Growth and Differences of Log-Normals

Robert Parham*

Abstract

Growth rates and net flows in economics are empirically heavy-tailed across settings as diverse as firms, cities, regional output, epidemics, and wages. This paper provides a simple unifying explanation and a *single* theoretically motivated distributional form that fits and organizes these disparate phenomena. The key observation is that many economic variables of interest are *net outcomes* shaped by two opposing multiplicative forces: sales and expenses, creation and destruction, inflows and outflows. When each side is itself the product of many small shocks, the multiplicative CLT implies each component is approximately log-Normal, while their net outcome follows a *Difference-of-Log-Normals* ($\mathcal{DLN}$) distribution.

I develop a CLT-based taxonomy of limiting distributions for economic data, and show how $\mathcal{DLN}$ variables admit a natural hyperbolic representation that decomposes outcomes into two interpretable separable components: *productive magnitude* and *productive efficacy*. I then test the distributional predictions in a large panel of U.S. public firms (1970–2019). Firm *magnitudes* are well described by Skew-Normals, with Normal upper tails. In contrast, firm cashflows, payouts, investment, key ratios, growth rates, and stock returns at multiple frequencies exhibit remarkable fit to the $\mathcal{DLN}$. Finally, I embed these findings in a tractable firm model via a *difference-of-log-linears* profit production function, which makes the (magnitude,efficacy) state space operational for estimation and counterfactuals.

JEL classifications: C46, D20, E23, G30, R10
Keywords: Growth, distributions, log-Normal, firm size, city size, COVID-19, income.

Current draft: Tuesday 27th January, 2026
Current status: R&R at ReStud

*University of Virginia (robertp@virginia.edu).

1 Introduction

What is the statistical distribution of growth? This can sound like a technical distraction until one notices how often economic questions hinge on the margins of the distribution. Selection and exit are tail events. Rare booms and collapses move aggregates in granular economies. Asset pricing and risk management live in the tails by construction. Growth is the law of motion that maps shocks into selection, reallocation, and aggregates.

The empirical fact motivating this paper is familiar: growth rates and net flows are heavy-tailed. Firm cashflows, investment, and payout flows exhibit frequent large deviations. City and regional output growth has occasional very large moves. Epidemic growth does too. These tails are not a curiosity. They are a central feature of the data-generating process.

Figure 1 provides the paper's punchline up front. Across settings as diverse as firm growth, regional GDP growth, COVID growth, and wage growth, the same distributional form fits strikingly well: the *Difference-of-Log-Normals* ($\mathcal{DLN}$). The goal of the paper is to explain why this distribution arises from first principles, to show that it organizes a wide range of economic objects, and to provide a tractable modeling and measurement framework that makes it operational.

The paper's starting point is deliberately mundane. Many variables we care about are *net outcomes*: profit is sales minus expenses; net investment is gross investment minus disinvestment; growth is often "creation minus destruction" in disguise. Combine that accounting fact with a second observation emphasized early by Gibrat (1931) and Roy (1951): proportional growth is multiplicative. When a quantity grows by $x\%$, it is multiplied by $(1+x)$, and many periods of such shocks push levels toward log-Normality by the multiplicative CLT. If both sides of a net object — benefits and costs — are themselves the product of many small multiplicative forces, then each side is approximately log-Normal, but the object of interest is their *difference*. That yields a sharp prediction: net outcomes, their induced intensities, and the growth rates built from them should be distributed as a *Difference-of-Log-Normals*.

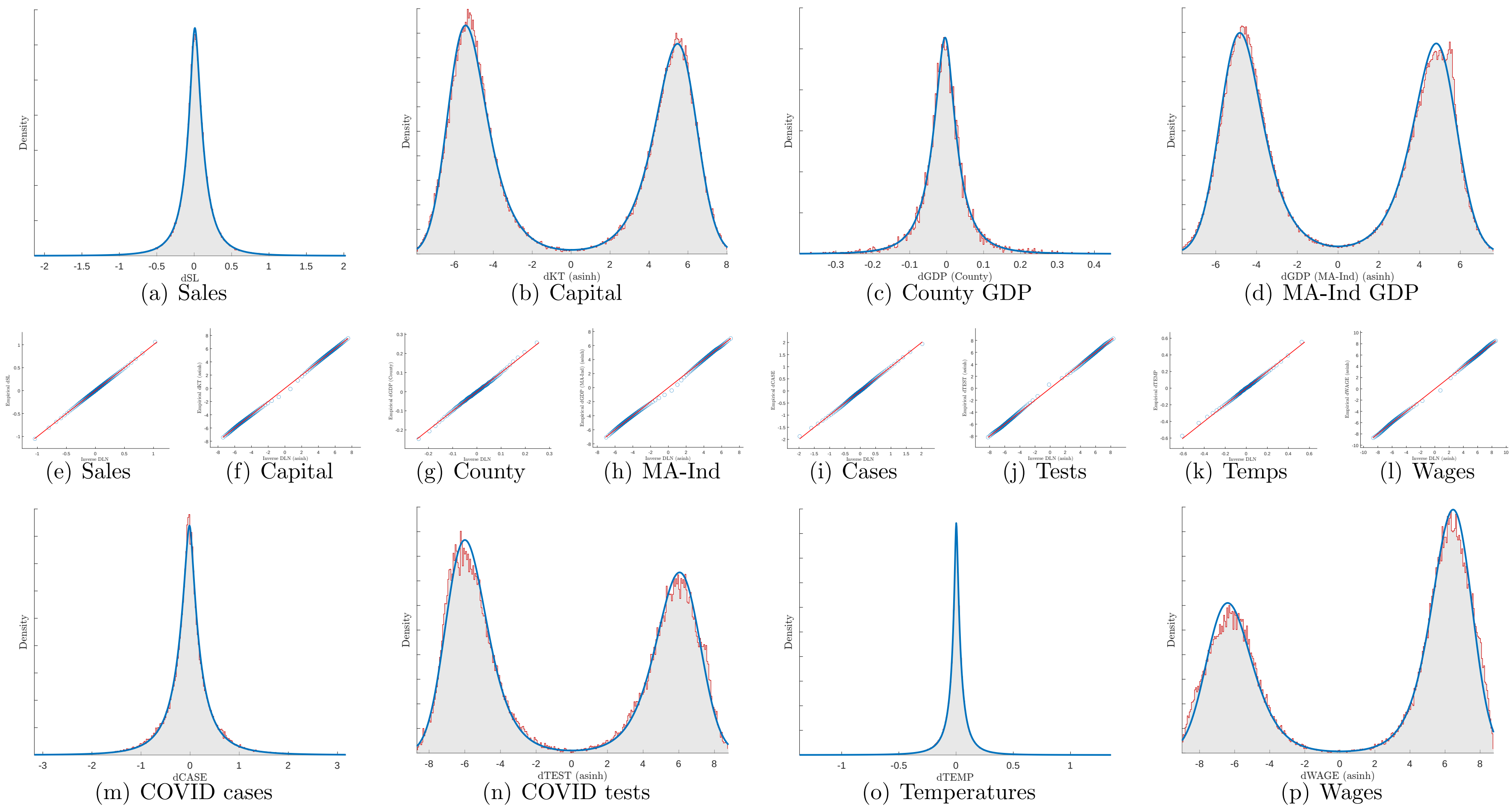


Fig. 1. Growth Distributions. This figure presents histograms of the growth rates of the phenomena listed (grey bars with red boundaries) along with a fitted Difference-of-Log-Normals distribution (blue solid line). Also presented are quantile-quantile plots comparing the empirical and theoretical distributions for each fit. Panels: (a),(b) are firm sales and capital growth (1970-2019, Compustat data); (c),(d) are GDP growth by county, and GDP growth by metropolitan area and industry (2017-2020 and 2001-2017, respectively, U.S. Bureau for Economic Analysis data); (m),(n) are daily COVID case and tests growth (3/2020-1/2022, Our World In Data data); (o) is daily temperature growth (1995-2020, Weather Project at the University of Dayton data); (p) is U.S. wage growth (2015-2020, Global Repository of Income Dynamics data). Panels (e)-(l) are the respective q-q plots. Odd panels present the data in log-points and linear x-axes, and even panels present the data in log-basepoints (1/1000 of a log-point) and asinh-scaled x-axes. See Section 3.3 for a discussion of the two canonical forms of the $\mathcal{DLN}$.

This perspective complements, and in some places reframes, a large literature on size, growth, and tails. Work on firm and city size has often emphasized Pareto and related power laws (e.g. Axtell, 2001; Gabaix, 2009), while dynamic models of firm growth and reallocation have delivered rich implications under approximately Gaussian growth (e.g. Klette and Kortum, 2004; Luttmer, 2007). More recent empirical work, including Arata (2019), has documented robust non-Normality in growth and returns. The contribution here is not another flexible heavy-tailed family, but a specific and economically interpretable distributional object implied by two banal facts: net accounting identities and multiplicative growth. In particular, the paper treats heavy tails as a *structural* implication of opposing multiplicative forces, rather than as evidence of "special" rare events.

A quick operational-leverage example makes the arithmetic hard to unsee. Suppose a firm has \$100 in sales and \$90 in expenses, so income is \$10. If both sales and expenses rise by 10%, income rises by 10%: \$110–\$99=\$11. But if sales rise by 10% while expenses fall by 10%, income becomes \$110–\$81=\$29: a 190% increase. Modest multiplicative movements in two positive components can generate very large movements in their difference. The distributional object implied by this arithmetic is not Normal, and it is not log-Normal either — it is a difference of two log-Normals.

Section 2 formally develops the statistical logic. I present a CLT-based taxonomy of limiting distributions that repeatedly arise in economic contexts: additive aggregation yields the Normal; mild constraints and selection yield Skew-Normal limits; multiplicative aggregation yields the log-Normal; its constrained analog yields the Log-Skew-Normal; and *opposing* multiplicative forces yield the Difference-of-Log-Normals ($\mathcal{DLN}$). The $\mathcal{DLN}$ has a useful and underappreciated structure. The identity

$$X - Y = \exp(x) - \exp(y) = 2\exp\left(\frac{x+y}{2}\right)\sinh\left(\frac{x-y}{2}\right)$$

is the hyperbolic analogue of moving from Cartesian to polar coordinates. It suggests a natural reparametrization into two distinct objects,

$$\lambda = \frac{x+y}{2} = \log\sqrt{XY}, \qquad \tau = \frac{x-y}{2} = \log\sqrt{X/Y},$$

which I interpret as *productive magnitude* λ and *productive efficacy* τ. Economically, λ captures scale of operations while τ captures the surplus wedge. The same identity provides a first-principles motivation for the asinh(·) transform: $\mathcal{DLN}$ objects are "generated" by the hyperbolic sine sinh(·), and its inverse asinh(·) is the natural map back to a log-like scale that remains well-defined at zero and for negative values. This is not an ad-hoc econometric trick; it is the native transformation of the object.

Section 3 takes the prediction to the data. The empirical focus is U.S. public firms over 1970–2019 (Compustat/CRSP, with higher-frequency CRSP for returns), where the accounting structure is clean and the measurement is rich. The core finding is that the $\mathcal{DLN}$ is not merely "a heavy-tailed candidate": it is the distributional form that repeatedly fits the objects economists actually use. Firm cashflows, payouts, and investment are well fit by $\mathcal{DLN}$, as are economically central intensities such as margins, yields, and investment rates. Growth rates and returns then follow naturally: once net flows and the intensities built from them have $\mathcal{DLN}$ structure, the corresponding growth and return objects inherit this structure across horizons. In parallel, the paper documents that firm *magnitudes* (log sizes) are well described by Skew-Normals, with approximately Normal behavior in the upper tail — a fact that is useful mainly because it clarifies what power-law claims imply *in logs* and where they fail.

Why should an economist care about getting the distribution right? Because distributional misspecification is not innocuous. As emphasized by Jaimovich, Terry, and Vincent (2023), the law of motion for growth is a first-order object: it governs selection and exit, reallocation, and the sensitivity of real decisions to wedges. A Normal approximation puts probability in the middle and muffles the margins; a $\mathcal{DLN}$ puts most mass on small moves with occasional very large moves while retaining finite moments. That shifts value toward exit margins and raises the elasticity of survival, hiring, and investment to small wedges. In models of industry dynamics, this changes inaction regions, trigger policies, and the implied response of entry and exit to taxes and subsidies. In urban and regional contexts, it changes how often large local shocks occur and how much aggregates load on local tail events. In public finance, it affects the mapping from micro tails to aggregate tax bases and the expected incidence of extreme outcomes. The point is not that tails are dramatic; it is that they are *structural*, and structural tails move structural objects.

This paper also squarely rejects the "black swan" agenda as an organizing principle for the data. Heavy tails do not imply ill-defined risk. The $\mathcal{DLN}$ generates high kurtosis and frequent large moves, but it retains finite moments of all orders. In the contexts studied here — including equity returns across horizons — the evidence supports heavy-tailed, finite-moment laws rather than infinite-variance rhetoric. The right lesson from the data is not that variance is meaningless; it is that the Normal is a poor approximation.

Section 4 embeds the distributional facts in a tractable economic framework and clarifies what kind of modeling is required. I propose a *difference-of-log-linears* profit production function: both demand (revenue) and cost are log-linear in quantity and state variables with Normal shocks, so each side is log-Normal in the limit and profit is their difference. The hyperbolic (λ, τ) representation becomes a natural state space for the firm. Besides its interpretability, this representation is practically useful: it turns two highly correlated processes (benefit and cost sides) into nearly uncorrelated magnitude and efficacy processes, which simplifies simulation, estimation, and counterfactual analysis. In particular, it makes explicit that matching the data is inherently a *two-factor* exercise. One-factor productivity models — from classic selection and industry-dynamics environments (e.g. Jovanovic, 1982; Hopenhayn, 1992) to more recent workhorse frameworks in the tradition of Klette and Kortum (2004) and Luttmer (2007) — have many virtues, but they generally deliver approximately Gaussian growth and they mechanically conflate "big firms" with "good firms."

In the data, magnitude and efficacy are close to orthogonal: high-efficacy firms exist at all scales, and large firms are generally *not* the highest-efficacy. In that sense, τ is the natural empirical and theoretical object for "how good is the firm," separable from size and therefore directly usable in models of selection, reallocation, and firm policies. This distinction matters for how we interpret heterogeneity, how we design empirical tests, and what we mean by "reallocation to more productive firms."

The same structure also strengthens measurement. Net outcomes can be negative, and standard growth measures break at zero. The $\mathcal{DLN}$ framework, together with its hyperbolic representation, motivates growth measures that remain coherent to and from negative values. A companion paper Parham (2023) develops the full dynamic model and estimation; here the goal is to show how a standard economic environment can generate the $\mathcal{DLN}$ law of motion and how the (λ, τ) state space makes it usable.

Two remarks about scope. First, while Section 3 concentrates on firms, Figure 1 illustrates that the same $\mathcal{DLN}$ logic appears far beyond firms — in regional output, pandemics, wages, and temperatures — suggesting a broader "opposing multiplicative forces" principle that may be empirically relevant across domains. Second, the paper intentionally takes a distribution-first approach: it pins down the relevant limiting objects, shows that they organize the data, and builds the minimal toolkit needed to embed them in economic models, estimation, and counterfactuals.

The rest of the paper is organized as follows. Section 2 presents the CLT-based taxonomy and the $\mathcal{DLN}$'s hyperbolic representation. Section 3 tests the resulting predictions in firm data, covering magnitudes, flows and intensities, growth, and returns. Section 4 proposes a general firm model with a difference-of-log-linears profit production function and discusses estimation and measurement. Section 5 concludes.

2 Distributions spanned by the CLT

2.1 The Normal — $\mathcal{N}$

What is so "normal" about *The Normal* distribution? The Normal, first described by de Moivre in 1733,[1] emerges from the Central Limit Theorem (CLT), one of the most fundamental results in probability theory. The CLT shows that *sums* of many random variables (RVs) will tend to distribute Normally, even if the variables themselves are not Normally distributed. Put differently, it states that a phenomenon in nature which is an additive combination of many latent random forces will tend to distribute Normally. Formally,

$$Y = \lim_{T\to\infty} \sum_{t=1}^{T} \varepsilon_t \sim \mathcal{N} \tag{1}$$

for $\varepsilon_t \sim \Omega$ under mild regularity conditions on Ω, the distribution of the noise terms, depending on the version of the CLT used.[2]

In economic context, this yields the well-known result that the ergodic distribution of a stationary AR(1) process is approximately Normal. A canonical form of that process is

$$Y_{i,t} = \rho \cdot Y_{i,t-1} + \varepsilon_{i,t}$$

with $|\rho| < 1$ and $\varepsilon_{i,t}$ i.i.d following some distribution Ω with finite variance and kurtosis. The distribution of Y_t (over all observations i) at time $t >> 1$ will then tend to be Normal, and the approximation will be better as $\rho \to 0$ or will be exact if Ω is itself the Normal distribution.

