Why are Firm Growth Distributions Heavy-tailed?

Robert Parham*

Thursday 12$^{\text{th}}$ September, 2024

Abstract

Firm income, 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 proposes novel 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 Ron Kaniel, Toni Whited, Chris Yung, William Wilhelm, Michael Gallmeyer, Mihail Velikov, Donald Bowen, Bradley Paye, Sara Easterwood, Jan Eeckhout, Sanket Korgaonkar, Ben McCartney, and seminar participants at the 2022 Commonwealth Finance Conference, 2023 Finance Down Under, 2023 Southern Finance Association, and 2024 Econometric Society Summer School on Dynamic Structural Estimation for helpful comments and suggestions. All remaining errors are my own.

"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 $\pm 10\%$ 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 extreme "winners" and "losers" are far more numerous and economically consequential than 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.

I show that a simple and intuitive modification to an otherwise standard q-theory model of the firm — separately accounting for sales and expenses rather than modelling income directly — is sufficient to yield heavy-tailed firm behavior, without appealing to time-varying volatility or factors external to the firm. Moreover, the economic model makes a specific prediction on the shape of several firm outcomes, predicting they should distribute as the difference-of-log-Normals.[1] This prediction is overwhelmingly supported by the data. The obscure difference-of-log-Normals distribution exhibits a remarkable fit to a plethora of firm outcomes, such as income, income growth, average product of capital, capital growth, and equity returns, as the model predicts. The fit with equity returns holds for daily, monthly, yearly, raw, and excess returns in a set of robustness tests.

The core mechanism at play is the impact of two opposing exponential forces, sales and expenses. Consider for example a firm with \$100 in sales and \$90 in expenses during year 1,

[1] Parham (2023) describes the emergence of the difference-of-log-Normals distribution in general economic data and fully characterizes it, deriving its PDF, CDF, central moments, and estimators for the distribution parameters given data.

yielding an income of $100 - 90 = 10$ dollars. What will be the *growth rate of income* from year 1 to year 2 if: (i) both sales and expenses increase by 10%? (ii) sales increase by 10% but expenses decrease by 10%? (iii) sales *decrease* by 10% but expenses *increase* by 10%? Case (i) is fairly simple, and yields income of $110 - 99 = 11$ dollars in period 2, or 10% higher income than in period 1. In case (ii), income is $110 - 81 = 29$ dollars, or 190% higher than in period 1. With such a large increase in income, one would expect e.g. firm value to increase significantly as well. This "operational leverage" effect is shown to be at the heart of the firm's heavy-tailed growth.[2]

Case (iii) is more interesting still. This is because firm income in period 2 is $90 - 99 = -9$ dollars, or a loss of \$9. Growth to (and from) *negative* values has hitherto been poorly defined, but is a necessary tool when discussing income growth, because of the occurrence of negative income values (i.e. losses) in the data. I consider below extensions to growth measures in the presence of negative values, and show why it is sensible to say that firm income in this case is 190% lower than in period 1. I further show how to extend log-point growth measures, overcoming the fact the log of negative values is a complex number.

Understanding the income generating process of firms is key to explaining corporate policies, asset pricing puzzles, and the distribution of firm outcomes, both across firms and time. The approach proposed here to modelling firm losses extends the models of e.g. Abel and Eberly (1994), Abel and Eberly (1996) and resolves the critique of Gorbenko and Strebulaev (2010); Strebulaev and Whited (2012) regarding the "highly unrealistic" lack of losses in dynamic firm models. I show that the 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. Specifically, it gives rise to extended and internally consistent definitions for common concepts such as firm

[2]Previous work (e.g. Carlson, Fisher, and Giammarino (2004), Sagi and Seasholes (2007), and especially Novy-Marx (2011, 2013)) related operating leverage to specific factor returns, such as value, momentum, and profitability. Here I instead establish a relation with the unconditional distribution of firm growth and of equity returns in the context of a simple q-theory model.

income scale, efficiency, growth, and returns-to-scale. In that, the paper follows in the tradition of, e.g., Epstein and Zin (1991); Campbell and Cochrane (1999) and Bansal and Yaron (2004).

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. The theoretically-implied difference-of-log-Normals distribution exhibits remarkable fit to the data and has finite moments of all orders.

My approach is by no means a first attempt at explaining the distribution of firm growth or its underlying mechanics. 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. Notable theoretical models 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 heavy-tailed growth but fail to fit the observed data distributions. The latter works further posit 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, exhibits remarkable fit and is a simple, intuitive, and straightforward extension of the workhorse q-theory model, lacking any such external assumptions.

The paper proceeds as follows: Section 2 presents q-theory models of firm dynamics 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) scale; (ii) income; (ii) income efficiency (iii) income growth; (iv) returns-to-scale; and (v) firm growth. Section 4 presents a structural estimation and 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 Model

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.[3] 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 subsections expand on these facets by presenting and discussing the relevant functional and stochastic forms. The only

[3]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 Bazdresch (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).

deviations from the standard q-theory literature are concerning the production function and its derived stochastic dynamics.

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 $Z_t > 0$.[4] The firm first chooses whether to remain for another period (denoted $\alpha_t = 1$) 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 function $\mathbf{I}()$ embeds any assumptions on depreciation, fixed and convex adjustment costs, irreversibility, etc. The firm produces income (sales net of all expenses and taxes) $Y_t = \mathbf{Y}(K_t, Z_t)$, 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

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

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

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

where $\mathbf{X}'_j()$ indicates the derivative of the function $\mathbf{X}()$ w.r.t its j^{th} argument. The R.H.S

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

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.[5]

Denote the investment policy function of a remaining firm prescribed by Equation 2 to be $K_{t+1} = \Psi_t = \mathbf{\Psi}(K_t, Z_t)$.[6] It is useful to define the exit (or bankruptcy) indicator of the 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 exit threshold (i.e. has a non-negative exit indicator $B_t \geq 0$) 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

$$\beta \cdot \mathbb{E}_t\left[\left(1 - \alpha_{t+1}\right) \cdot \mathbf{V}_1^{\text{exit}\prime}\left(\Psi_t, Z_{t+1}\right) + \alpha_{t+1} \cdot \left(\mathbf{Y}_1'\left(\Psi_t, Z_{t+1}\right) - \mathbf{I}_2'\left(\Psi_{t+1}, \Psi_t\right)\right)\right] = \mathbf{I}_1'\left(\Psi_t, K_t\right) \tag{4}$$

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.

Finally, 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.

[5]If the function $\mathbf{I}\left(\right)$ contains fixed costs of investment, then an inactivity region may arise as in Abel and Eberly (1994). I ignore this possibility below for ease of exposition, but the model can be easily extended to include such an assumption.

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

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 adjustment costs and depreciation. Throughout, I will be using the standard investment function common to the literature cited above, often written as

$$\mathbf{I}_{quad}\left(K_{t+1}, K_t\right) = \left(K_{t+1} - K_t\right) + \gamma \cdot \left(\frac{K_{t+1}}{K_t} - 1\right)^2 \cdot K_t + \delta \cdot K_t \tag{5}$$

with $\gamma \geq 0$ an adjustment parameter and $0 \leq \delta \leq 1$ the capital depreciation rate. It includes: (i) the cost of capital goods for a price-taking firm; (ii) a symmetric quadratic capital adjustment cost; and (iii) a constant rate of capital depreciation. I use this function, and especially the adjustment cost functional form, for simplicity and to demonstrate that the results do not depend on complex adjustment dynamics including limited reversibility, inaction regions, etc.

Note that when $\gamma \to 0$, $\mathbf{I}_{quad}$ 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 often be expressed analytically, without requiring value- or policy-function iterations.

2.3 Production function: Z-model

The canonical production function in the q-theory literature is a log-linear (LL, e.g. Cobb-Douglas) production function. I present and discuss it in this section before discussing its extension, the difference-of-log-linears production function that captures the interplay of sales and expenses, in the next section.

The canonical production function models firm income as

$$\mathbf{Y}_z\left(K_t, Z_t\right) = (1-\tau_z) \cdot Z_t \cdot K_t^{\theta_z} = (1-\tau_z) \cdot \exp\left(z_t + \theta_z \cdot k_t\right) \tag{8}$$

with returns to scale parameter $0 < \theta_z < 1$, expense parameter $0 \leq \tau_z < 1$, and with lower-case variables denoting log values as usual. The stochastic productivity process Z is assumed to follow the canonical AR(1) in logs, with Normal innovations. I.e. $z = \log(Z)$ follows

$$z_{t+1} = (1-\rho_z) \cdot \mu_z + \rho_z \cdot z_t + \epsilon_{t+1}^z \tag{9}$$

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

$$\epsilon^z \sim \mathbb{N}(0, \sigma_z^2) \tag{10}$$

with $\sigma_z > 0$ the standard deviation of ϵ^z. Models following this production function are denoted Z-models.

The q-theory models generally abstract from labor. Wages, materials, and other expenses are already accounted for, as $\mathbf{Y}_z\left(K_t, Z_t\right)$ models net income, i.e., sales minus expenses (including taxes). The expense parameter τ_z can be interpreted in different ways, depending on the relevant model calibration. Some models calibrate it to a fixed corporate tax rate and calibrate the remaining term $Z_t \cdot K_t^{\theta_z}$ to match firm EBIT. Others drop it altogether (i.e. set $\tau_z = 0$), and calibrate the remaining term $Z_t \cdot K_t^{\theta_z}$ to match firm net income. Alternatively, it can be set to a fixed expense ratio (e.g., the average ratio of firm expenses to firm sales in the data) and the remaining term $Z_t \cdot K_t^{\theta_z}$ then calibrated to match firm sales.

The main challenge with the first two interpretations of τ_z, and accordingly the two common calibrations of the $Z_t \cdot K_t^{\theta_z}$ term, is that while this term is non-negative by construction, neither EBIT nor net income are. Negative incomes (i.e., losses) are prevalent, and impact firm policies and behaviors in a variety of ways, not captured by Z-models. The third interpretation matches the non-negative term to firm sales, which are indeed a non-negative value. In doing so, however, the calibration ignores the entire cost structure of the firm, focusing only on the dynamics of firm sales. I return to this challenge shortly.

Before proceeding, it is worth contemplating the economic meaning of the exogenous productivity process Z. One way of framing Z is as the Solow residual, after the contribution of capital has been factored-out of firm income. As such, it is often thought of as the "productivity" of the firm (w.r.t capital). What determines this productivity? 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.

2.4 Production function: SX-model

Reconsider now the interpretation of the income function $\mathbf{Y}_z$ in which τ_z is a *fixed* expense ratio, and $Z_t \cdot K_t^{\theta_z}$ is calibrated to match firm sales. We can write firm income in the model as:

$$Y_t = \left(1 - \frac{\overline{\text{Expenses}}}{\text{Sales}}\right) \cdot \text{Sales}_t = \text{Sales}_t - \frac{\overline{\text{Expenses}}}{\text{Sales}} \cdot \text{Sales}_t \approx \text{Sales}_t - \text{Expenses}_t = \text{Income}_t \tag{11}$$

with τ_z explicitly spelled out as average expenses over sales. This decomposition highlights the economic definition of income — the difference between sales and expenses — possibly the most fundamental of accounting identities. It further demonstrates the difficulty of "single-factor" Z-models to capture the interplay of sales and expenses.

