Why are Firm Growth Distributions Heavy-tailed?

Robert Parham*

Friday 4th August, 2023

Abstract

Firm growth and return distributions are heavy-tailed. Accounting for the interplay of sales and expenses is sufficient to explain this fact without relying on time-varying volatility or factors external to the firm. Embedding the implied production function into a standard q-theory model yields novel and specific predictions regarding the distributions of income, growth, and returns. The predictions are supported by the data. The model is the first to correctly replicate the distribution of firm income and is hence useful as a foundational model for future work. It proposes extended definitions of firm income scale, efficiency, and growth.

JEL classifications: D21, G3, L11, C46

Keywords: Growth, firm size, heavy tails, q-theory.

*University of Virginia (robertp@virginia.edu). I thank Toni Whited, Ron Kaniel, Chris Yung, William Wilhelm, Mihail Velikov, Donald Bowen, Michael Gallmeyer, Bradley Paye, Sara Easterwood, Jan Eeckhout, Sanket Korgaonkar, and seminar participants at the 2022 Commonwealth Finance Conference and 2023 Finance Down Under for helpful comments and suggestions. All remaining errors are my own.

Robert Parham

— I have nothing to disclose

"There are only two ways to make money: increase sales and decrease costs"

— **Fred DeLuca**, founder of Subway

1 Introduction

We have known that the statistical distribution of firm growth is heavy-tailed since at least the work of Ashton (1926), who documents this for the growth of British textile businesses in the period 1884 – 1924. Most firms experience moderate growth rates, with about half of firms experiencing a yearly capital growth rate in the ±10% range. But some firms grow (or shrink) in large jumps. About 1.5% of firms more than double in size in a single year. These rare or extreme "disasters" and "winners" are far more numerous and economically consequential than the 0.00015% of firms a Normally distributed growth rate would predict. Yet we lack a first-principles explanation for the emergence of heavy-tailed firm growth or a clear statistical representation of its distribution. This work aims to fill this gap.

Much of the interest in the distribution of firm growth in the economic literature stems from the fact equity returns are themselves just another measure of firm growth. The heavy tails of growth and returns were studied by Mandelbrot (1960, 1961) and Fama (1963, 1965) who proposed the family of Stable distributions (also known as Stable-Paretian or Pareto-Lévy) as a statistical model of firm growth. This distribution was later rejected by Officer (1972), who concludes that "It may be that a class of fat-tailed distributions with finite second moments will be found [...] but as yet this remains to be clearly demonstrated." The shape of the return distribution is crucial for the coherence of modern portfolio theory because the Stable distribution lacks finite second moments (our ubiquitous measure of risk), for the predictions and accuracy of option pricing models, and for the fit of production-based asset-pricing models to the data.

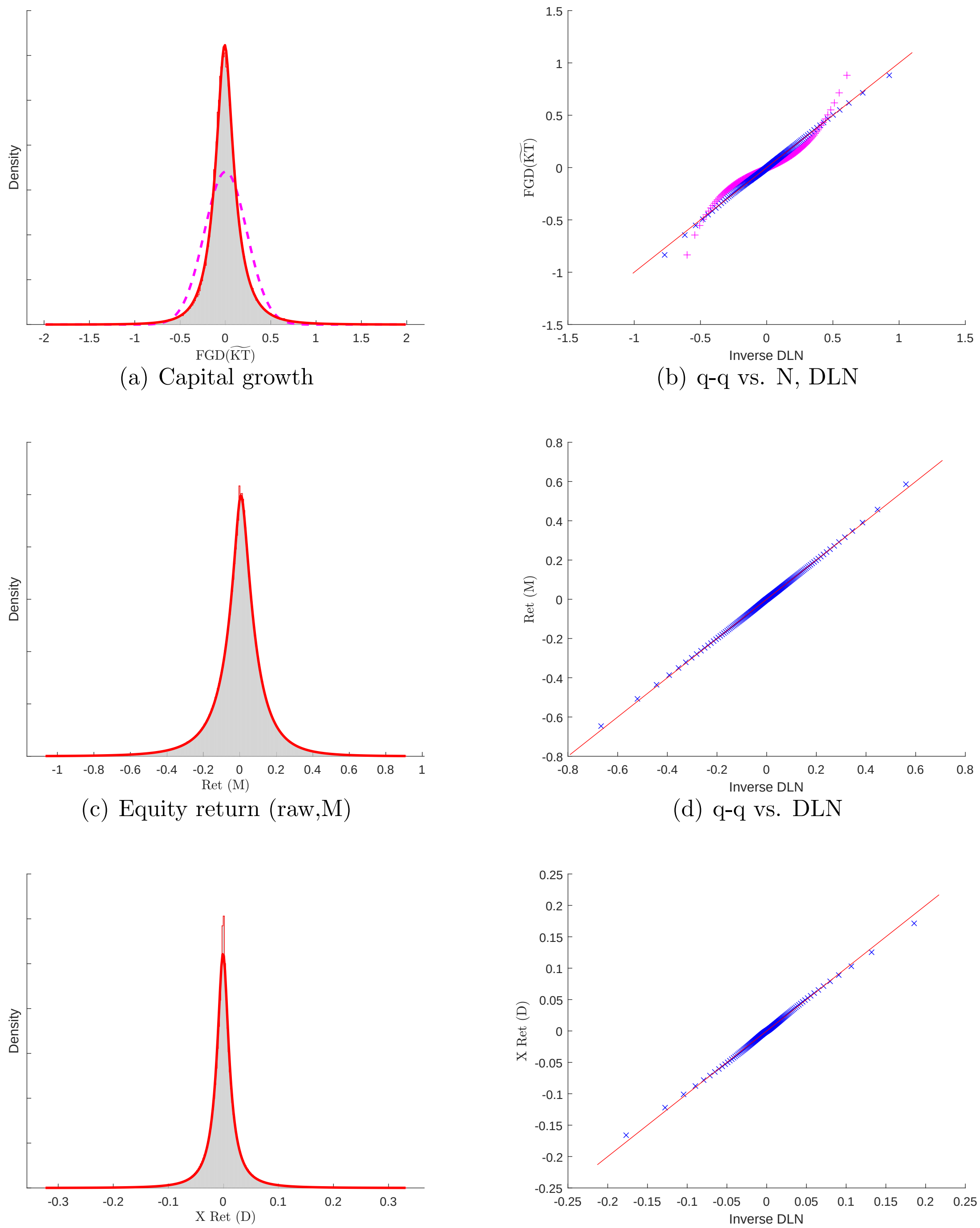
I show that a simple, intuitive modification to an otherwise standard q-theory model of the firm — separately accounting for sales and expenses — is sufficient to yield heavy-tailed firm growth, without appealing to time-varying volatility or factors external to the firm.

Moreover, the economic model makes a highly specific prediction on the shape of firm growth distributions, predicting they should distribute as the difference-of-log-Normals. I show that this prediction is supported by the data and that the obscure difference-of-log-Normals distribution exhibits a remarkable fit to a plethora of firm outcomes, such as income, capital growth, income growth, and equity returns, as the model predicts. The fit with equity returns is outstanding and holds for daily, monthly, yearly, raw, and excess returns in a set of robustness tests.

Figure 1 displays this fit using data on public US firms during the 50-year period 1970-2019. Panel (a) present the distribution of capital growth, along with two MLE-fitted distributions: a Normal and a difference-of-log-Normals, with Panel (b) presenting the corresponding q-q plot. Panels (c) and (e) present two equity return distributions: monthly raw and daily excess returns, respectively, again fitted with difference-of-log-Normals distributions. The q-q plots again exhibit a nearly perfect fit.

The contribution of the paper is not limited to its asset-pricing implications. I show that the novel production function implied by separately modeling sales and expenses — a difference-of-log-linears production function, itself a generalization of the log-linear (or Cobb-Douglas) production function — has several desirable properties beneficial for dynamic models of corporate finance. In that, the paper follows in the tradition of, e.g., Epstein and Zin (1991); Campbell and Cochrane (1999) and Bansal and Yaron (2004).

First, the model resolves the critique of Gorbenko and Strebulaev (2010); Strebulaev and Whited (2012) regarding the "highly unrealistic" lack of losses in dynamic firm models. Strebulaev and Whited (2012) note, in the context of transitory shocks leading to losses, that "mathematical problems [...] have not been solved satisfactorily even for the simplest cases." In contrast, the model presented here enables a simple and tractable treatment of losses that can be used when modeling firm leverage and bankruptcy decisions. Second, the difference-of-log-linears production function gives rise to extended and internally consistent definitions for common concepts such as firm scale, efficiency, returns-to-scale, and income



(a) Capital growth (b) q-q vs. N, DLN

(c) Equity return (raw,M) (d) q-q vs. DLN

(e) Equity return (FF3,D) (f) q-q vs. DLN

Fig. 1. Firm growth distributions. Panel (a) presents the growth of capital (total assets), with MLE-fitted Normal and DLN distributions, for a set of 143K firm-year observations from 1970-2019. Panel (b) presents the respective q-q plots. Panels (c) and (e) present the monthly raw and daily excess equity returns for 2M and 5M observations, respectively, with MLE-fitted DLN distributions. Panels (d) and (f) present the respective q-q plots.

growth. Third, I show that structural estimation of firm models based on this production function is significantly simplified relative to existing models of firm dynamics due to the observability of sales and expenses. For these reasons, I propose the model as a foundational building block when writing models exploring other aspects of the firm.

My approach is by no means a first attempt at explaining the distribution of firm growth. Gibrat (1931) introduces the log-Normal as the dominant distribution in measuring firm *size*, based on a simple argument, later named "the multiplicative Central Limit Theorem" (CLT). This prediction was confirmed for firms, cities, and other proportionally-growing entities. Gibrat, however, used the CLT to reason that firm growth should be Normally distributed and homoscedastic in scale — two predictions that have later been shown to fail in the data. Following the failure of the Stable distribution to describe firm growth, discussed above, the Laplace distribution was proposed (and later rejected as well) as a purely statistical model without an economic justification by Mantegna and Stanley (1995); Stanley, Amaral, Buldyrev, Havlin, Leschhorn, Maass, Salinger, and Stanley (1996); Bottazzi and Secchi (2003).

Prior theoretical models attempt to explain the mechanics of firm growth. Notable examples include Simon and Bonini (1958); Lucas (1978); Klette and Kortum (2004) and the more recent works of Bottazzi and Secchi (2006); Buldyrev, Growiec, Pammolli, Riccaboni, and Stanley (2007); Luttmer (2011). The early works counter-factually yield firms with Normal growth, while the latter works aim to replicate Laplace-distributed growth. They do so by positing an economy with a scarcity of opportunities, in which heavy-tailed growth stems from factors external to the firm. The model presented here, in contrast, is a simple, intuitive, and straightforward extension of the workhorse q-theory model, and it predicts a specific novel distribution confirmed by the data.

The paper proceeds as follows: Section 2 presents a q-theory model of firm dynamics, including entry and exit, using the traditional log-linear and the extended difference-of-log-linears production functions. Section 3 analyzes the theoretical implications of the models and confronts these implications with the data. The section analyzes implications for (i) income; (ii) scale and efficiency (iii) income growth; (iv) returns-to-scale; and (v) capital and value growth. Section 4 presents a structural estimation and a simulation of the models. The q-theory model with the new production function replicates the empirical firm data qualitatively (i.e., in distributional form) as well as quantitatively (i.e., the moments of said distributions). I provide concluding remarks in Section 5.

2 Models

Since the early works of Lucas (1967), Tobin (1969), Uzawa (1969), and especially the seminal work of Hayashi (1982), q-theory has become the canonical workhorse of firm modeling in the corporate finance literature.[1] The neo-classical q-theory model posits a value-maximizing firm facing a dynamic investment-dividend decision subject to adjustment costs. The value-maximizing firm invests up to the point where the marginal benefit of investment equals the marginal cost of investment, both then denoted marginal-q. The next subsection presents a general version of the q-theory model that includes entry and exit but abstracts from the specific forms of the: (i) investment function; (ii) production function; (iii) stochastic dynamics; and (iv) entry and exit mechanics. The following four subsections discuss each facet, presenting the relevant functional and stochastic forms.

2.1 The general q-theory model

At the beginning of every period, a representative value-maximizing firm observes its endogenous capital stock for the period $K_t > 0$ and exogenous (i.e., stochastic) productivity $Z_t > 0$.[2] The firm first chooses whether to remain for another period (denoted $\alpha_t = 1$)

[1] Recent examples include: Hennessy and Whited (2005), Hennessy and Whited (2007), Liu, Whited, and Zhang (2009), Livdan, Sapriza, and Zhang (2009), Riddick and Whited (2009), Bolton, Chen, and Wang (2011), DeAngelo, DeAngelo, and Whited (2011), Lin (2012), Belo, Lin, and Vitorino (2014), Nikolov and Whited (2014), Li, Whited, and Wu (2016), Belo, Li, Lin, and Zhao (2017), Michaels, Page, and Whited (2019), Sun and Xiaolan (2019), Falato, Kadyrzhanova, Sim, and Steri (2021).

[2] Both K and Z may be vectors in the general case.

or exit ($\alpha_t = 0$). Firm owners receive some non-negative payoff $\mathbf{V}^{\text{exit}}(K_t, Z_t) \geq 0$ upon exit. A remaining firm then chooses an investment level $I_t = \mathbf{I}(K_{t+1}, K_t)$ for the period, or equivalently an end-of-period capital level K_{t+1}, with negative investment values implying the proceeds from capital sale. The firm produces income (sales net of all expenses) $Y_t = \mathbf{Y}(K_t, Z_t)$, and dispenses $D_t = \mathbf{D}(K_{t+1}, K_t, Z_t) = Y_t - I_t$ to owners. All payoffs accrue at the beginning of the period for simplicity.

The value of the firm $V_t = \mathbf{V}(K_t, Z_t)$ is the expected present value of all dispensations. This value is recursively defined by the Bellman equation

$$V_t = \max_{K_{t+1}, \alpha_t} \left\{ (1 - \alpha_t) \cdot \mathbf{V}^{\text{exit}}(K_t, Z_t) + \alpha_t \cdot (\mathbf{D}(K_{t+1}, K_t, Z_t) + \beta \cdot \mathbb{E}_t [\mathbf{V}(K_{t+1}, Z_{t+1})]) \right\} \quad (1)$$

with $0 < \beta < 1$ the time discount parameter, such that $\beta = (1 + r)^{-1}$, and $r > 0$ is the cost of capital for the firm.