Two important facts about the Normal are worth noting. First, it is closed under addition and subtraction. That is, if N_1 and N_2 are two (possibly correlated) Normal random variables, then so are $N_1 + N_2$ and $N_1 - N_2$. Second, the Normal is the maximum entropy distribution for a specified mean and variance. In less technical terms, it means that assuming some data are Normally distributed is the simplest or least restrictive assumption in the

[1] And later popularized by Gauss, earning the name "Gaussian."

[2] E.g., Ω must have finite variance and negligibility via Lindeberg's condition for the Lindeberg–Feller CLT to hold.

"Occam's razor" sense.

2.2 The Skew-Normal — $\mathcal{SN}$

By way of a slight detour, it is worth presenting a little-known generalization of the Normal distribution, the Skew-Normal distribution, first described by Azzalini (1985). It is a three-parameter distribution $\mathcal{SN}(\mu, \sigma^2, \alpha)$ with $\alpha \in \mathbb{R}$ a skewness parameter. The $\mathcal{SN}$ collapses to the Normal distribution $\mathcal{N}(\mu, \sigma^2)$ when $\alpha = 0$. It is right-skewed (positive skewness) for positive values of α and left-skewed (negative skewness) for negative values. The Probability Density Function (PDF) for a Skew-Normal with $\mu = 0, \sigma = 1$ is given by

$$f(x) = 2 \cdot \phi(x) \cdot \Phi(\alpha \cdot x)$$

with ϕ and Φ the Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of the Normal, respectively.

The economic interest in the Skew-Normal stems from stochastic processes with frictions that depend on the current level of the process. This is true for many, if not most, real-world applications. Examples include processes subject to (size-dependent) selection, processes with boundary behavior, and heteroscedastic processes.[3] Consider for example the following process

$$Y_{i,t} = \rho \cdot Y_{i,t-1} + \varepsilon_{i,t} \qquad ; \qquad \varepsilon_{i,t} \sim \mathcal{N}\left(0, \frac{\sigma^2}{\exp(Y_{i,t-1})}\right) \tag{2}$$

For which the variance of the error term $\varepsilon_{i,t}$ (inversely) depends on the level of the process $Y_{i,t-1}$. A notable example of such a process is the growth of firms, for which Yeh (2023) documents an inverse relation between firm (log) size and growth variance (larger firms have lower growth variance). The distribution of (log) size will then tend to be Skew-Normal.

[3] See e.g. Anděl, Netuka, and Zvára (1984).

2.3 The Log-Normal — $\mathcal{LN}$

The Log-Normal distribution, first described by Galton in 1879, is the distribution that arises when one exponentiates a Normal random variable (RV). I.e., if $Y = \exp(N_1)$ where $N_1 \sim \mathcal{N}(\mu, \sigma^2)$, then $Y \sim \mathcal{LN}(\mu, \sigma^2)$. Conversely, if $Y \sim \mathcal{LN}$ then $\log(Y) \sim \mathcal{N}$. The Log-Normal was popularized by Gibrat (1931),[4] who described its emergence when considering positively constrained quantities that arise from multiplicative processes, such as wages, company sales, and city populations. Gibrat used a simple argument, later known as the *multiplicative* Central Limit Theorem, to show that a phenomenon in nature which is a *product* of many latent random forces will tend to distribute Log-Normally. Formally,

$$Y = \lim_{T \to \infty} \left(\prod_{t=1}^{T} \varepsilon_t \right) \sim \mathcal{LN}$$

for positive $\varepsilon_t \sim \Omega$ with mild regularity conditions as discussed above. Clearly, when one takes log of the multiplicative CLT, one backs out the additive CLT of (1).

Gibrat's observation is of note because growth itself is a multiplicative process — this is why we discuss it in percentage units.[5] When some quantity (a firm, a city, a wage, a price) grows by some percentage $x\%$, then the new value is the old value multiplied by $(1 + x)$. An early discussion of the implications and reasoning for assuming multiplicative impact in such quantities is the seminal work of Roy (1950, 1951). Roy empirically shows the distribution of employee earnings is approximately Log-Normal, and discusses why this must arise when one assumes productivity changes are multiplicative.

The canonical AR(1) process in logs, often used as the driving productivity process in Dynamic Stochastic General Equilibrium (DSGE) models, can be written as

$$y_{i,t} = \rho \cdot y_{i,t-1} + \varepsilon_{i,t} \quad ; \quad Y_{i,t} = \exp(y_{i,t}) = Y_{i,t-1}^{\rho} \cdot \exp(\varepsilon_{i,t})$$

with lower case y denoting $\log(Y)$ as usual. The productivity process $Y_{i,t}$ will then tend to distribute Log-Normally, because the log-process $y_{i,t}$ will tend to Normality, as discussed

[4] Earning it the name "Gibrat's distribution" for some time.

[5] "Sales increased by 12%"; "I got a 5% raise"; "Prices increased by 3%"; "The population grew by 4%".

above.

While the Normal and Skew-Normal are bi-directional distributions (i.e., are supported on all of $\mathbb{R}$ and take both positive and negative values), the Log-Normal is a uni-directional (always positive) distribution. Bi-directional distributions have both right- and left-tails, while uni-directional distributions only have a right-tail. To that end, I define a bi-directional distribution to be "heavy-tailed" if its tails are heavier than those of the Normal (i.e., high kurtosis) and a uni-directional distribution to be "heavy-tailed" if its tail is as heavy (or heavier) as that of the Exponential distribution.

Two important facts about the Log-Normal, mirroring the facts discussed above for the Normal, are worth noting. First, it is closed under multiplication and division. I.e., if L_1 and L_2 are two Log-Normal random variables, then so are $L_1 \cdot L_2$ and L_1/L_2. Second, the Log-Normal is the maximum entropy uni-directional distribution for a specified mean and variance. In less technical terms, it means that assuming always-positive data are Log-Normally distributed is the simplest or least-restrictive assumption. Such data are always positive, span several orders of magnitude,[6] and grow (or shrink) by percentage points — i.e. multiplicatively.

The Log-Normal is easy to work with, empirically, because it merely means one should consider not the size of phenomena (e.g., firm size, city population, etc.) but the *magnitude* of phenomena (defined hereafter to prevent confusion as the natural log of size). The line of reasoning leading to the emergence of the Log-Normal simply means one should expect magnitudes to distribute Normally, or Skew-Normally if any magnitude-dependent frictions exist, leading us to the next CLT-implied distribution.

2.4 The Log-Skew-Normal — $\mathcal{LSN}$

A natural and useful extension of the Log-Normal is the Log-Skew-Normal. As the name implies, the $\mathcal{LSN}$ is the uni-directional distribution that arises when one exponentiates a

[6]E.g., a \$2 pen vs. a \$20M jet; a factory producing 100 widgets vs. one producing 100M; a 10-person town vs. a 10M-person metropolis.

Skew-Normal random variable. I.e., if $Y = \exp(S_1)$ where $S_1 \sim \mathcal{SN}(\mu, \sigma^2, \alpha)$, then $Y \sim \mathcal{LSN}(\mu, \sigma^2, \alpha)$. Conversely, if $Y \sim \mathcal{LSN}$ then $\log(Y) \sim \mathcal{SN}$.

The economic interest in the Log-Skew-Normal, as the discussion above suggests, is due to it being the target distribution of phenomena sizes, when the underlying magnitude process is additive with frictions. The process

$$Y_{i,t} = \exp(y_{i,t}) \quad ; \quad y_{i,t} = \rho \cdot y_{i,t-1} + \varepsilon_{i,t} \quad ; \quad \varepsilon_{i,t} \sim \mathcal{N}(0, \frac{\sigma^2}{Y_{i,t-1}})$$

which merely exponentiates the $\mathcal{SN}$ process in (2), is an example that gives rise to $Y_{i,t} \sim \mathcal{LSN}$.

Testing whether a given data distribution is Log-Skew-Normal is as simple as taking logs and then testing whether the resulting magnitudes distribute Skew-Normal, which is a standard test in most common statistical packages. Parameter estimates given empirical data distributing $\mathcal{LSN}$ are similarly standardized.

2.5 The Difference-of-Log-Normals — $\mathcal{DLN}$

Finally, we reach the namesake distribution of this work, the Difference-of-Log-Normals distribution. As the name again implies, the distribution arises when one exponentiates two (possibly correlated) Normal random variables, and then takes the difference between these Log-Normal values.

Unlike the *sum* of Log-Normal RVs, which is approximately Log-Normal and has been used in several disciplines including telecommunication, actuary, insurance, and derivative valuation, the Difference-of-Log-Normals distribution is almost completely unexplored. At the time of writing, I was unable to find instances of using it anywhere in the sciences, and only two statistical works tangentially considering it: Lo (2012); Gulisashvili and Tankov (2016). Both papers concentrate on sums of Log-Normals but show their results hold for differences of Log-Normals as well, under some conditions. Nevertheless, this work shows both empirically and theoretically that the novel Difference-of-Log-Normals distribution is widespread in economic data.

Formally, define W such that

$$W = Y_p - Y_n = \exp(N_p) - \exp(N_n)$$

with (N_p, N_n) bivariate Normal, i.e. $(N_p, N_n)^T \sim \mathcal{N}(\mu, \Sigma)$ with parameters

$$\mu = \begin{bmatrix} \mu_p \\ \mu_n \end{bmatrix} \quad ; \quad \Sigma = \begin{bmatrix} \sigma_p^2 & \sigma_p \cdot \sigma_n \cdot \rho_{pn} \\ \sigma_p \cdot \sigma_n \cdot \rho_{pn} & \sigma_n^2 \end{bmatrix}$$

We say W follows the five-parameter Difference-of-Log-Normals distribution, and denote $W \sim \mathcal{DLN}(\mu_p, \sigma_p, \mu_n, \sigma_n, \rho_{pn})$.

Recall that we motivated the Normal and Log-Normal distributions based on the additive and multiplicative CLTs, concluding that phenomena in nature which are the sum or product of many latent random forces will tend to distribute Normally and Log-Normally, respectively. Consider now a natural phenomenon impacted by two main forces operating in opposite directions, i.e., $W = Y_p - Y_n$. If each of the two main forces is an additive combinations of many latent random forces, then the natural phenomenon W will tend to distribute Normally as well. In this case, the importance of modeling the two forces separately is diminished because aggregating them yields a model with similar distributional predictions.

The same is not true, however, if each force is a multiplicative combination. In this case, failing to explicitly model both forces will yield markedly different predictions, because the difference between two Log-Normal RVs does not collapse to a Log-Normal RV. For one, Log-Normal RVs are strictly positive (i.e., uni-directional), while Difference-of-Log-Normals RVs can clearly take any value on the real line $\mathbb{R}$ (i.e., bi-directional, similar to the Normal). Further, the Difference-of-Log-Normals exhibits Log-Normal tails in both the positive and negative directions, yielding a distributional shape quite different from the Normal "Gaussian bell curve."

From an economic perspective, the stochastic process giving rise to the Difference-of-Log-Normals distribution can be described as a VAR(1) process in logs

$$\begin{pmatrix} x_{i,t+1} \\ y_{i,t+1} \end{pmatrix} = R \cdot \begin{pmatrix} x_{i,t} \\ y_{i,t} \end{pmatrix} + \begin{pmatrix} \epsilon^x_{i,t} \\ \epsilon^y_{i,t} \end{pmatrix} \qquad ; \qquad \begin{pmatrix} \epsilon^x_{i,t} \\ \epsilon^y_{i,t} \end{pmatrix} \sim \mathcal{N}(0, \Sigma)$$

where R is the 2x2 autocorrelation matrix and the disturbance term is a mean-zero bivariate Normal with Σ its covariance matrix. The process

$$W_{i,t} = \exp(x_{i,t}) - \exp(y_{i,t}) = X_{i,t} - Y_{i,t}$$

will tend to distribute as the Difference-of-Log-Normals because each of $X_{i,t}, Y_{i,t}$ will tend to distribute Log-Normal as discussed above.

Two important facts about the Difference-of-Log-Normals are worth noting. First, it is closed under multiplication and division by a Log-Normal RV. I.e., if W is $\mathcal{DLN}$-distributed and L is $\mathcal{LN}$-distributed, then $W \cdot L$ and W/L will both distribute $\mathcal{DLN}$ as well — a property derived from the closure of Log-Normals under multiplication and division. Second, empirical work suggests that the $\mathcal{DLN}$ distribution subsumes the case of a difference between two Log-Skew-Normal random variables.[7] I.e., there appears to be no extra benefit from modeling a phenomenon as a Difference-of-Log-Skew-Normals over modeling it as a Difference-of-Log-Normals.

Empirical work with the Difference-of-Log-Normals is slightly more complicated than with the other distributions discussed so far, because no standard testing or estimation procedures for the distribution exist in common statistical packages. The online appendix describes how to construct these procedures, and provides a full suite of computer code implementing the PDF, CDF, estimation and testing for the distribution.

A second reason the Difference-of-Log-Normals is more complicated to deal with is the bi-directional log-tails (both positive and negative), "inherited" from the two Log-Normals composing it. The usual method of dealing with exponential data — taking logs — cannot be used because of the existence of negative values. The next section hence discusses the intimate relation between the Difference-of-Log-Normals distribution and Hyperbolic

[7] Because the definition of the $\mathcal{DLN}$ allows the two generating Normal RVs to be correlated.

Trigonometry, especially the Hyperbolic Sine function. It further demonstrates how this relation simplifies working with the distribution and proposes the Inverse Hyperbolic Sine as a natural "log-like" transform in both the positive and negative directions.

2.6 The Difference-of-Log-Normals and Hyperbolic Trigonometry

"Regular" trigonometry, also known as *circular* trigonometry, is based on the unit circle, given by the equation $X^2 + Y^2 = 1$. Hyperbolic trigonometry, in comparison, is based on the unit hyperbola, given by the equation $X^2 - Y^2 = 1$. The two are tightly related, as can be seen, e.g., by the definitions of the circular and hyperbolic Sine functions, given by:

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i} \qquad ; \qquad \sinh(x) = \frac{e^{x} - e^{-x}}{2}$$

and their inverses:

$$\mathrm{asin}(x) = \log\left(ix + \sqrt{1 - x^2}\right) \cdot i^{-1} \qquad ; \qquad \mathrm{asinh}(x) = \log\left(x + \sqrt{1 + x^2}\right)$$

such that $x = \mathrm{asin}(\sin(x)) = \mathrm{asinh}(\sinh(x))$ and $i = \sqrt{-1}$ is the imaginary unit.

Our interest in the hyperbolic functions stems from the fact that a function given by a difference of exponentials can be factored into an exponential multiplied by a hyperbolic Sine. That is,

$$W = X - Y = \exp(x) - \exp(y) = 2 \cdot \exp\left(\frac{x + y}{2}\right) \cdot \sinh\left(\frac{x - y}{2}\right) \tag{3}$$

is the hyperbolic equivalent of moving from cartesian (i.e., X,Y) coordinates to polar (i.e., r,θ) coordinates in circular trigonometry. The hyperbolic equivalents of the radius and angle are given by:

$$\lambda = \frac{x + y}{2} = \log\left(\sqrt{X \cdot Y}\right) \qquad ; \qquad \tau = \frac{x - y}{2} = \log\left(\sqrt{X/Y}\right) \tag{4}$$

in which λ is the equivalent of the radius and τ is the equivalent of the angle. The inverse mapping between (X, Y) and (λ, τ) is then given by:

$$X = \exp(\lambda + \tau) \qquad ; \qquad Y = \exp(\lambda - \tau) \tag{5}$$

with λ the mid-point between $x = \log(X)$ and $y = \log(Y)$, and τ the (equal) distance from λ to x and y, with the appropriate sign. Clearly, one can equivalently describe a $\mathcal{DLN}$ phenomenon by specifying its dynamics in the (X, Y) space or the (λ, τ) space. In what follows, we will see that the (λ, τ) representation is both more intuitive and more amenable to mathematical analysis than the (X, Y) representation.

Note that the identity in (3) provides a first-principles justification for using the Inverse Hyperbolic Sine (asinh) transform when encountering $\mathcal{DLN}$-distributed data, which is "generated" by the Hyperbolic Sine (sinh). The asinh transform has been the subject of recent econometric work, including Bellemare and Wichman (2020); Aihounton and Henningsen (2021); Mullahy and Norton (2024), mostly discussed in the context of ad-hoc transformations of non-negative values that include zeros. Here, it arises as the natural function to transform "bi-directional-Log-Normal" data, i.e. data that exhibit Log-Normal tails in both the positive and negative directions. From a practical perspective, the asinh transform has several desirable properties:

1. Differentiable and strictly increasing in x.
2. Odd function, such that $\text{asinh}(-x) = -\text{asinh}(x)$.
3. Zero based, such that $\text{asinh}(0) = 0$.
4. $\text{asinh}(x) \approx \text{sign}(x)(\log|x| + \log(2))$, with the approximation error rapidly vanishing as $|x|$ increases.

Hence, asinh is a bijection similar in flavor to the oft-used neglog transform:

$$\text{neglog}(x) = \text{sign}(x) \log(1 + |x|) \tag{6}$$

but with less distortion than the neglog around 0, at the cost of the bias term $\log(2) \approx 0.7$.