To capture this interplay of sales and expenses, I instead propose to model each one directly as both are observable in the data. Specifically, I use the following difference-of-log-linears production function,

$$\mathbf{Y}_{sx}\left(K_t, S_t, X_t\right) = \underbrace{S_t \cdot K_t^{\theta_s}}_{Sales \equiv \mathbb{S}_t} - \underbrace{X_t \cdot K_t^{\theta_x}}_{Expenses \equiv \mathbb{X}_t} = \exp\left(s_t + \theta_s \cdot k_t\right) - \exp\left(x_t + \theta_x \cdot k_t\right) \tag{12}$$

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.

The S, X process is 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 $-1 < \rho_{sx} < 1$. Models following this production function are denoted SX-models.

While the Z-model appears to be significantly more parsimonious than the SX-model, I show in Section 4 that all dynamic parameters in both the Z- and SX-models are strongly pinned-down by the data. In effect, the only meaningful increase in degrees-of-freedom is the move from one returns-to-scale parameter (θ_z) to two (θ_s, θ_x).

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.[7]

A difficulty with this modelling approach is the fact that, strictly speaking, capital does not "cause" expenses, as most firm expenses are in fact payments to labor and materials. One possible interpretation is that X merely tracks the projection (i.e. solow residual) of total expenses on firm capital, here used as an index of firm size. This approach has the benefit of being simple and intuitive. Another approach is to model such inputs directly. As an example, consider the following production function which tracks labor as well as capital, assuming labor, like capital, is quasi-fixed:

$$\begin{aligned} \mathbf{Y}\left(K_t, L_t, S_t, X_t\right) &= S_t \cdot K_t^{\theta_{s,k}} \cdot L_t^{\theta_{s,l}} - X_t \cdot K_t^{\theta_{x,k}} \cdot L_t^{\theta_{x,l}} \\ &= \exp\left(s_t + \theta_{s,k} \cdot k_t + \theta_{s,l} \cdot l_t\right) - \exp\left(x_t + \theta_{x,k} \cdot k_t + \theta_{x,l} \cdot l_t\right) \end{aligned} \tag{15}$$

[7] 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 $\rho_{sx} > 0$.

In such a production function one may, e.g., set $\theta_{x,k} \to 0, \theta_{x,l} \to 1$ and use $X_t > 0$ to track the wage rate in the economy, yielding $w_t \cdot L_t$.[8]

As in Abel and Eberly (1998), the wage rate (or more generally the price of inputs) is stochastic. This is because the model implicitly assumes the firm makes a production plan for period t during period $t-1$. The firm then uses this plan to decide the appropriate capital and labor levels for period t, K_t, L_t, which are decided at $t-1$. The firm hence makes the production plan based on the *expected* demand curve (and output price) for its output and *expected* supply curves (and input prices) for its inputs, with these expectations based on $t-1$ market conditions. The actual prices it ends up facing are hence stochastic. Adding consideration for labor does not, however, contribute much to the basic understanding of the production function and its impact on firm dynamics. I hence opt for the simple interpretation of X_t as a projection of expenses on capital, and remain within the q-theory framework which tracks only capital.

2.5 Production function: $\lambda\tau$-model

An important feature of the difference-of-log-linears 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. The Hyperbolic Sine function is defined as

$$\sinh(x) \equiv \frac{\exp(x) - \exp(-x)}{2} \tag{16}$$

for $x \in \mathbb{R}$.[9]

[8] The payment to capital $r \cdot K$ is explicit (i.e. residual) in the model, and hence does not require accounting in the cost function.

[9] Compare with the exponential definition of the "traditional" Sine function, $\sin(x) \equiv \frac{\exp(i \cdot x) - \exp(-i \cdot x)}{2 \cdot i}$.

It is simple to show that Y_{sx} can alternatively and equivalently be written as

$$\mathbf{Y}_{sx}\left(K_t, S_t, X_t\right) = \exp\left(s_t + \theta_s \cdot k_t\right) - \exp\left(x_t + \theta_x \cdot k_t\right) = 2 \cdot \exp\left(\lambda_t\right) \cdot \sinh\left(\tau_t\right)$$

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

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

which merits discussion.

The two new variables, $\lambda_t \in \mathbb{R}$ and $\tau_t \in \mathbb{R}$, are the hyperbolic equivalents of the "radius" and "angle" in Polar geometry, respectively. Both are observable given only firm sales and expenses, and do not require knowledge of the firm's capital stock to calculate. The variable λ_t, hereafter denoted *firm income scale*, is the (log) geometric mean of sales and expenses in period t. The variable τ_t, hereafter denoted *firm income efficiency*, is a transformation of the firm's operational efficiency ratio.[10] 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 $\text{Sales}_t = \exp\left(\lambda_t + \tau_t\right)$ and $\text{Expenses}_t = \exp\left(\lambda_t - \tau_t\right)$. Clearly, the sign of firm income depends on the sign of τ (i.e., a firm with negative τ suffers losses), and the magnitude of firm income primarily depends on λ, with a small role for τ.

As Equation 17 demonstrates, both λ_t and τ_t can be expressed in terms of s_t, x_t, k_t — the (log) state variables of the SX model. This allows us to consider a change-of-variables, and define two new stochastic state variables, $\widehat{\lambda}_t$ and $\widehat{\tau}_t$, and two new returns-to-scale parameters, θ_λ and θ_τ. These new stochastic variables are the Solow residuals of projecting income scale and efficiency on capital, rather than projecting sales and expenses as in the original SX-

[10] Operational efficiency is defined as the ratio of outputs to inputs in a production process, or here the ratio of sales to expenses. It is the inverse of the expense ratio from above, defined as the ratio of expenses to sales.

model. We can now recast the SX-model as:

$$\mathbf{Y}_{\lambda\tau}\left(k_t, \widehat{\lambda}_t, \widehat{\tau}_t\right) = 2 \cdot \exp\left(\widehat{\lambda}_t + \theta_\lambda \cdot k_t\right) \cdot \sinh\left(\widehat{\tau}_t + \theta_\tau \cdot k_t\right) \tag{18}$$

with $0 < \theta_\lambda < 1$ and $\theta_\tau \in \mathbb{R}$, and with the stochastic dynamics of $\widehat{\lambda}_t, \widehat{\tau}_t$ jointly-AR(1), as in Equations 13 and 14.

The $\lambda\tau$ formulation of the SX-model is useful for several reasons. First, it gives rise to the (plausibly meaningful) decomposition of income using the definitions of income scale and income (operational) efficiency described here. Second, it is considerably simpler to estimate and simulate because it transforms the highly-correlated variables of sales and expenses to the nearly-uncorrelated variables of scale and efficiency. Third, theoretical derivations and mathematical analysis of model behavior are often simpler under the $\lambda\tau$ formulation than under the SX formulation, as I show in the model analysis conducted in Section 3.

2.6 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

$$\mathbf{V}^{\text{exit}}\left(K_t, \cdots\right) = \nu \cdot K_t \tag{19}$$

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 current distribution of firm states in the simulation.

3 Analysis of models

The SX-model of the firm makes specific novel 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 log-linear production function, when appropriate.

The data analyzed cover 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. This removes both inflation and secular growth trend considerations. All reported results are robust to sample-period selection, and I verify they hold when limiting the sample to any single decade within the period.

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

$$\underbrace{\text{sales}}_{SL} - \underbrace{\text{expenses}}_{XS} = \underbrace{\text{income}}_{CF} = \underbrace{\text{total net dividends}}_{DI} + \underbrace{\text{total net investment}}_{IT} \tag{20}$$

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.

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

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

3.1 Scale

The firm size distribution has seen such intense research interest it has its own JEL classification: L11. Three common measures of firm size in the literature are total capital KT, equity value EQ, and total sales SL. To those, I add total expenses XS and total value VL. Because much of total expenses is payment to labor, XS is a noisy proxy of employee count, another common firm size measure. All those measures are non-negative and are often considered in logs. To prevent confusion, I will refer to the logged versions as "scales", e.g. $k_t = \log(K_t)$ is the firm's *capital scale* in period t. Finally, I consider firm income scale λ, as defined by Equation 17, along with the previous five firm scale measures.

A survey by Sutton (1997) concludes that the firm scale distribution (using various measures) is stable over time and approximately Normal. Later contributions, including Cooley and Quadrini (2001), Cabral and Mata (2003), Desai, Gompers, and Lerner (2003) and Angelini and Generale (2008) concentrate on the observed mild skewness of the distributions, and relate it to financial frictions hampering the growth of younger/smaller firms.

Panel (a) of Table 2 presents the descriptive statistics for these scale measures in the data. Perhaps unsurprisingly, they all have similar means, s.d., skewness, and kurtosis. They are also all highly correlated, as Panel (b) of Table 2 shows, and are in fact all co-integrated, as Panel (c) of Table 2 reports. Two notable factoids are that all have mild positive skewness (as discussed above) and all have kurtosis close to 3 (the kurtosis of the Normal distribution).

Empirical histograms of several of the scale measures are presented in Figure 1. They all indeed appear approximately Normal, with slight skewness. Each is overlaid with a fitted skew-Normal distribution yielding good visual fit.[11] The fit is further evident when viewing the accompanying quantile-quantile (q-q) plots in the figure.[12]

Formal statistical tests of each of the six scale measures versus the Normal and skew-

[11] The skew-Normal distribution is a 3-parameter distribution $\mathbb{SN}(\mu, \sigma^2, \alpha)$ with $\alpha \in \mathbb{R}$ a skewness parameter and $\mathbb{SN}(\mu, \sigma^2, 0) \sim \mathbb{N}(\mu, \sigma^2)$.

[12] A q-q plot presents the theoretical quantiles of a given distribution vs. the empirical quantiles in the observed data. When the quantiles match (i.e., the theoretical and empirical CDFs are identical), all points lie on the 45-degree line in the q-q plot.