The investment decision of a remaining firm can be characterized by equating the benefit and cost of a marginal unit of investment. This implies choosing K_{t+1} such that

$$\beta \cdot \mathbb{E}_t [\mathbf{V}'_1(K_{t+1}, Z_{t+1})] = -\mathbf{D}'_1(K_{t+1}, K_t, Z_t) = \mathbf{I}'_1(K_{t+1}, K_t) \quad (2)$$

where $\mathbf{X}'_j()$ indicates the derivative of the function $\mathbf{X}()$ w.r.t its j^{th} argument. The R.H.S of Equation 2 is the marginal cost today of one extra unit of next period capital, and the L.H.S the discounted expected marginal benefit of the extra unit. The value of both is the marginal-q of the firm at period t.

Denote the investment policy function of a remaining firm prescribed by Equation 2 to be $K_{t+1} = \Psi_t = \mathbf{\Psi}(K_t, Z_t)$.[3] It is useful to define the exit (or bankruptcy) threshold of the

[3] That the function $\mathbf{\Psi}()$ exists under mild conditions on the functions $\mathbf{Y}()$, $\mathbf{I}()$, and $\mathbf{V}^{\text{exit}}(K_t, Z_t)$ is a standard result. See e.g. Stokey, Lucas, and Prescott (1989).

firm in period t,

$$B_t = \mathbf{B}\left(K_t, Z_t\right) = \mathbf{D}\left(\Psi_t, K_t, Z_t\right) + \beta \cdot \mathbb{E}_t\left[\mathbf{V}\left(\Psi_t, Z_{t+1}\right)\right] - \mathbf{V}^{\text{exit}}\left(K_t, Z_t\right) \tag{3}$$

as the difference between the optimal values conditional on remaining and exiting. The firm's exit policy is to remain when it is above the bankruptcy threshold ($B_t \geq 0$) and exit otherwise.

We can now combine Equation 2 with the envelope condition to write the remaining firm's full first-order condition (f.o.c) for capital as

$$\beta \cdot \mathbb{E}_t\left[\left(1-\alpha_{t+1}\right) \cdot \mathbf{V}_1^{\text{exit}\prime}\left(\Psi_t, Z_{t+1}\right) + \alpha_{t+1} \cdot \left(\mathbf{Y}_1^{\prime}\left(\Psi_t, Z_{t+1}\right) - \mathbf{I}_2^{\prime}\left(\Psi_{t+1}, \Psi_t\right)\right)\right] = \mathbf{I}_1^{\prime}\left(\Psi_t, K_t\right) \tag{4}$$

which in turn characterizes the function $\mathbf{\Psi}\left(K_t, Z_t\right)$. The equation equates the cost of a marginal unit of extra capital with the discounted marginal benefits from higher exit value, higher production, and lower future investment costs.

Additionally, it is useful to define the function $\mathbf{\Phi}(Z_t)$ to be the fixed point of the function $\mathbf{\Psi}\left(K_t, Z_t\right)$ in the first input, such that $\Phi_t = \mathbf{\Phi}(Z_t) = \mathbf{\Psi}(\mathbf{\Phi}(Z_t), Z_t)$. I.e., Φ_t is the steady-state capital level corresponding to Z_t. As usual, the functions cannot be specified in closed form and require numerical evaluation.

2.2 Investment function

The investment function $\mathbf{I}\left(K_{t+1}, K_t\right)$ determines the investment level required to move from current capital level K_t to next-period capital level K_{t+1}. It embeds assumptions on capital depreciation and capital adjustment costs. Throughout, I will be using the investment function

$$\mathbf{I}_{exp}\left(K_{t+1}, K_t\right) = \left(K_{t+1} - K_t\right) \cdot \exp\left(\gamma \cdot \text{dlog}\left(K_{t+1}\right)\right) + \delta \cdot K_t \tag{5}$$

with $\text{dlog}(K_{t+1}) = \log(K_{t+1}) - \log(K_t) = k_{t+1} - k_t$, $0 \leq \delta \leq 1$ the capital depreciation rate, and $\gamma \geq 0$ an adjustment parameter.

Note that when $\gamma \to 0$, $\mathbf{I}_{exp}$ simplifies to the perpetual inventory formula with no adjustment costs

$$\mathbf{I}_{triv}\left(K_{t+1}, K_t\right) = \left(K_{t+1} - K_t\right) + \delta \cdot K_t = K_{t+1} - (1 - \delta) \cdot K_t \tag{6}$$

and the firm's f.o.c from Equation 4 simplifies to the well-known equality between the expected marginal product of capital and the user cost of capital, adjusted for exit

$$\mathbb{E}_t\left[(1 - \alpha_{t+1}) \cdot \mathbf{V}_1^{\text{exit}'}\left(K_{t+1}, Z_{t+1}\right) + \alpha_{t+1} \cdot \mathbf{Y}_1'\left(K_{t+1}, Z_{t+1}\right)\right] = r + \delta \tag{7}$$

Furthermore, without adjustment costs, the firm immediately adjusts to the optimal capital level $\mathbf{\Phi}(Z_t)$ every period, such that $\mathbf{\Phi}(Z_t) = \mathbf{\Psi}(K_t, Z_t) \;\; \forall K_t$, and marginal-q $\equiv 1$. In this simplified case that is nevertheless useful as a benchmark, the policy function of the firm can be found without value- or policy-function iterations. This fact will be useful when considering the implications of the models later.

Upon inspection, $\mathbf{I}_{exp}$ is closely related to the traditional quadratic adjustment form common to the literature cited above, often written as

$$\mathbf{I}_{quad}\left(K_{t+1}, K_t\right) = \left(K_{t+1} - K_t\right) + \gamma \cdot \left(\frac{K_{t+1}}{K_t} - 1\right)^2 \cdot K_t + \delta \cdot K_t \tag{8}$$

We can see the relation between the two functions by noting the parenthesis in Equation 8 is simply the per-period capital growth in *percentage terms*, while Equation 5 uses growth in *log-point terms*. The similarity and difference between the investment functional forms can be inspected in panels (a)-(c) of Figure 2.[4] In all panels, the current capital level of the firm is $K_t = 100$ and the depreciation rate is $\delta = 4\%$. The panels present I_{triv}, I_{quad}, and I_{exp} with low, medium and high adjustment costs ($\gamma \in \{0.1, 0.5, 1.5\}$), respectively.

[4] Also see https://kn.owled.ge/InvestFuncs for an interactive version.

2.3 Production functions

The main goal of this paper is to compare and contrast the canonical neo-classical log-linear (LL) production function with a difference-of-log-linears (DLL) production function that captures the interplay of sales and expenses. The canonical production function can be written as

$$\mathbf{Y}_z\left(K_t, Z_t\right) = Z_t \cdot K_t^{\theta_z} = \exp\left(z_t + \theta_z \cdot k_t\right) \tag{9}$$

with returns to scale parameter $0 < \theta_z < 1$ and with lower-case variables denoting log values as usual. The q-model generally abstracts from labor, assuming it is elastically adjustable within periods. Wages and other expenses are already accounted for, as $\mathbf{Y}_z\left(K_t, Z_t\right)$ models income, i.e., sales minus expenses. Models following this production function are denoted Z-models.

The second production function I consider models income explicitly using its economic definition — the difference between sales and expenses — possibly the most fundamental of accounting identities,

$$\mathbf{Y}_{sx}\left(K_t, S_t, X_t\right) = \underbrace{S_t \cdot K_t^{\theta_s}}_{Sales \equiv \mathbb{S}_t} - \underbrace{X_t \cdot K_t^{\theta_x}}_{Expenses \equiv \mathbb{X}_t} = \exp\left(s_t + \theta_s \cdot k_t\right) - \exp\left(x_t + \theta_x \cdot k_t\right) \tag{10}$$

with $0 < \theta_x, \theta_s < 1$ returns to scale parameters in sales and expenses, respectively. In a slight abuse of notation, firm sales during period t are denoted $\mathbb{S}_t$ and firm expenses $\mathbb{X}_t$. The function $\mathbf{Y}_{sx}()$ is now a function of three variables — the capital stock K_t and two stochastic exogenous variables, S_t and X_t, controlling the dynamics of sales and expenses. I.e., in this model, Z_t is vector-valued. Models following this production function are denoted SX-models.

Panels (d)-(f) of Figure 2 present examples of the $\mathbf{Y}_{sx}()$ functional form.[5] Panel (d) presents the (logs of) sales, expenses, income, and *net income* (income minus user-cost of capital) of a firm, as functions of the firm's (log) capital level k, at some given parameter

[5] Also see https://kn.owled.ge/ProdFuncs for an interactive version.

values. Panel (d) is hence in log-log scale. In this scale, sales and expenses are both linear in k, as is made clear by the last part of Equation 10. Their difference (i.e., income) is, however, negative until firm scale is high enough for sales to overtake expenses. Considering net income, Panel (d) demonstrates the firm is subject to Goldilocks conditions — it has positive net income for a limited range of capital levels and has negative net income otherwise.

Because negative incomes play an important role in our analysis, it is useful to change the Y-axis of Panel (d) from log to Inverse Hyperbolic Sine (asinh) scale, so negative values are better observed. Panel (e) hence repeats the presentation of Panel (d) but with asinh-transformed Y-axis. Finally, Panel (f) repeats the presentation of Panel (e), but for a higher value of x. I.e., it considers what will happen to the income of a firm if it faces a positive (un)productivity shock to the stochastic variable x. Geometrically, increasing the value of x simply shifts the expenses line (dotted red line) upwards in panel (e), as it controls the intercept in the (log-log) equation defining expenses. The outcome is to decrease income (and net income) at every capital level. In fact, the firm's net income is now negative at *all* capital levels.

2.4 Stochastic dynamics

We now define the stochastic dynamics of the productivity state variables Z and S, X, respectively. The Z process is assumed to follow the canonical AR(1) in logs. I.e. $z = \log(Z)$ follows

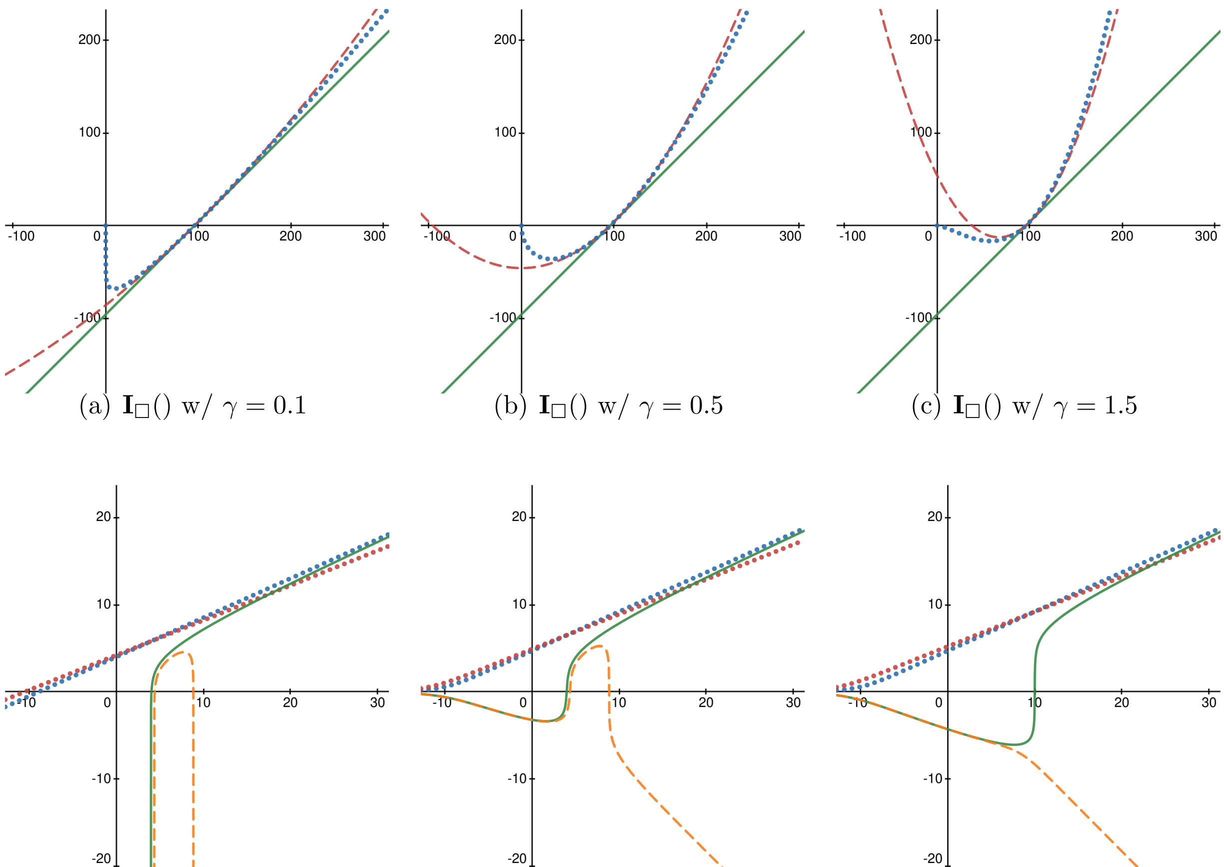
$$z_{t+1} = (1 - \rho_z) \cdot \mu_z + \rho_z \cdot z_t + \epsilon^Z_{t+1} \tag{11}$$

with persistence $0 < \rho_z < 1$ and mean μ_z. The i.i.d innovations follow

$$\epsilon^z \sim \mathbb{N}(0, \sigma_z^2) \tag{12}$$

with $\sigma_z > 0$ the standard deviation of ϵ^z.