2.7 Pareto, Zipf, Power-law, and Exponential distributions

A final set of distributions important for our discussion are the Power-law distributions and their transform, the Exponential distribution. This is because pivotal existing work on size distributions has focused on Pareto, Zipf, and Power-law distributions more generally, especially around the right tail (Gabaix, 1999a,b; Axtell, 2001; Luttmer, 2007; Gabaix, 2011). A Power-law distribution has the general form $f(x) \propto x^{-(\alpha+1)}$, with Pareto being its continuous archetype and Zipf its discrete analog.

Importantly, Power-law random variables (RVs) are transformed into Exponential RVs when taken in logs. That is, we have

$$X \sim \text{Pareto}(x_{\min}, \alpha), \quad Y = \log\left(\frac{X}{x_{\min}}\right) \quad \Rightarrow \quad Y \sim \text{Exponential}(\lambda = \alpha).$$

which in turn means that testing for size being Power-law distributed is equivalent to testing whether magnitude distributes Exponential.[8] Put differently, Power-law distributions can be referred to as *Log-Exponential*, in a similar way to the Log-Normal.

3 Empirical tests of distributional predictions

Equipped with a CLT-based theoretical framework proposing specific distributional forms and with mathematical tools to aid in the analysis, we next move on to testing these theoretical predictions in the data. In what follows, I concentrate on firm data and firm dynamics, but as Figure 1 exhibited, the analysis can be generalized to other phenomena, including cities, pandemics, wages, etc.

3.1 Firm Data

The firm data analyzed below cover all public U.S. firms in the 50-year period 1970-2019, with minimal filtering, resulting in 192K firm-year observations on 20K distinct firms. Data

[8]This is why Pareto/Zipf RVs are usually tested as a straight line on a log-log histogram: it is in effect a visual test of whether the log variable distributes Exponential.

are predominantly derived from the yearly CRSP/Compustat data set. For some tests related to equity returns I use higher-frequency CRSP data. Other minor data sources include the nominal and real GDP series from FRED and factor returns from Ken French's website.

Data variables are identified throughout by two capital letter mnemonics (e.g., SL for firm sales). Table 1 defines all data variables and provides a mapping to Compustat items used to construct them. I rely on the sources and uses identity,

$$\underbrace{\text{sales}}_{SL} - \underbrace{\text{expenses}}_{XS} = \underbrace{\text{income}}_{CF} = \underbrace{\text{total net disbursement}}_{DI} + \underbrace{\text{total net investment}}_{IT} \tag{7}$$

to calculate expenses as dissipated sales (i.e., sales - income). This guarantees all expenses, including labor, cost of goods, selling, general, administrative, taxes, and various "special" and "one-time" expenses are accounted for.[9]

I consider two approaches to adjusting historical dollar values to 2019 dollars: (i) using the (real) GDP-deflator to adjust to real 2019 dollars; and (ii) using nominal GDP as the deflator. The first approach only adjusts for inflation, while the second adjusts for both inflation and the economy-wide secular growth trend, thus yielding a stationary firm size distribution, an observation we will return to later. All results are reported using the nominal-GDP deflated data, i.e., in 2019 dollars with a 2019-sized economy, but I verify they all hold when using the real GDP deflator instead.[10]

Finally, note that a limitation of the analysis is the concentration on public firms, driven by data availability and quality. The selection into being public is not random, and many smaller firms, as well as some larger firms, are private and absent in the data. Nevertheless, the largest (and most economically consequential) firms are generally public. Hence, the results below discussing the largest firms, i.e., the much-debated right-tail of the firm size distribution, are unlikely to be overturned by the inclusion of private firm data.

[9]Using a traditional "top-down" definition of expenses does not materially change any of the results.

[10]Employee counts are similarly normalized by dividing each count by the total employee count for the year and then multiplying by the 2019 total employee count.

Table 1
Data definitions

This table defines all data items used. Data cover public U.S. firms in the 50-year period 1970-2019, as described in Section 3.1. 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 disbursement + total investment. 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 disbursement	DE	dvt + (prstkc - sstk)	dividends + net repurchase
Debt disbursement	DD	xint + (L.DB-DB)	interest paid + decrease in debt
Total disbursement	DI	DE + DD	to equity and debt
Physical capital	KP	ppent	PP&E, net of depreciation
Total capital	KT	at	total assets (tangible)
Depreciation	DP	dp	of physical capital
Physical investment	IP	KP - L.KP + DP	growth in physical 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
Employees	EM	emp	number of employees
Productive magnitude[1]	LM	$\sqrt{\mathrm{SL} \cdot \mathrm{XS}}$	$= \exp(\lambda)$
Productive efficacy[1]	TU	$\sqrt{\mathrm{SL}/\mathrm{XS}}$	$= \exp(\tau)$

[1] In exponentiated terms, for compatibility with the other values.

3.2 Firm size and magnitude

Which economic measure best captures firm size (and its log, firm magnitude)? The firm growth literature tends to use employees (EM) or sales (SL) as the preferred measures of size, the asset pricing literature tends to use market value of equity (EQ), and the corporate finance literature tends to use total capital stock (KT). To these, I add firm expenses (XS). While using expenses to measure firm size is not common, it is in-line with the common approach of using number of employees as a measure of firm size. Using expenses has the benefits of being widely available for all firms and being more holistic, i.e. neutralizing the "build vs. buy" decision of firms. Finally, I also add the firm productive magnitude measure, λ, defined per (4) to be the (log) geometric mean of firm sales and expenses. All analysis in this section is done using firm magnitudes (i.e., the natural logs of the relevant size measures).

Panel (a) of Table 2 reports the first four central moments of each magnitude measure, as well as their median and inter-quartile range (IQR). All dollar magnitudes have mean and median around 6.5 ($\approx$665M 2019$), and employee magnitude has mean and median around 7.5 ($\approx$2000 employees). All magnitudes have a standard deviation around 2.1, modest positive skewness around 0.2 and kurtosis very close to 3, the kurtosis of the Normal distribution, i.e., magnitudes exhibit no heavy tails (and if anything, present slightly lighter tails than the Normal, as kurtosis is lower than 3 for all of them).

That all magnitude measures have such close moments is not surprising, as Panels (b) and (c) of Table 2 show. In panel (b), we see that all magnitude measures are highly correlated, as expected, and in Panel (c) we see they are in fact all co-integrated. Because the cointegration tests of Pedroni (2004) and Westerlund (2005) require balanced panels, Panel (c) presents the tests by decade and includes all balanced panels available within the decade.[11] The hypothesis of no cointegration is strongly rejected by all tests for all decades. Intuitively,

[11] I exclude the productive magnitude measure λ as it is cointegrated with SL and XS by construction and will skew the test results in favor of finding cointegration.

we can say that all magnitude measures carry the same signal, with some measure-specific noise.

Visual evidence on the distribution of firm magnitudes is presented in Figure 2. Panels (a) and (b) present the distributions of firm market value of equity EQ and total assets KT, respectively, in the data. The figures are overlaid with fitted Normal, Skew-Normal, and Exponential distributions.[12] The fit to Skew-Normal is evident, as is the mismatch with the Exponential. This visual test between the Normal and Skew-Normal is confirmed by Panel (c), which presents quantile-quantile (q-q) plots for firm assets KT. The q-q plots compare the empirical vs. a specific theoretical distribution. If the two match, all quantiles will reside on the y=x line. The Normal presents tail-deviations, while the Skew-Normal presents excellent fit, even at the uppermost right tail of largest firms.

The above observation is however "unfair" to the Exponential distribution, as the Power-Law hypothesis specifically limits itself to the right tail of the firm size distribution. To that end, Panels (d)-(f) concentrate on the right tails of three other firm magnitude distributions: employees EM, productive magnitude LM, and firm sales SL, respectively. The figures present the upper decile (top 10%) of the respective magnitude distributions, overlaid with truncated versions of the Normal, Skew-Normal, and Exponential. The poor fit of the Exponential, especially for the largest of firms at the right tail, is evident in the corresponding q-q plots in Panels (g)-(i). That the upper-tail deviations are below the line indicates the Power-Law hypothesis inflates the expected number of very large firms relative to the data. Note that the fit lines for the Normal and Skew-Normal are indistinguishable from each other, proposing skewness is unnecessary to describe the right tail of magnitude distributions, and it is well described by a Normal tail.

To avoid relying on visual inference alone (and especially "ocular tests of straight lines," which Clauset, Shalizi, and Newman (2009) show are often misleading), formal statistical tests against the Normal, Skew-Normal, and Exponential are presented in Table 3. Following

[12]Fitted via Maximum Likelihood Estimation (MLE).

Clauset et al. (2009), I use three different distributional goodness-of-fit tests: Kolmogorov-Smirnov (K-S), Chi-square (C-2), and Anderson-Darling (A-D). 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 — I hence report results for all three tests. To account for the different number of parameters in each distribution, I also include the Akaike and Bayesian Information Criteria-based relative likelihood tests (AIC and BIC). The first six columns present tests for the full data, and the last six present tests for the upper decile tail.

At the left side of Panel (a), we can see that Normality is generally rejected, at the 5% confidence level, for the full data. At the right side, we can see that it is generally *not* rejected for the upper tail. The left side of Panel (b) indicates Skew-Normality is generally not rejected for the full data, and is similarly not rejected for the upper tail. In Panel (c), the Exponential distribution is overwhelmingly rejected for the full data, as expected, and is similarly strongly rejected for the upper tail. These results hold when considering instead the top 20%, 5%, or 1% of firms.

The relative likelihood tests in Panel (d) favor the Skew-Normal for the entire distribution, but imply that the Normal suffices for the tail. That is, Normality explains the tail sufficiently well, and the skewness parameter does not "carry its weight". We can hence conclude that the *size* distribution of firms in the data is *rejected* as being Power-law (Pareto, etc.) or Log-Normal, but is *not rejected* as being Log-Skew-Normal, either for the entire distribution or for the right tail, though for the right tail it is redundant and Log-Normality suffices. Even for the entire magnitude distribution, the estimated skewness parameters are modest, as are the deviations from Normality, and Log-Normality is a very good approximation for firm size across the entire data.

Economic implications. The prevailing view, following the Simon–Ijiri tradition and crystallized by Axtell (2001); Luttmer (2007); Gabaix (2009), is that the upper tail of the

Table 2
Magnitude - Descriptive statistics

Panel (a) of this table presents the first four central moments of firm magnitude, based on six alternative measures, as well as their Median and IQR. Variable definitions are in Table 1. Panel (b) presents the correlations between the measures. Panel (c) presents the results of three cointegration tests between all magnitude panels, 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 vs. the alternative that some panels are cointegrated. Tests are conducted by decade, on the available balanced sample of measures within each decade.

Panel (a): Magnitude = log(XX) moments

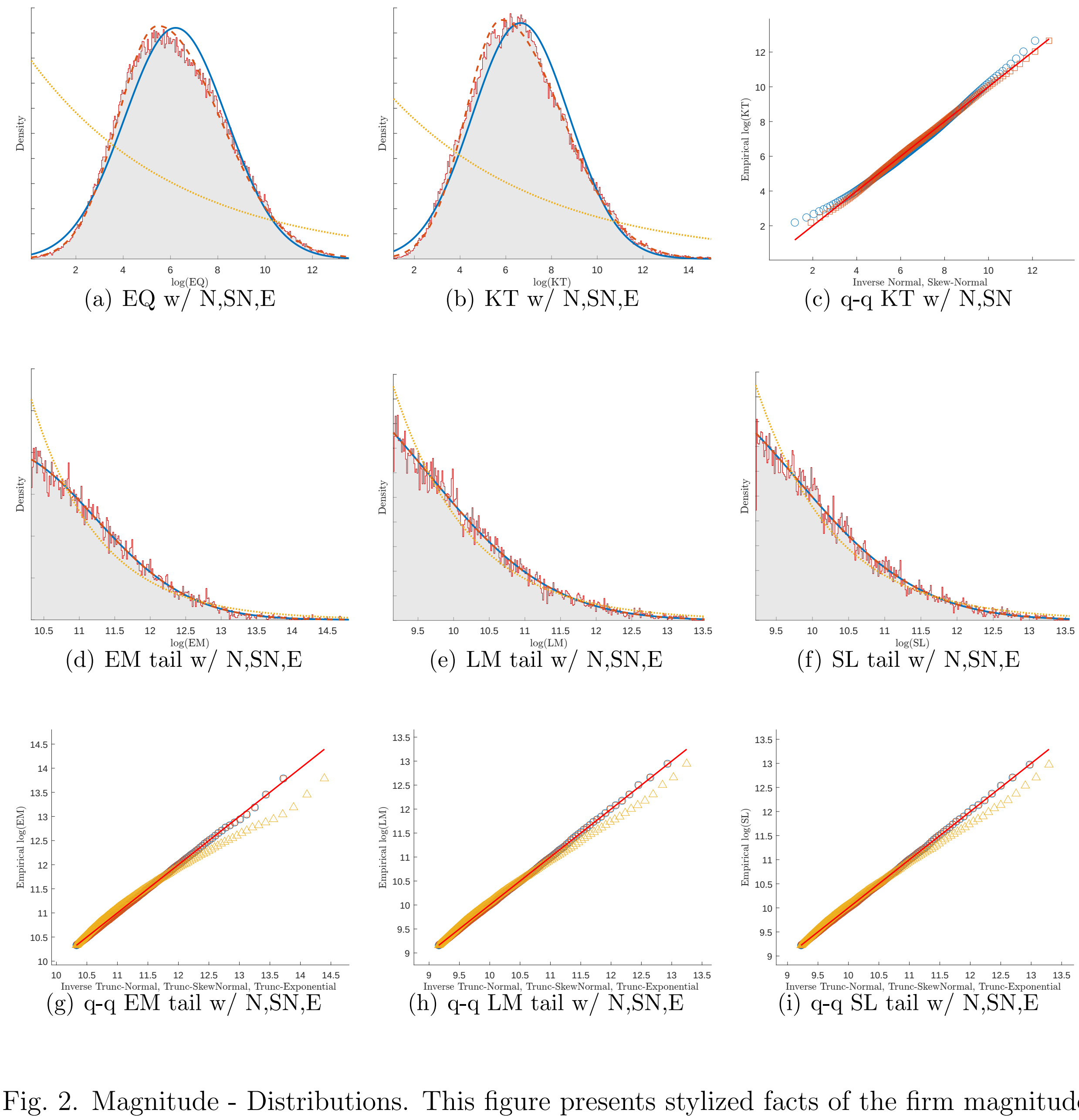
	EM	SL	EQ	KT	XS	LM
M_1 (mean)	7.59	6.41	6.21	6.65	6.34	6.37
M_2 (s.d.)	2.06	2.13	2.17	2.12	2.04	2.08
M_3 (skew)	0.06	0.07	0.26	0.37	0.22	0.17
M_4 (kurt)	2.75	2.86	2.73	2.98	2.77	2.75
Median	7.56	6.37	6.08	6.52	6.26	6.31
IQR	2.92	2.90	3.07	2.89	2.85	2.88

Panel (b): Magnitude correlations

	SL	EQ	KT	XS	LM
EM	.914	.732	.787	.916	.919
SL		.809	.885	.982	.996
EQ			.868	.800	.809
KT				.870	.881
XS					.995

Panel (c): Magnitude cointegration tests

	Phillips-Perron t	p-val	Dickey-Fuller t	p-val	Variance ratio	p-val
70's	38.97	<0.001	-47.15	<0.001	11.91	<0.001
80's	44.45	<0.001	-40.24	<0.001	18.21	<0.001
90's	48.71	<0.001	-50.40	<0.001	24.44	<0.001
00's	50.84	<0.001	-66.07	<0.001	20.40	<0.001
10's	49.73	<0.001	-42.67	<0.001	20.95	<0.001



(a) EQ w/ N,SN,E (b) KT w/ N,SN,E (c) q-q KT w/ N,SN

(d) EM tail w/ N,SN,E (e) LM tail w/ N,SN,E (f) SL tail w/ N,SN,E

(g) q-q EM tail w/ N,SN,E (h) q-q LM tail w/ N,SN,E (i) q-q SL tail w/ N,SN,E

Fig. 2. Magnitude - Distributions. This figure presents stylized facts of the firm magnitude distributions. Panels (a) and (b) present the full distributions of (log) equity value EQ and total capital KT respectively, overlaid with fitted Normal (solid blue), Skew-Normal (dashed red), and Exponential (dotted yellow) distributions. Panel (c) presents the q-q plot for the Normal and Skew-Normal fits. Panels (d),(e),(f) present the top 10% magnitude tails of employees (EM), productive magnitude (LM), and sales (SL), overlaid with truncated version of the Normal, Skew-Normal, and Exponential. Panels (g),(h),(i) present the respective q-q plots.

Table 3
Magnitude - Distributional tests

This table presents results of tests of distribution equality for firm magnitudes based on the measures described in Table 1. The first 6 columns pertain to the entire data, while the last 6 column pertain to the truncated top 10% of observations by size. 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 for the Normal, Skew-Normal, and Exponential, respectively. Panel (d) reports the AIC- and BIC-based relative likelihoods for each distribution.