Table 2
Scale - Descriptive statistics

Panel (a) presents the first four central moments of the firm scale measures (λ and the logs of capital, firm value, equity value, sales, and expenses). Variable definitions are in Table 1. Panel (b) presents the correlations between the various scale measures. Panel (c) presents the results of three cointegration tests between the 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 (a): Scale moments

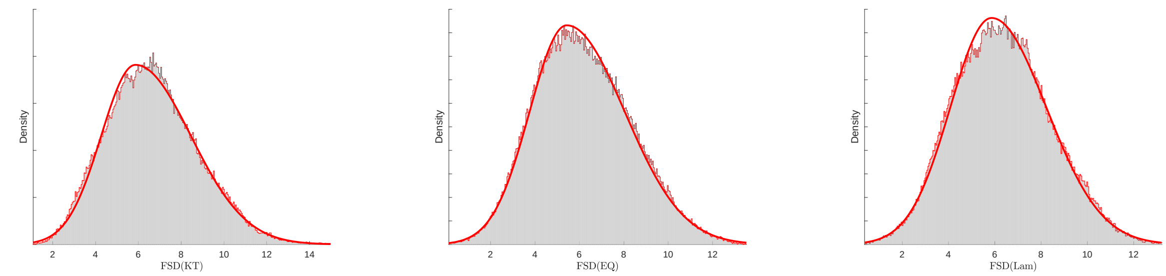
	λ	KT	VL	EQ	SL	XS
M_1 (mean)	6.33	6.65	6.98	6.19	6.36	6.30
M_2 (s.d.)	2.08	2.11	2.12	2.16	2.13	2.05
M_3 (skew)	0.19	0.36	0.34	0.27	0.09	0.23
M_4 (kurt)	2.76	2.99	2.91	2.75	2.86	2.78

Panel (b): Scale correlations

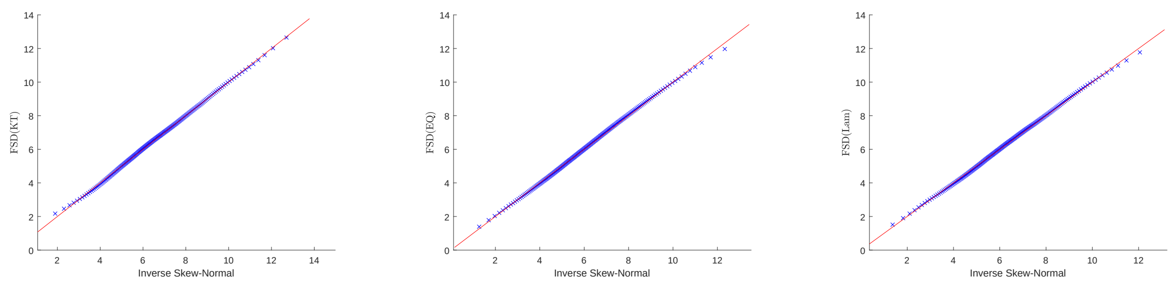
	λ	KT	VL	EQ	SL	XS
λ	—	.929	.880	.797	.995	.995
KT	.929	—	.961	.883	.928	.921
VL	.880	.961	—	.960	.878	.874
EQ	.797	.883	.960	—	.796	.792
SL	.995	.928	.878	.796	—	.982
XS	.995	.921	.874	.792	.982	—

Panel (c): 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



(a) Capital scale (b) Equity scale (c) Income scale λ



(d) q-q vs. SN (e) q-q vs. SN (f) q-q vs. SN

Fig. 1. Firm scale distributions. Panel (a) presents the histogram of capital scale (log total assets), overlaid with an MLE-fitted Skew-Normal distribution, for a set of 143K firm-year observations from 1970-2019. Panel (d) presents the respective q-q plot. Panels (b) and (c) present equity scale (log market value of equity) and income scale (λ = log geometric average of sales and expenses), respectively, with MLE-fitted Skew-Normal distributions. Panels (e) and (f) present the respective q-q plots.

Normal distributions are presented in Table 3. The three goodness-of-fit distributional tests I use are the Kolmogorov-Smirnov (K-S), the Chi-square (C-2), and the Anderson-Darling (A-D) tests. The three tests are sensitive to different distributional deviations — K-S has uniform power throughout, C-2 is more powerful around the center-mass, and A-D is more powerful around the tails — hence I report results of all three tests. Panel (a) shows that Normality is generally rejected across the board (at the 5% significance level). In contrast, Skew-Normality is not rejected for any scale measure by either test, as seen in Panel (b) of Table 3.

The better fit of the skew-Normal distribution might not be too surprising, given that it has an extra degree of freedom (i.e., an extra parameter). To account for the degrees of freedom, I use the relative likelihood test, derived from the AIC statistic of Akaike (1973). The relative likelihood is a non-nested version of the likelihood ratio test, accounting for the number of parameters.[13] I also report relative likelihood tests using the BIC statistic, which penalizes extra degrees of freedom more heavily. Panel (c) presents the relative likelihood tests of the Normal and the skew-Normal, showing that the skew-Normal is overwhelmingly favored even after penalizing for the extra parameter. These results imply the distribution of firm *size* is not rejected as being log-skew-Normal in the data, for any of the six size measures considered. Notably, this fit is also excellent at the upper tail of the size distribution, i.e., for the largest firms. Because the model presented above lacks financial frictions, I will generally ignore the mild skewness in what follows.

Consider now the predictions of the Z-model. The driving process of the model is the productivity process Z_t, which follows AR(1) in logs. Due to the Central Limit Theorem (CLT) and as a general property of AR(1) processes, the ergodic distribution of z is Normal and of Z is hence log-Normal. Being the only stochastic driving force in the model, we can expect this normality to be inherited by the observable measures of firm scale. Because firm capital K_t is a choice (and hence outcome) variable of the model, ascertaining its

[13]For a review of the information-theoretic approach to model selection see, e.g., Burnham and Anderson (2002).

Table 3
Distributional tests

This table presents the results of tests of distributional form for the six scale measures of Table 2. 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) and (b) report the test statistics and their p-values *rejecting the relevant distribution* for the Normal and Skew-Normal, respectively. Panel (c) reports the relative likelihoods for each distribution using the AIC and BIC.

	λ	KT	VL	EQ	SL	XS
Panel (a): Normal						
K-S	0.018	0.024	0.026	0.029	0.010	0.022
p-val	0.039	0.030	0.029	0.025	0.059	0.033
C-2	104.8	189.1	199.8	191.1	46.98	160.5
p-val	0.037	0.028	0.027	0.027	0.054	0.030
A-D	7.819	16.06	16.91	17.18	2.886	11.92
p-val	0.036	0.027	0.027	0.026	0.052	0.031
Panel (b): Skew-Normal						
K-S	0.010	0.009	0.008	0.011	0.006	0.010
p-val	0.060	0.061	0.066	0.056	0.086	0.058
C-2	33.62	39.69	30.51	22.47	22.26	41.73
p-val	0.063	0.059	0.066	0.077	0.077	0.057
A-D	1.918	1.667	1.639	2.163	0.681	2.252
p-val	0.059	0.061	0.062	0.057	0.081	0.056
Panel (c): Relative likelihood tests						
AIC R.L.:						
N	0.000	0.000	0.000	0.000	0.000	0.000
SN	1.000	1.000	1.000	1.000	1.000	1.000
BIC R.L.:						
N	0.000	0.000	0.000	0.000	0.000	0.000
SN	1.000	1.000	1.000	1.000	1.000	1.000

distributional form analytically in the absence of closed form solutions to the model policy and value functions is difficult.

To overcome this difficulty, consider the simplified case of no adjustment costs ($\gamma \to 0$). In this case it is easy to show that a remaining firm will find it optimal to choose $k_{t+1} = c_1 + c_2 \cdot z_t$, with c_1, c_2 constants depending on the parameter values. Put differently, capital scale is a linear function of z and hence Normal as well. Under the interpretation of τ_z as a fixed expense ratio, the Z-model represents sales scale as the linear function $z_t + \theta_z \cdot k_t$ in Equation 8. Because both are Normal, their linear combination is Normal as well, and sales scale is hence Normal in the Z-model. Expenses under this interpretation are taken as a fixed percentage τ_z of sales and expense scale is hence Normal too. Firm income scale λ_t can be written as the average of sales scale and expenses scale, and is hence Normally distributed as well. Finally, in the absence of adjustment costs firm value is fully correlated with firm capital, as in Hayashi (1982), and hence log value distributes Normally as well. While adjustment costs may somewhat alter these predictions, the corporate finance literature cited above generally finds adjustment costs to be quite small in magnitude (i.e., low γ) and these distributional assumptions to generally hold in such models.

Similar reasoning applies to the SX-model, albeit more tenuously. The two processes S_t, X_t are jointly-AR(1) in logs, and hence s_t, x_t each distribute Normally, similar to z_t. Analytically finding the distribution of the outcome variable k_t is again difficult, even in the simple no adjustment costs case. If we were to *assume* capital scale k_t in the SX-model is also Normally distributed (as in the simple Z-model and in the data), one could then deduce that sales and expense scales are Normally distributed, implying in turn that firm income scale λ is also Normally distributed in the SX-model. Using a similar no-adjustment condition as before implies value scale (or log firm value) is approximately Normally distributed as well. I revisit and test the maintained assumption that k_t is approximately Normally distributed in Section 4 when estimating and simulating the SX-model.

In summary, both models predict the distribution of firm *scale* to be approximately

Normal, a prediction supported by the data.

3.2 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. They are also both the means and ends of firm growth. Nevertheless, the statistical distribution of income has seen scant interest in the economic literature, especially in contrast with the heavily-studied firm size distribution.

While the Z- and SX-models agree on the shape of the size distribution (approximately log-Normal for both models), the two deviate when one considers the distribution of firm income CF. As noted prior, the Z-model 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. Specifically, the Z-model predicts income to be approximately log-Normally distributed as well. The SX-model, in contrast, predicts each of sales and expenses to be log-Normal, and their difference, firm income, to distribute as the difference between two correlated log-Normal RVs.

The difference-of-log-Normals distribution arises due to a simple set of statistical facts: (i) both the sum and difference of two Normal RVs are Normal; (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 difference-of-log-Normals 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."

