Scaling laws and a general theory for the growth of companies

Abstract
Companies are fundamental units of contemporary economies and markets and are important mechanisms through which humans interact with their environments. Understanding general properties that underlie the processes of growth in companies have long been of interest, yet fundamental debates about the effects of firm size on growth have persisted. Here we develop a scaling framework that focuses on company size as the critical feature determining a variety of tradeoffs, and use this to reveal novel systematic behavior across the diversity of publicly-traded companies. Using a large database of 31,553 companies over more than 70 years, we show how the dynamics of companies expressed as scaling relationships leads to a quantitative, predictive theory for their growth. We find that companies exhibit size-dependent changes in their financial composition. Most notably net income scales sublinearly with the assets of a company (i.e., size), while liabilities scale linearly. From these scaling relationships we derive an equation for the size-dependent growth of companies where, surprisingly and nontrivially, assets grow as a power law in time. These results illustrate that while companies are part of a larger class of growth phenomena driven by incomes and costs that scale with size, they are unique in that they grow without bound following a temporal power-law. This temporal growth sets companies apart from the scaling of organisms, and from other institutions, such as cities, nations, and markets, where growth over time is often exponential. The perspective we develop here highlights novel dynamics in the scaling of human economies.

1 Introduction
Companies are major drivers of economic growth, employment, and technological innovation and the markets they form are the primary mechanism through which resources, services, and wealth are generated and redistributed at all scales of society [1][2][3][4]. The total wealth created by the largest companies is often larger than many nations. For example, the total market capitalization of all firms currently on the U.S. stock market is over $49 trillion, which is more than twice the entire gross domestic product (GDP) of the USA, while the sales of Walmart and Amazon both exceed the GDP of Singapore [5][6]. Consequently, a quantitative understanding of the laws that govern the dynamics of companies is central to understanding the dynamics of the modern world and addressing its long-term sustainability.

From an ecological perspective, companies are individual entities that interact and compete with each other for finite resources within complex environments. From this perspective individual companies are complex adaptive systems governed by the same types of laws, rules, and dynamics governing individual organisms and their collectives ranging
from bacteria to hunter-gatherer groups and cities. Many of the most successful laws discovered for complex adaptive systems concern the regular patterns exhibited by large ensembles of entities. Indeed, companies have already been shown to display some of the statistical signatures of similar types of laws such as an exponential distribution of lifespans [7], and Zipfian rank-abundance distribution of sizes with Laplacian fluctuations (e.g. [8, 9, 10]).

Much like organisms and cities, the continuing challenge of adaptability, evolvability and growth in response to competitive forces requires companies to be scalable. The same generic underlying dynamical and organizational principles operate across multiple spatio-temporal scales leading to emergent laws manifested as empirical scaling laws. Scaling as the consequence of underlying dynamics and structure has been instrumental in illuminating universal principles and gaining deeper insights across the entire spectrum of science and technology from the unification of the fundamental forces of nature, to the building of airplanes and computers, the nature of tipping points, and the structure, dynamics and growth of organisms, ecosystems and cities [11].

Despite their extraordinary complexity and diversity, many of the most fundamental characteristics of organisms, such as metabolic rates, growth rates, lifespans and genome lengths, scale with size in a surprisingly simple and universal fashion ranging from cells to ecosystems [12]. Similarly, many characteristics of cities, including wages, patents, crime, and road lengths, scale systematically with size [13]. These scaling laws are a consequence of the generic organizational and dynamical properties of the underlying optimized networks that sustain these systems, such as circulatory systems, transport systems, and social networks, and lead to a quantitative, mechanistic, framework that captures many of their essential features.

Given the success of this framework, it is surprising that no such systematic analysis has as yet been performed for companies. Here we address this issue by asking to what extent, and in what sense, are small, medium, and large companies scaled versions of one another? Using data from over 30,000 US publicly traded companies, we show that the multiple metrics and indices characterising companies do indeed scale in a systematic way obeying, on average, simple non-linear power laws indicating that they are approximately self-similar and scale invariant.

The mechanisms traditionally suggested for understanding companies fall into three broad inter-related categories that are often treated separately. In the language developed for understanding organisms and cities, these are: (i) Organizational structure, which is the network system that conveys information, resources and capital to support, sustain and grow the company; (ii) Minimization of transaction costs to enhance economies of scale and maximize profits; (iii) Competitive forces inherent in the ecology of the market place. Properties such as these likely underlie the origin of the scaling laws [1, 2, 3, 4].

Historically, perhaps the most prominent theory of company growth is Gibrat’s hypothesis of the law of proportional effect [14]. This proposes that firm growth is dominated by random fluctuations so the distribution of company sizes converges to lognormal [15], implying that their proportional growth rate is independent of size. Empirically, however, this is not supported by data [16]. Among the many extensions of Gibrat’s statistical hypothesis are a class of models that assume that companies are composed of sub-units that grow according to a generalized preferential attachment model [17, 18, 19] that can account for the tent-like fluctuations in growth rates [10].

In contrast to these statistical approaches to growth, our framework focuses on the mechanisms that drive it, namely, the flows and utilisation of capital derived from profits and financing flowing through the company, which are governed by the scaling laws. We show how these lead to a universal equation for company growth whose predictions are in good quantitative agreement with data.