	EM	SL	EQ	KT	XS	LM	EM	SL	EQ	KT	XS	LM
Panel (a): Magnitudes vs. Normal												
K-S	0.017	0.009	0.029	0.026	0.021	0.016	0.005	0.006	0.008	0.014	0.006	0.007
p-val	0.040	0.065	0.026	0.028	0.035	0.042	0.095	0.082	0.067	0.047	0.085	0.079
C-2	108.3	38.25	>999	206.8	145.9	92.84	43.31	39.14	46.18	74.62	32.50	30.18
p-val	0.037	0.060	0.000	0.026	0.032	0.039	0.056	0.059	0.055	0.044	0.064	0.066
A-D	5.728	2.200	16.37	17.98	10.68	6.698	0.328	0.455	1.751	3.457	0.567	0.414
p-val	0.041	0.056	0.027	0.026	0.032	0.039	0.101	0.091	0.060	0.049	0.085	0.094
Panel (b): Magnitudes vs. Skew-Normal												
K-S	0.015	0.006	0.012	0.009	0.010	0.010	0.005	0.006	0.008	0.014	0.006	0.007
p-val	0.045	0.090	0.053	0.063	0.058	0.061	0.095	0.082	0.067	0.047	0.085	0.079
C-2	83.72	20.17	27.50	38.02	41.96	33.72	43.31	39.14	46.18	74.62	32.50	30.18
p-val	0.041	0.081	0.069	0.060	0.057	0.063	0.056	0.059	0.055	0.044	0.064	0.066
A-D	4.057	0.615	2.594	1.629	2.239	1.873	0.328	0.455	1.751	3.457	0.567	0.414
p-val	0.046	0.083	0.053	0.062	0.056	0.059	0.101	0.091	0.060	0.049	0.085	0.094
Panel (c): Magnitudes vs. Exponential												
K-S	0.391	0.341	0.332	0.353	0.351	0.344	0.062	0.048	0.035	0.035	0.046	0.047
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.014	0.021	0.021	0.015	0.014
C-2	>999	>999	>999	>999	>999	>999	501.6	260.3	216.2	149.7	244.5	243.3
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.023	0.026	0.031	0.024	0.024
A-D	>999	>999	>999	>999	>999	>999	82.89	44.66	29.98	23.28	41.09	41.15
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.016	0.020	0.023	0.017	0.017
Panel (d): Distribution comparison												
AIC R.L.:												
Normal	0.004	0.004	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000
Skew-N	1.000	1.000	1.000	1.000	1.000	1.000	0.368	0.368	0.368	0.368	0.368	0.368
Exp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BIC R.L.:												
Normal	0.145	0.143	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000
Skew-N	1.000	1.000	1.000	1.000	1.000	1.000	0.010	0.010	0.010	0.010	0.010	0.010
Exp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

firm size distribution is well described by a Pareto law. Later work (e.g., Kondo, Lewis, and Stella (2018)) revisits this with richer microdata and formal tail tests, finding the Log-Normal outperforms Pareto, consistent with the results above. Analogously, for income data, Azzalini, Cappello, and Kotz (2002) show that the Log-Skew-Normal dominates Pareto both overall and in the right tail (reviving the findings of Roy (1950)).

The tail shape matters for at least three active debates: (i) the *granular hypothesis* (Gabaix, 2011), where the tail index governs how much idiosyncratic shocks to very large firms survive aggregation; (ii) *optimal top marginal tax rates* for labor income, where sufficient-statistic formulas use the Pareto tail parameter of top incomes (e.g. Diamond and Saez, 2011; Piketty, Saez, and Stantcheva, 2014); and (iii) *firm dynamics and selection models* in the spirit of Luttmer (2007), where the stationary size distribution (and its tail exponent) is an equilibrium object pinned by entry, exit, and the structure of the model.

If the empirical tail for firm sizes is Log-Normal over economically relevant ranges, then (i) the contribution of "granular" shocks may be overstated; (ii) income-tax formulas remain valid for incomes; and (iii) quantitative firm-dynamics models should be calibrated/designed to reproduce Log-Normal (or Log-Skew-Normal) stationary size distributions, which can alter implied selection strength, misallocation wedges, and the mapping from micro shocks to macro outcomes.

3.3 Firm flows and ratios

After reviewing the distribution of firm magnitudes, which are inherently "stock" measures, we turn our attention to the three core "flow" measures of the firm, namely income, disbursement, and investment, connected by the firm's fundamental sources and uses identity in (7).

What is the statistical distribution of firm income? Firm income is of utmost importance in the theory of the firm as well as in both major branches of financial economics: corporate finance and asset pricing. Income is both the means to growth — providing money for investments, and the ends of growth — providing money for disbursements (e.g., dividends). Firm value is generally defined as discounted expected income. It is hence quite surprising that the statistical distribution of income has seen such scant interest in the economics and finance literature, and is hitherto unspecified.

Nevertheless, the fundamental sources and uses identity of the firm in (7) defines income as sales minus expenses. As the previous section showed, both sales and expenses are approximately Log-Normally distributed, with their generating Normals highly correlated. We can hence directly deduce that income (CF) should distribute as the Difference-of-Log-Normals.

The relevant figures for the three core firm flows are presented in Figure 3. Panels (a)-(c) present the distributions of income, disbursement, and investment, respectively, but truncated to the $[-\$50M, \$100M]$ range. All three exhibit clear log-tails in both the positive and negative directions. The common way of dealing with exponential tails – applying a log transform – hence cannot be used. As discussed above, the Inverse Hyperbolic Sine (asinh) transform allows us to overcome this problem, so Panels (d)-(f) present the entire data for each flow distribution, but with the X-axes transformed using asinh. The resulting two "Normal bell curves" are unmistakable. Panels (g)-(i) present the q-q plots of each vs. the Difference-of-Log-Normals, confirming the remarkable fit. The first three columns of Table 4 further confirm these predictions using formal statistical tests.

The two "Normal bells" in Panels (d)-(f) merit further discussion. In effect, they tell us that each of profits and losses is approximately Log-Normal, because the asinh transform acts as a log in both the positive and negative directions. They are outcomes of the observation that profit (and loss) are correlated with firm size: small firms make or lose small amounts of money, while large firms make or lose large amounts of money. Another way of conceptualizing it is by using the equation $\text{CF} = 2 \cdot \exp(\lambda) \cdot \sinh(\tau)$ and noting that the τ component mostly moderates the *sign* of cashflows, while the magnitude of cashflows is largely governed by λ. That is, the two Normals at the positive and negative directions are merely (near-exact) copies of the distribution of λ, mirrored.[13] This further lends credence to the observation that firm magnitude distributes approximately Normally.

Table 4
Flows and ratios - Distributional tests

This table presents results of tests of distribution equality to the Difference-of-Log-Normals for firm flows and ratios based on the measures described in Table 1. The first 3 columns pertain to the major flow variables income (CF), disbursement (DI), and investment (IT). The next 6 columns pertain to major firm ratios: IT/EM, CF/SL, CF/EQ, DI/EQ, CF/KT, IT/KT. 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.

	CF	DI	IT	IT/EM	CF/SL	CF/EQ	DI/EQ	CF/KT	IT/KT
K-S	0.005	0.004	0.008	0.044	0.006	0.012	0.007	0.011	0.008
p-val	0.098	0.121	0.069	0.016	0.089	0.053	0.075	0.055	0.068
C-2	13.66	11.63	30.70	247.9	17.45	39.14	29.50	38.51	33.87
p-val	0.100	0.111	0.066	0.024	0.087	0.051	0.067	0.052	0.063
A-D	0.588	0.344	1.119	17.33	0.465	2.364	1.084	2.109	1.791
p-val	0.084	0.099	0.069	0.026	0.090	0.051	0.070	0.054	0.060

Note that the closure of the Difference-of-Log-Normals to division by a Log-Normal implies that all "intensity" values for these flow variables should distribute $\mathcal{DLN}$ as well. This means any value of the form $\{CF, DI, IT\}/\{EM, SL, EQ, KT, XS, LM\}$ should have a Difference-of-Log-Normals distribution.[14] Such ratios include the oft-used: (i) investment per worker (aka capital deepening, IT/EM), (ii) cashflow margin (CF/SL), (iii) cashflow yield (CF/EQ), (iv) disbursement (aka dividend) yield (DI/EQ), (v) return on assets (CF/KT), (vi) investment rate (IT/KT), and many other lesser-used or hitherto undefined ratios (for the case of division by LM, which is a novel construct). The last six columns of Table 4 present formal statistical tests largely supporting this prediction, and Figure 4 presents visual evidence for three of the major ratios: CF/SL, DI/EQ, and IT/KT in Panels (a)-(c), along with the relevant q-q plots in Panels (d)-(f).

Finally, note the two typical "appearances" of the Difference-of-Log-Normals distribution, e.g., Panel (d) of Figure 3, with two "gaussian bells" vs. Panel (b) of Figure 4, with a single "spike". The two are equivalent, as can be seen by Panels (g)-(i) of Figure 4. For each of the

[13] This is because $\mathrm{Var}[\lambda] \gg \mathrm{Var}[\tau]$.

[14] With the denominators properly lagged to beginning of period values.

three ratios, these panels present the histogram of 1000 $\times$ ratio (with x-axes still in asinh scales), again exposing the gaussian bells.[15] The need to multiply by e.g. 1000 is an artifact of the scale dependence of the asinh transform, as discussed by Bellemare and Wichman (2020), and merely moves the two "bells" away from zero (where they are "compressed" to create a spike) so they can be observed separately.

Economic implications. The finding that all firm flow and intensity variables distribute $\mathcal{DLN}$ hands us a single, coherent statistical backbone for a big swath of firm fundamentals and theories. This resolves many quantitative "puzzles", which are in essence predicated on implicit or explicit assumptions regarding the Normality of intensity variables:

1. Large cash holdings are rationalized as buffers against occasional large losses or large investment needs without invoking extreme risk aversion

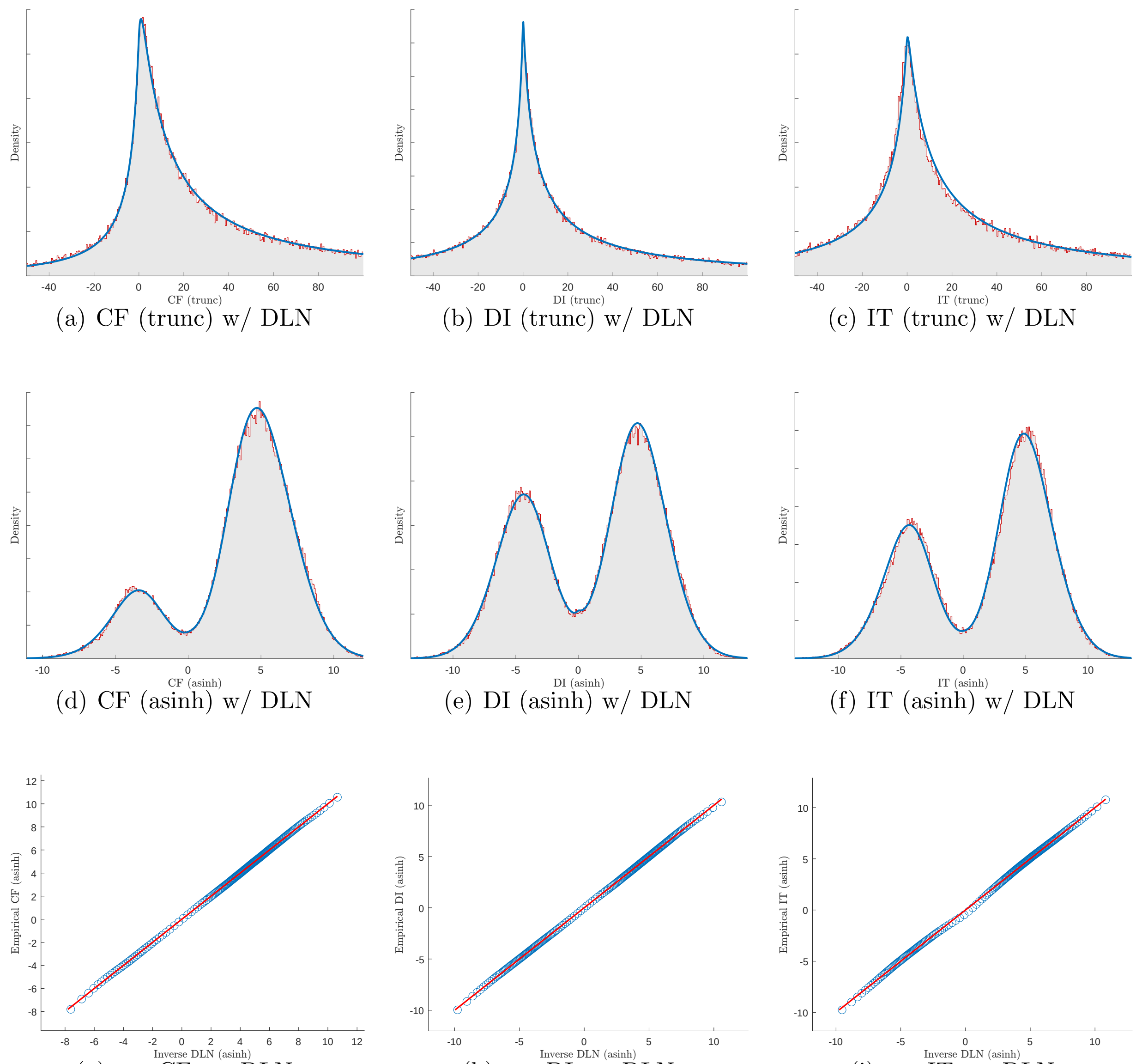
2. Investment per worker and related intensities are $\mathcal{DLN}$, implying the cross-section of technology adoption is heavy-tailed as well. This yields persistent right-tail firms with very high capital intensity, contributing to dispersion in wages and the labor share through composition effects.

3. Disbursement (i.e. dividend) yields distribute $\mathcal{DLN}$, implying asset pricing valuation ratios should be heavy-tailed and distribute $\mathcal{DLN}$ as well, as value is defined to be discounted expected dividends, thus eliminating the concept of "black swans."

4. In structural corporate finance, tail–aware earnings dynamics lower optimal leverage and widen no–issuance regions relative to Gaussian benchmarks while increasing recovery dispersion (Hennessy and Whited, 2005).

5. As in Gabaix (1999a, 2011), this distributional shape "helps constrain further theories" by giving us a target to aim at when constructing firm models. Section 4 does exactly

[15] I.e., instead of measuring growth in log-point units, we measure it in log basepoint units, defined to be 1/1000 of a log-point.

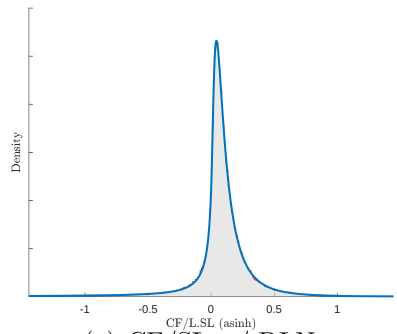


(a) CF (trunc) w/ DLN (b) DI (trunc) w/ DLN (c) IT (trunc) w/ DLN

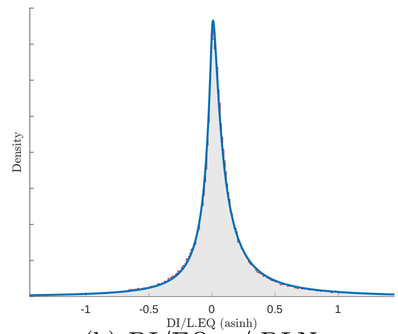
(d) CF (asinh) w/ DLN (e) DI (asinh) w/ DLN (f) IT (asinh) w/ DLN

(g) q-q CF vs. DLN (h) q-q DI vs. DLN (i) q-q IT vs. DLN

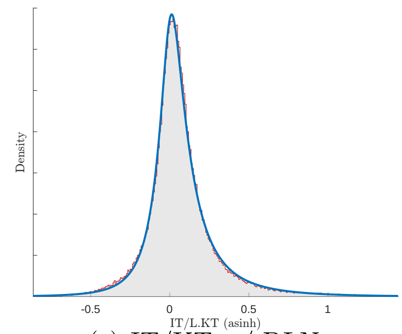
Fig. 3. Flows - distributions. This figure presents the distributions of the three major firm flows: income (CF), disbursement (DI), and investment (IT). Panels (a)-(c) present the truncated distribution of each flow in linear X-scale, between the values -50 and 100 in 2019 $M terms. Panels (d)-(f) present the full distributions, with asinh-scaled X-axes. Panels (g)-(i) present the respective q-q plots for the full distributions.