Before proceeding, it is worth contemplating the economic meaning of the exogenous



(a) $\mathbf{I}_{\square}()$ w/ $\gamma = 0.1$ (b) $\mathbf{I}_{\square}()$ w/ $\gamma = 0.5$ (c) $\mathbf{I}_{\square}()$ w/ $\gamma = 1.5$

(d) $\mathbf{Y}_{sx}()$ in log-log w/ low X (e) $\mathbf{Y}_{sx}()$ in asinh-log w/ low X (f) $\mathbf{Y}_{sx}()$ in asinh-log w/ high X

Fig. 2. Functional forms. This figure presents facts about the functional forms used for investment and production. Panels (a)-(c) graph the total investment (incl. adjustments) required to move from $K_t = 100$ to an arbitrary K_{t+1}, with $\delta = 4\%$. The solid green line is $\mathbf{I}_{triv}()$, the dashed red is $\mathbf{I}_{quad}()$, and the dotted blue is $\mathbf{I}_{exp}()$. Panels (a)-(c) present low ($\gamma = 0.1$), medium ($\gamma = 0.5$), and high ($\gamma = 1.5$) adjustment costs, respectively. Panels (d)-(f) graph the sales (dotted blue), expenses (dotted red), income (solid green), and net income (income minus user-cost of capital, dashed orange) of a firm as functions of its (log) capital stock k when using the $\mathbf{Y}_{sx}()$ production function. The X-axis in panels (d)-(f) is logarithmic, as is the Y-axis in panel (d). The Y-axis in panels (e)+(f) is in asinh scale. Panels (d)+(e) use the parameters $s = 4.83, x = 4.94$, while panel (f) modifies $x = 5.1$.

productivity process Z. What determines the factor-productivity of the firm? It is a function of the "skill, dexterity, and judgment with which labor is applied," as in Smith (1776), or of the firm's production technology, cost structure, managerial talent, market power, and a host of other components, including luck. In that sense, Z is partly endogenous. Of course, all firms would prefer to produce as much income as possible from a given amount of capital K. Put differently, all firms would like to have as high a Z as possible. Firms hence optimize the components of Z under their control, and as a result, achieve (log) productivity μ_Z on average. But firms differ in their ability to achieve a high Z, and the differences are persistent. Z_t hence represents the current productivity of the representative firm, given its optimizing behavior on the components of Z. In this way, Z is the usual measure of our ignorance regarding the firm, as in Abramovitz (1956).

The S, X process is similarly assumed to follow a joint-AR(1) in logs. I.e., $s = \log(S)$ and $x = \log(X)$ follow

$$\begin{aligned} s_{t+1} &= (1-\rho_s)\cdot\mu_s + \rho_s\cdot s_t + \epsilon^s_{t+1} \\ x_{t+1} &= (1-\rho_x)\cdot\mu_x + \rho_x\cdot x_t + \epsilon^x_{t+1} \end{aligned} \tag{13}$$

with persistence $0 < \rho_s, \rho_x < 1$ and mean μ_s, μ_x. The i.i.d innovations follow the bi-variate Normal

$$\begin{bmatrix} \epsilon^s_{t+1} \\ \epsilon^x_{t+1} \end{bmatrix} \sim \mathbb{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_s^2 & \sigma_{sx} \\ \sigma_{sx} & \sigma_x^2 \end{bmatrix} \right) \tag{14}$$

with $\sigma_{sx} = \rho_{sx}\cdot\sigma_s\cdot\sigma_x$ for $\sigma_s, \sigma_x > 0$ and $-1 < \rho_{sx} < 1$.

Let us again contemplate the economic meaning of S, X. Clearly, all firms would prefer $S \to \infty$ and $X \to 0$. On average, however, the representative firm achieves sales (log) productivity μ_s and expenses (log un)productivity μ_x, after taking all profitable moves to *jointly* optimize both S and X. Note that ϵ^s, ϵ^x are likely correlated. Consider e.g. a firm encountering a positive demand shock, and finding it profitable to increase sales by working a third shift in its factory to supply the newfound demand. The firm can increase S (the sales productivity of a unit of capital — here, the factory), but to do so, it will also need to increase X due to extra payments to labor for working a third shift and various other extra expenses. Hence, it is likely $\rho_{sx} > 0$.

2.5 Entry and exit

To close the Z- and SX-models, we still need to define the mechanics of entry and exit. For exit, I use the simple assumption

$$\mathbf{V}^{\text{exit}}\left(K_t, Z_t\right) = \mathbf{V}^{\text{exit}}\left(K_t, S_t, X_t\right) = \nu \cdot K_t \tag{15}$$

with a capital fire-sale rate $0 < \nu < 1$. This implies firms can fire-sell their capital stock for a share ν of its value and exit. To maintain a constant measure of firms when simulating the model, a new firm is "born" every time a firm exits. The new firm's state is drawn from the ergodic distribution of firm states in the simulation.

3 Analysis of models

The SX-model of the firm, using the DLL production function, makes specific predictions on various firm outcomes. In this section, I review these predictions and test them in the data. I also compare these predictions with those of the Z-model using the LL production function, when appropriate.

The data analyzed cover all public US firms in the 50-year period 1970-2019, and include $165,000$ firm-year observations. Data are predominantly derived from the yearly CRSP/Compustat data set. For some tests related to equity returns I use higher-frequency CRSP data. All dollar amounts are normalized by yearly nominal GDP, in 2019 terms.

Table 1 defines all data panels analyzed in terms of Compustat items. Each data panel is identified throughout with a two-letter mnemonic. I mainly rely on the sources and uses identity

$$\underbrace{\text{sales}}_{SL} - \underbrace{\text{expenses}}_{XS} = \underbrace{\text{income}}_{CF} = \underbrace{\text{total net dividends}}_{DI} + \underbrace{\text{total net investment}}_{IT} \quad (16)$$

to define expenses as dissipated sales (i.e., sales - income, SL-CF). This guarantees all expenses, including cost of goods, selling, general, administrative, taxes, and various other "special" and "one-time" expenses are fully accounted for. I verify all results with a traditional top-down definition as well.

Table 1
Data definitions

This table defines all data items used. The first column is the name of each data item and the second is the mnemonic used throughout. The third column is the mapping to Compustat items or previously defined mnemonics, and the fourth is a short description. The core accounting identity used is the sources and uses equation: income = sales - expenses = total dividends + total investment, with dividends broadly defined below. The last two data items are alternative definitions used for comparability with previous work. The "L." is the lag operator.

Name	XX	Definition	Description
Equity value	EQ	mve	market value, year end
Debt value	DB	lt	book total liabilities
Total value	VL	EQ + DB	equity + debt
Equity dividends	DE	dvt + (prstkc - sstk)	dividends + net repurchase
Debt dividends	DD	xint + (L.DB-DB)	interest paid + decrease in debt
Total dividends	DI	DE + DD	to equity and debt
Total capital	KT	at	total assets (tangible)
Depreciation	DP	dp	of tangible capital
Total investment	IT	KT - L.KT + DP	growth in net assets
Income	CF	DI + IT	bottom-up free cash flows
Sales	SL	sl	total sales
Expenses	XS	SL - CF	dissipated sales
Expenses (alt.)	XA	cogs + xsga + txt	top-down definition
Income (alt.)	CA	SL - XA	top-down definition

The following sub-sections review model predictions and data outcomes for (i) income; (ii) scale and efficiency; (iii) income growth; (iv) returns-to-scale; and (v) capital and value growth.

3.1 Income

What is the statistical distribution of income CF? Firm income, often called cashflows, is of utmost importance in both major branches of financial research: corporate finance and asset pricing. Cashflows are the departing point for corporate finance and production-based asset pricing models. It is hence quite surprising that the statistical distribution of income has seen such scant interest in the finance literature.

In typical corporate finance models, income is modeled using a LL production function in which z_t follows an AR(1) process with Normal innovations (i.e., they are Z-models). Recall that the ergodic distribution of any AR(1) process is Normal, under mild assumptions. This implies a log-Normal distribution of productivity Z_t, income $\mathbf{Y}_z$, and capital K_t in the Z-model. This modeling choice, however, counter-factually yields firms with strictly positive income. The lack of negative income in such models ignores a critical feature of the profit-and-loss mechanism of firm dynamics — namely, losses.

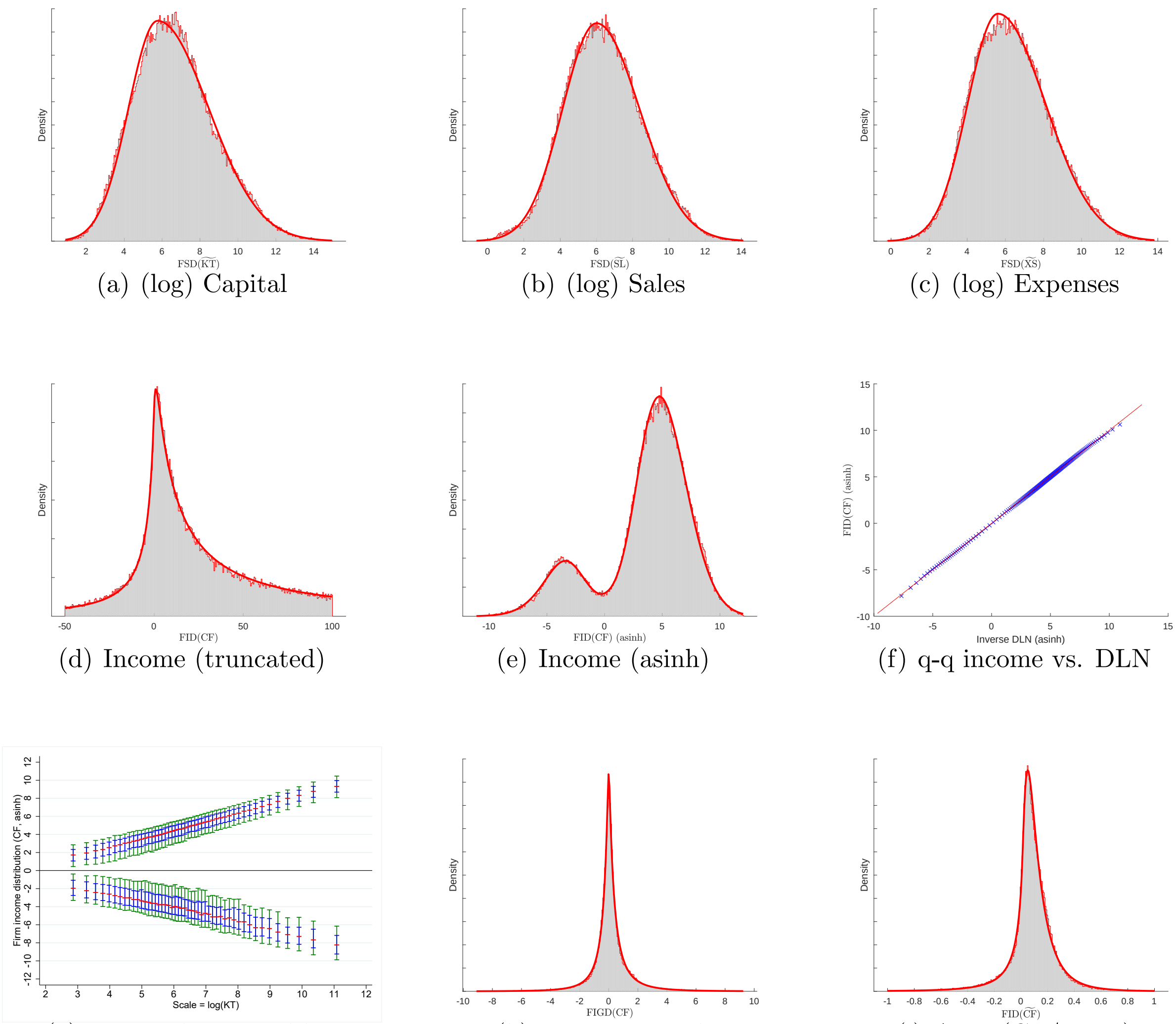
Conversely, here we model income as sales minus expenses using the DLL production function $\mathbf{Y}_{sx}$. The ergodic distributions of s_t and x_t, the stochastic processes governing sales and expenses, are similarly Normal, before considering the impact of exit. This implies that sales and expenses should have (approximately) log-Normal distributions, and in turn, implies that income should distribute as the difference between two correlated log-Normal RVs.

The difference-of-log-Normals (DLN) distribution arises due to a simple set of statistical facts: (i) both the sum and difference of two Normal RVs are generally Normal under mild assumption; (ii) the sum of two log-Normal RVs is best approximated by a log-Normal RV; and (iii) the difference of two log-Normal RVs is decidedly not log-Normal. For one, the log-Normal is strictly positive, while the difference-of-log-Normals is supported on the entire real line $\mathbb{R}$. Further, the DLN exhibits log-Normal (i.e., heavy) tails in both the positive and negative directions, yielding a distributional shape quite different from the Normal "Gaussian bell curve." Parham (2023) describes the emergence of the DLN distribution in general economic data and fully characterizes it, deriving its PDF, CDF, central moments, and estimators for the distribution parameters given data, as well as verifying them in extensive Monte-Carlo simulations.

Figure 3 presents the relevant data distributions. Panels (a)-(c) present the distributions of (log) capital KP, sales SL, and expenses XS in the data, fitted with skew-Normal distributions. The fit is excellent, and three different goodness-of-fit tests do not reject the skew-Normal for these firm outcomes. This result is in accordance with the Normal ergodic distribution of AR(1) processes. The most puzzling thing about Panels (a)-(c) is how unpuzzling they are, given the extensive literature on the distribution of firm size, here shown to simply be skew-Normal (in logs).

Panel (d) of Figure 3 then presents a truncated view of the income distribution, in the limited range between -50M and $+100$M. Income clearly presents exponential tails in both the positive and negative directions, explaining the need for truncation. The common way of dealing with exponential tails, applying a log transform, cannot be used due to the negative values involved. To deal with the double-exponential nature of the tails, Panel (e) then presents the *Inverse Hyperbolic Sine* (asinh) of income, untruncated. The asinh transform can simply be thought of as a log transform, but in both the positive and negative directions, and allows us to view the entire distribution. Panels (d) and (e) are also overlaid with MLE-fitted DLN distributions, exhibiting excellent fit, as do the q-q figure in Panel (f) and the formal goodness-of-fit tests in the CF and CA columns of Table 2. Income is not rejected as DLN using the goodness-of-fit tests. The Stable and Laplace distributions — the other distributions previously considered in the context of firm growth — are rejected by the data. The DLN also handily beats both in log-likelihood-based horse races for income, using the AIC and the BIC.

Finally, panel (g) of Figure 3 presents a closer look at the distribution of income, by considering the dependence of (asinh) income on firm (log) capital KT. I first split the data into 49 equal bins, based on firm capital, ignoring the top and bottom 1% of observations,



(a) (log) Capital (b) (log) Sales (c) (log) Expenses

(d) Income (truncated) (e) Income (asinh) (f) q-q income vs. DLN

(g) Income by capital bin (h) Income growth (i) APK (CF/L.KT)

Fig. 3. Firm size and income distributions. This figure presents firm size and income distributions. Panels (a)-(c) present (log) capital, sales, and expenses, overlaid with skew-Normal distributions. Panel (d) presents the (truncated) distribution of CF in linear scale while Panel (e) presents the untruncated distribution in asinh scale, both overlaid with MLE-fitted DLN distributions. Panel (f) presents the q-q plot corresponding to Panel (e). Panel (g) presents the dependence of income on capital, by presenting the $(10,25,50,75,90)^{th}$ percentiles of asinh(CF), conditional on the sign of CF, for 49 KT scale bins. Panel (h) presents income growth, given by Equation 22, and Panel (i) presents income intensity (average product of capital), both overlaid with MLE-fitted DLN distributions.