## 2 Theoretical Framework and Results

### Cash flows, stores, and the growth of companies

The size and growth of a company is ultimately determined by the flows and stores of money through the system. This is illustrated in Fig. 1 which shows the inter-relationship between the various quantities conventionally used to characterize the financial state of a company. We classify these variables into three categories: 1) income related quantities, which include sales, net income, and gross profit; 2) cost related quantities, which include cost of sales, total taxes, operating costs, and R&D; and 3) other size-related quantities, such as the number of employees, assets, liability, and the availability of cash. The incoming flows consist of two parts, sales and finance, whereas the outgoing flows represent expenses. We identify the amount of financial resources stored at a given time as a company’s assets which we use as the measure for its size.
Scaling the income, cost, and size of companies

Companies are, on average, self-similar and scale-invariant if, at any time \( t \), they obey power law scaling laws for their various measurable properties, \( X \), such as their sales, net income, profit or total liability. This is expressed mathematically as:

\[ X(A) = c_X A^{\beta_X} \]  

(1)

where \( A \) are their assets, \( c_X \) a normalization constant and \( \beta_X \) a scaling exponent; both \( c_X \) and \( \beta_X \) are time and scale invariant, but can depend on the sector or market in which companies operate.

We test these ideas and estimate the suite of parameters \( c_X \) and \( \beta_X \) using a dataset of 31,553 publicly-traded companies in US markets covering 1975-2018. Taking the logarithm of Eq. (1) yields the linear form \( \ln X = \ln c_X + \beta_X \ln A \), so a plot of \( \ln X \) vs. \( \ln A \) should yield a sequence of straight lines whose slopes are the exponents \( \beta_X \). As can be readily seen in Fig. 2, all income and cost variables do indeed exhibit significant power law scaling, albeit with fluctuations that are discussed in detail below.

Table I summarises the scaling exponents, \( \beta_X \), for selected variables averaged across the entire market as well as for various individual economic sectors. While there is some variation in their values, they are almost all sub-linear (that is, less than 1) indicating that, regardless of the sector, assets grow systematically faster than any company metric including sales, cost of sales and net income, as shown in Fig. 2. A notable exception is liability for which exponents converge to 1 indicating that it scales linearly with company size. The inset panels show the distributions of exponents across different sectors; the few outlying ones are likely due to sectors with only a small sampling of firms (see section S.1 in SI for details). These regularities and the close clustering of exponents for a given variable across different sectors suggests that similar underlying dynamics and network organizational principles are at play across companies regardless of their size, age or economic sector. An important observation that plays a crucial role in determining their growth trajectory is that income-related variables have comparatively larger exponents than cost-related ones.

Growth Equation

The detailed growth trajectory of any individual company is a complex response of its particular internal structure and the external conditions it confronts. Fig. 3a) shows the diversity of these trajectories for the 31,553 publicly-traded companies in US markets since 1975. Larger, more mature companies tend to grow slowly following the average growth of the market, whereas smaller, younger ones tend to grow more rapidly but with greater fluctuations. Nevertheless, underlying the stochasticity of these individual growth trajectories are, at any time \( t \), size-dependent regularities captured by scaling laws. We show how these time invariant constraints, reflecting the common underlying dynamics shared by all companies, can be used to derive an explicit growth equation that determines their generic growth trajectories. Indeed, we show that the wide diversity of these trajectories can be reduced to a single size-dependent universal growth process.

If \( S(t) \) is the total annual sales of a company and \( F(t) \) the total annual finances it raises from the market (see SI S3.1), then the total amount of money flowing into the company during a time interval \( \Delta t \) is \( [S(t) + F(t)] \Delta t \). As illustrated in Fig. 1 this total flux of incoming money is partially used to maintain the company by offsetting its total expenses, \( E(t) \Delta t \), during this time interval, where \( E(t) \) is the total annual expenses which include wages, taxes, cost of sales and equipment, etc. The remainder fuels the company’s growth by contributing to an increase in its assets, \( \Delta A(t) \); consequently, \( [S(t) + F(t)] \Delta t = E(t) \Delta t + \Delta A(t) \), so

\[ \Delta A(t) = [S(t) + F(t) - E(t)] \Delta t \]  

(2)

Finances raised from the market are the dominant contribution to the increase in a company’s total liability, \( L(t) \), thus, \( F(t) \Delta t \approx \Delta L(t) \)(see SI S4.1). Furthermore, the difference between sales \( S(t) \) and total expenses \( E(t) \) is net income \( I(t) \), i.e., \( S(t) - E(t) = I(t) \), so Eq. (2) can be re-expressed as \( \Delta A(t) = I(t) \Delta t + \Delta L(t) \). Taking the limit \( \Delta t \to 0 \) gives

\[ \frac{dA}{dt} = I + \frac{dL}{dt} \]  

(3)

This balance equation is exact at any time, \( t \). However, quantities such as sales, expenses and income, are not reported instantaneously; instead, they are reported annually, effectively making time discrete rather than continuous with one year being the minimum time unit. Consequently, when using results based on Eq. (3) in which \( \Delta t \to 0 \), and comparing them with data which are reported only annually, we inevitably introduce some unknown, though relatively small, error (see SI S4.1 for the details).
With scaling laws in mind and noting that \( \frac{dL}{dA} = \left( \frac{dL}{dA} \right) \frac{dA}{dt} \), it is prudent to consider \( I \) and \( L \) as functions of company size, \( A(t) \), in which case Eq. (3) can be re-expressed as
\[
\frac{dA}{dt} = \frac{I(A)}{1 - \frac{dL(A)}{dA}}.
\] (4)

Consequently, there are two separate strategies for ensuring positive growth (i.e., \( dA/dt > 0 \)):

- (i) having positive income \( (I > 0) \) coupled with the differential debt ratio \( dL/dA < 1 \);
- (ii) having negative income \( (I < 0) \) coupled with the differential debt ratio \( dL/dA > 1 \).

Note that, in general, it is the differential debt ratio, \( dL/dA \), rather than the conventional debt ratio, \( k(A) = L/A \), that is a determining factor for positive growth. In what follows we restrict the discussion and corresponding data to cases where net income is positive since these dominate the market; furthermore, on average, the inclusion of short-term losses does not appreciably affect long-term development of companies, nor the value of the exponents. We relegate the discussion of negative income to section S2.4 of the SI.