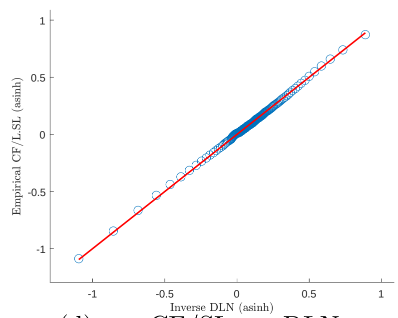
(a) CF/SL w/ DLN



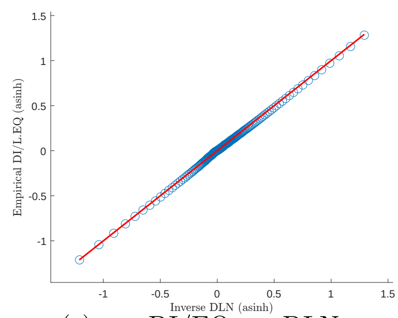
(b) DI/EQ w/ DLN



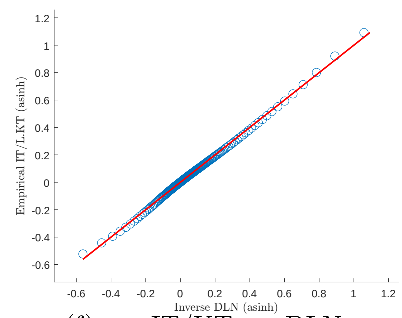
(c) IT/KT w/ DLN



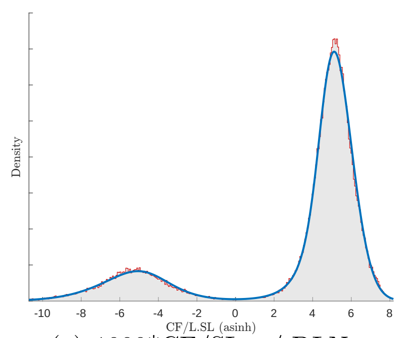
(d) q-q CF/SL vs. DLN



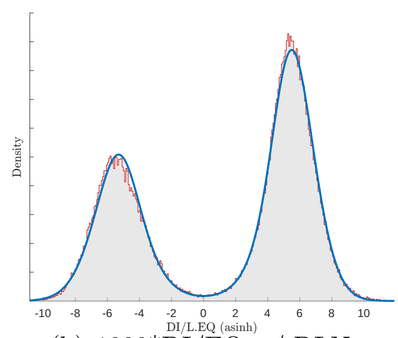
(e) q-q DI/EQ vs. DLN



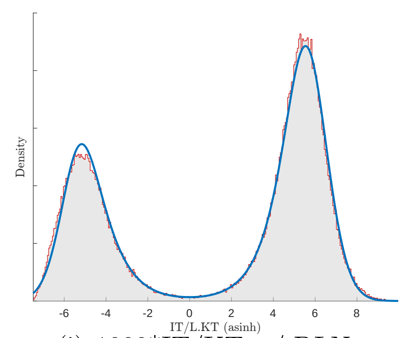
(f) q-q IT/KT vs. DLN



(g) 1000*CF/SL w/ DLN



(h) 1000*DI/EQ w/ DLN



(i) 1000*IT/KT w/ DLN

Fig. 4. Ratios - distributions. This figure presents the distributions of three major firm ratios: cashflow margin (CF/SL), disbursement yield (DI/EQ), and investment intensity (IT/KT). Panels (a)-(c) present the full distribution of each ratio. Panels (d)-(f) present the respective q-q plots. Panels (g)-(i) present the distribution of the transformed ratios, 1000 $\times$ ratio. The X-axes of all figures are in asinh-scale.

that, presenting a firm/industry model which yields $\mathcal{DLN}$ flows and ratios using a novel production function.

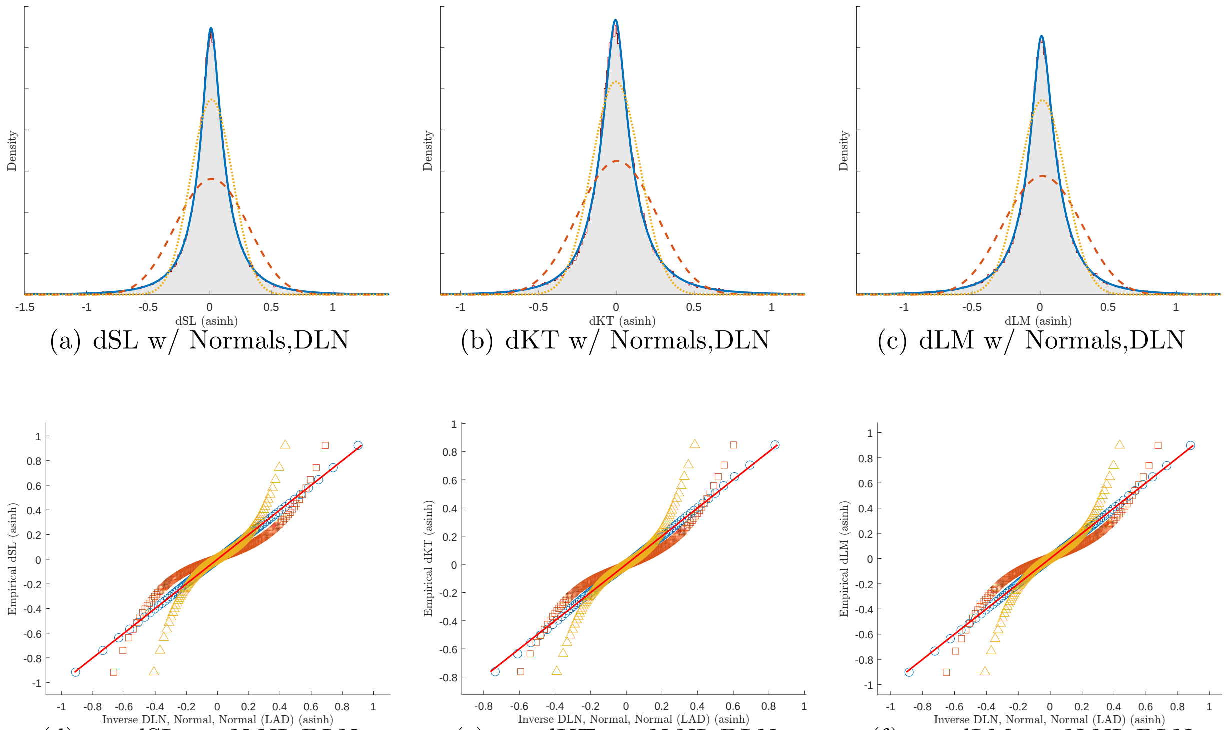
3.4 Firm growth and returns

After considering both firm stock and flow data and ascertaining that they follow the theoretical distributional predictions of Section 2, we turn our attention to the third category of firm data: firm growth. Firm growth is defined throughout as the change in firm magnitude between consecutive periods (which is simply the standard log-point growth in size measure). That is, for any size measure $XX \in \{EM, SL, EQ, KT, XS, LM\}$ and related magnitude measure log(XX), firm growth from period $t-1$ to period t is defined as

$$\mathrm{d}XX_t \equiv \log(XX_t) - \log(XX_{t-1})$$

Unlike firm cashflows, the statistical distribution of firm growth rates in the data has been of considerable interest to scholars, mainly concentrating on their observed deviations from Normality. Ashton (1926) is the first to document that the growth of British textile businesses in the period $1884-1924$, measured by the number of spindles employed, was heavy-tailed (i.e., non-Normal). Nevertheless, the assumption that growth rates are Normal has been ubiquitous in economic models (Lucas, 1978; Klette and Kortum, 2004) due to its simplicity, often referring back to the work of Gibrat (1931).

Figure 5 presents the empirical distributions of the growth in sales (dSL), capital (dKT), and productive magnitude (dLM) over the research period. Each of the Panels (a)-(c) is overlaid with a fitted Difference-of-Log-Normals distribution as well as two Normal distributions: one fitted via Maximum Likelihood (i.e., by matching the standard deviation of the data, as usual), and one fitted via Least-Absolute-Deviations (LAD, i.e., by matching the inter-quartile-range of the data). The fit of the $\mathcal{DLN}$ is remarkable, as can also be seen in the associated quantile-quantile plots in Panels (d)-(f). The Normals offer a poor fit, as the figures ascertain, and are unable to match the distributional shape or heavy tails using either method.



(a) dSL w/ Normals,DLN (b) dKT w/ Normals,DLN (c) dLM w/ Normals,DLN

(d) q-q dSL vs. N,NL,DLN (e) q-q dKT vs. N,NL,DLN (f) q-q dLM vs. N,NL,DLN

Fig. 5. Growth - distributions. This figure presents stylized facts of firm growth distributions. Panels (a)-(c) present the histograms of (log-point, yearly) growth in sales SL, capital KT, and productive magnitude LM, respectively. The panels are overlaid with MLE- and LAD-fitted Normal distributions (dashed red and dotted yellow, respectively), as well as a Difference-of-Log-Normals distribution (solid blue). Panels (d)-(f) present the respective q-q plots.

Two prominent non-Normal alternatives extensively examined in the literature are the *Stable* and *Asymmetric Laplace* distributions. The early contributions of Mandelbrot (1960, 1961) and Fama (1963, 1965) advanced the four-parameter Stable (or Stable–Paretian) family as a candidate model for stock returns and, by extension, for firm growth. Although subsequent empirical analyses (e.g., Officer, 1972) rejected the Stable as an exact description of return data, it has remained a workhorse approximation in the modeling of financial returns. The persistence of this specification reflects its analytical tractability and its capacity to capture the heavy tails (i.e. high kurtosis) observed in empirical return data.

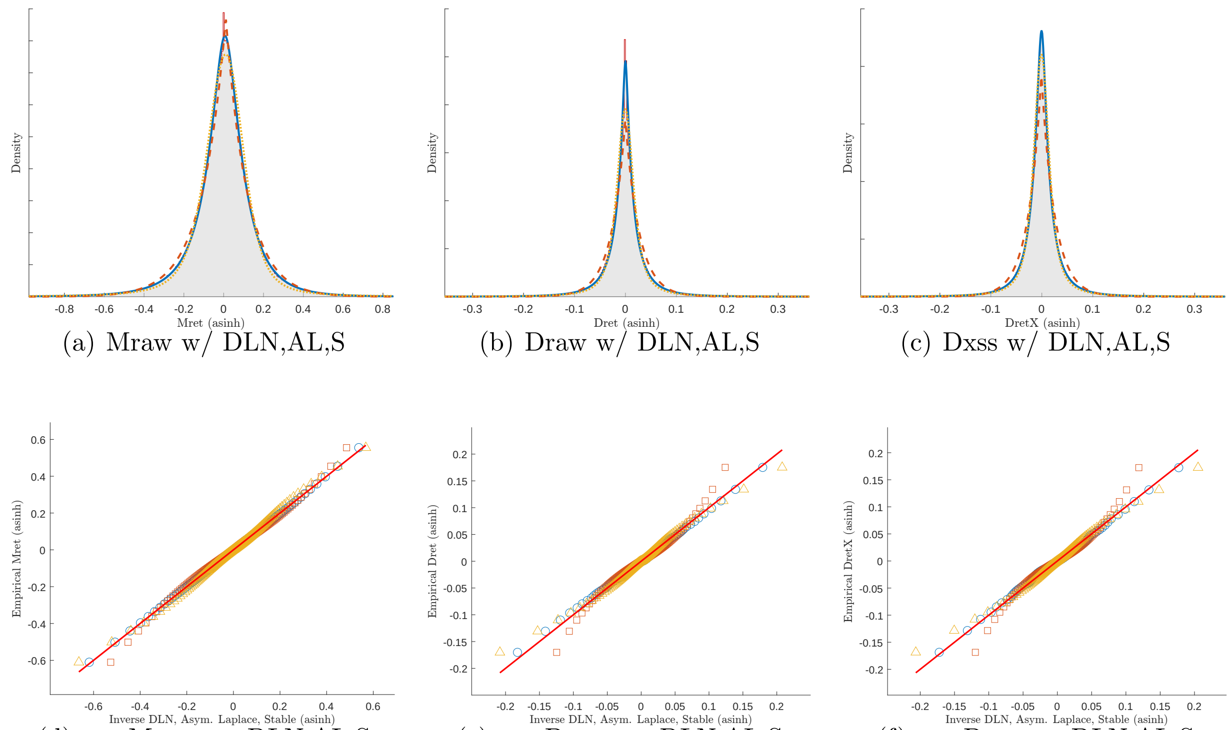
A parallel strand of research, initiated by Stanley, Amaral, Buldyrev, Havlin, Leschhorn, Maass, Salinger, and Stanley (1996), proposed the three-parameter Asymmetric Laplace as an alternative heavy-tailed distribution. A large body of subsequent work demonstrates that it provides a good approximation to growth data in a variety of domains,[16] although subsequent analysis (e.g., Bottazzi, Coad, Jacoby, and Secchi, 2011; Arata, 2019) formally rejected it under standard goodness-of-fit criteria as well. Nevertheless, both firm models (Alfarano and Milaković, 2008; Luttmer, 2011) and econometric work (Toda, 2012; Toda and Walsh, 2015) used the Laplace as a "target distribution" informing modeling choices.

The economic literature's sustained interest in the distribution of firm growth stems in part from the recognition that equity returns constitute a high-frequency index of firm growth, when firm size is taken as market value of equity (EQ). Indeed, much of the early work on the distribution of firm growth has been done using stock (and other asset) returns as a natural laboratory with high-quality and high-frequency data — while most measures of size are usually observable to the econometrician only at the yearly frequency, asset prices are observable per month, day, or even second.

Figure 6 hence presents empirical stock return distributions. Return data are from the CRSP dataset, minimally filtered, and cover the same research period 1970-2019. Panels (a) and (b) of the figure present the monthly and daily "raw" return distributions, overlaid with Difference-of-Log-Normals, Asymmetric-Laplace, and Stable MLE-fitted distributions. Panel (c) again displays the daily return distribution, but this time considers the daily *excess* return relative to the Fama-French 3-factor model. The fit of the Laplace and Stable is much closer than for Normals, as can be expected. Nevertheless, the quantile-quantile plots in panels (d)-(f) exhibit deviations around the tails for both, while no such deviations are apparent for the Difference-of-Log-Normals.

To again avoid depending on visual straight-line tests, Table 5 presents comprehensive statistical tests for the growth of each of the six size measures used throughout, as well as

[16] E.g. Canning, Amaral, Lee, Meyer, and Stanley (1998), Bottazzi and Secchi (2003, 2006), Gabaix, Gopikrishnan, Plerou, and Stanley (2006), Buldyrev, Growiec, Pammolli, Riccaboni, and Stanley (2007).



(a) Mraw w/ DLN,AL,S (b) Draw w/ DLN,AL,S (c) Dxss w/ DLN,AL,S

(d) q-q Mraw vs. DLN,AL,S (e) q-q Draw vs. DLN,AL,S (f) q-q Dxss vs. DLN,AL,S

Fig. 6. Return - distributions. This figure presents stylized facts of higher-frequency firm return distributions. Panel (a) presents the histogram of monthly raw returns for firms in the CRSP universe during the period 1970-2019. Panel (b) presents daily raw returns, and Panel (c) presents daily excess returns (relative to the FF3 model). The panels are overlaid with MLE-fitted Difference-of-Log-Normals (solid blue), Asymmetric-Laplace (dashed red), and Stable (dotted yellow) distributions. Panels (d)-(f) present the respective q-q plots.

all six stock return flavors $\{Yearly, Monthly, Daily\}X\{raw, xss\}$, vs. the three candidate distributions. Neither growth nor returns are generally rejected as being Difference-of-log-Normals, regardless of specific definitions, as the high p-values in Panel (a) of Table 5 demonstrate. The same data are generally strongly rejected as being Asymmetric-Laplace or Stable, as Panels (b) and (c), with very low p-values, demonstrate. The test most sensitive to tail deviations, the Anderson-Darling test, rejects none of the 12 data series as being $\mathcal{DLN}$-distributed. It however rejects 11 of the 12 as being Asymmetric-Laplace or Stable, at the 5% level. The other two tests are similarly conclusive. The $\mathcal{DLN}$ is also overwhelmingly favored when conducting an AIC/BIC-based likelihood ratio "horse-race" between the three distributions.

An intuitive explanation of why growth rates are $\mathcal{DLN}$-distributed stems from a stock vs. flow analysis. If the flow of value created is $\mathcal{DLN}$-distributed, as shown above both theoretically and empirically, and the magnitude of this flow of value is cointegrated with all other magnitude measures, as shown above as well, then a simple stock-flow analysis will immediately yield that the growth in each stock must distribute similar to the flow-intensity distribution. The simple intuitive example is a firm with trivial investment-disbursement decision, which accumulates all its earning into increased capital (or decreased, if earnings at any period are negative).[17] Such a firm will have $\mathcal{DLN}$-distributed growth in capital, and via the endogenous co-integration channel, $\mathcal{DLN}$-distributed growth in all other size measures.