Table 2
Distributional tests

This table presents the results of tests of distributional form for income (CF), alt. income (CA), income growth (dCF), APK (CF/L.KT), capital growth (dKT), yearly total value growth adjusted for cash dispensations (dVT), yearly raw equity returns (dEQ_Y), monthly excess equity returns (dEQ_M^{ex}), daily excess equity returns (dEQ_D^{ex}), and daily excess equity returns w/ time and scale f.e. (dEQ_D^{adj}). K-S is a Kolmogorov–Smirnov test; C-2 is a binned χ^2 test with 50 bins; A-D is an Anderson-Darling test. Panels (a)-(c) report the test statistics and their p-values *rejecting the distribution* for the Stable, Laplace, and DLN, respectively. Panel (d) reports the relative likelihoods for each distribution using the AIC and BIC.

	CF	CA	dCF	APK	dKT	dVT	dEQ_Y	dEQ_M^{ex}	dEQ_D^{ex}	dEQ_D^{adj}
Panel (a): Stable										
K-S	0.034	0.038	0.014	0.028	0.011	0.015	0.017	0.015	0.014	0.015
p-val	0.022	0.019	0.046	0.027	0.055	0.044	0.040	0.045	0.046	0.043
C-2	647	>999	120	379	91	179	236	195	201	204
p-val	0.012	0.000	0.035	0.018	0.040	0.028	0.024	0.027	0.027	0.027
A-D	19.48	50.97	5.86	19.99	3.79	6.50	8.54	6.67	6.16	6.70
p-val	0.025	0.014	0.041	0.025	0.047	0.039	0.035	0.039	0.040	0.039
Panel (b): Laplace										
K-S	0.364	0.386	0.073	0.288	0.027	0.017	0.027	0.026	0.051	0.027
p-val	0.000	0.000	0.006	0.000	0.027	0.041	0.028	0.028	0.012	0.027
C-2	>999	>999	>999	>999	290	118	134	244	761	182
p-val	0.000	0.000	0.002	0.000	0.022	0.035	0.033	0.024	0.010	0.028
A-D	>999	>999	105.75	>999	18.72	7.75	9.21	19.38	62.69	17.40
p-val	0.000	0.000	0.007	0.000	0.025	0.037	0.034	0.025	0.012	0.026
Panel (c): DLN										
K-S	0.003	0.003	0.007	0.005	0.006	0.004	0.004	0.003	0.013	0.007
p-val	0.138	0.149	0.074	0.104	0.087	0.108	0.113	0.148	0.050	0.080
C-2	8	12	59	15	17	14	10	5	60	21
p-val	0.142	0.111	0.049	0.096	0.089	0.099	0.123	0.353	0.048	0.078
A-D	0.21	0.18	1.11	0.39	0.50	0.45	0.38	0.10	1.74	0.47
p-val	0.117	0.123	0.070	0.095	0.089	0.091	0.096	0.148	0.061	0.090
Panel (d): Relative likelihood tests										
AIC R.L.:										
Stable	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Laplace	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
DLN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BIC R.L.:										
Stable	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Laplace	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
DLN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

such that each bin contains 2% of the observations. For each bin, Panel (g) plots the $(10, 25, 50, 75, 90)^{th}$ percentiles of (asinh) income, separately for positive and negative values. Larger firms earn and lose more money than smaller firms, so the "middle" of the panel hollows as firm scale rises. This closer look at income is an excellent segue to the next set of model implications regarding firm scale and efficiency.

3.2 Scale and efficiency

An important feature of the DLL production function is that it can be factored into the multiplication of an exponential function and a Hyperbolic Sine (sinh) function — the hyperbolic equivalent of moving from Cartesian to Polar coordinates.[6] We can hence write

$$\mathbf{Y}_{sx}\left(K_t, S_t, X_t\right) = 2 \cdot \exp\left(\lambda_t\right) \cdot \sinh\left(\tau_t\right)$$

$$\lambda_t = \frac{s_t + x_t}{2} + \frac{\theta_s + \theta_x}{2} \cdot k_t = \widehat{\lambda}_t + \theta_\lambda \cdot k_t = \log(\sqrt{\text{Sales} \cdot \text{Expenses}}) \tag{17}$$

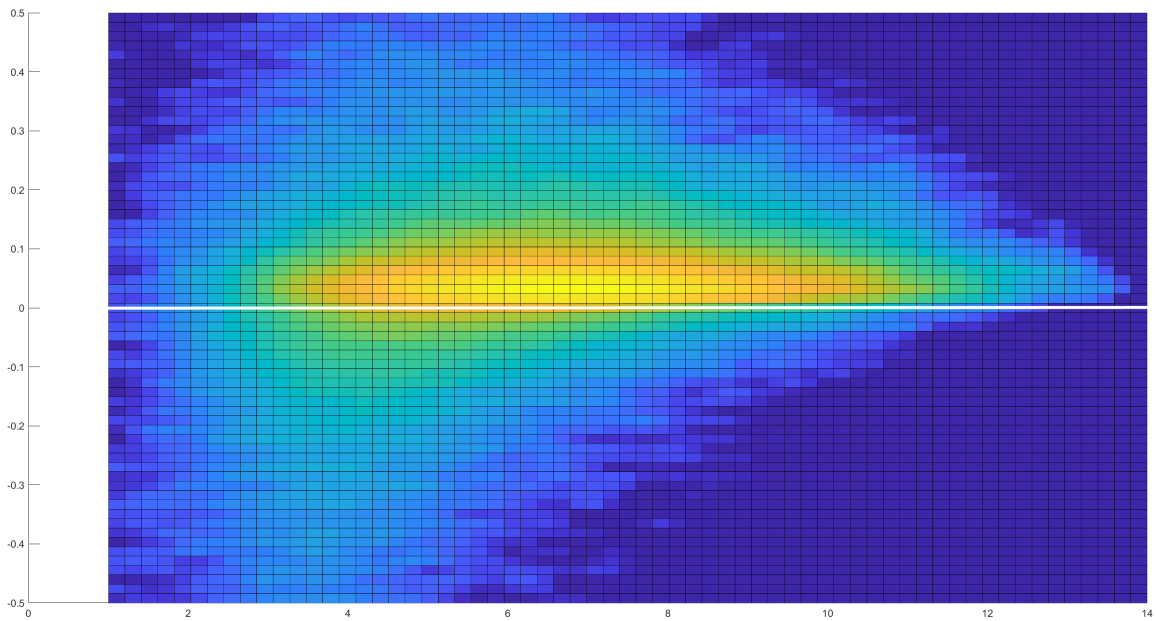
$$\tau_t = \frac{s_t - x_t}{2} + \frac{\theta_s - \theta_x}{2} \cdot k_t = \widehat{\tau}_t + \theta_\tau \cdot k_t = \log(\sqrt{\text{Sales}/\text{Expenses}})$$

which defines the *scale* $\lambda \in \mathbb{R}$ and the *efficiency* $\tau \in \mathbb{R}$ of a firm's income. Note that λ is the mid-point between log sales and log expenses, and τ is the (equal) distance from λ to log sales and log expenses. The inverse mapping is hence sales $= \exp\left(\lambda + \tau\right)$ and expenses $= \exp\left(\lambda - \tau\right)$. Clearly, the sign of firm income depends on the sign of τ, and the magnitude of firm income primarily depends on λ, with a small role for τ.

Figure 4 presents a heat map of the income scale and efficiency for all observations in the data. We can see the vast majority of firm observations (about 86%) have efficiency in the -0.1 to 0.2 range, with a clear ridge around $\tau = 0.033$. Scale is approximately Normally distributed and centered around $\lambda = 6.5$. The profit/loss line at $\tau = 0$ appears

[6] This also rationalizes the ad-hoc use of the asinh transform above.

to significantly impact firms, as we would expect. We can further see that the location (though not the dispersion) of efficiency τ is nearly independent of scale λ. When efficiency is 0, expenses equal sales and income is zero, for all scale values. When scale is $\lambda_t = 6.5$ and efficiency is $\tau_t = 0.033$, the firm has sales of $\exp(6.5 + 0.033) = \$687$M, expenses of $\exp(6.5 - 0.033) = \$644$M, and income of \$43M.



(a) Firm scale and efficiency heat-map

Fig. 4. Firm scale and efficiency. This figure presents a heat map (two-dimensional histogram) of the scale and efficiency of US public firms in the 50-year period 1970-2019. The horizontal axis depicts firm scale $\lambda_t = \log(\sqrt{\text{Sales} \cdot \text{Expenses}})$, and the vertical axis presents firm efficiency $\tau_t = \log(\sqrt{\text{Sales}/\text{Expenses}})$. Zero efficiency (i.e., the profit/loss line) is marked by the white horizontal line.

This decomposition explains the "hollow middle" pattern in panel (g) of Figure 3. Large firms make or lose large amounts of money, but seldom small amounts of money, due to the magnifying power of scale λ. The model further implies firm scale should be correlated and co-integrated with firm (log) capital and other measures of firm scale such as (the logs of) equity value and sales. Panels (a) and (b) of Table 3 show this is indeed the case, with high correlations between the various co-integrated scale measures. We can hence see the commonly used ratio of income to capital (i.e., return on assets ROA, or the average product of capital APK) as a proxy for the easily calculable firm efficiency τ.

Firm efficiency is, in fact, the core difference between the Z- and SX-models. Note that Equation 10, defining the SX production function, collapses into Equation 9, defining the Z production function, if (i) the returns-to-scale in capital are equal for sales and expenses ($\theta_s = \theta_x$). In that case, we can write $Z_t = S_t - X_t$ and return to the formulation in Equation 9, though without the assumption $Z_t > 0$. This assumption can be maintained if we further assume (ii) expenses are always lower than sales ($S_t > X_t\ \forall t$). The canonical Z-model production function implicitly makes both assumptions about the dynamics of the firm.

We can make these assumptions explicit by rewriting the Z-model in Equation 9 to represent income as a share of sales,

$$\mathbf{Y}_{\widetilde{s}}\left(K_t, \widetilde{S}_t\right) = (1-\widetilde{\tau}) \cdot \widetilde{S}_t \cdot K_t^{\theta_{\widetilde{s}}} = \exp\left(\log\left(1-\widetilde{\tau}\right) + \widetilde{s}_t + \theta_{\widetilde{s}} \cdot k_t\right) \tag{18}$$

with expense ratio $0 \leq \widetilde{\tau} \leq 1$. The expense ratio measures how much of firm sales is diffused as expenses, or the expected $\mathbb{X}/\mathbb{S}$ ratio, which is simply a transformation of firm efficiency τ.[7] Put differently, the SX-model can be viewed as endogenizing τ, and allowing $\tau < 0$ or $\widetilde{\tau} > 1$, rather than assuming a fixed value for τ as is done in the Z-models.

3.3 Income growth

The traditional definition of growth (e.g., the difference in consecutive logged values) has hitherto been poorly defined when applied to income due to the existence of negative income values that cannot be logged. Consider: What was the income growth of a firm with \$100M of losses last year and \$120M of profits this year? Parham (2023) extends the instantaneous growth definition of Barro and Sala-I-Martin (2003) to RVs possibly taking negative values, yielding $\frac{dY_t/dt}{|Y_t|}$. The absolute value in the definition of growth is necessary to maintain the

[7] With $\widetilde{\tau} = \exp\left(-2 \cdot \tau\right)$.

Table 3

Scale - Descriptive statistics

Panel (a) presents the correlations between the various firm scale measures (λ and the logs of capital, firm value, equity value, sales, and expenses). Panel (b) presents the results of three cointegration tests between all scale measures, with the first two tests from Pedroni (2004), and the third from Westerlund (2005). The first two test the null of no cointegration vs. the alternative that all panels are cointegrated while the third tests the null vs. the alternative that some panels are cointegrated. Tests are conducted by decade, on the available balanced sample of firms within each decade. Panel (c) presents regressions of income growth (CF_{t+1} - CF_t)/abs(CF_t) on changes in firm scale $d\lambda = \lambda_{t+1} - \lambda_t$, changes in firm efficiency $d\tau = \tau_{t+1} - \tau_t$, and percent changes in firm efficiency $\%\tau = (\tau_{t+1} - \tau_t)/\tau_t$. All regressions include firm and year fixed-effects, w/ N=165K.

Panel (a): Scale correlations

	λ	KT	VL	EQ	SL	XS
λ	—	.929	.880	.797	.995	.995
KT	.929	—	.961	.883	.928	.921
VL	.880	.961	—	.960	.878	.874
EQ	.797	.883	.960	—	.796	.792
SL	.995	.928	.878	.796	—	.982
XS	.995	.921	.874	.792	.982	—

Panel (b): Scale cointegration tests

	Phillips-Perron t	p-val	Dicky-Fuller t	p-val	Variance ratio	p-val
70's	51.12	<0.001	-65.76	<0.001	17.42	<0.001
80's	57.38	<0.001	-57.41	<0.001	21.32	<0.001
90's	58.97	<0.001	-65.87	<0.001	22.05	<0.001
00's	62.51	<0.001	-75.13	<0.001	21.67	<0.001
10's	58.97	<0.001	-60.87	<0.001	22.57	<0.001

Panel (c): determinants of income growth

	(1)	(2)	(3)	(4)
$d\lambda$	2.264	-5.19		
s.e	.8550	3.920		
$d\tau$	16.14		21.68	
s.e	1.430		6.557	
$\%\tau$	.5036			.5036
s.e	.0007			.0007
within-R^2	.8022	.0000	.0001	.8020

direction of growth when beginning from negative values (i.e., positive growth will lead to more profit or at least fewer losses).

The first way to use this equation is to apply it to firm income directly, measuring the generalized percentage growth in income

$$\frac{dY_t/dt}{|Y_t|} \approx (Y_{t+1} - Y_t)/|Y_t| \tag{19}$$

with the approximation stemming from using the forward discrete difference for the time derivative. A second way is to assume Y follows the LL production function (i.e., the Z-model), in which case $Y_t \geq 0$ and

$$\frac{dY_t/dt}{|Y_t|} = \frac{Y_t \cdot \left(\frac{dz_t}{dt} + \theta_z \cdot \frac{dk_t}{dt}\right)}{Y_t} \approx (z_{t+1} + \theta_z \cdot k_{t+1}) - (z_t + \theta_z \cdot k_t) = \log(Y_{t+1}) - \log(Y_t) \tag{20}$$

yielding the familiar difference-in-logs growth measure.