Introducing the scaling relationships \( I(A) = c_I A^{\beta_I} \) and \( L(A) = c_L A^{\beta_L} \) into Eq. (4) leads to
\[
\frac{dA}{dt} = \frac{c_I A^{\beta_I}}{1 - c_L A^{\beta_L - 1}}.
\] (5)

This can be straightforwardly integrated to give
\[
\frac{A^{1-\beta_I}}{c_I(1-\beta_I)} \left[ 1 - \left( \frac{1-\beta_I}{\beta_L - \beta_I} \right) c_L \beta_L A^{\beta_L - 1} \right] \equiv f(A) = t
\]
where we have imposed the initial condition \( A(0) = 0 \) at \( t = 0 \) (see S3.3 of SI for the complete expression, and S3.2 for the solution when \( A(0) \neq 0 \)). In general, it is not possible to invert Eq. (6) and obtain an analytic expression for \( A(t) = f^{-1}(t) \). However, we saw in Fig. 2 and Table 1 that, averaged across the market, \( \beta_L \to 1 \) (i.e., liability is a constant fraction of assets), in which case Eq. (6) reduces to a simple tractable analytic solution for the averaged growth trajectory of companies:
\[
A(t) = \left[ \frac{c_I(1-\beta_I)}{1-c_L} t \right]^{1/(1-\beta_I)}.
\] (7)

More generally, the exact solution to Eq. (6) for asymptotically large \( t \) and \( A(t) \), i.e., for large mature companies, is given by \( A(t) = [c_I(1-\beta_I)t]^{1/(1-\beta_I)} \), provided \( \beta_L < 1 \), a condition satisfied for most sectors.

To a good approximation, therefore, the theory predicts that companies, on average, grow following a simple power law in time:
\[
A(t) \approx c\tau^\gamma
\] (8)

with both \( c \) and \( \gamma \) predicted by the theory. We refer to Eq. (5) as the universal growth equation and Eq. (8) as the universal growth curve with \( t \) being the universal age. Since \( \beta_I \) is close to 1 (\( \approx 0.85 \)), the leading order solution, Eq. (7), predicts that \( A(t) \) grows rapidly with a relatively large exponent, \( \gamma = 1/(1-\beta_I) \approx 6.7 \), though not as fast as an exponential. It is straightforward to calculate the leading corrections to Eq. (7) by expanding \( A(t) \) in Eq. (6) in a Taylor series around \( \beta_L = 1 \). The predominant behaviour is still a power law as in Eq. (8), but with a logarithmic correction leading to deviations for smaller companies; the explicit formulae and details are presented in SI S3.4.

The refined prediction gives \( \gamma \approx 5.8 \).

The power law relationship (Eq. 8) is confirmed by the data as shown in Fig. 3(b) in terms of a semi-logarithmic plot, in which case the leading order prediction is a logarithmic curve (the solid line). Shown in the inset of Fig. 3(b) is a log-log plot of the data, in which case the leading order prediction is a straight line with the slope \( \gamma \approx 5.8 \); importantly, this plot also confirms the deviation from the strict power predicted by the full theory for smaller companies, Eq. (6). Compared to the large diversity of individual company growth trajectories across all sectors shown in Fig. 3(a) the collapse of the data to a single universal growth curve is illustrated by Fig. 3(b).

In confronting the universal growth model from Eq. (8) with data it is important to note that there are many companies whose first year of operation predates the reported time period. To deal with this we can infer an effective initial year, \( t^*_0 \), that best aligns the data with the universal curve. This is done by solving for the initial time as \( t^*_0 = \left( \frac{A^*_0}{c} \right)^{1/\gamma} \) using the initial assets, \( A^*_0 \), and then plotting the universal growth curve as
\[
\hat{A}^\tau = c \cdot \left( \tau + t^*_0 \right)^\gamma.
\] (9)

Here, \( \tau \) is the years we observe the company in the data starting from the inferred age of \( t^*_0 \), and \( \hat{A}^\tau \) is the predicted total assets of a company in year \( \tau \). Fits of \( t^*_0 \) for each individual company reveal an impressive universal curve in
Figs. 3(b). We also validate the method for inferring \( t_0 \) by a group of companies with randomly generated initial startup year in S4.2 of SI.

A more direct test of our theory is to solve Eq. (5) using the observed initial size of each individual company (see S3.2 of SI) in which case we find an impressively good fit to the data without employing any free parameters other than the cross-company scaling exponents and constants. Fig. 3(c) shows four examples for comparison between theoretical prediction and real growth trajectories of individual companies. Three show excellent agreement between prediction and data, while the fourth example is an outlier showing over-performance of a company (Apple) relative to the idealised prediction. Notice, however, that this trajectory still approximates a power law, albeit with a higher exponent. In general, Fig. 3(d) shows good agreement between the predicted value of the assets compared to their actual values for all companies in all years. This is equivalent to the universal growth curve because, as discussed in S3.3 of the SI, ideally each individual curve can be regarded as a part of the universal growth curve.

Analogous universal growth curves can be derived for individual sectors as shown in Fig. 4. Note, however, that for the relatively few sectors where \( \beta_L \) deviates significantly from 1, the full equation, Eq. (6), must be used rather than (8).

Distribution of Deviations

Our universal growth curve provides a natural baseline against which to quantify growth fluctuations. Although most companies conform well to the universal growth behaviour, some companies deviate substantially. For example, Apple has grown faster than the prototypical growth curve, whereas American Plastics & Chemicals has grown significantly slower. In contrast, Coca-Cola has followed the curve relatively closely, as illustrated in Fig. 3(b).