Economic implications. As recently discussed by Jaimovich et al. (2023), the non-Normal features of the return (and generally growth) distributions have first-order implications to economic questions, including: the sensitivity of aggregate outcomes and policy responses to micro-level shocks, the dynamics of firm exit and investment decisions, and the propagation of idiosyncratic shocks into aggregate fluctuations. This is because fat tails, leptokurtosis, and non-Gaussian transition dynamics radically alter the relationship between

[17]That is, $KT_{t+1} = KT_t + IT_t = KT_t(1 + IT_t/KT_t) \rightarrow dKT_{t+1} = log(1 + IT_t/KT_t) \approx IT_t/KT_t \sim \mathcal{DLN}$.

Table 5

Growth - Distributional tests

This table presents results of tests of distribution equality for firm growth and stock return measures. The first six columns pertain to log-point growth in the six magnitude measures of Table 2, while the last six columns pertain to log-point stock returns at the Yearly, Monthly, and Daily frequency, considering both raw and excess (relative to the Fama-French 3 factor model) returns. 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 for the Difference-of-Log-Normals, Asymmetric-Laplace, and Stable distributions, respectively. Panel (d) reports the AIC- and BIC-based relative likelihoods for each distribution.

	EM	SL	EQ	KT	XS	LM	Yraw	Yxss	Mraw	Mxss	Draw	Dxss
Panel (a): Growth / return vs. Difference-of-Log-Normals												
K-S	0.009	0.003	0.003	0.007	0.003	0.002	0.005	0.002	0.003	0.001	0.017	0.009
p-val	0.067	0.154	0.181	0.080	0.191	0.206	0.105	0.213	0.174	0.908	0.040	0.063
C-2	46.64	7.989	6.986	18.89	5.191	5.646	10.47	4.922	4.203	0.750	115.0	21.61
p-val	0.054	0.148	0.171	0.084	0.313	0.246	0.119	0.376	0.607	1.000	0.036	0.078
A-D	0.969	0.131	0.145	0.547	0.114	0.081	0.384	0.086	0.085	0.021	1.717	1.520
p-val	0.072	0.136	0.131	0.086	0.144	0.163	0.096	0.160	0.160	0.369	0.061	0.063
Panel (b): Growth / return vs. Asymmetric- Laplace												
K-S	0.043	0.037	0.012	0.026	0.044	0.033	0.019	0.022	0.017	0.022	0.039	0.044
p-val	0.016	0.020	0.052	0.028	0.016	0.022	0.038	0.033	0.040	0.033	0.019	0.016
C-2	436.6	359.3	62.79	274.9	566.6	317.6	85.57	104.0	101.0	145.0	551.9	542.1
p-val	0.016	0.019	0.047	0.022	0.013	0.020	0.041	0.037	0.038	0.032	0.014	0.014
A-D	33.38	29.03	2.325	17.74	45.89	23.47	4.699	6.793	6.865	11.17	39.92	48.20
p-val	0.019	0.020	0.055	0.026	0.015	0.023	0.044	0.038	0.038	0.032	0.017	0.015
Panel (c): Growth / return vs. Stable												
K-S	0.015	0.012	0.013	0.011	0.010	0.012	0.017	0.016	0.012	0.011	0.017	0.011
p-val	0.044	0.052	0.048	0.055	0.062	0.053	0.041	0.043	0.054	0.055	0.040	0.057
C-2	146.5	103.0	96.96	94.37	78.94	104.6	197.0	185.7	102.7	100.6	316.2	124.2
p-val	0.032	0.038	0.039	0.039	0.043	0.037	0.027	0.028	0.038	0.038	0.020	0.034
A-D	6.697	4.061	4.261	3.909	3.021	4.037	6.966	6.524	3.969	3.741	6.649	4.082
p-val	0.039	0.046	0.045	0.047	0.051	0.046	0.038	0.039	0.046	0.047	0.039	0.046
Panel (d): Distribution comparison												
AIC R.L.:												
D-L-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
A-L	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
S	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BIC R.L.:												
D-L-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
A-L	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
S	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

current and future firm value, firm exit rates, and macroeconomic responsiveness compared to standard Gaussian AR(1) assumptions.

The high kurtosis of the growth distribution concentrates lifetime values near the continuation–exit margin in heterogeneous–firm models, so small wedges move many firms across it. Entry subsidies raise competitive pressure and improve selection while operating subsidies keep low–efficacy firms alive and worsen selection. The prevailing Normal calibrations systematically mis-rank these policies and understate exit elasticities. (e.g., Jovanovic, 1982; Hopenhayn, 1992; Luttmer, 2007). Furthermore, heavier but finite tails widen inaction regions and shift investment thresholds in firm investment models such as Dixit and Pindyck (1994); Abel and Eberly (1994). Models that match levels yet impose Normal innovations misstate churn, the dispersion–productivity link, and the speed of selection.

The young age of large firms is also rationalized when one considers a $\mathcal{DLN}$ growth rate, as some firms may enjoy large positive shocks, quickly catapulting them to large sizes contrary to the predictions of Normal growth rates. A simple Monte-Carlo simulation with Normal and $\mathcal{DLN}$-distributed growth, both calibrated to the data, shows large firms emerge quickly (decades) with $\mathcal{DLN}$ growth vs. slowly (centuries) with Normal growth.

In an asset pricing context, the Stable family remains in use in parts of empirical finance even though it lacks finite second and higher moments, the very objects risk measurement targets, thus providing a precarious basis for Modern Portfolio Theory (MPT). The Stable was later rejected by Officer (1972), who shows that the second moment of returns in the data is well-behaved and 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 $\mathcal{DLN}$ has finite moments of all orders, and provides a remarkable match to raw and factor–adjusted returns. Finally, because inference on factor residuals and "alpha" becomes fragile when tails are mis-specified, tail robust estimators (i.e., L1-norms rather than L2-norms) should be used when estimating common asset pricing models. (Fama and French, 1993; Harvey, Liu, and Zhu, 2016).

Finally, with leptokurtic growth, idiosyncratic shocks survive aggregation more than Normal models imply, strengthening granular transmission to macro volatility and reallocation as in Gabaix (2011). Hence, while the findings above regarding the Normality of the magnitude distribution *weaken* the "granular hypothesis", the findings herein regarding the $\mathcal{DLN}$ growth distribution *strengthen* it.

4 A general firm model

4.1 Model Setup

State vector. Collect every observable shifter of the firm including, if desired, rivals' outputs and lags, prices, etc, as well as the firm's choice variable, assumed here to be quantity Q_t, in

$$X_t \;=\; \big[\,1,\; \log(Q_t),\; X_{1,t}, \ldots, X_{k,t}\big] \in \mathbb{R}^{k+2}, \qquad Q_t > 0.$$

It is generally assumed that $X_{i,t}$ are logged when appropriate, e.g. log rivals' quantities, log prices, etc. In addition, assume that each $X_{i,t}$ carries an additive normal error $u_{i,t}$. This can arise via classical measurement error, expectational error on future values, AR(1) dynamics (which we will return to), etc. Note that e.g. if prices follow a geometric Brownian motion, as generally assumed, then log prices have an unexpected additive Normal component. Assume that the $u_{i,t}$ are jointly Normal, and possibly correlated between factors i, though independent across time t.

Log–link primitives. Demand D_t and cost C_t, both in dollar terms, are given by

$$\log(D_t) = d_t = \theta_d \cdot X_t + \varepsilon_{d,t}, \qquad \theta_d = [\alpha_d,\; \beta_d,\; \gamma_d], \tag{8}$$

$$\log(C_t) = c_t = \theta_c \cdot X_t + \varepsilon_{c,t}, \qquad \theta_c = [\alpha_c,\; \beta_c,\; \gamma_c]. \tag{9}$$

and note the parameter vectors θ_d, θ_c encompass all information on the firm's production technology and cost structure. Taken together, these specifications both generalize and unify the textbook iso-elastic and power–cost cases, respectively, which remain the workhorses in applied micro. On the demand side, the iso-elastic quantity law $Q_t = A\,P_t^{-\eta}$ implies dollar demand $D_t = P_t Q_t = A\,P_t^{1-\eta}$, so the coefficient on $\log P_t$ in (8) satisfies $\beta_d = 1-\eta$. On the cost side, (9) is the cost dual of a Cobb–Douglas technology, yielding log-linear expenditure in prices (Varian, 1992); more generally, Constant Elasticity of Substitution (CES) "power-cost" forms $C_t = \kappa\,P_t^* Q_t^\zeta$ (with P_t^* an input-price index) imply the coefficient on $\log Q_t$ equals the cost–output elasticity $\beta_c = \zeta$ ($\zeta = 1$ under constant returns). Empirically, these log-link specifications are ubiquitous in IO/marketing, energy demand, and trade/gravity.

Stochastic structure. The sum of scaled Normals remains Normal, so all noise terms $u_{i,t}$ are absorbed into $\varepsilon_{d,t}, \varepsilon_{c,t}$, such that $[\varepsilon_{d,t},\ \varepsilon_{c,t}] \sim \mathcal{N}(0, \Sigma_{dc})$, with

$$\Sigma_{dc} = \begin{bmatrix} \sigma_d^2 & \sigma_{dc} \\ \sigma_{dc} & \sigma_c^2 \end{bmatrix}, \qquad \sigma_{dc} = \rho_{dc}\,\sigma_d\sigma_c, \qquad -1 < \rho_{dc} < 1, \tag{10}$$

is a bivariate Normal error term. Note the special case in which we assume all observables $X_{i,t}$ are modeled as a single VAR(1) process, i.e.

$$\begin{bmatrix} X_{1,t} \\ \vdots \\ X_{k,t} \end{bmatrix} = (I_k - R)\cdot M \ + \ R\cdot \begin{bmatrix} X_{1,t-1} \\ \vdots \\ X_{k,t-1} \end{bmatrix} \ + \ u_t, \qquad \begin{gathered} u_t \sim \mathcal{N}(0, \Sigma_x), \\ R, \Sigma_x \in \mathbb{R}^{k\times k},\ M \in \mathbb{R}^k \end{gathered}$$

with $u_{i,t}$ stemming from the VAR(1) process. In this case, (log) demand and (log) cost d_t, c_t explicitly inherit a VAR(1) structure such that

$$\begin{bmatrix} d_t \\ c_t \end{bmatrix} = (I_2 - R_{dc})\,M_{dc} \ + \ R_{dc} \begin{bmatrix} d_{t-1} \\ c_{t-1} \end{bmatrix} \ + \ v_t, \qquad \begin{gathered} v_t \sim \mathcal{N}(0, \Sigma_{dc}), \\ R_{dc}, \Sigma_{dc} \in \mathbb{R}^{2\times 2},\ M_{dc} \in \mathbb{R}^2 \end{gathered}$$

with $R_{dc}, \Sigma_{dc}, M_{dc}$ functions of the underlying dynamic structure of X_t and the parameter vectors θ_d, θ_c. This allows us to see explicitly how d_t, c_t will tend to Normality, and how D_t, C_t will then tend to be $\mathcal{LN}$-distributed in this model. It will also be useful when taking the model to data.

Profit. The firm's period-t profit is

$$\Pi_t(Q_t) \;=\; D_t - C_t \;=\; \exp(d_t) \;-\; \exp(c_t) \tag{11}$$

Conditional on X_t, (d_t, c_t) are jointly Normal with means $\theta_d{\cdot}\, X_t$ and $\theta_c{\cdot}\, X_t$, respectively, and covariance matrix Σ_{dc} given by (10). Hence, profit distributes $\mathcal{DLN}$ in this general model. Furthermore, using Log-Normal moments, we have

$$\mathbb{E}[\Pi_t \mid X_t] \;=\; \exp\!\big(\theta_d \cdot X_t + \tfrac{1}{2}\sigma_d^2\big) \;-\; \exp\!\big(\theta_c \cdot X_t + \tfrac{1}{2}\sigma_c^2\big). \tag{12}$$

and for a risk-neutral firm the first-order condition (FOC) on expected profit yields the closed-form optimum (interior if $0 < \beta_d < \beta_c$),

$$Q_t^\star \;=\; \left(\frac{\beta_d\, \exp\!\big(\alpha_d + \sum_{i=1}^k \gamma_{d,i} \cdot X_{i,t} + \frac{1}{2}\sigma_d^2\big)}{\beta_c\, \exp\!\big(\alpha_c + \sum_{i=1}^k \gamma_{c,i} \cdot X_{i,t} + \frac{1}{2}\sigma_c^2\big)}\right)^{1/(\beta_c - \beta_d)} \tag{13}$$

Note that the covariance term σ_{dc} and equivalently the correlation term ρ_{dc} do not enter the expected profit or optimal quantity calculation in this case, as an artifact of the risk-neutrality assumption. The correlation ρ_{dc} instead impacts profit volatility (or "riskiness") $\mathrm{Var}[\Pi_t \mid X_t]$. The higher ρ_{dc}, the less volatile profit becomes.

The model is a proper extension, as under the iso–elastic/power–cost restriction $\beta_d = 1-\eta$ and $\beta_c = \zeta$, the expression in (13) collapses to the textbook (Varian/Tirole) expression for a single-product monopolist.

Finally, note the model highlights the idea that adjustment costs, time to build, and expectational errors are core components in faithfully representing the dynamics of firms. If firms can freely and flexibly adjust their inputs within period and predict next-period prices / economic conditions with certainty, then no firm should ever suffer a loss.

Market structure. The model is general enough to accommodate a variety of market structures using different assumptions on the components of X_t:

- Monopoly benchmark: omit rivals' (log) quantities $\log(Q_{-i,t})$ from X_t.
- Residual demand / Static Cournot: include $\log(Q_{-i,t})$ in X_t. The FOC characterizes the firm's best response to the residual demand taking rival's quantities as fixed, or solve the system of equations given by each firm's FOC to get the static Cournot-Nash equilibrium.
- Dynamic oligopoly: include observable lags of $\log(Q_{-i,t-1})$ in X_t to calculate a Markov–perfect equilibrium as in e.g. Ericson and Pakes (1995).

4.2 Hyperbolic representation of the model

Define the transformed objects $b_t, a_t, \theta_b, \theta_a, \varepsilon_{b,t}, \varepsilon_{a,t}$,

$$b_t = \tfrac{d_t+c_t}{2}, \quad a_t = \tfrac{d_t-c_t}{2}, \quad \theta_b = \tfrac{\theta_d+\theta_c}{2}, \quad \theta_a = \tfrac{\theta_d-\theta_c}{2}, \quad \varepsilon_{b,t} = \tfrac{\varepsilon_{d,t}+\varepsilon_{c,t}}{2}, \quad \varepsilon_{a,t} = \tfrac{\varepsilon_{d,t}-\varepsilon_{c,t}}{2}$$

which yields the expressions for hyperbolic radius and angle,

$$b_t = \theta_b \cdot X_t + \varepsilon_{b,t} \qquad a_t = \theta_a \cdot X_t + \varepsilon_{a,t} \tag{14}$$

with $[\varepsilon_{b,t},\ \varepsilon_{a,t}] \sim \mathcal{N}(0, \Sigma_{ba})$. We then have profit given by

$$\Pi_t(Q_t) = 2\ \exp\big(b_t\big)\sinh\big(a_t\big), \tag{15}$$

and expected profit given by

$$\mathbb{E}[\Pi_t \mid X_t] = 2\exp\Big(\theta_b \cdot X_t + \tfrac{1}{2}(\sigma_b^2 + \sigma_a^2)\Big)\ \sinh\Big(\theta_a \cdot X_t + \sigma_{ba}\Big). \tag{16}$$

which yields the closed-form optimality condition (for an interior solution $|\beta_b/\beta_a| < 1$)

$$\tanh(\theta_a \cdot X_t + \sigma_{ba}) = -\frac{\beta_b}{\beta_a} \tag{17}$$

or equivalently

$$Q_t^\star = \exp\Big(-\frac{\operatorname{atanh}(\beta_b/\beta_a) + \alpha_a + \sum_{i=1}^k \gamma_{a,i} \cdot X_{i,t} + \sigma_{ba}}{\beta_a}\Big).$$

which then collapses to (13) when one transforms back to (d_t, c_t) space. Note (17) can be interpreted as requiring that the straight line given by varying $\log(Q_t)$ (a line in (b_t, a_t) space, whose slope is β_b/β_a), will be tangent to the iso-expected-profit hyperbolic contour at the optimum $Q_t^\star$. It is evident from (15) that the sign of profit is fully controlled by a_t, with $a_t > 0$ yielding profit and $a_t < 0$ yielding loss, and the magnitude of profit/loss is largely controlled by b_t.

4.3 An estimable q-theory variant

So far, we have assumed the firm chooses the quantity of its single product, Q_t. This is difficult to take to data in a multi-product world and absent data on quantities. Instead, I will now follow the q-theory tradition and assume the firm chooses its capital level K_t. That is, the firm faces an investment-disbursement decision and chooses how much to invest vs. how much to disburse to its owners (with negative values representing capital sale and cash infusions from owners, respectively). The firm makes the decision with the goal of maximizing firm value, defined as expected discounted disbursements.