The Z-model implies income growth is approximately Normally distributed. The log-point growth of income in the Z-model can be written as

$$\text{dlog}(Y_{t+1}) = \log(Y_{t+1}) - \log(Y_t) = (1 - \rho_z) \cdot (\mu_z - z_t) + \theta_z \cdot (k_{t+1} - k_t) + \epsilon^z_{t+1} \tag{21}$$

with z_t distributing Normally as a property of the AR(1) process, ϵ^z_{t+1} distributing Normally by Equation 12, and $\text{dlog}(K_{t+1}) = k_{t+1} - k_t$ difficult to pin down analytically in the general case. But prior work, as well as steady-state analysis and simulation results discussed below, indicate $\text{dlog}(K_{t+1})$ distributes Normally as well in the Z-model. This implies income growth is Normally distributed in the Z-model. Note the $\text{dlog}(Y_{t+1})$ income growth measure derived from the Z-model's LL production function fails when one of the periods has negative income.

The third way to define income growth is to assume Y follows the SX-model's DLL production function, in which case

$$\frac{dY_t/dt}{|Y_t|} \approx \frac{\mathbb{S}_t \cdot \text{dlog}\left(\mathbb{S}_{t+1}\right) - \mathbb{X}_t \cdot \text{dlog}\left(\mathbb{X}_{t+1}\right)}{|\mathbb{S}_t - \mathbb{X}_t|} \tag{22}$$

which expresses income growth as a weighted average of sales growth and expenses growth. A fourth way, equivalent to the third but yielding considerably more intuition, is to define income growth using scale and efficiency, in which case

$$\begin{aligned} \frac{dY_t/dt}{|Y_t|} &= \frac{2 \cdot \exp\left(\lambda_t\right) \cdot \left(\frac{d\lambda_t}{dt} \cdot \sinh\left(\tau_t\right) + \frac{d\tau_t}{dt} \cdot \cosh\left(\tau_t\right)\right)}{2 \cdot \exp\left(\lambda_t\right) \cdot |\sinh\left(\tau_t\right)|} \\ &= \text{sgn}(\tau_t) \cdot \left[\frac{d\lambda_t}{dt} + \frac{d\tau_t}{dt} \cdot \frac{1}{\tanh\left(\tau_t\right)}\right] \approx \text{sgn}(\tau_t) \cdot \left[\left(\lambda_{t+1} - \lambda_t\right) + \frac{\tau_{t+1} - \tau_t}{\tau_t}\right] \end{aligned} \tag{23}$$

with sgn() the sign function. The approximation is now due to two reasons: the forward discrete difference, as usual, and replacing $\tanh(\tau_t)$ with τ_t, which is valid because firm efficiency in the data is clustered tightly in the region where $\tanh(\tau_t) \approx \tau_t$.

Put differently, the expression for the growth of firm income in the SX-model includes both an expression for log-point growth in firm scale and an expression for the *percent growth* in firm efficiency, added to the log-point growth in scale. Explosive income growth (or the heavy tails of income growth) occurs due to operational leverage, or a low "base rate" in τ (i.e., τ_t close to zero), leading to high measured growth in income. The correlation in the data between the SX-based growth measure from Equation 23 and the generalized percentage growth of income from Equation 19 is above 0.97. Furthermore, nearly all variation in income growth in the data stems from the dynamics of τ and the percentage growth in τ term, rather than from the dynamics of λ, as Panel (c) of Table 3 unequivocally demonstrates.

As an example of the impact of operational leverage, consider a firm with $\$1B$ in sales and $\$950M$ in expenses during period t. Firm scale is then $\lambda_t = 6.88$ and firm efficiency is $\tau_t = 0.026$, both close to the median values observed in the data. First, assume that in period $t+1$ the firm increases both sales and expenses by 10% to $\$1.1B$ and $\$1.045B$, respectively. This means $\lambda_{t+1} = 6.98$, 0.1 log-units higher, and $\tau_t = 0.026$ is the same. Equation 23 will yield income growth of 0.1, the same as percentage income growth $55/50-1 = 10\%$. Alternatively, assume that in period $t+1$ the firm increases sales by 10% to $\$1.1B$, but decreases expenses by 10% to $\$855M$. This means now $\lambda_{t+1} = 6.88$, the same as λ_t, but $\tau_{t+1} = 0.126$ is 0.1 log-units higher. Equation 23 yields income growth of $(0.126-0.026)/0.026 = 3.9$ log-units, equal to the percent growth of income at $245/50-1 = 390\%$. With income exhibiting heavy-tailed growth, we would expect firm value to exhibit heavy-tailed growth as well. Firm value is simply the NPV of future income, so rapid growth in income should propagate to rapid growth in value.

It is difficult to analytically pin down the heavy-tailed distribution of income growth resulting from the SX-model. Nevertheless, Panel (h) of Figure 3 presents the distribution of income growth, as defined by Equation 23, in the data along with an MLE-fitted DLN. It is easy to see that income growth is not Normally distributed, while the fit to the DLN distribution is excellent, as the dCF column of Table 2 confirms. A horse race between the DLN, Stable, and Laplace again decisively favors the DLN.

3.4 Returns-to-scale

What are the returns-to-scale (RTS) implications of the different production functions? The RTS of income w.r.t capital is simply defined in terms of the elasticity of $\mathbf{Y}()$ w.r.t K, or the marginal product of capital (MPK) relative to the average product of capital (APK).

First, applying this to the Z-model production function $\mathbf{Y}_z()$ yields

$$\text{RTS}^z = \frac{\partial \mathbf{Y}\left(K_t, Z_t\right)}{\partial K_t} \Big/ \frac{\mathbf{Y}\left(K_t, Z_t\right)}{K_t} = \theta_z \tag{24}$$

or the well-known result that all firms, regardless of their state, always have $\text{RTS}^z = \theta_z$.

Applying the same definition to $\mathbf{Y}_{sx}()$, in contrast, yields

$$\mathrm{RTS}^{sx} = \frac{\theta_s \cdot \mathbb{S}_t - \theta_x \cdot \mathbb{X}_t}{\mathbb{S}_t - \mathbb{X}_t} \tag{25}$$

Importantly, the model no longer implies constant RTS for all firms. Firms may have different RTS^{sx} depending on their current sales and expenses, even if all firms in the economy share the same θ_s and θ_x parameters. Note that both the numerator and denominator in the definition of RTS^{sx} are DLN-distributed, as they are weighted differences of log-Normally distributed values (sales and expenses).

Better intuition can again be gleaned by equivalently writing Equation 25 in terms of λ, τ using the $\mathbf{Y}_{\lambda\tau}$ production function. In this case, we have:

$$\mathrm{RTS}^{\lambda\tau} = \theta_\lambda + \frac{\theta_\tau}{\tanh(\tau_t)} \approx \theta_\lambda + \frac{\theta_\tau}{\tau_t} \tag{26}$$

with $\theta_\lambda, \theta_\tau$ given by Equation 17. The model proposes a two-part schedule for RTS: a constant term and a term inversely related to efficiency τ. Because $\mathrm{RTS}^{\lambda\tau}$ explodes to $\pm\infty$ when $|\tau_t| \to 0$, we again have a base-rate effect, similar to the base-rate effect in income growth above. The significant mass of firms around $\tau = 0$ then implies a heavy-tailed distribution of RTS in the data.

While RTS is unobservable, due to the unobservability of MPK, we can nevertheless observe its other component, the average product of capital APK. Because CF is DLN and KT is approximately log-Normal, we can predict APK to be DLN as well. This is because dividing a DLN RV by a log-Normal RV yields another DLN RV. That APK is DLN is confirmed in Panel (i) of Figure 3 and in the APK column of Table 2.

3.5 Capital and value growth

What are the model implications on the growth in firm capital and on the growth in firm value (i.e., buy-and-hold returns when also accounting for cash dispensations)? Because the model lacks closed-form solutions to the value and policy function, it is difficult to ascertain those directly. Despite that, some progress can be made in two avenues: by considering the simplified case in which $\gamma \to 0$, and via model estimation and simulation. I consider the first method here before moving on to simulations in the next section.

For the simplified case of $\gamma \to 0$, MPK is the core determinant of firm size and growth. The firm simply sets its next period capital K_{t+1} to maintain $\mathbb{E}_t\left[\text{MPK}\right] = r + \delta$, regardless of K_t, and all capital growth just follows MPK changes. For the Z-model,

$$\begin{aligned} MPK_t &= \theta_z \cdot \exp(z_t - (1-\theta_z)\cdot k_t) \\ K_{t+1} &= \exp\left(\frac{C_z + \rho_z \cdot z_t}{1-\theta_z}\right) \\ C_z &= \log\left(\frac{\theta_z}{r+\delta}\right) + (1-\rho_z)\cdot \mu_z + \frac{\sigma_z^2}{2} \\ \text{dlog}(K_{t+1}) &= \frac{\rho_z}{1-\theta_z}\cdot(z_{t+1} - z_t) \end{aligned} \tag{27}$$

with C_z a constant depending on the parameters of the model. Capital growth in this case is Normal, based on the assumption regarding the Normal dynamics of z in Section 2.4.

For the SX-model, we have

$$\begin{aligned} &MPK_t = \theta_s \cdot \exp\left(s_t - (1-\theta_s)\cdot k_t\right) - \theta_x \cdot \exp\left(x_t - (1-\theta_x)\cdot k_t\right) \\ &\underbrace{\exp(C_s + \rho_s \cdot s_t - (1-\theta_s)\cdot k_{t+1})}_{g_s(s_t, k_{t+1})} - \underbrace{\exp(C_x + \rho_x \cdot x_t - (1-\theta_x)\cdot k_{t+1})}_{g_x(x_t, k_{t+1})} = 1 \end{aligned} \tag{28}$$

with C_s and C_x defined analogously to C_z, and the second equation the non-separable equation defining K_{t+1} in terms of s_t, x_t and the model parameters. The equation determines K_{t+1} such that the LHS, itself a DLL function, equals 1. Depending on the values of s_t, x_t and the model parameters, this equation may not have a solution, implying no capital level is capable of equating MPK to $r + \delta$ because the firm is losing money even in tiny scales (e.g., $k_t < 0$). Nevertheless, we can use the implicit function theorem to write capital growth, when it exists, as

$$\begin{aligned}\frac{dK_t/dt}{K_t} &= \frac{\frac{\partial K_t}{\partial s_t}\cdot\frac{ds_t}{dt}+\frac{\partial K_t}{\partial x_t}\cdot\frac{dx_t}{dt}}{K_t} \\ &\approx \frac{\rho_s\cdot g_s(s_t,k_{t+1})\cdot(s_{t+1}-s_t)-\rho_x\cdot g_x(x_t,k_{t+1})\cdot(x_{t+1}-x_t)}{(1-\theta_s)\cdot g_s(s_t,k_{t+1})-(1-\theta_x)\cdot g_x(x_t,k_{t+1})}\end{aligned} \tag{29}$$

with the approximation due to the forward discrete difference, as usual. Both the first and second terms of the numerator, as well as the entire denominator, are again DLL functions. This (weakly) indicates that capital growth is heavy-tailed in the SX-model, especially when the denominator $\to 0$.

With both capital and income exhibiting heavy-tailed growth, we would expect firm value in the SX-model to exhibit heavy-tailed growth as well. This is simply because firm value is the NPV of future income, so rapid growth in income should propagate to rapid growth in value. Figure 1 has already presented the distributions of capital and value growth in the data, showing that they appear DLN-distributed. Table 2 reports the formal distributional tests. For the growth in capital and for five measures of value growth (or equivalently, returns) the DLN is not rejected, and it wins every horse race against the other distributional candidates. The five value growth measures are yearly change in firm value adjusted for dispensations, yearly raw equity returns, monthly excess returns from an FF3 model, daily excess returns from an FF3 model, and daily excess returns w/ time and scale fixed effects.

4 Estimation and simulation of models

This section describes taking four models to data via indirect inference. The models are the Z-, $\widetilde{\text{S}}$-, SX-, and $\lambda\tau$-models, with the first and second two models using the LL and DLL production functions, respectively. Estimation is relatively straightforward, owing to the observability of sales, expenses, and capital. I then simulate the estimated models and consider the distributions of firm outcomes in the models vs. the data. The DLL-based models replicate the DLN-distributed firm outcomes and fit several important firm moments considerably better than the LL-based models.

4.1 Identification and initial parameter values

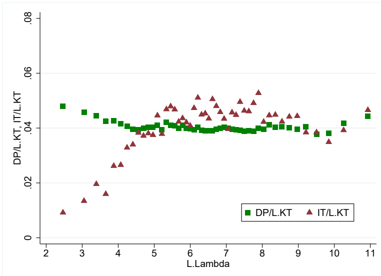
The Z-, $\widetilde{\text{S}}$-, SX-, and $\lambda\tau$-models are described by the respective parameter vectors

$$\begin{aligned}
\Theta_z &= \{r, \delta, \theta_z, \rho_z, \mu_z, \sigma_z, \gamma, \nu\} \\
\Theta_{\widetilde{s}} &= \{r, \delta, \theta_{\widetilde{s}}, \rho_{\widetilde{s}}, \mu_{\widetilde{s}}, \sigma_{\widetilde{s}}, \widetilde{\tau}, \gamma, \nu\} \\
\Theta_{sx} &= \{r, \delta, \theta_s, \rho_s, \mu_s, \sigma_s, \theta_x, \rho_x, \mu_x, \sigma_x, \rho_{sx}, \gamma, \nu\} \\
\Theta_{\lambda\tau} &= \{r, \delta, \theta_\lambda, \rho_\lambda, \mu_\lambda, \sigma_\lambda, \theta_\tau, \rho_\tau, \mu_\tau, \sigma_\tau, \rho_{\lambda\tau}, \gamma, \nu\}
\end{aligned} \tag{30}$$

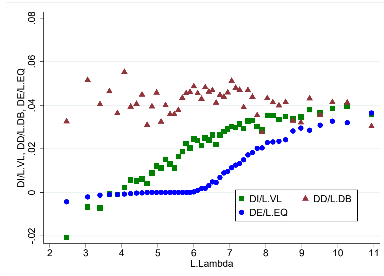
which we aim to estimate using the method of simulated moments. This task is considerably simplified by noting the following three facts: (i) r, δ are relatively easy to pin down; (ii) initial guesses for the $\theta_\square$ values $(\theta_z, \theta_{\widetilde{s}}, \theta_s, \theta_x, \theta_\lambda, \theta_\tau)$ can be derived from steady-state arguments regarding returns-to-scale (RTS), conditional on r, δ; and (iii) $z_t, \widetilde{s}_t, s_t, x_t, \widehat{\lambda}_t, \widehat{\tau}_t$ are observable, conditional on $\theta_\square$, allowing us to estimate their dynamic parameters directly.