To characterize how an individual company \( i \) deviates from the universal growth curve, we introduce its average relative deviation over its lifespan, \( T^i \):

\[
\epsilon^i = \sigma \left( \sum_{t=1}^{T^i} \epsilon^i(t) \right) \sum_{t=1}^{T^i} |\epsilon^i(t)| \frac{1}{T^i} \tag{10}
\]

Here \( \sigma(x) \) is the sign function (i.e., \( \sigma(x) = +1 \) if \( x > 0 \), and \( -1 \) if \( x < 0 \)); \( \epsilon^i(t) \equiv \ln A^i(t) - \ln \hat{A}^i(t) \equiv |A^i(t) - \hat{A}^i(t)|/\hat{A}^i(t) \) is the relative deviation of the predicted total assets at time step \( t \), \( A^i \) is the observed data, and \( \hat{A}^i \) is the prediction. Note that by including the \( \sigma \) function this measure gives the sign of the average deviation, either above or below the universal growth curve, potentially providing a metric for a company’s over- or under-growth performance relative to the expectation for its size.

Fig. 5 shows the distribution of the \( \epsilon^i \) for the entire dataset as well as for selected sectors. This clearly exhibits a tent-like shape characteristic of a Laplace distribution, defined by

\[
p(\epsilon) = \frac{1}{2b} \exp\left(\frac{-|\epsilon - \mu|}{b}\right), \tag{11}
\]

where \( \mu \) is the displacement parameter and \( b \) the dispersion. For comparison, we also show in the figure a shadow Gaussian distribution, which is clearly a very poor fit being unable to accommodate the larger deviations.

The main point here, however, is that, like scaling laws, the universal growth curve, when coupled with the deviation analysis, provides a principled science-based baseline metric for assessing a company’s performance. The analysis of deviations and fluctuations in company growth has a long history [17]. Typically, this has focused on the distribution of growth rates, which were originally argued to follow a Laplace distribution (e.g. [10, 20]). However, these are known to have curvature and asymmetry away from the Laplace distribution for large and small growth rates (e.g. [21]), which may be due to the fact that they effectively used a Gibratian perspective of size-independent growth as a base-line. In our work it is the deviations in time from the universal growth curve that follow a Laplace distribution, which only results because of the size-dependent growth implicit in the scaling laws.

Growth, equity and debt

The equity of a company is defined as the difference between its assets and liability: \( Q \equiv A - L \). Thus, \( dQ/dt = dA/dt - dL/dt \) which, from Eq. (3), is \( I(t) \) so \( dQ/dt = I(t) \), i.e., the rate of increase of equity is just the net income. Similarly, since \( dQ/dA = 1 - dL/dA \), the growth equation (4) can be expressed in a very compact form:

\[
\frac{dA}{dt} = \frac{I(A)}{dQ/dA} \tag{12}
\]

Conceptually implicit in using power law scaling for both \( I(A) \) and \( L(A) \) in Eq. (5) is that these are the prime quantities companies, on average, strive to optimise. In this sense, equity, \( Q = A - L \), is a “secondary” derived quantity since the
difference between two power laws cannot itself be a power law. Consequently, if equity is not directly optimised it
cannot scale as a power law. It would therefore be technically incorrect to use power laws for both \( I(A) \) and \( Q(A) \)
in Eq. (12), even though it might be a useful approximation. On the other hand, it is straightforward to consider the
alternative scenario where equity is presumed to be the prime quantity companies try to optimise, rather than income
and/or liability, and so obeys power law scaling. In that case, Eq. (8) is the exact solution for \( A(t) \) with \( \gamma = (\beta_Q - \beta_I)^{-1} \)
and \( c = [(c_I/c_Q)(1 - \beta_I/\beta_Q)]((\beta_Q - \beta_I)^{-1} \), where \( \beta_Q \) and \( c_Q \) are the scaling exponent and normalization constant of \( Q \)
with respect to \( A \), respectively.

It is instructive to consider the realistic case when \( \beta_L \approx 1 \), which holds for the the market as a whole as well as for many
individual sectors. In that case, \( c_L = k \), the approximately constant debt ratio: \( k = L(A)/A \) and \( Q(A) = (1 - k)A \)
(i.e., \( \beta_Q = 1 \) and \( c_Q = 1 - k \)). Consequently, the solution for \( A(t) \) reduces precisely to Eq. (7). The ratio \( A/Q \) is
the financial leverage ratio conventionally referred to as the equity multiplier, \( r \) \( \approx \) \( (1 - k)^{-1} \), in terms of which Eq.
(12) becomes \( dA/dt \approx rI \). That is, the growth rate of a company is its net income modulated by the equity multiplier
leverage ratio.

Note that when the equity multiplier, \( r = 1, k = 0 \), so \( L = 0 \) and \( Q = A \) and all assets owned by a company are held
in stockholder equity and none are funded by debt, in which case company growth is determined solely by capital in the
form of net income. In our data we find \( r \approx 2 \), so \( k \approx 0.5 \) implying that, on average, about 50% of a company’s equity
is funded by debt, or equivalently, that the total assets of the average company are twice those held in stockholder equity.
Perhaps counter-intuitively, this implies that, holding all else constant, issuing debt increases growth by magnifying
the returns from net income (i.e., retained earnings): debt is an additional inflow of finances thereby increasing the
total assets of a company, and because growth is a function of assets, companies with debt will grow faster than those
without. On the other hand, issuing debt incurs risks and costs; higher debt potentially leads to increased debt expenses,
higher leverage, lower flexibility, lower investor confidence, and the greater risk of bankruptcy. However, at low interest
rates debt compensates for short term losses and provides additional capital to finance growth where the additional
interest costs can be offset by higher returns on capital invested in the company. Consequently, the ability to manage
debt and its associated risk is a crucial factor in the long-term growth of a company.

It follows that the asset growth rate is the return on equity - the ratio of net income to stockholder equity (net assets) -
which is a measure of the ability of a company to generate profits from the assets it holds;

\[
\frac{1}{A} \frac{dA}{dt} \approx r \frac{I}{A} = \frac{I}{Q} = rc_I A^{\beta_I-1}.
\]

Because \( \beta_I < 1 \), Eq. (13) captures the decreasing returns on net assets with increasing company size. This prediction is
quantitatively confirmed by data as shown in Fig. 6 and shows explicitly why Gibrat’s assumption of a constant relative
growth rate is incorrect.