The firm's *profit production function* is given by

$$\Pi_t(k_t, \lambda_t, \tau_t) = 2\,\exp\big(\lambda_t + \beta_\lambda \cdot k_t\big)\sinh\big(\tau_t + \beta_\tau \cdot k_t\big), \tag{18}$$

with k_t log capital and λ_t, τ_t two exogenous stochastic processes given by the VAR(1),

$$\begin{bmatrix}\lambda_t \\ \tau_t\end{bmatrix} = (I_2 - R_{\lambda\tau})\,M_{\lambda\tau} \;+\; R_{\lambda\tau}\begin{bmatrix}\lambda_{t-1} \\ \tau_{t-1}\end{bmatrix} \;+\; \varepsilon_t, \qquad \begin{matrix}\varepsilon_t \sim \mathcal{N}(0, \Sigma_{\lambda\tau}), \\ R_{\lambda\tau}, \Sigma_{\lambda\tau} \in \mathbb{R}^{2\times 2},\; M_{\lambda\tau} \in \mathbb{R}^2\end{matrix} \tag{19}$$

which collapses the terms $\alpha_b + \sum_{i=1}^{k} \gamma_{b,i} \cdot X_{i,t} + \varepsilon_{b,t}$ into λ_t, and the equivalent a-terms into τ_t.

Importantly, the firm's revenue D_t, expenses C_t, and capital K_t are usually easily observable. If we assume the two scalars $\beta_\lambda, \beta_\tau$ are known (e.g., due to a first-step estimation), then one can observe λ_t, τ_t by using

$$\lambda_t + \beta_\lambda \cdot k_t = \log(\sqrt{D_t \cdot C_t}) = \frac{d_t + c_t}{2} \quad ; \quad \tau_t + \beta_\tau \cdot k_t = \log(\sqrt{D_t / C_t}) = \frac{d_t - c_t}{2} \tag{20}$$

and then one can extract the entire VAR(1) structure in (19) via simple MLE. That is, this setup allows us to easily estimate the two "productivity processes" governing the firm, without explicitly knowing the values of $X_{i,t}$ or of the relevant α, γ parameters. Note that with knowledge of λ_t, τ_t, one can now regress them on candidate observables $X_{i,t}$ (with proper IVs for ε_t) to uncover the structure of the revenue and cost functions – or equivalently the productive magnitude and efficacy functions – of the firm.

All that's left is to embed this production function into a standard q-theory model of the firm, with endogenous entry and exit, convex capital adjustment costs, and accounting for capital depreciation, and then estimate it. The companion paper Parham (2023) does precisely that, showing in detail how to estimate such models and confirming the model yields the expected distributions discussed above. The move from the (d,c)-space to the (b,a)-space is quite useful in this estimation process as well. This is because it turns the two highly correlated stochastic processes of revenue "productivity" and cost "dis-productivity" into the nearly uncorrelated productive magnitude and efficacy processes, making them much simpler to simulate and estimate.

The companion paper also shows that nearly all cashflow growth in the data is driven by changes in the efficacy τ of firms, rather than changes in their magnitude λ. This implies that the dynamics of τ — a novel empirical and theoretical object — are of first order importance to understanding firm dynamics and growth.

Interestingly, the companion paper also shows that the "return to scale" (or profit elasticity w.r.t capital) of the profit production function in Equation 18 is given by:

$$\text{RTS} = \beta_\lambda + \frac{\beta_\tau}{\tanh\left(\tau_t\right)} \approx \beta_\lambda + \frac{\beta_\tau}{\tau_t} \tag{21}$$

explicitly making RTS a state-dependent and firm-specific object, which is nearly linearly dependent on $1/\tau$ and can "explode" to $\pm\infty$ when τ is close to 0 (very common in the data).

5 Concluding remarks

This paper's core message is simple: many growth phenomena in economics are best understood as the balance of two opposing *multiplicative* forces. When each force is itself the product of many small shocks, the (multiplicative) Central Limit Theorem (CLT) pushes each component toward a log-Normal limit. The economically relevant object, however, is typically a *net* outcome: sales minus expenses; agglomeration minus congestion; births minus deaths. This net object then inherits a *Difference-of-Log-Normals* ($\mathcal{DLN}$) structure. This is not a modeling outcome. It is an accounting fact combined with a CLT fact.

The paper contributes along three dimensions.

First, it provides a CLT-based taxonomy of limiting distributions that repeatedly arise in economic contexts: Normal and Skew-Normal limits for additive aggregation with mild asymmetries; Log-Normal and Log-Skew-Normal limits for multiplicative aggregation; and $\mathcal{DLN}$ limits for net outcomes driven by opposing multiplicative forces. The $\mathcal{DLN}$ admits a natural hyperbolic representation that decomposes a net outcome into *productive magnitude* λ and *productive efficacy* τ, and it delivers a first-principles justification for the use of the asinh($\cdot$) transform when analyzing bi-directional log-tailed data. In practice, this representation turns two highly correlated log processes (e.g. log sales and log expenses) into two nearly uncorrelated objects, namely magnitude and efficacy, making both interpretation and econometrics cleaner.

Second, the paper documents that these predictions organize a broad set of data. In firm panels, magnitudes are well described by Skew-Normals, with Normal upper tails. Upper tails are inconsistent with log-Exponential, i.e. power-law, behavior over economically relevant ranges. In contrast, net firm flows (cashflows, disbursements, investment), their key ratios, growth rates, and stock returns at multiple horizons exhibit strong $\mathcal{DLN}$ fit and are generally rejected as being Normal, Asymmetric-Laplace, or Stable under standard goodness-of-fit tests, including tail-sensitive tests. Empirically, the $\mathcal{DLN}$ is not merely a "heavy-tailed" candidate: it is a disciplined, theory-implied, finite-moment alternative that puts probability mass where the data are (lots of modest moves and occasional very large moves) without requiring infinite-variance "black swans," time-varying volatility, or explanations outside the firm. Unlike stochastic-volatility or mixture explanations, the $\mathcal{DLN}$ arises even with homoskedastic Normal primitives. More broadly, Figure 1 suggests the same logic reaches well beyond firms, organizing growth in regional output, pandemic spread, wages, temperatures, and other settings in which net outcomes reflect the difference of two multiplicative components.

Third, the paper shows how to build and estimate economic models that generate $\mathcal{DLN}$ outcomes transparently. The difference-of-log-linears profit production function, a generalization of the Cobb-Douglas (or log-linear) production function, embeds $\mathcal{DLN}$ net outcomes into familiar firm environments and makes the (λ, τ) state space operational. Beyond interpretability, it yields testable implications: it links heavy-tailed growth to the interaction of magnitude and efficacy, implies state-dependent returns to scale through $\tanh(\tau)$, and suggests that empirically plausible heterogeneity in efficacy can translate into large and time-varying differences in effective scale elasticities. The framework also motivates a coherent extension of growth measurement to sometimes-negative variables, as described in the companion paper Parham (2023). When the object of interest is $\mathcal{DLN}$ and generated by a difference-of-log-linears process, hyperbolic geometry provides a natural generalization of log-point and percent growth that remains well-defined across zero (i.e., growth *to and from negative values*), thus allowing us to discuss the growth of profits and losses without ad hoc fixes.

Distributional misspecification is not innocuous. A Normal approximation puts probability in the middle and mutes the margins; a $\mathcal{DLN}$ allocates substantial probability to large moves while retaining finite moments. This shifts value toward exit margins and raises the elasticity of survival, hiring, and investment to small wedges. These are all mechanisms that can change quantitative conclusions and even policy rankings (entry subsidies versus operating subsidies versus taxes) in calibrated or estimated models. Recent work by Jaimovich et al. (2023) emphasizes that the non-Normal features of growth and return distributions can be first-order objects for macro and welfare; the $\mathcal{DLN}$ provides a tractable, principled, and analytic way to take that message seriously without giving up well-behaved moments or resorting to using empirical distributions. If growth is the foundation, then the law of motion that governs it is a first-order object; the evidence here suggests that "two-factor" thinking, and its $\mathcal{DLN}$ implications, should be part of the economist's default toolkit.

Finally, there is a sense in which this paper is simply an extension of Gibrat (1931). Gibrat's core insight was that proportional growth makes the log-Normal a fundamental object for sizes. Gibrat's goal in his 1931 book was to "*convince his readers that this was a statistical regularity sufficiently sharp to provide a basis for serious mathematical modeling*" (Sutton, 1997). The core insight here is that opposing proportional forces make the Difference-of-Log-Normals a fundamental object for net outcomes and growth, in the sense that it arises in disparate settings where net outcome and growth are concerned, similar to the repeated occurrence across disciplines of its better-known peers, the Normal and Log-Normal.

References

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

Aihounton, G. B. D., Henningsen, A., 2021. Units of measurement and the inverse hyperbolic sine transformation. The Econometrics Journal 24, 334–351.

Alfarano, S., Milaković, M., 2008. Does classical competition explain the statistical features of firm growth? Economics Letters 101, 272–274.

Anděl, J., Netuka, I., Zvára, K., 1984. On Threshold Autoregressive Processes. Kybernetika pp. 89–106.

Arata, Y., 2019. Firm growth and Laplace distribution: The importance of large jumps. Journal of Economic Dynamics and Control 103, 63–82.

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

Axtell, R. L., 2001. Zipf Distribution of U.S. Firm Sizes. Science 293, 1818–1820.

Azzalini, A., 1985. A Class of Distributions Which Includes the Normal Ones. Scandinavian Journal of Statistics 12, 171–178.

Azzalini, A., Cappello, T. D., Kotz, S., 2002. Log-Skew-Normal and Log-Skew-t Distributions as Models for Family Income Data. Journal of Income Distribution 11, 2–2.

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

Bottazzi, G., Coad, A., Jacoby, N., Secchi, A., 2011. Corporate growth and industrial dynamics: Evidence from French manufacturing. Applied Economics 43, 103–116.

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

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

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

Cambanis, S., Huang, S., Simons, G., 1981. On the theory of elliptically contoured distributions. Journal of Multivariate Analysis 11, 368–385.

Canning, D., Amaral, L. A. N., Lee, Y., Meyer, M., Stanley, H. E., 1998. Scaling the volatility of GDP growth rates. Economics Letters 60, 335–341.

Clauset, A., Shalizi, C. R., Newman, M. E. J., 2009. Power-law distributions in empirical data. SIAM Review 51, 661–703.

Diamond, P., Saez, E., 2011. The Case for a Progressive Tax: From Basic Research to Policy Recommendation. Journal of Economic Perspectives 25, 165–190.

Dixit, A. K., Pindyck, R. S., 1994. Investment Under Uncertainty. Princeton University Press.

Ericson, R., Pakes, A., 1995. Markov-perfect industry dynamics: A framework for empirical work. Review of Economic Studies 62, 53–82.

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.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E. F., Roll, R., 1971. Parameter Estimates for Symmetric Stable Distributions. Journal of the American Statistical Association 66, 331–338.

Fang, K.-T., Kotz, S., Ng, K. W., 1990. Symmetric Multivariate and Related Distributions. CRC Press.

Gabaix, X., 1999a. Zipf's Law and the Growth of Cities. The American Economic Review 89, 129–132.

Gabaix, X., 1999b. Zipf's Law for Cities: An Explanation. The Quarterly Journal of Economics 114, 739–767.

Gabaix, X., 2009. Power Laws in Economics and Finance. Annual Review of Economics 1, 255–294.

Gabaix, X., 2011. The Granular Origins of Aggregate Fluctuations. Econometrica 79, 733–772.

Gabaix, X., Gopikrishnan, P., Plerou, V., Stanley, H. E., 2006. Institutional Investors and Stock Market Volatility. The Quarterly Journal of Economics 121, 461–504.

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

Gubner, J., 2006. A New Formula for Lognormal Characteristic Functions. IEEE Transactions on Vehicular Technology 55, 1668–1671.

Gulisashvili, A., Tankov, P., 2016. Tail behavior of sums and differences of log-normal random variables. Bernoulli 22, 444–493.

Harvey, C. R., Liu, Y., Zhu, H., 2016. … and the Cross-Section of Expected Returns. Review of Financial Studies 29, 5–68.

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

Hopenhayn, H. A., 1992. Entry, Exit, and firm Dynamics in Long Run Equilibrium. Econometrica 60, 1127–1150.

Jaimovich, N., Terry, S., Vincent, N., 2023. The Empirical Distribution of Firm Dynamics and Its Macro Implications. NBER Working Paper .

Jovanovic, B., 1982. Selection and the Evolution of Industry. Econometrica 50, 649–670.

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

Kondo, I. O., Lewis, L. T., Stella, A., 2018. On the U.S. Firm and Establishment Size Distributions. Mimeo .

Lo, C. F., 2012. The Sum and Difference of Two Lognormal Random Variables. Journal of Applied Mathematics 2012, 1–13.

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

Luttmer, E. G. J., 2007. Selection, Growth, and the Size Distribution of Firms. The Quarterly Journal of Economics 122, 1103–1144.

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.

Miller, M. B., 2013. Mathematics and Statistics for Financial Risk Management. Wiley.

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.

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

Parham, R., 2023. Why Are Firm Growth Distributions Heavy-Tailed? Mimeo .

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.

Piketty, T., Saez, E., Stantcheva, S., 2014. Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. American Economic Journal: Economic Policy 6, 230–271.

Roy, A. D., 1950. The Distribution of Earnings and of Individual Output. The Economic Journal 60, 489–505.

Roy, A. D., 1951. Some Thoughts on the Distribution of Earnings. Oxford Economic Papers 3, 135–146.

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

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

Toda, A. A., 2012. The double power law in income distribution: Explanations and evidence. Journal of Economic Behavior & Organization 84, 364–381.

Toda, A. A., Walsh, K., 2015. The Double Power Law in Consumption and Implications for Testing Euler Equations. Journal of Political Economy 123, 1177–1200.

Varian, H. R., 1992. Microeconomic Analysis. Norton, New York, third ed.

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

Yeh, C., 2023. Revisiting the Origins of Business Cycles with the Size-Variance Relationship. The Review of Economics and Statistics pp. 1–28.

OA Online Appendix

This appendix fully characterizes the $\mathcal{DLN}$ distribution, deriving its Cumulative Distribution Function (CDF) and Probability Density Function (PDF). It further derives the moments of a $\mathcal{DLN}$-distributed Random Variable (RV) and discusses a procedure for parameter estimation given data using Maximum Likelihood Estimation (MLE). Finally, it discusses how to generalize the uni-variate $\mathcal{DLN}$ distribution to a multi-variate version using elliptical distribution theory. To my knowledge, it is the first such treatment of the distribution in the literature. A complete suite of computer code implementing these procedures for working with $\mathcal{DLN}$ distributions is provided as well.

We define the basic structure of a $\mathcal{DLN}$ RV,

$$W = Y_p - Y_n = \exp(X_p) - \exp(X_n) \quad \text{with} \;\; X = (X_p, X_n)^T \sim \mathcal{N}(\mu, \Sigma)$$

and denote $W \sim \mathcal{DLN}(\mu_p, \sigma_p, \mu_n, \sigma_n, \rho_{pn})$. The next section begins by characterizing its PDF and CDF, as well as discussing the simplified case when $\rho_{pn} = 0$ and the PDF can be derived as a simple convolution using a Fourier transform.

OA.1 PDF and CDF

The PDF for the Bi-Variate Normal (BVN) RV X is well-known to be

$$f_{BVN}(x) = \frac{|\Sigma|^{-\frac{1}{2}}}{2\pi} \cdot \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) = \frac{|\Sigma|^{-\frac{1}{2}}}{2\pi} \cdot \exp\left(-\frac{1}{2}||x-\mu||_\Sigma\right) \quad \text{(OA.1)}$$

with $|\Sigma|$ the determinant of Σ and $||x||_\Sigma$ the Euclidean norm of x under the Mahalanobis distance induced by Σ.

The PDF for a Bi-Variate Log-Normal (BVLN) RV $Y = (Y_p, Y_n)^T$ can be obtained by using the multivariate change of variables theorem. If $Y = g(X)$ then

$$f_Y(y) = f_X(g^{-1}(y)) \cdot ||J_{g^{-1}}(y)|| \tag{OA.2}$$

with $J_{g^{-1}}$ the Jacobian matrix of $g^{-1}(\cdot)$ and $||J_{g^{-1}}||$ the absolute value of its determinant. Applying the theorem for $Y = g(X) = (\exp(X_p), \exp(X_n))^T$ we have $g^{-1}(y) = (log(y_p), log(y_n))^T$ and $||J_{g^{-1}}(y)|| = (y_p \cdot y_n)^{-1}$. The PDF of a BVLN RV is then

$$f_{BVLN}(y) = \frac{|\Sigma|^{-\frac{1}{2}}}{2\pi y_p y_n} \exp\left(-\frac{1}{2}||\log(y) - \mu||_\Sigma\right) \tag{OA.3}$$

We can now define the cumulative distribution function (CDF) of the $\mathcal{DLN}$ distribution using the definition of the CDF of the difference of two RV

$$\begin{aligned} F_{DLN}(w) &= P[W \le w] = P[y_p - y_n \le w] = P[y_p \le y_n + w] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{y_n + w} f_{BVLN}(y_p, y_n) dy_p dy_n \end{aligned} \tag{OA.4}$$

which can be differentiated w.r.t w to yield the PDF

$$f_{DLN}(w) = \int_{-\infty}^{\infty} f_{BVLN}(y + w, y) dy = \int_{-\infty}^{\infty} f_{BVLN}(y, y - w) dy \tag{OA.5}$$

but because $f_{BVLN}(y)$ is non-zero only for $y > 0$, we limit the integration range

$$f_{DLN}(w) = \int_{\max(0,w)}^{\infty} f_{BVLN}(y, y - w) dy \tag{OA.6}$$

which yields the PDF of the $\mathcal{DLN}$ distribution.