Pinning down δ is easy because firms generally report their depreciation expenses. Panel (a) of Figure 5 presents the binned median depreciation rate DP/L.KT and investment rate IT/L.KT in the data as functions of firm scale L.λ. Throughout nearly the entire scale distribution, both are tightly packed around 0.04. I hence set $\delta = 4\%$.

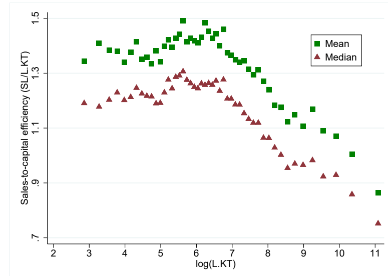
Pinning down r is slightly more complicated and raises curious questions. A good guess for r, the firm's cost of capital, is the median payout ratio DI/L.VL. Panel (b) of Figure 5 presents the binned median ratio in the data. It also presents the binned medians of its components, the debt payout ratio DD/L.DB and the equity payout ratio DE/L.EQ, all as functions of firm scale. The median payout ratio is smoothly increasing from 0 to 4%, the median debt payout ratio is roughly constant at 4%, and the median equity payout ratio is flat at 0 for the lower half of the scale distribution and then smoothly rises up to 4% for the upper half. Panel (c) presents the mean and median of the sales-to-capital ratio SL/L.KT, by scale, to complete the picture.



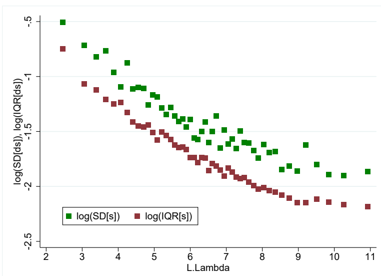
(a) depreciation



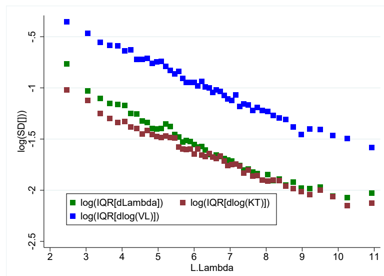
(b) dispensations



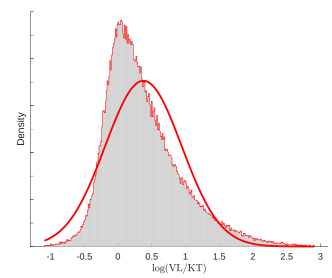
(c) sales



(d) DDWS s growth



(e) DDWS firm growth



(f) log average-q

Fig. 5. Stylized facts of firm scale. For each of 49 scale bins, panel (a) presents the binned median of depreciation rate DP/L.KT and investment rate IT/L.KT. Panel (b) presents the binned medians of total, debt, and equity dispensation rates DI/L.VL DD/L.DB DE/L.EQ, respectively. Panel (c) presents the binned mean and median of sales intensity SL/L.KT. Panel (d) presents the binned log(SD[]) and log(IQR[]) of the growth in the stochastic variable s, controlling sales, while Panel (e) repeats for log(IQR[]) of the growth in firm scale λ, assets KT, and value VL. Panel (f) presents the distribution of log average-q log(VL/KT), overlaid with a fitted Normal.

The stylized facts in Panels (b) and (c) of Figure 5 are somewhat puzzling. First, the slope of median DI/L.VL is the opposite of that implied by the SMB factor — large firms have a higher cost of capital than small firms (though the portfolios are not tradable, and this finding might be spurious). Second, the finding regarding a structural break around the median scale of 6.5 (or $\approx 660M$ in 2019 \$) is, to my knowledge, new to the literature. Nevertheless, exploring these scale-dependent stylized facts steers us away from our main interest and is left for future work, discussed in the concluding remarks. In light of the evidence in Figure 5, I set $r = 2\%$, which is both approximately the median DI/L.VL in the data and the median DI/L.VL for firms around median scale. Setting $r = 4\%$ yields similar qualitative conclusions.

Moving on to initial guesses for $\theta_\square$ — note that for firms close to steady state, or when $\gamma \to 0$, we have MPK $= r + \delta$ from Equation 7. Hence, we can write:

$$RTS^* = \frac{r+\delta}{\text{APK}} = \begin{cases} \theta_z & \text{if } Y_t = \mathbf{Y}_z() \\ \theta_{\widetilde{s}} & \text{if } Y_t = \mathbf{Y}_{\widetilde{s}}() \\ \frac{\theta_s \cdot \mathbb{S}_t - \theta_s \cdot \mathbb{X}_t}{Y_t} & \text{if } Y_t = \mathbf{Y}_{sx}() \\ \theta_\lambda + \frac{\theta_\tau}{\tanh(\tau_t)} & \text{if } Y_t = \mathbf{Y}_{\lambda\tau}() \end{cases} \tag{31}$$

Because APK is observable, We can use this equation to estimate initial guesses for $\theta_\square$, when considering firms plausibly in steady-state. Doing so yields initial values: $\theta_z = \theta_{\widetilde{s}} = 0.7, \theta_s = 0.28, \theta_x = 0.25, \theta_\lambda = 0.268, \theta_\tau = 0.015$.

The marked difference between the initial RTS guesses of the LL-based models and the DLL-based models is notable. It stems from the fact the median RTS of around 0.7, a much-used value in the relevant literature, arises from the *interaction* between the lower RTS of sales and expenses. The two-part schedule in the $\lambda\tau$-model at the last line of Equation 31 is particularly useful in seeing this. With τ in the data clustered around 0.033, the two-part schedule and the estimates above imply the typical firm has an RTS = 0.268 + 0.015/0.033 = 0.72, close to the initial value found for θ_z. The dynamics of τ are thus critical to understanding the dynamics of RTS in the data.

Finally, the quasi-observability of $z_t, \widetilde{s}_t, s_t, x_t, \widehat{\lambda}_t, \widehat{\tau}_t$ can be seen by rewriting the definitions of $\mathbf{Y}_z()$, $\mathbf{Y}_{\widetilde{s}}()$, $\mathbf{Y}_{sx}()$ and $\mathbf{Y}_{\lambda\tau}()$ in Equations 9, 10, and 18 as

$$\begin{aligned}
z_t &= \log\left(Y_t\right) - \theta_z \cdot \log\left(K_t\right) \\
\widetilde{s}_t &= \log\left(\left(1-\tilde{\tau}\right) \cdot \mathbb{S}_t\right) - \theta_{\widetilde{s}} \cdot \log\left(K_t\right) \\
s_t &= \log\left(\mathbb{S}_t\right) - \theta_s \cdot \log\left(K_t\right) \\
x_t &= \log\left(\mathbb{X}_t\right) - \theta_x \cdot \log\left(K_t\right) \\
\widehat{\lambda}_t &= \lambda_t - \theta_\lambda \cdot \log\left(K_t\right) \\
\widehat{\tau}_t &= \tau_t - \theta_\tau \cdot \log\left(K_t\right)
\end{aligned} \tag{32}$$

and noting that income, sales, expenses, scale, efficiency, and capital $(Y_t, \mathbb{S}_t, \mathbb{X}_t, \lambda_t, \tau_t, K_t)$ are all observable.

Two main challenges arise when using this equation. The first challenge is when taking the log of income in the imputation of z_t, because about 20% of income observations in the data are negative (i.e., losses). The log of income is undefined, z_t cannot be imputed for these observations, and I ignore them in the Z-model. The $\widetilde{\mathrm{S}}$-model is included in the estimation as a second way of overcoming this challenge because it derives its stochastic dynamics from sales, a strictly positive value, and merely explicitly assumes a fixed τ or $\widetilde{\tau}$.

Using the imputed values for the stochastic variables, we can estimate initial guesses for the parameters controlling their dynamics, namely $\rho_\square, \mu_\square, \sigma_\square$. The estimation of $\sigma_\square$, the standard deviation of the innovations to each stochastic variable, brings us to the second challenge: the decreasing dispersion (i.e. heteroscedasticity) with scale (DDWS) of the innovations to the stochastic variables. Note that for, e.g., s_t, the stochastic dynamics in Section 2.4 yield $\mathrm{SD}\left[s_{t+1} - \rho_s \cdot s_t\right] = \mathrm{SD}\left[\epsilon^s_{t+1}\right] \equiv \sigma_s \forall s$, with SD[] the standard deviation operator. Put differently, the model assumes innovations to the stochastic variables are homoscedastic.

Panel (d) of Figure 5 presents the binned log(SD[]) and log(IQR[]) of the innovations to the stochastic variable controlling sales s, by scale, using the initial guess for θ_s above to extract s. The panel presents systematic DDWS, which also applies to innovations in λ and to the outcome variables of the firm, the growth in capital and the growth in value, whose binned log(IQR[]) are presented in Panel (e). Here again, exploring these scale-dependent stylized facts steers us away from our main interest and is left for future work. In light of the evidence in Panels (d) and (e) of Figure 5, I set $\sigma_{\square}$ to match the dispersion of innovations around the median scale in the data, $\lambda = 6.5$. Finally, I was unable to find a natural initial guess for the adjustment cost parameter γ. Prior work (cited above) generally finds fairly low values for γ, around 0.01-0.1.

The two remaining parameters are set to constants and are not estimated. The parameter $\widetilde{\tau}$ of the $\widetilde{\mathrm{S}}$-model is set to a value of exp(-2*0.033) = 0.936, the (transformed) value of the efficiency "ridge" in Figure 4. The parameter ν, controlling exit in all models, determines the average-q VL/KT at which firms exit because firms with $V_t < V_t^{exit} = \nu \cdot K_t$ will find it more profitable to exit. Throughout the analysis, I present and use the log of average-q, log(VL/KT) rather than "simple" average-q VL/KT, because we already established that both value and capital are approximately log-Normal. The ratio of two log-Normal RVs is itself log-Normal, implying a less distorted way of measuring the highly skewed and always positive average-q is measuring it in log terms. The distribution of (log) average-q in the data is presented in Panel (f). The censoring below log(VL/KT)=0 (i.e., VL=KT) is evident, leading to a deviation from the predicted Normal shape. I set ν to the median average-q conditional on it being < 1, which is 0.85 in the data.

The recipe for initial guesses is also used to determine which moments we should match in the estimation procedure. For θ_*, I use the median of RTS* from Equation 31, along with its IQR for the SX-models. Note this is equivalent to matching the location and dispersion of APK. For the AR(1) dynamics parameters, I use their direct data counterparts (or the heavy-tail-robust versions thereof) as identifying moments. Finally, to identify the adjustment cost parameter γ, I match the persistence of capital growth dlog(K_{t+1}) between periods. When $\gamma \to 0$, firms immediately adjust to the optimal capital level every period and hence capital is independent between periods, leading to zero capital growth persistence. But as γ increases, firms adjust slowly towards their optimal capital level and we observe increasing capital growth persistence. The initial values for γ are set such that this persistence is matched when holding all other parameters at their initial values.

4.2 Estimation

The estimation procedure is two-step: in the first step, I guess (i.e., grid-search) $\theta_{\Box}$ values, and in the second step, I conduct a full method of simulated moments (MSM) estimation conditional on the $\theta_{\Box}$ values. I then choose the parameter values minimizing the Mahalanobis distance between the simulated and data moments. The two-step procedure is necessary due to our reliance on the observability of the stochastic variates and their dependence on $\theta_{\Box}$, leading to a dependence of their moments on $\theta_{\Box}$ as well. Table 4 summarizes the estimation and presents for each model: the initial and estimated parameter values; the identifying moments at the initial and estimated parameter values; the identifying moments in the data; and the t-value on the difference between the data and simulated moments at the estimated parameters. Throughout, the model uses the robust estimates of scale and dispersion, median MED[] and inter-quartile range IQR[], with the IQR divided by 1.35 to make it comparable to the standard deviation of a Normal distribution.

Panel (a) presents the estimation results for the Z-model. While the initial guess for θ_z is 0.7, it leads (in conjunction with the other initial values dependent on it) to a higher median RTS* than in the data. Lowering θ_z to around 0.6 enables the model to replicate the identifying data moments very well. Note that the estimated values for the parameters controlling the AR(1) dynamics of z are identical to their data counterparts and to the simulated values, implying quasi-observability works well as an identification strategy. The capital adjustment cost parameter, γ, is estimated to be 0.016 and allows the model to replicate the persistence of capital growth — its identifying moment.

Recall that the Z-model is difficult to work with due to the existence of negative income.

Results of estimating the $\widetilde{S}$-model, a flavor of the Z-model which overcomes this problem, are presented in Panel (b). Here, the initial value for $\theta_{\widetilde{s}}$ overshoots MED[RTS*] only mildly, and a slight decrease from 0.7 to 0.685 is sufficient to match this identifying moment. The AR(1) parameters are again identical to their data counterparts and their simulated values. The persistence of $\widetilde{s}$ is much higher than that of z, and the standard deviation is much lower. The $\widetilde{S}$-model matches the identifying moments very well too.

Panel (c) moves on to the SX-model. A main challenge of estimating the SX-model is that s and x, the stochastic variables controlling sales and expenses, are highly correlated in the data (with correlation coeff. > 0.95), although their innovations are less correlated (around 0.5). Estimating models with highly correlated stochastic variates is notoriously difficult. While the SX-model is able to replicate most of its identifying moments very well, it is far from capturing the dispersion of RTS*, one of the two moments identifying θ_s, θ_x.

This difficulty is resolved by considering the $\lambda\tau$-model, the flavor of the SX-model which tracks $\widehat{\lambda} = (s + x)/2$ and $\widehat{\tau} = (s - x)/2$. This transformation naturally resolves the problem of high correlation between s and x, as $\widehat{\lambda}$ and $\widehat{\tau}$ are nearly uncorrelated. Panel (d) of Table 4 presents the estimation results for the $\lambda\tau$-model, showing that it matches its identifying moments very well. The estimated value for θ_λ is slightly higher than its initial value (0.3 vs. 0.27), but with the estimated values the model matches both the location and dispersion of RTS*, the identifying moments for $\theta_\lambda, \theta_\tau$. The AR(1) coefficients are again well-identified and matched to their data and simulation counterparts, with the exception of μ_λ, estimated to a value of 2.9 vs. a value of 4.7 in the data and simulation. The source of this discrepancy appears to be exit-induced selection — firms with low $\widehat{\lambda}$ exit, such that the ergodic distribution of $\widehat{\lambda}$ of remaining firms matches the data. The capital adjustment parameter γ is somewhat higher at 0.06 but still within the 0.01-0.1 range of previous works, and the persistence of growth is again well-matched by the model.

