017\)](#) is justified by numerical stability.

⁵An alternative to out-of-sample simulations is to use the SMC sampling method in the simulations, as in ABBL. This latter approach delivers a qualitative similar ordering across models, and a much better fit to the data for model (V). As such, our out-of-sample simulations can be seen as conservative.

⁶These restrictions can find justification through a procedure that accounts for path dependence in simulating firms’ growth. This procedure would assign a history-dependent probability of converging towards a long-run firm-value. Given our data, we expect the probability of diverging to decrease, with history-dependent rates, as firms keep growing.

E Additional estimation results

Departures from normality and the ability of the model to match the data can be visualized looking at how conditional moments of the distribution of changes of log productivity residuals vary with firm sales. Figures A.12a, A.12b, and A.12c report the patterns of conditional dispersion, skewness, and kurtosis in the data, as documented in Section 3, and in the model. The model does a very good job at matching conditional dispersion and kurtosis, and a fair job with regard to skewness too, although it misses the mildly negative skewness for productive firms. Figure A.13 plots the histogram of the persistent and transitory components of log productivity as estimated by the model discussed in Section 4. We then obtain the histograms of the estimated components of log productivity by past quantiles of the distribution of simulated TFP and report them for the top and bottom quantiles in Figure A.14.

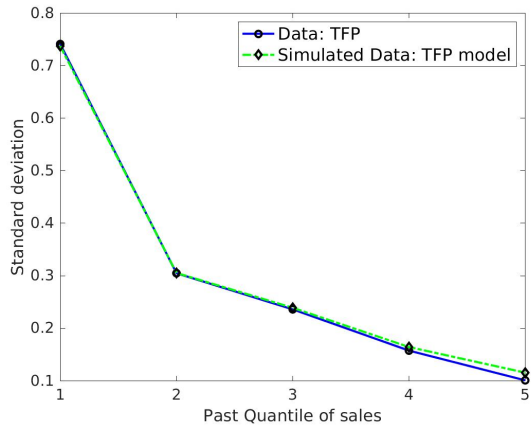
In Figure A.15, we show the evolution of TFP and its estimated persistent component for a specific large and productive firm, Apple. In 2006, Apple’s productivity grew as much as in the top 5% of all firms in the sample for that year. The model estimates that this productivity boost was mostly driven by the persistent component and led to a subsequent sustained productivity growth. In 2012, Apple’s productivity dropped sharply.⁷ Again, the model assigns that even to the persistent component. In line with the results showed in Figure 6c of Section 4, the model suggests that negative shocks hitting productive firms will erase part of the persistence of histories. Indeed, Apple’s productivity growth has been sluggish thereafter.

F Some insights on the effects of the model’s features on the aggregates

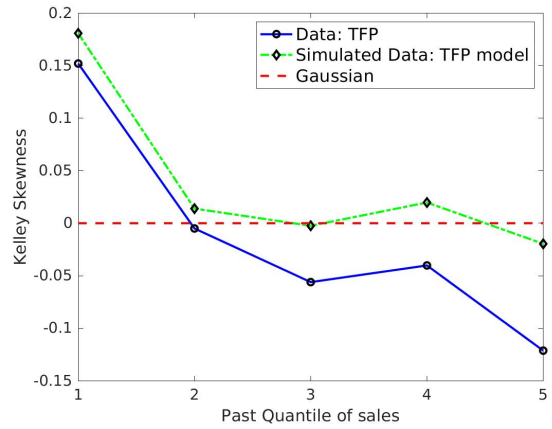
Consider the following linear AR(1) model for the TFP of a firm i at time t , with $i = 1, \dots, N$: $o_{i,t+1} = \rho o_{i,t} + \nu_{i,t+1}$, where innovations ν are normally distributed with variance σ_ν^2 . Define the aggregated TFP $\tilde{G}_t = \sum_i o_{i,t} - \sum_i o_{i,t-1} = \sum_i (o_{i,t} - o_{i,t-1})$. We are interested in assessing the importance of ρ and σ_ν^2 on the variance and autocorrelation of \tilde{G}_t . Let us assume that $o_{i,t}$ are i.i.d., which is an unappealing assumption that we make here only to illustrate the

⁷That productivity change was in the bottom 5% of the distribution for that year.