3 Discussion

The theory of company growth developed here is similar in spirit to the dynamics of growth processes of other entities
well-described by scaling phenomena, such as populations, cities, nations, mammals, plants, or bacterial colonies
[11]. Our growth equation is part of a general class of models which captures how the scaling of inputs and outputs
determines the specific form of growth across this diverse range of entities and institutions. For example, in mammals,
income is the total available metabolic energy while costs are the energy required for repair and maintenance, both of
which scale systematically with body size. Because the latter scales more quickly than metabolic rate - the exponents
are 1 and 3/4 respectively - body size reaches a maximum where maintenance consumes all of the incoming energy
and growth ceases, leading to a stable adult size [22]. There is no indication of a similar maximum size in companies
based on the growth curves or on the scaling laws for liabilities and net income which predict the observed open-ended
power-law growth in time. Another example is cities whose metabolism scales superlinearly with population (i.e.,
with an exponent greater than 1) whereas its per-capita costs are approximately constant [13]. In this scenario growth
is open-ended as it is for companies; however, rather than growing as a power law in time, cities grow faster than
exponentially (super-exponentially) leading to a finite time singularity. Bacteria exhibit a similar scaling pattern to
cities, namely, superlinear “incomes” and linear costs [23, 24]. Companies are yet another variation on this theme
in which the sublinear scaling of net income and the linear scaling of liabilities with assets leads to sub-exponential
growth. Consequently, companies are open-ended growers whose size is dependent on their ability to deploy assets
in response to external market conditions, rather than pre-determined by a genetic code, such as in the ontogeny of
mammals. However, the ability of a company to deploy assets to promote growth is size-dependent.

A mechanism seemingly unique to companies and not found in other complex adaptive systems is the ability to raise
capital from markets by issuing liability or equity to manage growth. While there may be biological precursors to
offsetting immediate growth demands by either borrowing resources from kin for mutual benefit (i.e., gestation, followed by survival and eventual reproduction), or buffering against future uncertainty (i.e., storage, hibernation, or estivation, for example), arguably, the ability to raise capital by selling somatic control (i.e., stockholder equity), or by borrowing against perceived future value (i.e., debt) is unique to, and perhaps definitive of, markets. Debt is a fundamental component in the evolution of human economies, as it is a mechanism that allows for economic transactions to be deferred and to remain incomplete for a finite period of time through negotiation [25]. This ability to raise capital by borrowing from the future means the growth trajectories of companies are, in many ways, more flexible than organisms or cities, as runs of bad years can be temporarily mitigated. However, capital investments are not free and so this mitigation can only offset, not replace, productivity [26]. Moreover, this has interesting implications for theories of the firm: not only do companies reduce transaction costs, aggregate information, and achieve economies of scale, they also provide mechanisms for buffering by being able to displace immediate shortfalls into the future. This buffering is much harder to sustain on an individual level, especially as the primary mechanism for storing and building wealth (and other buffering mechanisms) is the market itself. Public companies and markets only exist through flows of finance with the expectation of future growth, and the ability of public companies to perform this role ultimately determines their fate.

The unique power-law growth with time we find here has its origins in the specific scaling of financial quantities in companies, where liabilities, on average, scale with assets with an exponent consistent with unity. This approximate linearity implies that debt is independent of size, and so the liability accessible per unit of assets does not depend on the total assets of the company [27]. Note, however, that for small companies there are deviations from this, as reflected in the curvature away from the power law at the lower end of the plots in Fig. 2 and the inset of Fig. 3b).

These results highlight the importance of the value of the exponent for net income, \( \beta \), since, from Eqs. (7) and (8), the time exponent governing growth, \( \gamma \approx 1/(1 - \beta) \). Consequently, small changes in \( \beta \) have potentially large effects on \( \gamma \) and therefore on growth trajectories. For instance, if \( \beta \) were 1, and not 0.85, then companies would grow exponentially in time, whereas an exponent of 0.9 would give a temporal scaling of \( t^{10} \), while 0.8 would give \( t^{0.8} \). The value of the overall normalisation constant, \( c \), is equally sensitive to small changes in the scaling parameters. The sensitivity of both \( \gamma \) and \( c \) is closely related to the well-known sensitivity of marginal returns in determining firm survival in general.

Developing mechanistic theory to explain why we see these exponents will be of considerable importance given the consequences for understanding scaling of company growth over time. In both biological systems and cities scaling exponents have their origin in the dynamics and geometry of internal network structures that transport energy, resources and information, such as vasculature, roads, and social networks. For firms there have been several proposals of how internal structure leads to growth dynamics [20,10,17]. While these models do not predict scaling exponents nor the growth dynamics we predict and observe, perhaps similar inherent constraints of company structure lead to the sublinear scaling we report. Additionally, the extreme competitiveness of markets at short time scales and how larger firms feel the size of the market also constrain internal function and these effects of the the overall ecology could likely be the driver of the sublinear scaling.

Importantly, Gibratian growth is the special case of the theory we develop here, in which net income scales linearly with assets leading to purely exponential growth. However, the inherently nonlinear scaling of net income reveals that company growth cannot be a Gibrat process. This result also suggests a paradigmatic shift in how we should think about how companies grow over time: In exponential systems, we are conditioned to think about compound interest and growth as a constant proportion of size, and thus exponential in time. However, we have shown that the expected growth rate of a company is a sublinear function of its size, and so the capacity for growth decreases as a company increases in size. As a company becomes larger a decreasing proportion of its assets can be deployed to promote growth: The more assets a company deploys toward maintaining market share and meeting production, the less flexible it is in responding to external fluctuations or diversifying it’s production [26].