Panel (a) of Figure 2 presents a truncated view of the income distribution, in the limited range between −50M and +100M. Income clearly presents exponential tails in both the positive and negative directions, explaining the need for truncation in Panel (a). The common way of dealing with exponential tails, applying a log transform, cannot be used due to the negative values involved. Two candidate transforms that can deal with the double-exponential nature of the tails are

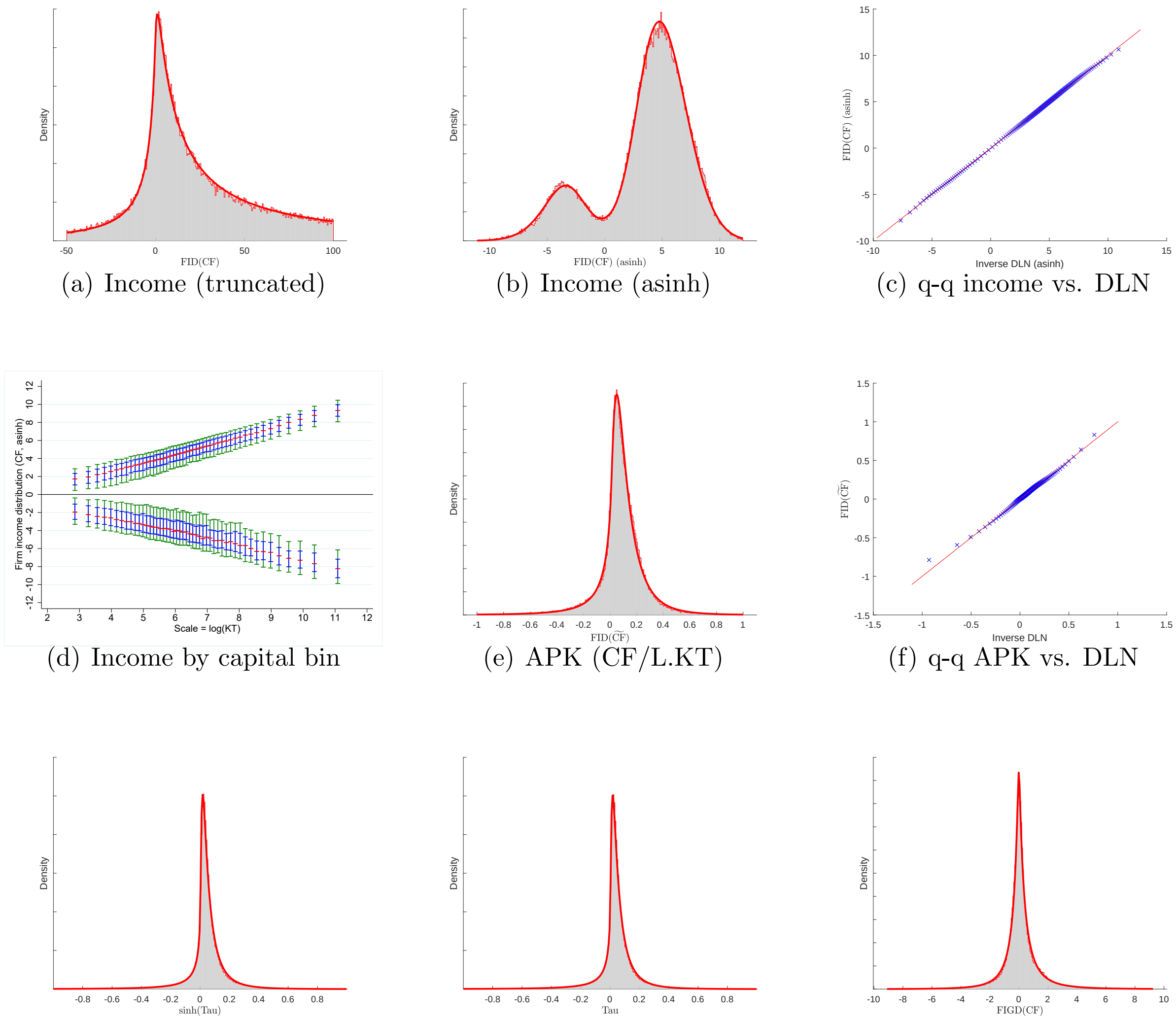
$$\begin{aligned} \text{neglog}(x) &= \text{sgn}(x) \cdot \log(1 + |x|) \xrightarrow[|x| \gg 0]{} \text{sgn}(x) \cdot \log(x) \\ \text{asinh}(x) &= \log(x + \sqrt{1 + x^2}) \quad \xrightarrow[|x| \gg 0]{} \text{sgn}(x) \cdot \left(\log(x) + \sqrt{0.5}\right) \end{aligned} \tag{21}$$

with the asinh transform the inverse of the Hyperbolic Sine function of Equation 16, and sgn() the sign function. The asinh transform, also known as IHS, has recently seen increased interest as an ad-hoc method of transforming non-positive economic values.[14] Here it instead arises as a natural and theoretically-based transform from the $\lambda\tau$-model.

Panel (b) hence presents the untruncated income distribution, with the x-axis under asinh-transform, or equivalently displaying the asinh of income CF. The two Normal distributions, one at the positive side of zero and one at the negative side, are evident. The panels are also overlaid with MLE-fitted difference-of-log-Normals distributions, exhibiting excellent fit. The fit is also seen in the appropriate q-q plot in Panel (c). The formal distributional tests reported in Table 4 support this conclusion. Income CF is not rejected as distributing difference-of-log-Normals using either test. The same is true for the top-down income definition CA, adding to the robustness of this result. It is worth noting that the SX-model theoretically predicted a specific, novel statistical distribution for cashflows which we have now confirmed in the data.

To further inspect the dependence of income on scale, panel (d) of Figure 2 presents a closer look at the distribution of income. I first split the data into 49 equal bins, based on firm capital KT, ignoring the top and bottom 1% of observations, such that each bin contains 2% of the observations. For each bin, Panel (d) plots the $(10, 25, 50, 75, 90)^{th}$ percentiles of (asinh) income, separately for positive and negative income values. Larger firms earn and

[14] See e.g. Bellemare and Wichman (2020); Chen and Roth (2024); Mullahy and Norton (2024) for a review. The asinh trades off fixed bias at large $|x|$ with less bias near $x = 0$, and is also differentiable everywhere.



(a) Income (truncated) (b) Income (asinh) (c) q-q income vs. DLN

(d) Income by capital bin (e) APK (CF/L.KT) (f) q-q APK vs. DLN

(g) Income efficiency (sinh) (h) Income efficiency (i) Income growth

Fig. 2. Firm income distributions. Panel (a) presents the (truncated) distribution of CF in linear scale while Panel (b) presents the untruncated distribution in asinh scale. Panel (c) presents the q-q plot corresponding to Panel (b). Panel (d) 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 (e) presents income intensity (average product of capital), and panel (f) the corresponding q-q plot. Panels (g),(h) present the distributions of sinh(τ) and firm income efficiency τ, respectively. Panel (i) presents income growth, given by Equation 27. Panels (a),(b),(e),(g),(h),(i) are overlaid with MLE-fitted difference-of-log-Normals distributions, which is the distribution tested against in the q-q plots of Panels (c),(f).

Table 4
Distributional tests

This table presents the results of tests of distributional form for income (CF), alt. income (CA), APK (CF/L.KT), transformed and raw income efficiency (sinh(τ) and τ), income growth (dCF), alt. income growth (dCA), sales growth (dSL), capital growth (dKT), and firm value growth (dVL). 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) and (b) report the test statistics and their p-values *rejecting the relevant distribution* for the Normal and difference-of-log-Normals, respectively. Panel (c) reports the relative likelihoods for each distribution using the AIC and BIC.

	CF	CA	APK	sinh(τ)	τ	dCF	dCA	dSL	dKT	dVL
Panel (a): Normal										
K-S	>0.5	>0.5	0.207	0.260	0.254	0.262	0.259	0.229	0.211	0.202
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C-2	>999	>999	>999	>999	>999	>999	>999	>999	>999	>999
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A-D	>999	>999	774.8	>999	>999	>999	>999	>999	>999	>999
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Panel (b): Difference-of-log-Normals										
K-S	0.003	0.003	0.005	0.011	0.009	0.007	0.005	0.006	0.003	0.008
p-val	0.144	0.156	0.104	0.056	0.066	0.074	0.095	0.087	0.155	0.069
C-2	8.493	11.75	15.18	22.47	20.23	59.03	25.62	17.44	7.633	39.91
p-val	0.140	0.110	0.096	0.077	0.078	0.049	0.073	0.089	0.171	0.059
A-D	0.219	0.186	0.390	2.163	1.632	1.114	0.691	0.505	0.152	1.080
p-val	0.115	0.121	0.095	0.057	0.068	0.070	0.080	0.089	0.130	0.070
Panel (c): Relative likelihood tests										
AIC R.L.:										
N	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.:										
N	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

lose more money than smaller firms, so the "middle" of the panel hollows as firm scale rises. Consider the following thought experiment: if you knew a firm had \$1B of profits last year, but this year it suffered an unknown loss — what would be your guess regarding the magnitude of the loss? A single-factor thinking would imply the best guess is some small negative loss, close to zero. The data imply the correct intuition is that the firm had losses of about \$800M this year, stemming from the interaction of scale λ (setting the magnitude) and operational efficiency τ (setting the sign).

3.3 Efficiency

Consider next an important property of the difference-of-log-Normals distribution: it is closed under multiplication and division by a log-Normal RV — a property it inherits from the closure of the Normal distribution to addition and subtraction. I.e., if L_1, L_2, L_3 are three (possibly correlated) log-Normal RVs, then $(L_1 - L_2) \sim \text{DLN}$ (by definition) but also $(L_1 - L_2) \cdot L_3 \sim \text{DLN}$ and $(L_1 - L_2)/L_3 \sim \text{DLN}$.

An example is the average product of capital APK in the SX-model:

$$\begin{aligned} \text{APK} = \frac{Y_t}{K_t} &= \frac{\exp\left(s_t + \theta_s \cdot k_t\right) - \exp\left(x_t + \theta_x \cdot k_t\right)}{\exp\left(k_t\right)} \\ &= \exp\left(s_t + (\theta_s - 1) \cdot k_t\right) - \exp\left(x_t + (\theta_x - 1) \cdot k_t\right) \sim \text{DLN} \end{aligned} \tag{22}$$

Put differently, the SX-model predicts the APK to distribute difference-of-log-Normals, whereas a similar analysis will yield a prediction that the APK distributes log-Normally in the Z-model (and again barring negative APK values). Panel (e) of Figure 2 presents the APK in the data, with MLE-fitted difference-of-log-Normals distribution, and Panel (f) presents the matching q-q plot. The appropriate column in Table 4 presents the formal statistical tests. Normality is strongly rejected but the fit of the APK to the difference-of-log-Normals distribution is both visually and statistically sound.