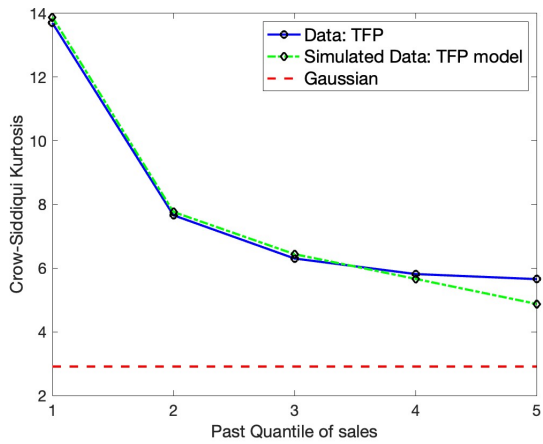
Figure A.12: Conditional moments of residualized log TFP: model vs data



(a) Conditional standard deviation



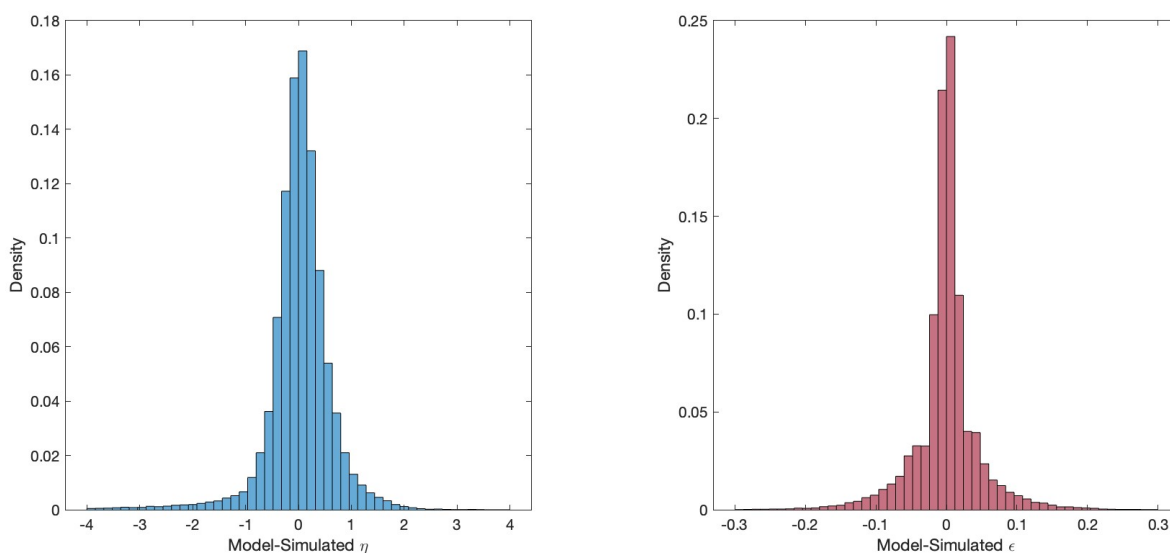
(b) Conditional skewness



(c) Conditional kurtosis

Notes: Blue lines refer to the Compustat data, while dashed-dotted green lines to data simulated according to the nonlinear model, with parameters set to their estimated values. Dashed red lines show the moments if the process was Gaussian.

Figure A.13: Distribution of productivity components in the model

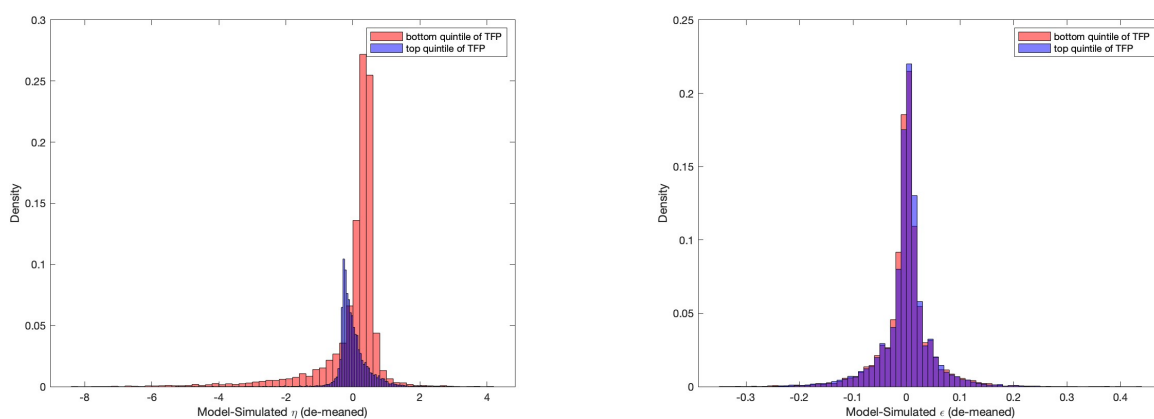


(a) Persistent component η

(b) Transitory component ϵ

Notes: Model estimated on the baseline sample of manufacturing firms, as discussed in Section 4. Histogram of model-simulated persistent – left panel – and transitory – right panel – components of log productivity.