It is well-known, however, that the integral in equation OA.6 does not have a closed-form solution. The accompanying code suite evaluates it numerically, and also numerically evaluates the CDF using its definition

$$F_{DLN}(w) = \int_{-\infty}^{w} f_{DLN}(y)dy \tag{OA.7}$$

For the simpler case with difference of uncorrelated log-Normals, i.e. $\rho_{pn} = 0$, we can derive the PDF of the $\mathcal{DLN}$ via a characteristic function (CF) approach as well. In this case, we can write the CF of the $\mathcal{DLN}$ as $\varphi_{DLN}(t) = \varphi_{LN}(t) \cdot \varphi_{LN}(-t)$ with $\varphi_{LN}(t)$ the CF of the log-Normal. Next, we can apply a Fourier transform to obtain the PDF,

$$f_{DLN}(w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i \cdot t \cdot w} \cdot \varphi_{DLN}(t)dt \tag{OA.8}$$

Unfortunately, the log-Normal does not admit an analytical CF, and using Equation OA.8 requires a numerical approximation for $\varphi_{LN}(t)$ as well. Gubner (2006) provides a fast and accurate approximation method for the CF of the log-Normal which I use in the calculation of $f_{DLN}(w)$ when using this method.

OA.2 Moments

OA.2.1 MGF

The moment generating function (MGF) of the $\mathcal{DLN}$ can be written as

$$M_W(t) = \mathbb{E}\left[e^{tW}\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{tw} f_{BVLN}(y + w, y)dydw \tag{OA.9}$$

but this formulation has limited usability due to the lack of closed-form solution for the integrals. Instead, it is useful to characterize the moments directly, as we can obtain them in closed-form.

OA.2.2 Mean and variance

Using the definitions of μ and Σ for the Bi-Variate Normal, define the mean and covariance of the Bi-Variate Log-Normal RV, $\hat{\mu}$ and $\hat{\Sigma}$ (element-wise) as

$$\begin{aligned}\hat{\mu}_{(i)} &= \exp\left(\mu_{(i)} + \frac{1}{2}\Sigma_{(i,i)}\right)\\ \hat{\Sigma}_{(i,j)} &= \exp\left(\mu_{(i)} + \mu_{(j)} + \frac{1}{2}\left(\Sigma_{(i,i)} + \Sigma_{(j,j)}\right)\right) \cdot \left(\exp\left(\Sigma_{(i,j)}\right) - 1\right)\end{aligned} \tag{OA.10}$$

Note that if Σ is diagonal (i.e., X_p and X_n are uncorrelated) then $\hat{\Sigma}$ will be diagonal as well. We are however interested in the general form of the $\mathcal{DLN}$ distribution. The identities regarding the expectation and variance of a sum of RV yield

$$\mathbb{E}[W] = \mathbb{E}[Y_p] - \mathbb{E}[Y_n] = \hat{\mu}_{(1)} - \hat{\mu}_{(2)} = \exp(\mu_p + \frac{\sigma_p^2}{2}) - \exp(\mu_n + \frac{\sigma_n^2}{2}) \tag{OA.11}$$

and

$$\begin{aligned}\mathrm{Var}[W] &= \mathbb{C}[Y_p, Y_p] + \mathbb{C}[Y_n, Y_n] - 2 \cdot \mathbb{C}[Y_p, Y_n] = \hat{\Sigma}_{(1,1)} + \hat{\Sigma}_{(2,2)} - 2 \cdot \hat{\Sigma}_{(1,2)}\\ &= \exp\left(2\mu_p + \sigma_p^2\right) \cdot \left(exp\left(\sigma_p^2\right) - 1\right) + \exp\left(2\mu_n + \sigma_n^2\right) \cdot \left(exp\left(\sigma_n^2\right) - 1\right)\\ &\quad - 2\exp\left(\mu_p + \mu_n + \frac{1}{2}(\sigma_p^2 + \sigma_n^2)\right) \cdot \left(\exp\left(\sigma_p \sigma_n \rho_{pn}\right) - 1\right)\end{aligned} \tag{OA.12}$$

with $\mathbb{C}$ the covariance operator of two general RV U_1, U_2

$$\mathbb{C}[U_1, U_2] = \mathbb{E}[(U_1 - \mu_1)(U_2 - \mu_2)] \tag{OA.13}$$

OA.2.3 Skewness and kurtosis

Skewness and kurtosis of the $\mathcal{DLN}$ can similarly be established using coskewness and cokurtosis — see e.g. Miller (2013) for an overview. Coskewness of three general RV

U_1, U_2, U_3 is defined as

$$\mathbb{S}\left[U_1, U_2, U_3\right] = \frac{\mathbb{E}\left[(U_1-\mu_1)(U_2-\mu_2)(U_3-\mu_3)\right]}{\sigma_1\sigma_2\sigma_3} \tag{OA.14}$$

and cokurtosis of four general RV U_1, U_2, U_3, U_4 is defined as

$$\mathbb{K}\left[U_1, U_2, U_3, U_4\right] = \frac{\mathbb{E}\left[(U_1-\mu_1)(U_2-\mu_2)(U_3-\mu_3)(U_4-\mu_4)\right]}{\sigma_1\sigma_2\sigma_3\sigma_4} \tag{OA.15}$$

with the property that $\mathbb{S}\left[U,U,U\right] = \text{Skew}\left[U\right]$ and $\mathbb{K}\left[U,U,U,U\right] = \text{Kurt}\left[U\right]$. More importantly, it is simple to show that

$$\text{Skew}\left[U-V\right] = \frac{\sigma_U^3\mathbb{S}\left[U,U,U\right] - 3\sigma_U^2\sigma_V\mathbb{S}\left[U,U,V\right] + 3\sigma_U\sigma_V^2\mathbb{S}\left[U,V,V\right] - \sigma_V^3\mathbb{S}\left[V,V,V\right]}{\sigma_{U-V}^3} \tag{OA.16}$$

and similarly

$$\begin{aligned}\text{Kurt}\left[U-V\right] = \frac{1}{\sigma_{U-V}^4}[&\sigma_U^4\mathbb{K}\left[U,U,U,U\right] - 4\sigma_U^3\sigma_V\mathbb{K}\left[U,U,U,V\right] \\ &+ 6\sigma_U^2\sigma_V^2\mathbb{K}\left[U,U,V,V\right] - 4\sigma_U\sigma_V^3\mathbb{K}\left[U,V,V,V\right] + \sigma_V^4\mathbb{K}\left[V,V,V,V\right]]\end{aligned} \tag{OA.17}$$

with $\sigma_{U-V} = \text{Var}\left[U-V\right]^{\frac{1}{2}}$ calculated using Equation OA.12. Evaluating the operators $\mathbb{S}$ and $\mathbb{K}$ for the case of $\mathcal{DLN}$ requires evaluating expressions of the general form $\mathbb{E}\left[Y_p^i Y_n^j\right]$, which can be done via the MGF of the BVN distribution

$$\mathbb{E}\left[Y_p^i Y_n^j\right] = \mathbb{E}\left[e^{iX_p}e^{jX_n}\right] = \text{MGF}_{BVN}\left(\left[\begin{smallmatrix} i \\ j \end{smallmatrix}\right]\right) = \mathbb{E}\left[Y_p^i\right]\mathbb{E}\left[Y_n^j\right]e^{ij\Sigma_{(1,2)}} \tag{OA.18}$$

with $\mathbb{E}\left[Y_p^i\right] = \exp\left(i\mu_p + \frac{1}{2}i^2\sigma_p^2\right)$. This concludes the technical details of the derivation.

The method presented can be extended to higher central moments as well. The accompanying code suite includes functions that implement the equations above and use them to calculate the first five moments of the $\mathcal{DLN}$ given the parameters $(\mu_p, \sigma_p, \mu_n, \sigma_n, \rho_{pn})$.

OA.3 Estimation

Given data $D \sim \mathcal{DLN}(\Theta)$ with $\Theta = (\mu_p, \sigma_p, \mu_n, \sigma_n, \rho_{pn})$, we would like to find an estimate $\hat{\Theta}$ to the parameter vector Θ. Experiments show that given an appropriate initial guess, the MLE estimates of Θ perform well in practice. The main parameter of difficulty is ρ_{pn}. This parameter is akin to the shape parameter in the Stable distribution, which plays a similar role and is similarly difficult to estimate, see e.g. Fama and Roll (1971). It hence requires special care in the estimation.

The estimation code provided minimizes the negative log-likelihood of the data w.r.t the $\mathcal{DLN}$ PDF using a multi-start algorithm. The starting values for the first four parameters are fixed for all start points as:

$$\begin{bmatrix} \mu_p \\ \sigma_p \\ \mu_n \\ \sigma_n \end{bmatrix} = \begin{bmatrix} \text{Median}\left[\log\left(D\right)\right] \quad \text{for} \quad D > 0 \\ \text{IQR}\left[\log\left(D\right)\right]/1.35 \quad \text{for} \quad D > 0 \\ \text{Median}\left[\log\left(-D\right)\right] \quad \text{for} \quad D < 0 \\ \text{IQR}\left[\log\left(-D\right)\right]/1.35 \quad \text{for} \quad D < 0 \end{bmatrix} \tag{OA.19}$$

while the initial guesses for ρ_{pn} are $(-0.8, -0.3, 0, 0.3, 0.8)$. The estimator $\hat{\Theta}$ is then the value which minimizes the negative log-likelihood in the multi-start algorithm. The estimator inherits asymptotic normality, consistency, and efficiency properties from the general M-estimator theory, as the dimension of $\hat{\Theta}$ is fixed, the likelihood is smooth, and is supported on $\mathbb{R}\ \forall\hat{\Theta}$. A better estimation procedure for the parameters of the $\mathcal{DLN}$ might be merited, but is left for future work.

OA.4 The elliptical multi-variate $\mathcal{DLN}$

Some practical applications of the $\mathcal{DLN}$ require the ability to work with multi-variate DLN RVs. I hence present an extension of the $\mathcal{DLN}$ to the multi-variate case using elliptical distribution theory, with the standard reference being Fang, Kotz, and Ng (1990).

The method of elliptical distributions requires a symmetric baseline distribution. We will therefore focus our attention on the symmetric $\mathcal{DLN}$ case in which $\mu_p = \mu_n \equiv \mu$ and $\sigma_p = \sigma_n \equiv \sigma$, yielding the three parameter uni-variate symmetric distribution $\text{Sym}\mathcal{DLN}(\mu, \sigma, \rho) = \mathcal{DLN}(\mu, \sigma, \mu, \sigma, \rho)$. I begin by defining a standardized N-dimensional elliptical $\mathcal{DLN}$ RV using Sym$\mathcal{DLN}$ and the spherical decomposition of Cambanis, Huang, and Simons (1981), and later extend it to a location-scale family of distributions.

Let U be an N-dimensional RV distributed uniformly on the unit hyper-sphere in $\mathbb{R}^N$ and arranged as a column vector. Let $R \geq 0$ be a univariate RV independent of U with PDF $f_R(r)$ to be derived momentarily, and let $Z = R \cdot U$ be a standardized N-dimensional elliptical $\mathcal{DLN}$ RV. A common choice for U is $\widehat{U}/||\widehat{U}||_2$ with $\widehat{U} \sim MVN(0_N, 1_N)$. U captures a direction in $\mathbb{R}^N$, and we have $\sqrt{U^T \cdot U} = ||U||_2 \equiv 1$, which implies $\sqrt{Z^T \cdot Z} = ||Z||_2 = R$. We further know that the surface area of an N-sphere with radius R is given by

$$S_N(R) = \frac{2 \cdot \pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)} \cdot R^{N-1} \tag{OA.20}$$

and can hence write the PDF of Z as

$$f_Z(z) = \frac{f_R(||z||_2)}{S_N(||z||_2)} = \frac{\Gamma\left(\frac{N}{2}\right) \cdot f_R(||z||_2)}{2 \cdot \pi^{\frac{N}{2}} \cdot ||z||_2^{N-1}} \tag{OA.21}$$

We require $f_R(r)$ and $f_Z(z)$ to be valid PDFs, which yields the conditions

$$\begin{aligned} & f_R(r) \geq 0 \;\forall\; r \in \mathbb{R} \\ & f_Z(z) \geq 0 \;\forall\; z \in \mathbb{R}^N \\ & \int_{-\infty}^{\infty} f_R(r)\; dr = 1 \\ & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_Z(z)\; dz_{(N)} \cdots dz_{(1)} = 1 \end{aligned} \tag{OA.22}$$

to those, we can add the condition that the properly normalized distribution of $f_R(r)$ will be Sym$\mathcal{DLN}$,

$$f_R(r) = \widetilde{M}_N(r) \cdot f_{DLN}(r) \tag{OA.23}$$

with $\widetilde{M}_N(r)$ chosen such that the conditions in Equation OA.22 hold. Solving for this set of conditions yields

$$f_R(r) = \frac{r^{N-1}}{\int_0^\infty \widetilde{r}^{N-1} \cdot f_{DLN}(\widetilde{r})\ d\widetilde{r}} \cdot f_{DLN}(r) \tag{OA.24}$$

and

$$f_Z(z) = \frac{\Gamma\left(\frac{N}{2}\right)}{2 \cdot \pi^{\frac{N}{2}} \cdot \int_0^\infty \widetilde{r}^{N-1} \cdot f_{DLN}(\widetilde{r})\ d\widetilde{r}} \cdot f_{DLN}(||z||_2) = M_N \cdot f_{DLN}(||z||_2) \tag{OA.25}$$

with M_N a normalization constant depending only on the dimension N and the parameters of the baseline Sym$\mathcal{DLN}(\mu, \sigma, \rho)$ being used. We can use Z's CDF definition to write

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{z_{(1)}} \cdots \int_{-\infty}^{z_{(N)}} f_Z(\widehat{z})\ d\widehat{z}_{(N)} \cdots d\widehat{z}_{(1)} \\ &= \int_{-\infty}^{z_{(1)}} \cdots \int_{-\infty}^{z_{(N)}} M_N \cdot f_{DLN}(||z||_2)\ d\widehat{z}_{(N)} \cdots d\widehat{z}_{(1)} \end{aligned} \tag{OA.26}$$

which concludes the characterization of the standardized N-dimensional elliptical $\mathcal{DLN}$ RV.

Extending the standardized N-dimensional $\mathcal{DLN}$ to a location-scale family of distributions is now straightforward. Let $\widetilde{\mu} = (\mu_1, \mu_2, ..., \mu_N)^T$ be a column vector of locations and let $\widetilde{\Sigma}$ be a positive-semidefinite scaling matrix of rank N. Define

$$W = \widetilde{\mu} + \widetilde{\Sigma}^{\frac{1}{2}} \cdot Z \tag{OA.27}$$

with $\widetilde{\Sigma}^{\frac{1}{2}}$ denoting the eigendecomposition of $\widetilde{\Sigma}$. The PDF of W is then given by

$$\begin{aligned} f_W(w) &= |\widetilde{\Sigma}|^{-\frac{1}{2}} \cdot f_Z\left(\widetilde{\Sigma}^{-\frac{1}{2}} \cdot (w - \widetilde{\mu})\right) \\ &= |\widetilde{\Sigma}|^{-\frac{1}{2}} \cdot M_N \cdot f_{DLN}\left(\sqrt{(w - \widetilde{\mu})^T \cdot \widetilde{\Sigma}^{-1} \cdot (w - \widetilde{\mu})}\right) \\ &= |\widetilde{\Sigma}|^{-\frac{1}{2}} \cdot M_N \cdot f_{DLN}\left(||w - \widetilde{\mu}||_{\widetilde{\Sigma}}\right) \end{aligned} \tag{OA.28}$$

The CDF of W can similarly be written as

$$F_W(w) = |\widetilde{\Sigma}|^{-\frac{1}{2}} \cdot M_N \cdot \int_{-\infty}^{w_{(1)}} \cdots \int_{-\infty}^{w_{(N)}} f_{DLN}\left(||w - \widetilde{\mu}||_{\widetilde{\Sigma}}\right) \, d\widehat{w}_{(N)} \cdots d\widehat{w}_{(1)} \tag{OA.29}$$

which characterizes a general elliptical multi-variate $\mathcal{DLN}$ RV.

Finally, note that the scaling matrix $\widetilde{\Sigma}$ is not the covariance matrix of W due to the heavy-tails of W, similar to other heavy-tailed elliptical distributions such as the multi-variate Stable, t, or Laplace distributions. Further note that the normalization integral in Equation OA.24 is numerically unstable for high values of N (e.g., $N \geq 5$), and care should be taken when deriving the PDF of high-dimensional $\mathcal{DLN}$ RVs in practice.