A similar analysis applies to firm income efficiency τ, slightly transformed. While in the

Z-model firm efficiency is constant, note that in the SX-model we can rewrite Equation 17 as,

$$\sinh(\tau_t) = \frac{Y_t}{2 \cdot \exp(\lambda_t)} \sim \text{DLN} \tag{23}$$

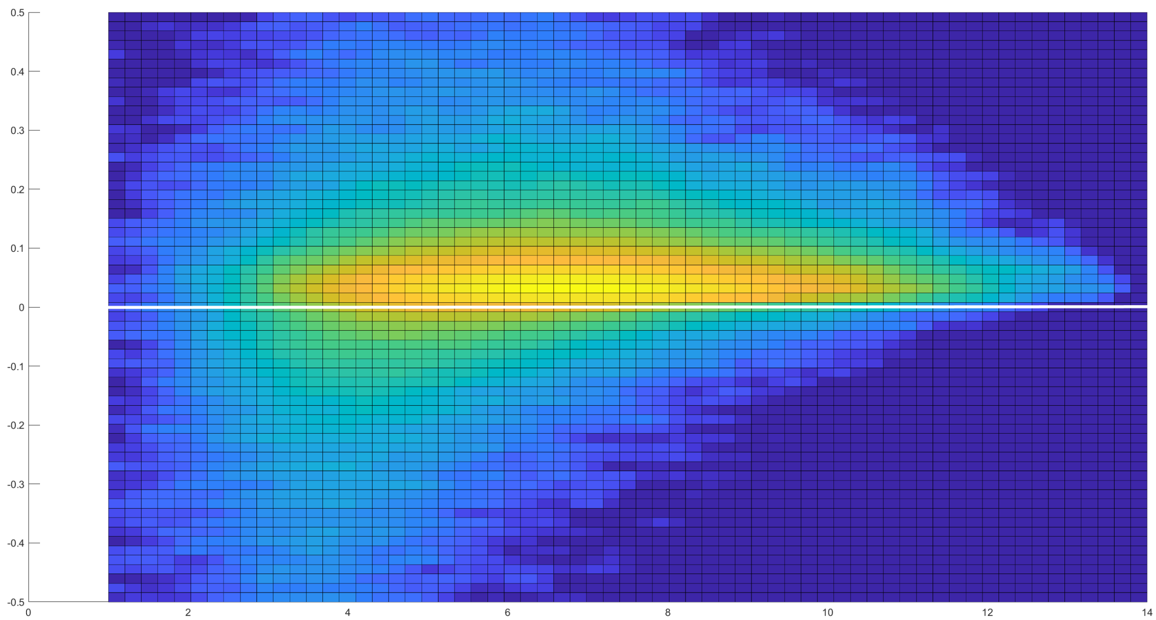
and the result follows because firm scale λ_t in the SX-model is Normal, as discussed above. Panel (g) of Figure 2 ascertains this prediction, and the appropriate column of Table 4 confirms with formal tests.

Note that the prediction that $\sinh(\tau_t) \sim$ DLN *does not* necessarily imply $\tau \sim$ DLN, because the sinh() of a Normal RV distributes difference-of-log-Normals as well.[15] Interestingly, Panel (h) of Figure 2, as well as the formal tests in Table 4, indicate the un-transformed value τ distributes difference-of-log-Normals in the data as well. This is a deviation from the prediction of the $\lambda\tau$-model, as currently specified. The model predicts τ should distribute Normally, as it follows an AR(1) with Normal innovations. I return to this (intentional) discrepancy in Section 4.4.

Figure 3 presents a heat map of the income scale and efficiency for all observations in the data. We can see the 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 $\lambda = 6.5$, as seen prior. The profit/loss line at $\tau = 0$ appears 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.[16] The systematic dependence of τ's dispersion on λ (i.e. the decreasing dispersion of τ as scale λ increases) is another salient feature of the data not captured by the model, and I return to it in Section 4.4 as well.

[15] If $n_1 \sim \mathbb{N}$, then $\sinh(n_1) = 0.5 \cdot (\exp(n_1) - \exp(-n_1)) \sim$ DLN.

[16] I.e., the coefficient of regressing τ_t on λ_t is 0.005, with a within-$R^2 < 0.01$.



(a) Firm scale and efficiency heat-map

Fig. 3. 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.

3.4 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 (stemming e.g. from \$1B in sales and \$1.1B in expenses) and \$200M of profits this year (stemming e.g. from \$1.2B in sales and \$1B in expenses)? Parham (2023) discusses extending the instantaneous growth definition of Barro and Sala-I-Martin (2003), who use the definition $\frac{dY_t/dt}{Y_t}$, to RVs possibly taking negative values, by using the definition $\frac{dY_t/dt}{|Y_t|}$. The absolute value in the definition of growth is necessary to maintain the 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 "signed" percentage growth in income

$$\frac{dY_t/dt}{|Y_t|} \approx \left(Y_{t+1} - Y_t\right)/|Y_t| \tag{24}$$

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

$$\begin{aligned} \frac{dY_t/dt}{|Y_t|} = \frac{Y_t \cdot \left(\frac{dz_t}{dt} + \theta_z \cdot \frac{dk_t}{dt}\right)}{Y_t} &\approx \left(z_{t+1} + \theta_z \cdot k_{t+1}\right) - \left(z_t + \theta_z \cdot k_t\right) \\ &= \log(Y_{t+1}) - \log(Y_t) \equiv \text{dlog}(Y_{t+1}) \end{aligned} \tag{25}$$

yielding the familiar difference-in-logs growth measure, denoted dlog(). The approximation again arises from the forward discrete difference.

The log-point growth of income in the Z-model can also be written as

$$\text{dlog}(Y_{t+1}) = (1 - \rho_z) \cdot (\mu_z - z_t) + \theta_z \cdot (k_{t+1} - k_t) + \epsilon^z_{t+1} \tag{26}$$

with z_t distributing Normally as a property of the AR(1) process, ϵ^z_{t+1} distributing Normally by Equation 10, 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 the simple no-adjustment case 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, which is derived from the Z-model's log-linear 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 production function, which yields the result

$$\frac{dY_t/dt}{|Y_t|} \approx \frac{\mathbb{S}_t \cdot \text{dlog}\left(\mathbb{S}_{t+1}\right) - \mathbb{X}_t \cdot \text{dlog}\left(\mathbb{X}_{t+1}\right)}{|\mathbb{S}_t - \mathbb{X}_t|} \tag{27}$$

that expresses income growth as a weighted sum of (log-point) sales growth and expenses growth. This equation is useful for calculating income growth when one only has the values of sales and expenses (and their lag), and is robust to negative income values. I.e., it provides a proper growth measure that can answer the question posited in the opening paragraph of this section. The equation is not, however, conducive to analytically exploring the resulting distribution of income growth.

A fourth way to define income growth, equivalent to the third but yielding considerably more intuition and analytical ease, is to define income growth using the $\lambda\tau$-model. In this case,

$$\frac{dY_t/dt}{|Y_t|} = \text{sgn}(\tau_t) \cdot \left[\frac{d\lambda_t}{dt} + \frac{d\tau_t}{dt} \cdot \frac{1}{\tanh\left(\tau_t\right)}\right] \approx \text{sgn}(\tau_t) \cdot \left[(\lambda_{t+1} - \lambda_t) + \frac{\tau_{t+1} - \tau_t}{\tau_t}\right] \tag{28}$$

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 $\tanh(\tau_t) \approx \tau_t$.[17]

Consider the two terms in the large brackets. The first term is the difference in firm income scales. Recall that firm income scale in the $\lambda\tau$-model is a scale (i.e., logged) value, and hence this term is similar to the dlog() measure from the Z-model, which also measures the difference between "income scales" (i.e., logs of income). The second term is novel, and captures the *percent growth* in firm efficiency. The implications of the second term, and especially the fact that it captures percent growth in efficiency rather than simple difference in efficiency, are notable. Explosive income growth (or the heavy tails of income growth) occurs due to *operational leverage*, or a "low base" in τ (i.e., τ_t close to zero). This then leads to high growth in income.

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 28 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 28 yields income growth of $(0.126 - 0.026)/0.026 = 3.9$ log-units, equal to the percent growth of income at $245/50 - 1 = 390\%$. The low-base phenomenon described here is especially salient given our prior observation that τ is tightly concentrated close to 0 in the data.

The correlation in the data between the $\lambda\tau$-based growth measure from Equation 28 and the "signed" percentage growth of income from Equation 24 is above 0.97. Furthermore,

[17]Note also that $-1 \leq \tanh() \leq 1$.

Table 5 ascertains that nearly all variation in income growth in the data stems from the dynamics of percentage growth in the τ term, rather than from the dynamics of differences in τ or differences in λ. This is another novel prediction of the model confirmed by the data.

Table 5
Determinants of income growth

This tale presents regressions of income growth (CF_{t+1} - CF_t)/abs(CF_t) on changes in firm scale $\text{d}\lambda = \lambda_{t+1} - \lambda_t$, changes in firm efficiency $\text{d}\tau = \tau_{t+1} - \tau_t$, and percent changes in firm efficiency $\%\tau = (\tau_{t+1} - \tau_t)/\tau_t$. Four specifications are presented in the respective columns. All regressions include firm and year fixed-effects, w/ N=165K.

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

Even using the $\lambda\tau$-formulation, it is difficult to analytically pin down the distribution of income growth resulting from the SX-model, beyond the conclusion that it must be heavy-tailed rather than Normal due to the low-base effect. Nevertheless, Panel (i) of Figure 2 presents the distribution of income growth, as defined by Equation 28, in the data along with a fitted difference-of-log-Normal distribution. Here again, Normality is strongly rejected, while the fit to the difference-of-log-Normals distribution is excellent, as Table 4 confirms for the growth of both CF and the alternate definition CA.

3.5 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).

For the Z-model, each of APK and MPK are log-Normally distributed, and in fact, $\text{MPK} = \theta_z \cdot \text{APK}$, yielding the well-known results that all firms, regardless of their state, always have $\text{RTS} = \theta_z$.

For the SX-model, in contrast, we have both APK and MPK distributing difference-of-log-Normals, and their ratio, RTS, given by:

$$\text{RTS}_t = \frac{\theta_s \cdot \mathbb{S}_t - \theta_x \cdot \mathbb{X}_t}{\mathbb{S}_t - \mathbb{X}_t} \tag{29}$$

or a form of "weighted average" of the two scale parameters of the model, θ_s and θ_x. Notably, the SX-model no longer implies constant RTS for all firms. Firms have different RTS depending on their current sales and expenses, even if all firms in the economy share the same scale parameters.

Better intuition regarding the mechanics and distribution of RTS can again come from equivalently writing Equation 29 in terms of λ, τ using the $\mathbf{Y}_{\lambda\tau}$ production function. In this case, we have:

$$\text{RTS} = \theta_\lambda + \frac{\theta_\tau}{\tanh(\tau_t)} \approx \theta_\lambda + \frac{\theta_\tau}{\tau_t} \tag{30}$$

with $\theta_\lambda, \theta_\tau$ the scale parameters given by Equation 17. The model proposes a two-part schedule for RTS. The first, constant term θ_λ, is the counterpart of the constant θ_z in the Z-model, while the second term is novel and dependant on (inverse) τ. The division by τ_t causes RTS to explode to $\pm\infty$ when $|\tau_t| \to 0$. This low-base effect, along with the large mass of firms around $\tau = 0$ in the data, then imply a heavy-tailed distribution of RTS. RTS is however unobservable, due to the unobservability of MPK, so these predictions cannot easily be tested.

3.6 Firm growth

The last data and model implications we consider are regarding firm growth. The definition of firm growth naturally depends on the firm size measure used — the growth of *what*? — e.g. sales, capital, or value growth. For any measure of firm size M_t , firm growth is defined as the log-point difference between periods (possibly corrected for dividends): $dM_t \equiv \text{dlog}(M_t) = \log(M_t) - \log(M_{t-1})$. Put differently, firm growth is the increase (or decrease) in firm scale.