Table 4
Estimation results

Panels (a)-(d) present the results of estimating the Z-model, $\widetilde{\text{S}}$-model, SX-model, and $\lambda\tau$-model, respectively. Init is the initial parameter guess. Estim is the estimated value for the parameter. Moment is the corresponding identifying moment used in estimation. Dmom is the value of the moment in the data, Imom is its simulated value at Θ = Init, and Smom is its simulated value at Θ = Estim. t-val is the t-statistic on (Dmom-Smom). MED, IQR, RHO, and COR are the median, inter-quartile range (divided by 1.35), persistence, and correlation operators. The stochastic variables are defined by Equation 32.

Panel (a): Z-model

		Value at:			Value at:			
	Name	Init	Estim	Moment	Init	Estim	Data	t-val
θ_z	RTS	0.699	0.592	MED[RTS*][a]	0.825	0.693	0.699	0.755
ρ_z	z pers.	0.691	0.771	RHO[z][a]	0.691	0.772	0.771	-0.133
μ_z	z mean	-0.185	0.542	MED[z][a]	-0.185	0.543	0.542	-0.098
σ_z	dz std.	0.430	0.424	IQR[dz][b]	0.430	0.423	0.424	0.282
γ	Cap. adj.	0.018	0.016	RHO[dk][b]	0.301	0.300	0.295	-0.486

Panel (b): $\widetilde{\text{S}}$-model

		Value at:			Value at:			
	Name	Init	Estim	Moment	Init	Estim	Data	t-val
$\theta_{\widetilde{s}}$	RTS	0.699	0.685	MED[RTS*][a]	0.713	0.699	0.699	-0.021
$\rho_{\widetilde{s}}$	$\widetilde{s}$ pers.	0.956	0.959	RHO[$\widetilde{s}$]	0.957	0.959	0.959	-0.307
$\mu_{\widetilde{s}}$	$\widetilde{s}$ mean	2.141	2.232	MED[$\widetilde{s}$]	2.141	2.232	2.232	-0.037
$\sigma_{\widetilde{s}}$	$d\widetilde{s}$ std.	0.127	0.126	IQR[$d\widetilde{s}$][b]	0.127	0.126	0.126	0.243
γ	Cap. adj.	0.006	0.006	RHO[dk][b]	0.303	0.298	0.295	-0.356

[a] For firms with $dk \in \text{IQR}[dk]$ and $CF \geq 1$.
[b] For firms around median scale $\lambda \in \text{IQR}[\lambda]$.

Table 4
Estimation results

Panel (c): SX-model

	Name	Value at: Init	Estim	Moment	Value at: Init	Estim	Data	t-val
θ_s	S RTS	0.283	0.616	MED[RTS*]^a	0.301	0.692	0.698	0.635
θ_x	X RTS	0.252	0.583	IQR[RTS*]^a	0.017	0.118	0.359	21.500
ρ_s	s pers.	0.990	0.968	RHO[s]	0.990	0.968	0.968	-0.211
μ_s	s mean	4.832	2.676	MED[s]	5.648	2.687	2.676	-0.804
σ_s	ds std.	0.118	0.123	IQR[ds]^b	0.118	0.123	0.123	0.042
ρ_x	x pers.	0.982	0.957	RHO[x]	0.982	0.957	0.957	-0.266
μ_x	x mean	4.938	2.797	MED[x]	4.824	2.791	2.797	0.362
σ_x	dx std.	0.127	0.125	IQR[dx]^b	0.127	0.125	0.125	0.018
ρ_{sx}	ds, dx cor	0.493	0.491	COR[ds, dx]	0.494	0.491	0.491	0.039
γ	Cap. adj.	0.012	0.012	RHO[dk]^b	0.208	0.289	0.295	0.679

Panel (d): $\lambda\tau$-model

	Name	Value at: Init	Estim	Moment	Value at: Init	Estim	Data	t-val
θ_λ	λ RTS	0.268	0.302	MED[RTS*]^a	0.560	0.689	0.698	0.636
θ_τ	τ RTS	0.015	0.016	IQR[RTS*]^a	0.223	0.356	0.359	0.253
ρ_λ	$\widehat{\lambda}$ pers.	0.990	0.989	RHO[$\widehat{\lambda}$]	0.990	0.989	0.989	0.113
μ_λ	$\widehat{\lambda}$ mean	4.882	2.866	MED[$\widehat{\lambda}$]	5.627	4.698	4.674	-1.238
σ_λ	$d\widehat{\lambda}$ std.	0.119	0.118	IQR[$d\widehat{\lambda}$]^b	0.119	0.118	0.118	0.065
ρ_τ	$\widehat{\tau}$ pers.	0.563	0.562	RHO[$\widehat{\tau}$]	0.562	0.553	0.562	0.708
μ_τ	$\widehat{\tau}$ mean	-0.068	-0.074	MED[$\widehat{\tau}$]^b	-0.068	-0.073	-0.074	-1.026
σ_τ	$d\widehat{\tau}$ std.	0.022	0.022	IQR[$d\widehat{\tau}$]^b	0.022	0.022	0.022	0.194
$\rho_{\lambda\tau}$	$d\widehat{\lambda}, d\widehat{\tau}$ cor	-0.126	-0.123	COR[$\widehat{\lambda}, \widehat{\tau}$]	-0.127	-0.128	-0.123	0.349
γ	Cap. adj.	0.020	0.060	RHO[dk]^b	0.294	0.294	0.295	0.153

4.3 Simulation

While the models are able to match their identifying moments, the core questions in this work revolve around un-targeted moments. With estimated models in hand, we can now simulate the models and observe their ability to match the distributional forms and moments not targeted by the MSM procedure — most importantly those pertaining to the heavy tails of income and growth.

Table 5 presents the values of some un-targeted moments in the data and the four models. The table also includes the standard errors on the data moments (obtained via block-bootstrap), which are mostly very low as the moments are well-measured in the data. Both Z-models yield kurtosis values close to 3 (the kurtosis of the Normal distribution) for all growth measures (growth in income, capital, value, scale, and efficiency), as predicted in Sections 3.3 and 3.5. The same is not true for the two SX-models. As predicted, the kurtosis of income growth, capital growth, and value growth are all significantly higher than 3, and the $\lambda\tau$-model (the better estimated of the two) matches the kurtosis values in the data fairly well, even *without having any moments regarding the kurtosis targeted in the estimation.* The SX-models, capturing the interaction between sales and expenses, indeed yield heavy-tailed growth.

A visual comparison of these results is provided in Figure 6. The figure presents, for the data and the simulations of the $\widetilde{S}$- and $\lambda\tau$-models, histograms of (asinh) income cf, income growth dcf, capital growth dk, and value growth (i.e. returns) dv. The data and $\lambda\tau$ simulation histograms are overlaid with MLE-fitted DLN distributions, while the $\widetilde{S}$ simulation is overlaid with MLE-fitted Normal distributions. The visual fit of the data distributions to DLN is excellent, as previously ascertained in Table 2. The $\widetilde{S}$-model distributions again appear Normal and exhibit no heavy tails. The visual fits of the $\lambda\tau$-model distribution to the DLN (and to the data), especially for cf, dcf, and dk, are quite striking. The $\lambda\tau$-model yields the now-familiar double-peaked income distribution, capturing both profit and loss. It also captures the peaked, non-Normal distributions of income, capital, and value growth,

Table 5
Estimation results

This table presents moments of the Data, Z-model, $\widetilde{\text{S}}$-model, SX-model, and $\lambda\tau$-model, respectively. The moment values for each model are at the estimated parameter values of Table 4. The operator and stochastic variable definitions are from the same table. KUR is the kurtosis operator. s.e. is the std. err. of the data moment.

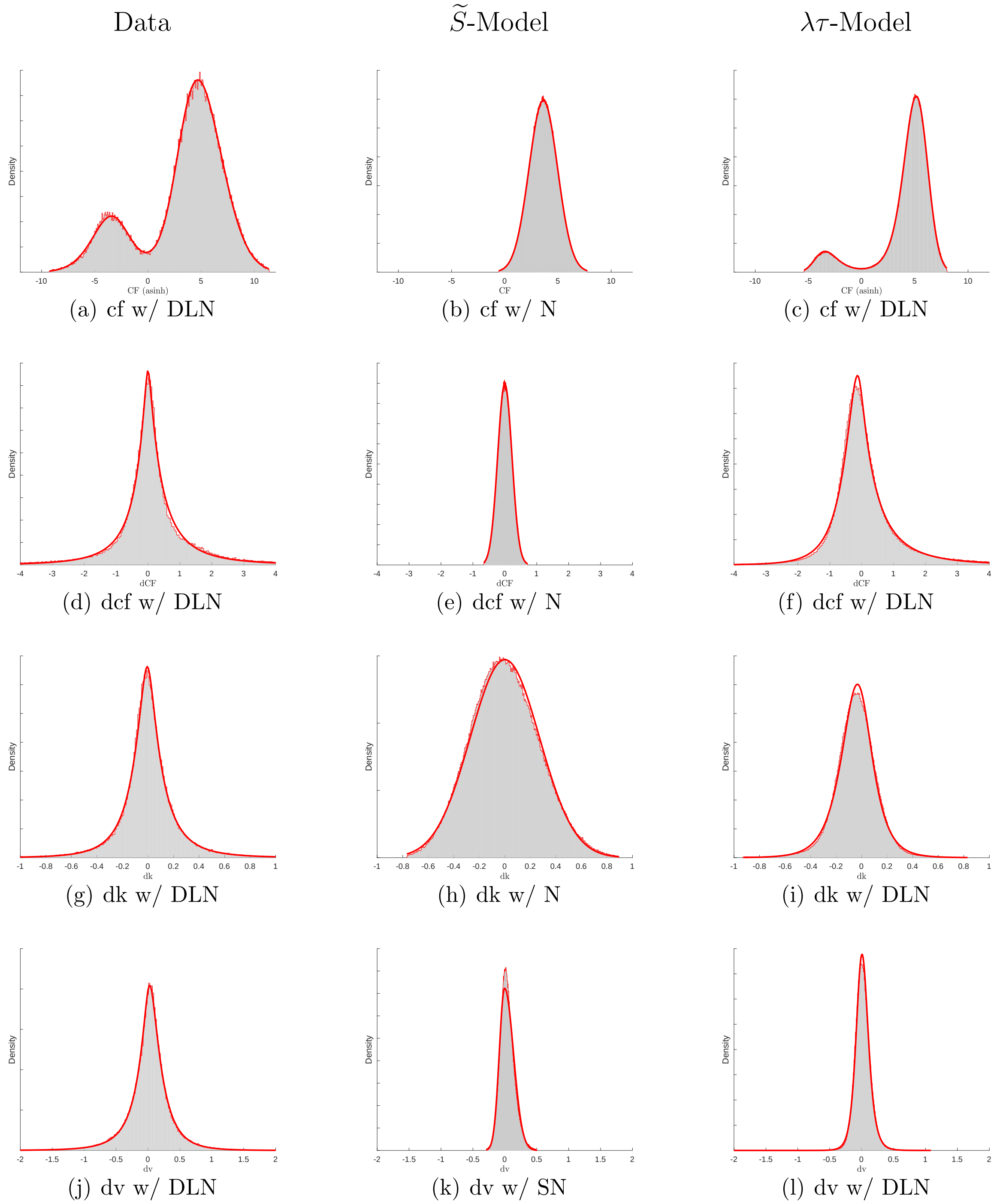
Moment	Data	Z	$\widetilde{S}$	SX	$\lambda\tau$	s.e.
MED[cf]	4.441	4.828	3.622	8.749	4.291	0.030
IQR[dcf]	0.436	0.464	0.222	0.613	0.561	0.004
KUR[dcf]	8.114	3.003	2.989	7.322	6.240	0.319
MED[k]	6.527	7.237	6.037	8.884	6.658	0.029
IQR[dk]	0.130	0.404	0.270	0.908	0.135	0.002
KUR[dk]	14.426	3.181	2.998	4.542	11.751	1.152
MED[v]	6.850	7.956	6.572	10.910	6.925	0.031
IQR[dv]	0.256	0.213	0.193	0.347	0.153	0.002
KUR[dv]	6.901	3.426	3.003	3.248	7.522	0.705
MED[λ]	6.626	7.546	6.334	7.934	6.800	0.028
IQR[$d\lambda$]	0.132	0.464	0.222	0.577	0.125	0.002
KUR[$d\lambda$]	20.539	3.003	2.989	4.347	3.856	1.863
MED[τ]	0.033	0.033	0.033	0.092	0.037	0.001
IQR[$d\tau$]	0.023	N/A	N/A	0.065	0.022	0.001
KUR[$d\tau$]	72.509	N/A	N/A	3.026	2.988	9.900
MED[$v-k$]	0.187	0.730	0.527	1.813	0.254	0.006
IQR[$v-k$]	0.449	0.182	0.117	1.931	0.151	0.006
MED[RTS*]	0.698	0.693	0.698	0.692	0.679	0.006
IQR[RTS*]	0.359	0.087	0.028	0.118	0.356	0.010

[a] For firms with $dk \in$ IQR[dk] and $CF \geq 1$.
[b] For firms around median scale $\lambda \in$ IQR[λ].

though the visual fit to the DLN is less than perfect.

Formal distributional tests for the simulated cf, dcf, dk, and dv in the $\widetilde{S}$- and $\lambda\tau$-models, vs. the Normal, skew-Normal, and DLN are reported in Table 6. For the $\widetilde{S}$-model, none of the variables is rejected as a Normal, which in turn implies they are also not rejected as skew-Normal or DLN. The relative likelihood tests, designed to choose the most parsimonious model, prefer the Normal for income cf, and value growth dv, but the skew-Normal for income growth dcf and capital growth dk. These are all in line with the expected approximate Normality of the Z-models. For the $\lambda\tau$-model, Normality and skew-Normality are strongly rejected for all four, while the DLN is not rejected for any of the four. The relative likelihood test again overwhelmingly prefers the DLN over the Normal and skew-Normal.