Figure A.14: Distribution of productivity components in the model by top and bottom quantiles of TFP

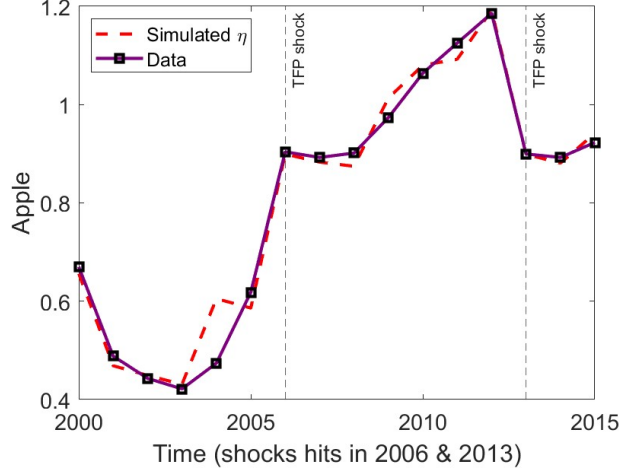


(a) Persistent component η

(b) Transitory component ϵ

Notes: Model estimated on the baseline sample of manufacturing firms, as discussed in Section 4. Histogram of model-simulated persistent – left panel – and transitory – right panel – components of log productivity, by top and bottom past quantiles of the simulated TFP distribution.

Figure A.15: Evolution of productivity components for Apple



mechanisms. Let us also assume that $\rho < 1$. For the variance of \tilde{G}_t , we get:

$$\begin{aligned}
 Var(\tilde{G}_t) &= Var\left(\sum_i (o_{i,t} - o_{i,t-1})\right) = \sum_i Var(o_{i,t} - o_{i,t-1}) \\
 &= \sum_i Var(-(1-\rho)o_{i,t-1} + \nu_{i,t}) = N \left((1-\rho)^2 \frac{\sigma_\nu^2}{1-\rho^2} + \sigma_\nu^2 \right) \\
 &= 2N \frac{\sigma_\nu^2}{1+\rho}
 \end{aligned}$$

For the first order autocovariance of \tilde{G}_t , we get:

$$\begin{aligned}
 Cov(\tilde{G}_t, \tilde{G}_{t-1}) &= Cov\left(\left(\sum_i (o_{i,t} - o_{i,t-1})\right), \left(\sum_i (o_{i,t-1} - o_{i,t-2})\right)\right) \\
 &= NCov\left(\left(- (1-\rho)\rho o_{i,t-2} - (1-\rho)\nu_{i,t-1} + \nu_{i,t}\right), \left(- (1-\rho)o_{i,t-2} + \nu_{i,t-1}\right)\right) \\
 &= N \left(\frac{- (1-\rho)\sigma_\nu^2}{1+\rho} \right)
 \end{aligned}$$

Second, we consider a model that also features a Gaussian transitory component as follows: $o_{i,t+1} = \eta_{i,t+1} + \varepsilon_{i,t+1}$, where $\eta_{i,t+1} = \rho\eta_{i,t} + \nu_{i,t+1}$, and innovations ν and ε are normally distributed with respective variances σ_ν^2 and σ_ε^2 . Under the same assumptions considered

above, it is possible to show that now:

$$\begin{aligned}
Var(\tilde{G}_t) &= Var\left(\sum_i(o_{i,t} - o_{i,t-1})\right) = \sum_i Var(o_{i,t} - o_{i,t-1}) \\
&= \sum_i Var(-(1 - \rho)\eta_{i,t-1} + \nu_{i,t} + \varepsilon_{i,t} - \varepsilon_{i,t-1}) \\
&= 2N\left(\frac{\sigma_\nu^2}{1 + \rho} + \sigma_\varepsilon^2\right)
\end{aligned}$$

Thus, as expected ε increases the volatility of \tilde{G}_t . As for the first order autocovariance of \tilde{G}_t , we now get:

$$\begin{aligned}
Cov(\tilde{G}_t, \tilde{G}_{t-1}) &= Cov\left(\left(\sum_i(o_{i,t} - o_{i,t-1})\right), \left(\sum_i(o_{i,t-1} - o_{i,t-2})\right)\right) \\
&= N\left((1 - \rho)^2 \rho \frac{\sigma_\nu^2}{1 - \rho^2} - (1 - \rho)\sigma_\nu^2 - \sigma_\varepsilon^2\right) \\
&= N\left(\frac{-(1 - \rho)\sigma_\nu^2}{1 + \rho} - \sigma_\varepsilon^2\right)
\end{aligned}$$

Thus, as expected ε decreases the autocorrelation of \tilde{G}_t .