For example, using the data mnemonics from Table 1, firm capital growth in the data is defined as $dKT_t = \log(KT_t) - \log(KT_{t-1})$, or the increase in capital scale between periods (in the models $k_t - k_{t-1}$). Firm equity value growth however corrects for dividends and is defined as $dEQ_t = \log(EQ_t + DE_t) - \log(EQ_{t-1})$, with EQ_t market value of equity at the end of year t and DE_t distributions to equity holders during year t. Note that dEQ_t is the dividend-adjusted buy-and-hold return on firm equity. This definition highlights the fact that the distribution of firm growth is possibly one of the most-studied distributions in economics, owing to the fact equity returns are merely one measure of firm growth.

Consider first the data. Figure 4 presents the distributions of sales growth, capital growth, and value growth for the sample, along with the respective q-q plots. All three are notably heavy-tailed, and all three present remarkable fit to the difference-of-log-Normals distribution. Normality is easily rejected, but difference-of-log-Normals is not rejected for any of the distributions, as the relevant columns of Table 4 show. Equity returns are a special type of firm growth metric, because data are available at higher frequency than annually. Figure 5 hence presents the distributions of yearly, monthly, and daily stock returns for the sample. The fit to the difference-of-log-Normals is again remarkable, as the q-q plots of Figure 5 and the formal tests of Table 6 show.[18] The yearly returns presented in Figure 5 are raw, but monthly and daily returns are the excess returns from a Fama-French 3-factor model. This is to demonstrate that these results hold even for factor-model residuals.

Given our observations so far regarding the difference-of-log-Normals or heavy-tailed distributions of firm income, income growth, APK, and MPK in the data and the SX-model, the results above for firm growth may not be surprising. These economic magnitudes

[18] In unreported results, I find that the Stable and Laplace distributions - two candidate return distributions discussed in the literature, are both strongly rejected.

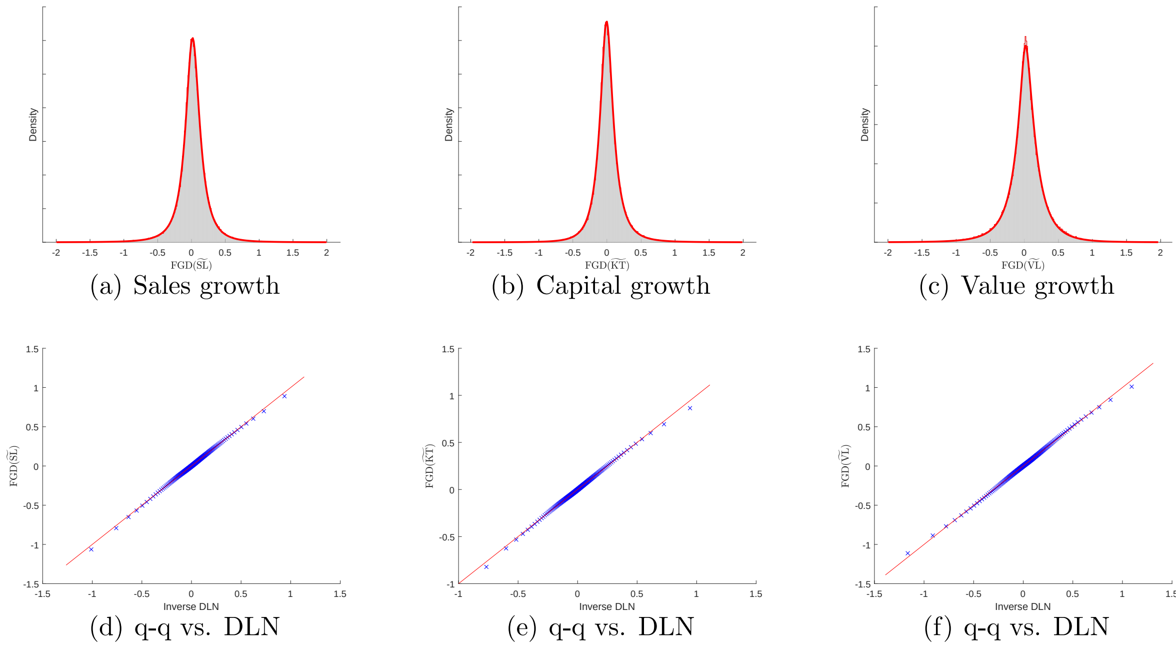


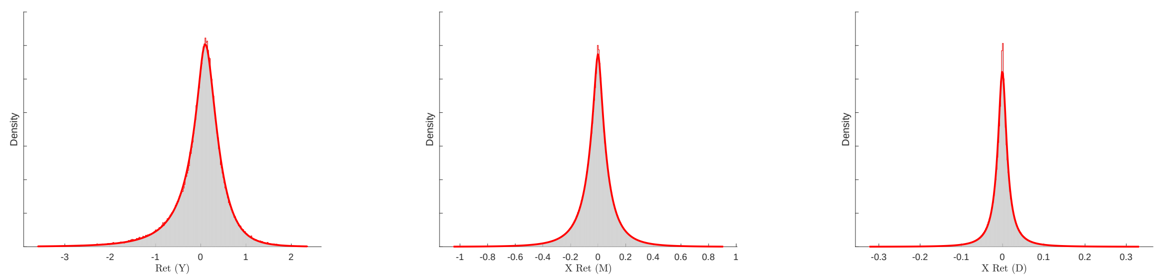
Fig. 4. Firm growth distributions. Panels (a)-(c) present the histograms of sales (SL) growth, capital (KT) growth, and firm value (VL) growth, respectively, for a set of 143K firm-year observations from 1970-2019. The histograms are overlaid with MLE-fitted difference-of-log-Normals distributions. Panels (d)-(f) present the respective q-q plots.

Table 6

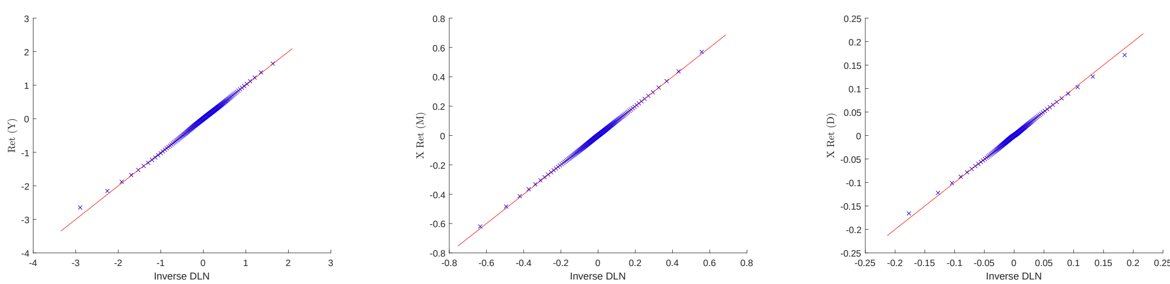
Equity returns - Distributional tests

This table presents results of tests of distributional form vs. the difference-of-log-Normals distribution for equity returns at the yearly, monthly and daily frequencies. The table presents both raw and excess returns relative the the Fama-French 3-factor model. 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.

	Yearly		Monthly		Daily	
	Raw	FF3	Raw	FF3	Raw	FF3
K-S	0.002	0.003	0.002	0.003	0.006	0.007
p-val	0.256	0.166	0.277	0.148	0.084	0.080
C-2	3	6	2	9	18	21
p-val	0.962	0.199	1.000	0.136	0.086	0.078
A-D	0.03	0.13	0.05	0.19	0.28	0.47
p-val	0.297	0.138	0.199	0.120	0.105	0.090



(a) Raw (Y) w/ DLN (b) FF3 (M) w/ DLN (c) FF3 (D) w/ DLN



(d) q-q Raw (Y) vs. DLN (e) q-q FF3 (M) vs. DLN (f) q-q FF3 (D) vs. DLN

Fig. 5. Equity return distributions. Panels (a),(c),(e) present the histograms of yearly, monthly, and daily equity returns (in log-point units). Yearly returns are raw while monthly and daily returns are excess returns relative to the Fama-French 3-factor model. The panels are overlaid with fitted difference-of-log-Normals distributions. Panels (b),(d),(f) present the respective q-q plots.

are at the heart of the dynamic program of the firm, and it is intuitive that their impact will propagate to the distribution of firm growth. Nevertheless, the finding above that the heavily-studied equity returns distribution cannot be rejected as a difference-of-log-Normals — a novel distribution predicted by the SX-model — merits highlighting.

Directly ascertaining the models' implications for firm growth is again difficult because they lack closed-form solutions to the value and policy function. Despite that, some progress can be made in two avenues: by analytically considering the simplified case in which $\gamma \to 0$, and via model estimation and simulation. I discuss the first method here before moving on to simulations in the next section.

Because capital is the index of size in the model, it is useful to consider the case of capital growth. Note that for $\gamma \to 0$ case, MPK is the core determinant of firm size and growth in the general q-theory model. 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.

The Z-model, unsurprisingly, predicts approximately Normal capital growth, as:

$$\begin{aligned} MPK_t &= \theta_z \cdot \exp(z_t - (1-\theta_z) \cdot k_t) \\ K_{t+1} &= \exp\left(\frac{C_z + \rho_z \cdot z_t}{1-\theta_z}\right) \\ C_z &= \log\left(\frac{\theta_z}{r+\delta}\right) + (1-\rho_z) \cdot \mu_z + \frac{\sigma_z^2}{2} \\ \text{dlog}(K_{t+1}) &= \frac{\rho_z}{1-\theta_z} \cdot (z_{t+1} - z_t) \end{aligned} \tag{31}$$

with C_z a constant depending on the parameters of the model. Capital growth is a linear function of changes in the stochastic state variable z_t, which are Normally distributed in the Z-model based on the assumption regarding the Normal dynamics of z in Section 2.3.

The SX-model, in contrast, yields:

$$\begin{aligned} MPK_t &= \theta_s \cdot \exp\left(s_t - (1-\theta_s)\cdot k_t\right) - \theta_x \cdot \exp\left(x_t - (1-\theta_x)\cdot k_t\right) \\ &\underbrace{\exp(C_s + \rho_s \cdot s_t - (1-\theta_s)\cdot k_{t+1})}_{g_s(s_t,k_{t+1})} - \underbrace{\exp(C_x + \rho_x \cdot x_t - (1-\theta_x)\cdot k_{t+1})}_{g_x(x_t,k_{t+1})} = 1 \end{aligned} \tag{32}$$

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 difference of log-linear functions, equals 1. Even with Normal innovations to s_t and x_t, as the SX-model posits, their exponential opposing forces then yield difference-of-log-Normals growth in capital.

To further analyze the distributions governing the firm without the limiting assumption of no adjustment costs ($\gamma \to 0$), I next estimate and simulate the models, yielding numeric value and policy functions that can then be tested.