Considering the dynamics of firm scale λ and firm efficiency τ, we can observe puzzling deviations from the assumptions of our model in Table 5. Recall that we assumed all stochastic innovations are Normal in Section 2.4. Specifically, we have $\epsilon_\lambda, \epsilon_\tau \sim$ Normal, implying $d\lambda$ and $d\tau$ should have Normal tails and kurtosis of 3. This is far from the case in the data, and the innovations to both are exceedingly heavy-tailed. Our current model cannot explain this stylized fact. This fact, however, explains some of the deviations we observe between outcome variables in the data and SX-model — with heavy-tailed innovations, we would expect heavier-tailed growth, especially in dv, as well as wider (i.e. higher IQR) distribution of (log) median-q $v - k$. I return to this puzzle shortly, in the conclusion, along with the other puzzles identified in the paper. Finally, note that the $\lambda\tau$-model is the only one capable of even coming close to matching the values of MED$[v - k]$ and IQR[RTS*].



(a) cf w/ DLN (b) cf w/ N (c) cf w/ DLN

(d) dcf w/ DLN (e) dcf w/ N (f) dcf w/ DLN

(g) dk w/ DLN (h) dk w/ N (i) dk w/ DLN

(j) dv w/ DLN (k) dv w/ SN (l) dv w/ DLN

Fig. 6. Data and Model distributions. This figure presents the histograms of several firm variables in the data and the $\widetilde{S}$ and $\lambda\tau$ models. The variables presented are (asinh) income cf, income growth dcf, capital growth dk, and value growth (i.e., return) dv. The figures are overlaid with MLE-fitted distributions as indicated.

Table 6
Distributional tests

This table presents the results of tests of distributional form for (asinh) income cf, income growth dcf, capital growth dk, and value growth dv, in the $\widetilde{S}$ and $\lambda\tau$-models. K-S is a Kolmogorov–Smirnov test; C-2 is a binned χ^2 test with 50 bins; A-D is an Anderson-Darling test. Panels (a)-(c) report the test statistics and their p-values *rejecting the distribution* for the Normal, skew-Normal, and DLN, respectively. Panel (d) reports the relative likelihoods for each distribution using the AIC and BIC.

	$\widetilde{S}$-model				$\lambda\tau$-model			
	cf	dcf	dk	dv	cf	dcf	dk	dv
Panel (a): Normal								
K-S	0.002	0.006	0.011	0.004	0.043	0.143	0.174	0.034
p-val	0.483	0.081	0.054	0.116	0.016	0.000	0.000	0.022
C-2	1.162	11.12	32.69	25.12	459	>999	>999	262
p-val	1.000	0.114	0.064	0.096	0.016	0.000	0.000	0.023
A-D	0.017	0.934	2.951	2.493	41.55	412	662	35.04
p-val	0.436	0.073	0.051	0.057	0.016	0.000	0.000	0.018
Panel (b): skew-Normal								
K-S	0.002	0.001	0.002	0.004	0.027	0.145	0.190	0.033
p-val	0.571	0.687	0.254	0.138	0.027	0.000	0.000	0.023
C-2	1.205	1.002	1.704	11.19	142	>999	>999	222
p-val	1.000	1.000	1.000	0.112	0.032	0.000	0.000	0.025
A-D	0.020	0.022	0.078	1.179	12.54	365	600	31.43
p-val	0.391	0.364	0.166	0.082	0.030	0.000	0.000	0.019
Panel (c): DLN								
K-S	0.002	0.002	0.003	0.004	0.002	0.010	0.008	0.009
p-val	0.194	0.219	0.149	0.112	0.287	0.058	0.068	0.064
C-2	3.717	2.768	3.982	5.339	5.585	110	44.29	40.61
p-val	0.803	1.000	0.694	0.287	0.253	0.036	0.056	0.058
A-D	0.121	0.158	0.310	0.333	0.117	2.670	1.887	2.013
p-val	0.280	0.128	0.102	0.100	0.142	0.053	0.059	0.058
Panel (d): Relative likelihood tests								
AIC R.L.:								
Normal	1.000	0.004	0.000	1.000	0.000	0.000	0.000	0.000
skew-Normal	0.379	1.000	1.000	0.115	0.000	0.000	0.000	0.000
DLN	0.018	0.005	0.002	0.007	1.000	1.000	1.000	1.000
BIC R.L.:								
Normal	1.000	0.148	0.000	1.000	0.000	0.000	0.000	0.000
skew-Normal	0.010	1.000	1.000	0.062	0.000	0.000	0.000	0.000
DLN	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000

5 Conclusion

This work begins with possibly the most fundamental of accounting identities, income = sales - expenses. It uses this identity to motivate a novel production function for the firm, the difference-of-log-linears (DLL). Together with the CLT-implied fact that AR(1) processes have Normal ergodic distributions, the DLL production function predicts a little-known distribution for firm income and consequently firm growth, the difference-of-log-Normals (DLN) distribution. These theoretical predictions are confirmed by the data in statistical tests, horse races, and simulation exercises. Because equity returns are themselves one measure of firm growth, the DLN arises as the distribution of returns as well. These results are achieved without using: time-varying volatility, factors external to the firm, mixture-of-Normals assumptions, or non-standard stochastic processes. Thus, this paper provides an intuitive and simple answer to the question posed in its title: "Why are firm growth distributions heavy-tailed?"

The theoretical analysis yields two new magnitudes for characterizing firms — firm scale and firm efficiency, both defined in terms of firm sales and expenses. Both measures are observable and easy to calculate and interpret. Firm income scale is tightly correlated with other measures of firm scale, and firm efficiency changes are shown to be the main driver of income growth. I show that the source of heavy-tailed growth can be traced to a base-rate effect in firm efficiency and that for most firms, firm efficiency is indeed remarkably close to zero, yielding rampant base-rate effects. The DLL production function further enables new and coherent definitions of income growth and returns to scale, among others.

While the question this paper considers may seem somewhat aloof from practical considerations, the findings have many downstream uses. Models based on the SX-model can: (i) Replicate the distribution of firm income — the departing point for corporate finance and production- or consumption-based asset pricing models; (ii) Replicate the distribution of equity returns — an object of intense interest in financial economics and specifically in asset pricing; (iii) Provide models with both rare disasters and rare winners — i.e. models with heavy-tailed growth; (iv) Allow consideration and modeling of loser firms — as standard models cannot model firms experiencing losses; (v) Enable straightforward models of exit and entry — thus enabling investigation of dynamism within the work-horse q-theory model. For example, consider the unobservable value of the marginal product of capital - the core driver of firm investment and an object of considerable interest in the theory of the firm. The model proposes a simple estimate of MPK, because RTS = MPK/APK is quasi-observable, as Equation 26 shows, and APK is observable. The DLL production function also informs production-based asset pricing models such as Delikouras and Dittmar (2021). The idea that investment return equals stock return from Cochrane (1991) is pre-disposed on the assumption of a linear-homogeneous production function. This work convincingly establishes this is not the case for firms and that the deviations from the assumption have significant implications.

Several data puzzles were identified in the paper, including the decreasing dispersion of growth rates with scale (DDWS) and the fact that the growth of the stochastic variables is DLN rather than Normal, as the model posits. While I leave a full consideration of these puzzles for future work, it is worth noting that both can be rationalized by appealing to the internal structure of the firm. Consider the firm as composed of sub-units, each behaving according to the SX-models above, and the firm as their simple agglomeration. In this case, the decreasing dispersion with scale is a direct outcome of portfolio theory, similar to how a portfolio of more stocks has a lower variance. The same assumption is also sufficient to yield DLN growth in the aggregate stochastic variables, even if each sub-unit's stochastic growth is Normal, due to the intervening impact of heavy-tailed capital growth in each sub-unit.

Finally, because firms comprise the productive side of the economy, and dynamic stochastic general equilibrium (DSGE) models ubiquitously include firms as the source of all individual income, embedding the DLL production function in DSGE models allows for a micro-foundation of heavy-tailed income growth in a succinct and tractable manner. Thus, heavy-tailed growth can be embedded in "upstream" economic models as well.

References

Abramovitz, M., 1956. Resource and Output Trends in the United States since 1870. In: *Resource and Output Trends in the United States since 1870*, NBER, pp. 1–23.

Ashton, T. S., 1926. The Growth of Textile Businesses in the Oldham District, 1884-1924. Journal of the Royal Statistical Society 89, 567–583.

Bansal, R., Yaron, A., 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. The Journal of Finance 59, 1481–1509.

Barro, R. J., Sala-I-Martin, X. I., 2003. Economic Growth. The MIT Press, Cambridge, Mass, second ed.

Belo, F., Li, J., Lin, X., Zhao, X., 2017. Labor-Force Heterogeneity and Asset Prices: The Importance of Skilled Labor. The Review of Financial Studies 30, 3669–3709.

Belo, F., Lin, X., Vitorino, M. A., 2014. Brand capital and firm value. Review of Economic Dynamics 17, 150–169.

Bolton, P., Chen, H., Wang, N., 2011. A Unified Theory of Tobin's q, Corporate Investment, Financing, and Risk Management. The Journal of Finance 66, 1545–1578.

Bottazzi, G., Secchi, A., 2003. Why are distributions of firm growth rates tent-shaped? Economics Letters 80, 415–420.

Bottazzi, G., Secchi, A., 2006. Explaining the distribution of firm growth rates. The RAND Journal of Economics 37, 235–256.

Buldyrev, S. V., Growiec, J., Pammolli, F., Riccaboni, M., Stanley, H. E., 2007. The Growth of Business Firms: Facts and Theory. Journal of the European Economic Association 5, 574–584.

Campbell, J. Y., Cochrane, J. H., 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. Journal of Political Economy 107, 205–251.

Cochrane, J. H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. The Journal of Finance 46, 209–237.

DeAngelo, H., DeAngelo, L., Whited, T. M., 2011. Capital structure dynamics and transitory debt. Journal of Financial Economics 99, 235–261.

Delikouras, S., Dittmar, R., 2021. Do Investment-Based Models Explain Equity Returns? Evidence from Euler Equations. The Review of Financial Studies p. hhab099.

Epstein, L. G., Zin, S. E., 1991. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis. Journal of Political Economy 99, 263–286.

Falato, A., Kadyrzhanova, D., Sim, J., Steri, R., 2021. Rising Intangible Capital, Shrinking Debt Capacity, and the US Corporate Savings Glut. The Journal of Finance forthcoming.

Fama, E. F., 1963. Mandelbrot and the Stable Paretian Hypothesis. The Journal of Business 36, 420–429.

Fama, E. F., 1965. The Behavior of Stock-Market Prices. The Journal of Business 38, 34–105.

Gibrat, R., 1931. Les Inégalités Économiques. Librairie du Recueil Sirey, Paris.

Gorbenko, A. S., Strebulaev, I. A., 2010. Temporary versus Permanent Shocks: Explaining Corporate Financial Policies. The Review of Financial Studies 23, 2591–2647.

Hayashi, F., 1982. Tobin's Marginal q and Average q: A Neoclassical Interpretation. Econometrica 50, 213–224.

Hennessy, C. A., Whited, T. M., 2005. Debt Dynamics. The Journal of Finance 60, 1129–1165.

Hennessy, C. A., Whited, T. M., 2007. How Costly Is External Financing? Evidence from a Structural Estimation. The Journal of Finance 62, 1705–1745.

Klette, T. J., Kortum, S., 2004. Innovating firms and aggregate innovation. Journal of Political Economy 112, 986–1018.

Li, S., Whited, T. M., Wu, Y., 2016. Collateral, Taxes, and Leverage. The Review of Financial Studies 29, 1453–1500.

Lin, X., 2012. Endogenous technological progress and the cross-section of stock returns. Journal of Financial Economics 103, 411–427.

Liu, L. X., Whited, T. M., Zhang, L., 2009. Investment-Based Expected Stock Returns. Journal of Political Economy 117, 1105–1139.

Livdan, D., Sapriza, H., Zhang, L., 2009. Financially Constrained Stock Returns. The Journal of Finance 64, 1827–1862.

Lucas, R. E., 1967. Optimal Investment Policy and the Flexible Accelerator. International Economic Review 8, 78.

Lucas, R. E., 1978. On the Size Distribution of Business Firms. The Bell Journal of Economics 9, 508–523.

Luttmer, E. G. J., 2011. On the Mechanics of Firm Growth. The Review of Economic Studies 78, 1042–1068.

Mandelbrot, B., 1960. The Pareto-Lévy Law and the Distribution of Income. International Economic Review 1, 79–106.

Mandelbrot, B., 1961. Stable Paretian Random Functions and the Multiplicative Variation of Income. Econometrica 29, 517–543.

Mantegna, R. N., Stanley, H. E., 1995. Scaling behaviour in the dynamics of an economic index. Nature 376, 46.

Michaels, R., Page, B., Whited, T. M., 2019. Labor and Capital Dynamics under Financing Frictions. Review of Finance 23, 279–323.

Nikolov, B., Whited, T. M., 2014. Agency Conflicts and Cash: Estimates from a Dynamic Model. The Journal of Finance 69, 1883–1921.

Officer, R. R., 1972. The Distribution of Stock Returns. Journal of the American Statistical Association 67, 807–812.

Parham, R., 2023. The Difference-of-Log-Normals Distribution: Properties, Estimation, and Growth. arXiv:2302.02486 [stat.ME] .

Pedroni, P., 2004. Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Econometric Theory 20.

Riddick, L. A., Whited, T. M., 2009. The Corporate Propensity to Save. The Journal of Finance 64, 1729–1766.

Simon, H. A., Bonini, C. P., 1958. The Size Distribution of Business Firms. The American Economic Review 48, 607–617.

Smith, A., 1776. An Inquiry into the Nature and Causes of the Wealth of Nations. W. Strahan and T. Cadell, London.

Stanley, M. H. R., Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M. A., Stanley, H. E., 1996. Scaling behaviour in the growth of companies. Nature 379, 804.

Stokey, N. L., Lucas, R. E., Prescott, E. C., 1989. Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, MA.

Strebulaev, I. A., Whited, T. M., 2012. Dynamic Models and Structural Estimation in Corporate Finance. Foundations and Trends in Finance 6, 1–163.

Sun, Q., Xiaolan, M. Z., 2019. Financing intangible capital. Journal of Financial Economics 133, 564–588.

Tobin, J., 1969. A General Equilibrium Approach To Monetary Theory. Journal of Money, Credit and Banking 1, 15–29.

Uzawa, H., 1969. Time Preference and the Penrose Effect in a Two-Class Model of Economic Growth. Journal of Political Economy 77, 628–652.

Westerlund, J., 2005. New Simple Tests for Panel Cointegration. Econometric Reviews 24, 297–316.