It is interesting to note that while the ensemble of companies grows following a power law of time, the overall market grows exponentially. In fact, an average company will achieve their fastest rate of growth at intermediate times out-performing the average growth of the market, but only for a finite period. Over this period, companies may either merge or be acquired by competitors, or eventually succumb to mounting maintenance costs. An open question then is why the market grows exponentially while the ensemble average of companies does not. The answer is likely to do with the constant and rapid turnover of companies entering and exiting markets through birth and death processes.

4 Methods

Dataset

The dataset contains the financial information obtained from the income statements and balance sheets of publicly traded companies from 1950 to 2018. There are 31,553 companies in total and most of them are from North American
and overseas American Depository Receipt firms. The data is from Compustat North America and Compustat Historical databases compiled by Standard & Poor’s [5].

The classification for companies is according to the North American Industry Classification System (NAICS) standard which is developed by the statistical agencies of Canada, Mexico and the United States. A six digits code is assigned to each company except 1,839 companies. We classify the companies into 447 economic sectors according to the code. The list of the sectors are provided in S7 of SI.

**Deflation**

Most financial variables are measured by U.S. dollars, therefore the inflation factor should be considered. Especially when we compare the same indicator for different years, the inflation factor plays a significant role. In this paper, we deflate each financial variable in the same way. Take sales as an example. At first, we select the last year as the base year \( T \), then all the sales data will be converted and measured by the money in year \( T \) as shown in equation 14.

\[
S(t) = S_0(t)e^{r(t)}
\]

\( S(t) \) is the deflated sales at year \( t \), \( S_0(t) \) is the raw data of sales at \( t \). \( r(t) \) is the inflation rate relative to \( T \), it can be computed as:

\[
r(t) = \sum_{\tau=1}^{T} \pi_{\tau}
\]

Where, \( \pi_{\tau} \) is the inflation rate of time \( \tau \) to the previous year. Thus, all the values shown in the main text are normalized in this way.

**Fitting Methods**

In the main text, we use two different methods to fit the data on log-log coordinate to obtain the scaling laws. For the 423 individual economic sectors, we used the linear mixed-effect models (LMMs) to fit and estimated the exponents and coefficients of the scalings. For the data set of all sectors mixed together, we at first flat all the data points for all companies in all years together, and we used the ordinary least square (OLS) method to fit the data and estimated parameters. The comparing results of the two methods on estimating scaling exponents and coefficients are shown in Section S2.1 of the SI.

LMMs are extensions of linear regression models for data that are collected in groups. The assumption behind LMM is that the coefficients can vary with respect to one or more grouping variables. In our case, for any company within the sector that we considered, and for any aggregated-level variable \( X \), we assumed that each data point \((\ln A_{t,i}, \ln X_{t,i})\) for company \( i \) at year \( t \) follows a scaling law, but the scaling exponent and coefficient vary with the company \( i \). That is,

\[
\ln X_{t,i} = \ln(c_X) + \beta_X \ln A_{t,i} + \xi_{t,i}
\]

for the company \( i \) at time \( t \), where, \( \xi_{t,i} \sim N(0, \sigma^2) \) is an independent random number, and \( \sigma \) is a constant. And \( \ln(c_X), \beta_X \) are the company dependent intercept and slope respectively, and satisfy:

\[
\ln(c_X) = \ln(c_X) + \xi_c^i,
\]

and:

\[
\beta_X = \beta_X + \xi_{\beta}^i,
\]

where, \( \xi_c^i \sim N(0, \sigma_0^2) \) and \( \xi_{\beta}^i \sim N(0, \sigma_1^2) \), and \( \sigma_0, \sigma_1 \) are constants.

We then fit the linear mixed-effect model to obtain the fixed effect exponent \( \beta_X \) and the coefficient \( c_X \) as the estimations.

As shown in SI Table S1 and S2, the exponents estimated by LMM method are similar with the exponents estimated by OLS method.

However, we give up using the LMM method to estimate the exponent and the coefficient for all sectors. The reason is that the exponent will not be used for estimating the relation between \( X \) and \( A \) but also for deriving the growth equation, and to better fit the data through time, we adopt the OLS method to estimate the exponent and the coefficient.

**Definition of Natural Age**

The first year a company \( i \) is born is not reported in our data, and so we treat the first year that the assets of company \( i \) appears in the data as it’s first year. We define the natural age of \( i \) as the time span from its first year to the current year (i.e., the number of years \( i \) appears in the data).
Estimation for the leverage $r$

Because $\beta_L \approx 1$ for the whole market, therefore $L \approx k \cdot A$, where $k \approx c_L$. Since $Q = A - L$, thus, $Q \approx r \cdot A = (1 - k) \cdot A$.

To estimate $k$ for all companies, we cannot use $c_L$ estimated by equation 1 directly because this is not a linear relationship. We use the equation below instead of algebraic average value to estimate because the size range is very large.

$$k = \exp \left \langle \ln L - \ln A \right \rangle,$$

(19)

where the average $\langle \cdot \rangle$ is taken over all data points. This is equivalent to the OLS estimation of $k$ by taking the slope is exactly 1 from the relationship $\ln L = \ln k + \ln A$. Therefore, $r$ is calculated by $1 - k$.

5 Figures and Tables

Figure 1: The cash flow and capital structure of financial variables influence the size and growth of a company. Cash flow is the flux of financial resources through a company generated by sales. Sales are generated from transactions of goods and services outside the company within the market. Net income (i.e., net profit/loss) is the difference between sales and expenses (i.e., costs released back into the market). Once dividends have been paid to shareholders the remainder constitutes the retained earnings which accumulate as assets. The capital structure of a company is the structure of assets which are apportioned into equity (i.e., issued stock) and liabilities (i.e., issued debt). The ability of a company to deploy assets to generate cash flow constitutes the return on assets and is a common measure of the financial efficiency of a company.
Figure 2: The scaling laws of annual sales (a), annual cost of sales (b), annual net incomes (c), and total liability (d) as a function of total assets for all 31,553 companies over all years across the SIC economic sectors. Different colors represent different SIC sectors (colors are assigned according to the average assets of the sector). The solid black lines are fitted scaling functions using ordinary least square (OLS) method over all sectors. Colored lines are linear mixed-effect model (LMM) fits for each sector, and the distributions of the fitted exponents are shown in the inset figures where the red vertical lines represent the OLS fits. Further statistical results are available in Section S2 of the Supplementary Materials.