4 Estimation and simulation of models

This section describes taking the Z- and $\lambda\tau$-models to data via indirect inference. I concentrate on the $\lambda\tau$ rather than the equivalent SX formulation of the SX-model because it resolves the challenge of estimating models with highly correlated stochastic variables.[19] After presenting the identification strategy and how to obtain initial parameter values, I provide the estimation results. Estimation is relatively straightforward, owing to the observability of sales, expenses, and capital. I then simulate the estimated models in Section 4.3 and consider the distributions of firm outcomes in the models vs. the data. The two-factor $\lambda\tau$-model is able to replicate the difference-of-log-Normals distribution of firm outcomes, and also fits several important un-targeted firm moments considerably better than the Z-model.

[19] The stochastic variables controlling sales and expenses in the SX-model, s and x, are highly correlated in the data as expected (with correlation coeff. > 0.95). Using $\widehat{\lambda} = (s+x)/2$ and $\widehat{\tau} = (s-x)/2$ resolves the problem and $\widehat{\lambda}$, $\widehat{\tau}$ are nearly uncorrelated.

4.1 Identification and initial parameter values

The Z- and $\lambda\tau$-models are described by the respective parameter vectors

$$\begin{aligned}\Theta_z &= \{r, \delta, \tau_z, \theta_z, \rho_z, \mu_z, \sigma_z, \gamma, \nu\} \\ \Theta_{\lambda\tau} &= \{r, \delta, \theta_\lambda, \rho_\lambda, \mu_\lambda, \sigma_\lambda, \theta_\tau, \rho_\tau, \mu_\tau, \sigma_\tau, \rho_{\lambda\tau}, \gamma, \nu\}\end{aligned} \tag{33}$$

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 θ values $(\theta_z, \theta_\lambda, \theta_\tau)$ can be derived from steady-state arguments regarding returns-to-scale (RTS), conditional on r, δ; and (iii) The stochastic state variables $z_t, \widehat{\lambda}_t, \widehat{\tau}_t$ are observable, conditional on their respective θ values, allowing us to estimate their dynamic AR(1) parameters directly.

Pinning down δ is easy because firms generally report their depreciation expenses. The observed values of both the depreciation rate DP/L.KT and investment rate IT/L.KT are tightly packed around 0.04 in the data. I hence set $\delta = 4\%$. Similarly, I set $r = 4\%$ because the observed values of both dispensation yield DI/L.VL and debt payout ratio DD/L.DB are tightly packed around 0.04 as well. I also set the expense parameter of the Z-model τ_z to a value of exp(-2*0.033) = 0.936, the (transformed) value of the efficiency "ridge" in Figure 3, or equivalently the median ratio of XS/SL in the data.

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

$$RTS^* = \frac{r+\delta}{\text{APK}} = \begin{cases} \theta_z & \text{if } Y_t = \mathbf{Y}_z() \\ \theta_\lambda + \frac{\theta_\tau}{\tanh(\tau_t)} & \text{if } Y_t = \mathbf{Y}_{\lambda\tau}() \end{cases} \tag{34}$$

Because APK and τ_t are observable and we have already established values for r, δ, we can use this equation to estimate initial guesses for the θ values, when considering firms plausibly in steady-state. Doing so yields initial values: $\theta_z = 0.7, \theta_\lambda = 0.268, \theta_\tau = 0.015$. Equivalently, this implies $\theta_s = 0.283, \theta_x = 0.253$.

The marked difference between the initial RTS guesses of the Z-model and the $\lambda\tau$-model 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. With τ in the data clustered around 0.033, the two-part schedule in the $\lambda\tau$-model 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 the stochastic state variables $z_t, \widehat{\lambda}_t, \widehat{\tau}_t$ can be seen by rewriting the definitions of $\mathbf{Y}_z()$ and $\mathbf{Y}_{\lambda\tau}()$ in Equations 8 and 17 as

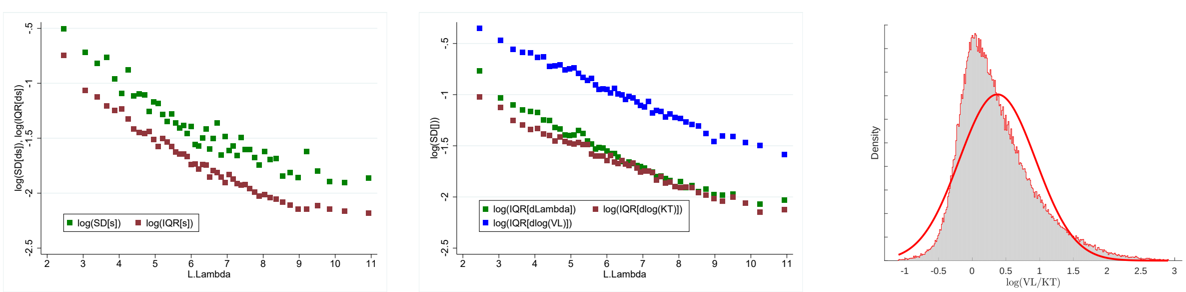
$$
\begin{aligned}
z_t &= \log\left((1-\tau_z)\cdot \mathbb{S}_t\right) - \theta_z \cdot \log\left(K_t\right) \\
\widehat{\lambda}_t &= \lambda_t - \theta_\lambda \cdot \log\left(K_t\right) \\
\widehat{\tau}_t &= \tau_t - \theta_\tau \cdot \log\left(K_t\right)
\end{aligned}
\tag{35}
$$

and noting that sales and expenses (and hence scale and efficiency) as well as capital $(\mathbb{S}_t, \mathbb{X}_t, \lambda_t, \tau_t, K_t)$ are all observable. This means knowledge (or a guess) of the θ values allows us to observe all state variables in the two models. Using the imputed values for the stochastic state variables, we can then 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, raises a challenge. The models assume innovations to the stochastic variables (i..e, $\epsilon_z, \epsilon_\lambda, \epsilon_\tau$) are homoscedastic — have dispersion independent of firm scale — such that e.g. $\text{SD}\left[s_{t+1} - \rho_s \cdot s_t\right] = \text{SD}\left[\epsilon^s_{t+1}\right] \equiv \sigma_s \forall s$. The data however exhibits clear decreasing dispersion (i.e. heteroscedasticity) with scale of the innovations to the stochastic variables, similar to our findings prior about the decreasing dispersion of τ with scale.

Panel (a) of Figure 6 presents the binned log(SD[]) and log(IQR[]) of the innovations to the stochastic variable controlling sales s, by scale. The panel presents systematic decreasing dispersion, 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 (b). I return to this puzzling finding as well at Section 4.4 below. In light of the evidence in Panels (a) and (b), I set $\sigma_{\square}$ to match the dispersion of innovations around the median scale in the data, $\lambda = 6.5$.

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 than remain. 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 (c). The exit-induced 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 approximately 0.85 in the data.



(a) DDWS s growth (b) DDWS firm growth (c) log average-q

Fig. 6. Estimation stylized facts. For each of 49 scale bins, panel (a) presents the binned (log) dispersions, log(SD[]) and log(IQR[]), of the growth in the stochastic variable s, controlling sales. Panel (b) repeats for log(IQR[]) of the growth in: firm scale λ, assets KT, and value VL. Panel (c) presents the distribution of log average-q log(VL/KT), overlaid with a fitted Normal.

The recipe for initial guesses is also used to determine which moments we should match in the estimation procedure. For θ_z, θ_λ, I use the median of RTS* from Equation 34, along with its IQR to identify θ_τ. 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. I was unable to find a natural initial guess for the adjustment cost parameter γ, though the prior work (cited above) generally finds fairly low values for γ, around 0.01-0.1. The initial values for γ are hence set such that capital growth 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) θ_z or $\theta_\lambda, \theta_\tau$ values, and in the second step, I conduct a full method of simulated moments (MSM) estimation conditional on the θ 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 θ, leading to a dependence of their moments on θ as well. Table 7 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. It is notable that the estimated parameter values are very close to their initial guesses, and that the identifying moments are matched very well too. Furthermore, the AR(1) parameters are also very close to the values of their identifying moments, implying quasi-observability works well as an identification strategy. The RTS parameter θ_z is estimated to be close to 0.7, a common value in such models.

Panel (b) of Table 7 moves on to the estimation results of the $\lambda\tau$-model. The model 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 work, and the persistence of growth is again well-matched by the model. Overall, both exactly-identified models are able to match their identifying moments and appear to be well-identified.

4.3 Simulation

While the models are able to match their identifying moments, the core questions in this work revolve around un-targeted moments and distributions. 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 8 presents the values of some un-targeted moments in the data and the two 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. The

Table 7
Estimation results

Panels (a) and (b) present the results of estimating the Z-model and $\lambda\tau$-model, respectively. Parameters' initial and estimated values are reported in the first four columns. The values of the corresponding identifying moments at the initial and estimated parameter values, as well as in the data, are reported in the next four columns. t-val is the t-statistic on the moment's error (Estim-Data). 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 35.

Panel (a): Z-model

		Value at:			Value at:			
	Name	Init	Estim	Moment	Init	Estim	Data	t-val
θ_z	RTS	0.699	0.685	MED[RTS*][a]	0.713	0.699	0.699	-0.021
ρ_z	z pers.	0.956	0.959	RHO[z]	0.957	0.959	0.959	-0.307
μ_z	z mean	2.141	2.232	MED[z]	2.141	2.232	2.232	-0.037
σ_z	dz std.	0.127	0.126	IQR[dz][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

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

		Value at:			Value at:			
	Name	Init	Estim	Moment	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

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

Z-model yields 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 3.4 and 3.6. The same is not true for the $\lambda\tau$-model. As predicted, the kurtosis of income growth, capital growth, and value growth are all considerably greater than 3, and the $\lambda\tau$-model matches the kurtosis values in the data fairly well, even without having *any moments* of kurtosis targeted in the estimation.

Table 8
Estimation results

This table presents moments of the Data, Z-model, and $\lambda\tau$-model, respectively. The moment values for each model are at the estimated parameter values of Table 7. 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	$\lambda\tau$	s.e.
MED[cf]	4.441	3.622	4.291	0.030
IQR[dcf]	0.436	0.222	0.561	0.004
KUR[dcf]	8.114	2.989	6.240	0.319
MED[k]	6.527	6.037	6.658	0.029
IQR[dk]	0.130	0.270	0.135	0.002
KUR[dk]	14.426	2.998	11.751	1.152
MED[v]	6.850	6.572	6.925	0.031
IQR[dv]	0.256	0.193	0.153	0.002
KUR[dv]	6.901	3.003	7.522	0.705
MED[λ]	6.626	6.334	6.800	0.028
IQR[$d\lambda$]	0.132	0.222	0.125	0.002
KUR[$d\lambda$]	20.539	2.989	3.856	1.863
MED[τ]	0.033	0.033	0.037	0.001
IQR[$d\tau$]	0.023	N/A	0.022	0.001
KUR[$d\tau$]	72.509	N/A	2.988	9.900
MED[$v-k$]	0.187	0.527	0.254	0.006
IQR[$v-k$]	0.449	0.117	0.151	0.006
MED[RTS*]	0.698	0.698	0.679	0.006
IQR[RTS*]	0.359	0.028	0.356	0.010

A visual comparison of these results is provided in Figure 7. The figure presents, for the data and the simulations of the Z- and $\lambda\tau$-models, histograms of (asinh) income cf, income growth dcf, capital growth dk, and adjusted value growth (i.e. returns) dv. The data and $\lambda\tau$

simulation histograms are overlaid with MLE-fitted difference-of-log-Normal distributions, while the Z simulation is overlaid with MLE-fitted Normal distributions. The visual fit of the data distributions to the difference-of-log-Normal is excellent, as previously ascertained in Table 4. The Z-model distributions again appear Normal and exhibit no heavy tails, while the $\lambda\tau$-model distribution are difference-of-log-Normal. 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.

Formal distributional tests for the simulated cf, dcf, dk, and dv in the Z- and $\lambda\tau$-models, vs. the Normal, skew-Normal, and difference-of-log-Normals are reported in Table 9. For the Z-model, none of the variables is rejected as a Normal. 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-model. For the $\lambda\tau$-model, Normality and skew-Normality are rejected for all four, while the difference-of-log-Normals is not rejected for any of the four. The relative likelihood test again overwhelmingly prefers the difference-of-log-Normals 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 8. Recall that we assumed all stochastic innovations are Normal in Section 2.5. 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 data factoid shortly, along with the other puzzling factoids 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*].

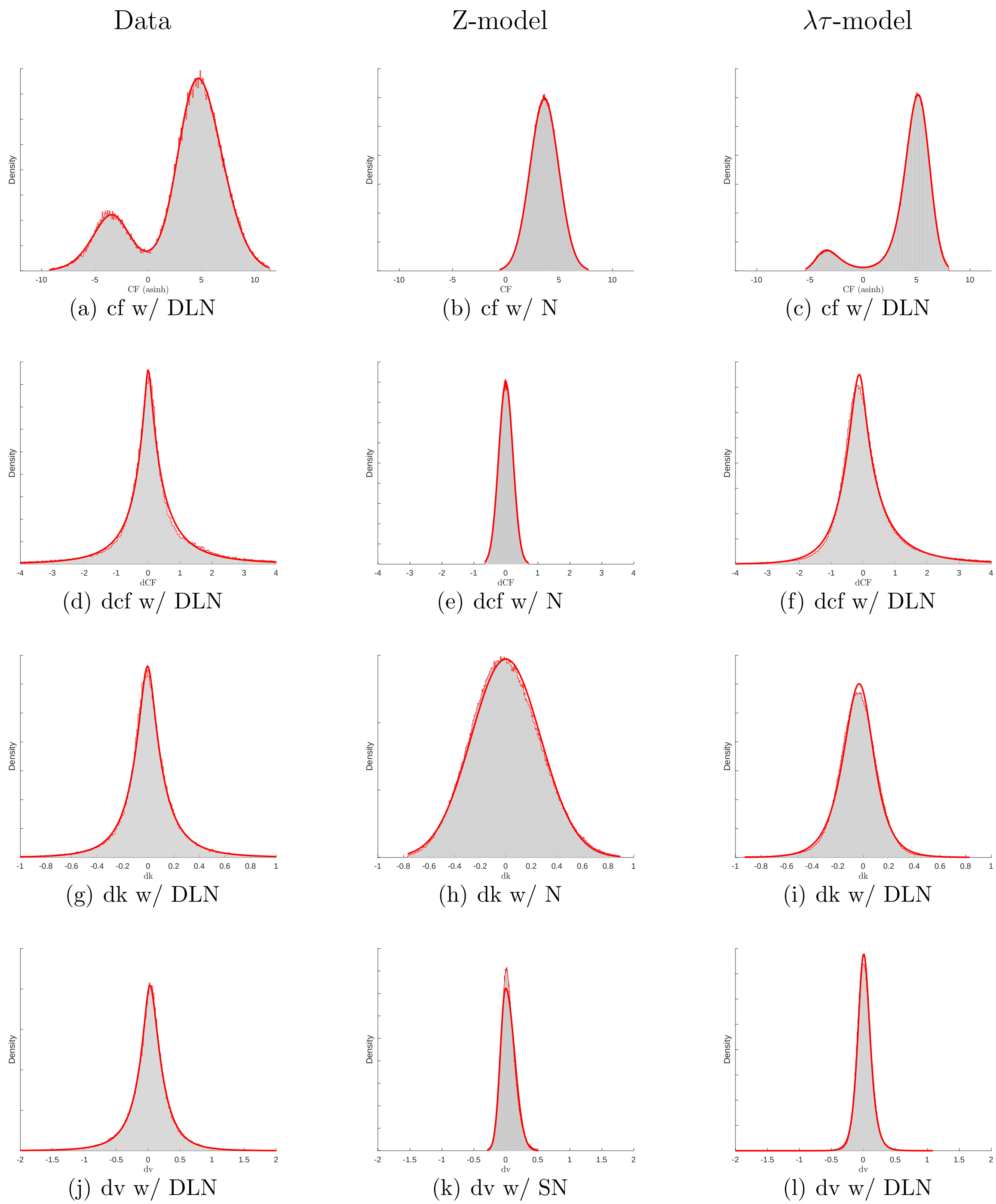


Fig. 7. Data and Model distributions. This figure presents the histograms of several firm variables in the data and the Z and $\lambda\tau$ models. The variables presented are (asinh) income cf, income growth dcf, capital growth dk, and value growth dv. The figures are overlaid with MLE-fitted distributions as indicated.

Table 9
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 Z 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 difference-of-log-Normals, respectively. Panel (d) reports the relative likelihoods for each distribution using the AIC and BIC.

	Z-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): Difference-of-log-Normals								
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.:								
N	1.000	0.004	0.000	1.000	0.000	0.000	0.000	0.000
SN	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.:								
N	1.000	0.148	0.000	1.000	0.000	0.000	0.000	0.000
SN	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

4.4 A few puzzles

During the analysis in Sections 3 and 4, we have encountered several puzzling results. They were: (i) the decreasing dispersion with scale of firm efficiency τ in Figure 3; (ii) the decreasing dispersion with scale of the innovations to the stochastic variables and of the growth in capital and the growth in value documented in Figure 6; (iii) the difference-of-log-Normals distributions of $\tau, d\lambda, d\tau$, all of which are predicted to be Normal by the model, as documented in Table 4 and Table 9.

While I leave a full consideration of these puzzles for future work, it is worth noting that all of them 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-model 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. Larger firms will tend to have more sub-units, leading to lower growth variance. The same assumption is also sufficient to yield difference-of-log-Normals growth in the aggregate stochastic variables and outcome 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.

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 production function which generalizes the log-linear (or Cobb-Douglas). Together with the CLT-implied fact that AR(1) processes have Normal ergodic distributions, the production function predicts a little-known distribution for firm income and consequently firm growth, the difference-of-log-Normals 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 difference-of-log-Normals 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?" Namely — operational leverage and the interaction of sales and expenses in the firm.

The theoretical analysis yields two new magnitudes for characterizing firms — firm income scale and efficiency, both defined in terms of 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 low-base effect in firm efficiency and that for most firms, firm efficiency is indeed remarkably close to zero, yielding rampant low-base effects. Separately accounting for sales and expenses further enables new and coherent definitions of income growth and returns to scale, among others. Another contribution of the analysis is the treatment of growth from negative values and the development of measures of growth robust to such values.

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 extreme winners and losers — 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.

The production function presented 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 establishes this is far from the case for firms and that the deviations from the assumption have important implications.

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 production function in DSGE models allows for a micro-foundation of heavy-tailed income growth in a succinct and tractable manner. This follows e.g. the work of Guvenen, Ozkan, and Song (2014); Guvenen, Karahan, Ozkan, and Song (2021). Thus, heavy-tailed growth can be embedded in "upstream" economic models as well.

References

Abel, A. B., Eberly, J. C., 1994. A Unified Model of Investment Under Uncertainty. The American Economic Review 84, 1369–1384.

Abel, A. B., Eberly, J. C., 1996. Optimal Investment with Costly Reversibility. The Review of Economic Studies 63, 581–593.

Abel, A. B., Eberly, J. C., 1998. The mix and scale of factors with irreversibility and fixed costs of investment. Carnegie-Rochester Conference Series on Public Policy 48, 101–135.

Akaike, H., 1973. Information Theory and an Extension of the Maximum Likelihood Principle. Second International Symposium on Information Theory pp. 267–281.

Angelini, P., Generale, A., 2008. On the Evolution of Firm Size Distributions. American Economic Review 98, 426–438.

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.

Bellemare, M. F., Wichman, C. J., 2020. Elasticities and the Inverse Hyperbolic Sine Transformation. Oxford Bulletin of Economics and Statistics 82, 50–61.

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., Bazdresch, S., 2014. Labor Hiring, Investment, and Stock Return Predictability in the Cross Section. Journal of Political Economy 122, 129–177.

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., 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.

Burnham, K. P., Anderson, D. R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, New York, second ed.

Cabral, L. M. B., Mata, J., 2003. On the Evolution of the Firm Size Distribution: Facts and Theory. American Economic Review 93, 1075–1090.

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.

Carlson, M., Fisher, A., Giammarino, R., 2004. Corporate Investment and Asset Price Dynamics: Implications for the Cross-section of Returns. The Journal of Finance 59, 2577–2603.

Chen, J., Roth, J., 2024. Logs with Zeros? Some Problems and Solutions. The Quarterly Journal of Economics 139, 891–936.

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

Cooley, T. F., Quadrini, V., 2001. Financial Markets and Firm Dynamics. American Economic Review 91, 1286–1310.

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.

Desai, M., Gompers, P., Lerner, J., 2003. Institutions, Capital Constraints and Entrepreneurial Firm Dynamics: Evidence from Europe. Mimeo .

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.

Guvenen, F., Karahan, F., Ozkan, S., Song, J., 2021. What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics? Econometrica 89, 2303–2339.

Guvenen, F., Ozkan, S., Song, J., 2014. The Nature of Countercyclical Income Risk. Journal of Political Economy 122, 621–660.

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.

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

Mullahy, J., Norton, E. C., 2024. Why Transform Y? The Pitfalls of Transformed Regressions with a Mass at Zero. Oxford Bulletin of Economics and Statistics 86, 417–447.

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

Novy-Marx, R., 2011. Operating Leverage. Review of Finance 15, 103–134.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics 108, 1–28.

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.

Sagi, J. S., Seasholes, M. S., 2007. Firm-specific attributes and the cross-section of momentum$. Journal of Financial Economics .

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.

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.

Sutton, J., 1997. Gibrat's Legacy. Journal of Economic Literature 35, 40–59.

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.