Table 1: The scaling exponents $\beta$ and the 95% CI for selected variables and economic sectors.

| Variables        | Air Transp. | Biolog. Prod. | Compu. Sys. | Motor Veh. Part | All Sectors |
|------------------|-------------|---------------|-------------|-----------------|-------------|
| Income related   |             |               |             |                 |             |
| Sales            | 0.88[0.80,0.96] | 0.68[0.62,0.74] | 0.82[0.77,0.86] | 0.88[0.84,0.92] | 0.90[0.90,0.90] |
| Net Incomes      | 0.89[0.84,0.93] | 0.76[0.68,0.84] | 0.83[0.78,0.88] | 0.75[0.70,0.80] | 0.85[0.85,0.85] |
| EBITDA           | 0.86[0.82,0.90] | 0.94[0.86,1.01] | 0.89[0.85,0.93] | 0.87[0.82,0.92] | 0.94[0.93,0.94] |
| Gross Profit     | 0.87[0.82,0.91] | 0.84[0.77,0.91] | 0.82[0.77,0.86] | 0.86[0.81,0.90] | 0.85[0.85,0.85] |
| Dividends        | -           | -             | -           | -               | 0.56[0.55,0.58] |
| Retained Earnings| 1.01[0.90,1.12] | 1.11[0.97,1.26] | 1.01[0.93,1.10] | 0.97[0.84,1.10] | 0.90[0.90,0.90] |
| Cost related     |             |               |             |                 |             |
| Cost of Sales    | 0.82[0.75,0.88] | 0.59[0.56,0.63] | 0.77[0.73,0.82] | 0.90[0.84,0.95] | 0.85[0.85,0.85] |
| Total Tax        | 0.80[0.69,0.90] | 0.80[0.66,0.93] | 0.96[0.87,1.05] | 0.83[0.77,0.94] | 0.93[0.93,0.94] |
| Operating costs  | 0.80[0.74,0.87] | 0.58[0.56,0.60] | 0.74[0.70,0.77] | 0.83[0.81,0.90] | 0.79[0.79,0.79] |
| R&D              | 0.41[0.13,0.69] | 0.62[0.59,0.64] | 0.63[0.57,0.69] | 0.81[0.70,0.92] | 0.71[0.70,0.71] |
| Size related     |             |               |             |                 |             |
| Employee         | 0.65[0.58,0.71] | 0.52[0.49,0.54] | 0.68[0.64,0.72] | 0.81[0.76,0.86] | 0.75[0.75,0.75] |
| Cash             | 0.85[0.73,0.97] | 0.93[0.90,0.95] | 0.96[0.90,1.01] | 0.92[0.83,1.00] | 0.82[0.82,0.82] |
| Total Liability  | 1.00[0.94,1.06] | 0.59[0.55,0.63] | 0.74[0.69,0.78] | 1.01[0.95,1.07] | 0.99[0.99,0.99] |
| Equity           | 0.87[0.82,0.92] | 1.06[1.04,1.08] | 1.09[1.06,1.12] | 0.93[0.87,0.98] | 0.90[0.90,0.91] |
Figure 3: The growth trajectories of all 31,553 companies in terms of total assets over time and the universal growth predictions. (a) The raw growth trajectories of all companies in our dataset from 1975 to 2018. Each curve represents a single company. Three representative companies are highlighted: Apple Inc. (green), Coca-Cola Co. (blue), and American Plastics & Chem. (gray). These companies were chosen as they reflect companies out-performing predictions (Apple Inc.), meeting predictions (Coca-Cola Co.), or under-performing expectations (American Plastics & Chem.). (b) The predicted universal growth curves and the translated growth trajectories of each company to the corresponding $t_0$ position in a semi-log plot, highlighting that companies grow as a power law of time. The inset shows the same data (gray points cloud) with their log binned averages (the red points) and the same predicted universal growth curve (black line) on a log-log pot. The predicted power law exponent $\gamma$ and the coefficient $(\ln c)$ are also shown. (c) Real growth trajectories (blue circles) and their corresponding predicted growth (red solid lines) obtained by solving Eq. 5 with the observed initial size for four selected companies; time is the natural year. To compare, we also plot the exponential growth by fitting the tail of the universal growth curve by a dashed gray line. (d) The comparison of assets from real growth data with the growth predicted by Eq. 5 showing excellent agreement, as manifested by the red straight line having a slope of 1. Across the panel (a)-(b) companies are consistently colored according to their lifespans in the data (defined as the difference between their first and last appearance in the data set, which does not necessarily correspond with their actual births and deaths). Other forms of the growth equation are also shown to compare in S5 of the SI.
Figure 4: The universal growth curves (black solid lines) from Eq. (6) for selected industries, and the growth trajectories of the individual companies in those sectors, where each company is colored by their longevity. The parameters of $\beta_I, c_I, \beta_L, c_L$ are also shown for each sector. Notice that $\beta_L$ for the sectors of biological products, computer integrating systems design and semiconductors are significantly less than 1 so that the approximate solution Eq. (8) is not strictly valid and the full equation should be used.
Figure 5: The normalized distributions of deviations $\epsilon$ for selected economic sectors (colored data) and for all sectors (black dots). The selected economic sectors are the same as in Fig. 4. The solid colored lines for each representative sector and the black line for all sectors are Laplacian distributions and the grey dashed line is a Gaussian distribution fit to all sectors (black data points) included for comparison. The scale parameters $\mu$ and $b$ of Laplacian distributions are also shown for the representative sectors.
Figure 6: The dependence of growth rate on total assets for all companies (main figure) and selected SIC sectors (inset) over all years. Positive growth rates are blue and negative growth rates are orange. To show negative growth rates on the logarithmic scale, absolute values are plotted. To clarify the relationship between growth rate and total assets we divided the total assets of companies into 100 logarithmic bins and calculated the average growth rate in each bin, plotted as the black data points. The red line is the theoretical prediction from our model \(\frac{dA}{dt} = c_I \cdot A^{\beta_I - 1} / (1 - c_L \cdot A^{\beta_L - 1})\), and the dashed blue line is the approximation \(\frac{dA}{dt} \approx c_I \cdot A^{\beta_I - 1} / (1 - c_L) = r \cdot c_I \cdot A^{\beta_I - 1}\) when \(\beta_L \approx 1\). The inset shows the same logarithmic binned results for the selected sectors.
References

[1] R. H. Coase, “The nature of the firm,” *Economica*, vol. 4, no. 16, pp. 386–405, 1937.

[2] E. T. Penrose, *The theory of the growth of the firm*. Oxford, UK: Oxford/Blackwell, 1959.

[3] O. Williamson and S. Winter, *The nature of the firm: origins, evolution, and development*. Oxford, UK: Oxford University Press, 1993.

[4] G. R. Carroll and M. T. Hannan, *The demography of corporations and industries*. Princeton, NJ: Princeton University Press, 2004.

[5] “Standard & poor’s. COMPUSTAT (north america). see www.compustat.com.”

[6] “The World Bank: https://data.worldbank.org” 2021.

[7] M. I. Daepp, M. J. Hamilton, G. B. West, and L. M. Bettencourt, “The mortality of companies,” *Journal of The Royal Society Interface*, vol. 12, no. 106, p. 20150120, 2015.

[8] R. L. Axtell, “Zipf distribution of US firm sizes,” *science*, vol. 293, no. 5536, pp. 1818–1820, 2001.

[9] M. H. Stanley, S. V. Buldyrev, S. Havlin, R. N. Mantegna, M. A. Salinger, and H. E. Stanley, “Zipf plots and the size distribution of firms,” *Economics Letters*, vol. 49, no. 4, pp. 453–457, 1995.

[10] M. H. Stanley, L. A. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, and H. E. Stanley, “Scaling behaviour in the growth of companies,” *Nature*, vol. 379, no. 6568, pp. 804–806, 1996.

[11] G. B. West, *Scale: the universal laws of growth, innovation, sustainability, and the pace of life in organisms, cities, economies, and companies*. New York, NY: Penguin, 2017.

[12] G. B. West and J. H. Brown, “The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization,” *Journal of Experimental Biology*, vol. 208, no. 9, pp. 1575–1592, 2005.

[13] L. M. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, “Growth, innovation, scaling, and the pace of life in cities,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 17, pp. 7301–7306, 2007.

[14] R. Gibrat, *Les inégalités économiques: applications aux inégalités des richesses, à la concentration des entreprises... d’une loi nouvelle, la loi de l’effet proportionnel*. Paris, France: Librairie du Recueil Sirey, 1931.

[15] K. Shimizu and L. C. Edwin, “History, genesis, and properties,” in *Lognormal distributions*, pp. 1–25, Abingdon-on-Thames, UK: Routledge, 2018.

[16] E. Santarelli, L. Klomp, and A. R. Thurik, “Gibrat’s law: An overview of the empirical literature,” *Entrepreneurship, Growth, and Innovation*, pp. 41–73, 2006.

[17] D. Fu, F. Pammolli, S. V. Buldyrev, M. Riccaboni, K. Matia, K. Yamasaki, and H. E. Stanley, “The growth of business firms: Theoretical framework and empirical evidence,” *Proceedings of the National Academy of Sciences*, vol. 102, no. 52, pp. 18801–18806, 2005.

[18] S. V. Buldyrev, F. Pammolli, M. Riccaboni, K. Yamasaki, D.-F. Fu, K. Matia, and H. E. Stanley, “A generalized preferential attachment model for business firms growth rates,” *The European Physical Journal B*, vol. 57, no. 2, pp. 131–138, 2007.

[19] M. Riccaboni, F. Pammolli, S. V. Buldyrev, L. Ponta, and H. E. Stanley, “The size variance relationship of business firm growth rates,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 50, pp. 19595–19600, 2008.

[20] L. A. N. Amaral, S. V. Buldyrev, S. Havlin, M. A. Salinger, and H. E. Stanley, “Power law scaling for a system of interacting units with complex internal structure,” *Physical Review Letters*, vol. 80, no. 7, p. 1385, 1998.

[21] R. Perline, R. Axtell, and D. Teitelbaum, *Volatility and asymmetry of small firm growth rates over increasing time frames*. Washington, D.C.: SBA Office of Advocacy, 2006.

[22] G. B. West, J. H. Brown, and B. J. Enquist, “A general model for ontogenetic growth,” *Nature*, vol. 413, no. 6856, pp. 628–631, 2001.

[23] J. P. DeLong, J. G. Okie, M. E. Moses, R. M. Sibly, and J. H. Brown, “Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 29, pp. 12941–12945, 2010.

[24] C. P. Kempes, S. Dutkiewicz, and M. J. Follows, “Growth, metabolic partitioning, and the size of microorganisms,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 2, pp. 495–500, 2012.

[25] D. Graeber, *Debt: The first 5000 years*. London, UK: Penguin, 2012.
[26] J. Roberts, *The modern firm: Organizational design for performance and growth*. Oxford, UK: Oxford university press, 2007.

[27] F. Modigliani and M. H. Miller, “The cost of capital, corporation finance and the theory of investment,” *The American economic review*, vol. 48, no. 3, pp. 261–297, 1958.