Fat tails in private equity fund returns: The smooth double Pareto distribution

How to cite:

© 2022 Henry Lahr

https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Version of Record

Link(s) to article on publisher’s website:

oro.open.ac.uk
Fat tails in private equity fund returns: The smooth double Pareto distribution

Henry Lahr

Department of Accounting and Finance, The Open University Business School, Walton Hall, Milton Keynes MK7 6AA, UK
Centre for Business Research, University of Cambridge, Trumpington Street, Cambridge CB2 1AG, UK

ARTICLE INFO

JEL codes:
C46
D31
G24
G32
L11
R12

Keywords:
Size distribution
Income distribution
Pareto law
Power-law distribution
Fat tails
Private equity
Venture capital
Financial returns

ABSTRACT

Does the distribution of private equity returns have fat tails? A new smooth double Pareto distribution can explain the stationary distribution of private equity funds’ valuation multiples. This distribution emerges from a random growth model with lognormally distributed initial fund valuations. This model endogenously generates power-law tails in the stationary cross-section. The new distribution fits the data better than competing distributions. Fat tails are particularly pronounced in venture capital funds and suggest returns with infinite variance over the lifetime of the fund. The smooth double Pareto distribution has wide applicability to growth processes with a random initial value.

1. Introduction

The purpose of this paper is to explain the stationary distribution of valuation multiples of private equity funds by introducing the smooth double Pareto distribution, which can describe random growth processes with a random initial size. The stationary distribution of such valuation multiples can be defined as the steady-state distribution that can be observed at any point in time once the process of fund birth, growth and death has reached a stationary state. It should not be confused with the distribution of instantaneous growth, which governs the growth of individual valuation multiples from one point in time to the next. Analysing the stationary distribution allows inferences about the underlying growth process, including the distribution of its instantaneous growth rate, which is relatively opaque in private equity compared with public markets. The new smooth double Pareto distribution has Pareto (i.e., power-law) tails and a central body that smoothly connects the tails. In addition to private equity, it can be applied to a wide range of growth processes including people’s income and wealth, or populations and the size of human settlements.

Many variables studied in finance, economics and other disciplines are found to be distributed according to a power law. Perhaps the most famous example is the frequency distribution of words in a natural language, studied by Zipf (1949), who popularised the eponymous law. It states that the frequency of any word is inversely proportional to its rank in the frequency table. For any word W, its probability of exceeding length x is \( P(W > x) = k/x^\alpha \) for some constant k. The special case of \( \alpha = 1 \) is called Zipf’s law. Earlier, Pareto (1896) had found a similar relationship in the upper-tail distribution of the number of people with an...
income or wealth $W$ greater than a large $x$, where $\alpha$ is some positive real number. Research has established power laws in firm sizes, city sizes, short-term stock price fluctuations and trading volume (Gabai, 2009). The tail of the distribution of the productivity of ideas and innovations, as well as the number of patent citations, has been found to follow a power law (Ghiglino, 2012). In entrepreneurship, many financial and non-financial measures can be described as a power law (Crawford, Aguinis, Lichtenstein, Davidson, & McKelvey, 2015). Beyond the field of economics, power laws are found in continuous and discrete variables as diverse as the frequency of occurrence of unique words in the novel *Moby Dick* and peak gamma-ray intensity of solar flares (Clauset, Shalizi, & Newman, 2009).

Mechanisms known to generate power laws are preferential attachment (Simon, 1955; Yule, 1925), random growth with a reflective barrier (Champernowne, 1953; Gabai, 1999) or death (Gabai, 2009; Luttmert, 2007) to stabilise the distribution, self-organized criticality (Sak, Chen, Scheinkman, & Woodford, 1993), optimisation (Mandelbrot, 1953), random observation (Huberman & Adamic, 1999; Huberman & Adamic, 2000), and exponential technological progress combined with exponential technology diffusion (Hilbert, 2013). Mitzenmacher (2003), Newman (2005) and Gabai (2009) provide overviews of such generative processes.

Returns of private equity funds – and venture capital (VC) funds in particular – are often hypothesised by industry insiders to have a right power-law tail (Masters & Thiel, 2014). Power-law tails would be an attractive feature for both entrepreneurs and investors, because they can offer extraordinary gains from a modest investment. The continued search for “unicorns”, start-up companies with a value of greater than $1bn, through seed and early-stage investments could then be justified by rational investors even if historical samples suggest poor average returns. If the power-law exponent of the upper tail is sufficiently low, then higher moments converge slowly and can even be infinite, and history will be an imprecise guide for future returns. Prencipe (2017) finds evidence in support of a power-law distribution of VC investments that may have an infinite variance. In his sample of VC investments backed by the European Investment Fund, a power law fits the data but cannot be statistically distinguished from a lognormal distribution.

A challenge for statistical estimation and tests of a right-tail distribution is to find a cut-off point above which the power law is presumed to hold (Clingingsmith, 2017; Nicolau & Rodrigues, 2019). If part of the distribution cannot be explained by the power law to be tested, this threshold can be arbitrary. More importantly, many distributions can fit the data if only part of the distribution needs fitting. For example, a lognormal distribution can be stretched sufficiently by assuming a small mean and large variance to make it look linear in a log-log diagram over large intervals and indistinguishable from a power law. It is thus desirable to avoid having to specify a cut-off point by finding a theoretically justified distribution that fits all the data, including the distribution’s body and tails.

Returns to private equity investments are notoriously difficult to estimate, because the underlying assets are traded infrequently or not at all. Several authors have proposed solutions to the problem of estimating return characteristics from investor cash flows or fund valuations. If interim return observations are not available, an assumption must be made regarding the path of returns between observations. The standard assumption in studies of private equity returns is that growth in value between subsequent periods is lognormally distributed (Ang, Chen, Goetzmann, & Phalippou, 2018; Cochrane, 2005; Driessen, Lin, & Phalippou, 2012; Korteweg & Nagel, 2016). The stationary distribution of a growth process can be used to test this assumption without the need to observe fine-grained time-series returns.

This paper makes three contributions. Firstly, this paper explains the stationary (steady-state) distribution of private equity returns, which has not been modelled before, and tests whether stationary returns follow a power law. Its main objective is to model the stationary cross-section of private equity returns in its entirety, instead of a tail distribution above an arbitrary cut-off value, taking into account the life cycle of private equity funds and potential power-law behaviour in the tails. This theoretically driven approach allows, for the first time, to explain the stationary distribution starting from first principles. Secondly, fitting the new smooth double Pareto distribution to private equity data allows for inferences about the growth process through which the stationary distribution was generated. In particular, the common assumption of lognormally distributed time-series returns in private equity can be tested. Finally, the smooth double Pareto distribution as derived from a growth process with constant death and random initial size can yield new insights into growth phenomena in nature and society due to its wide applicability to processes involving birth and death.

To mirror the life cycle and investment process of private equity funds, I derive the stationary return distribution from a standard diffusion process that is augmented by a lognormally distributed initial value. The starting point and best existing candidate for a cross-sectional distribution of private equity returns is the double Pareto distribution first described by Reed (2001) and extended by Gabai (2009) and Gabai, Lasry, Lions, Moll, and Qu (2016). However, this distribution features a discontinuity at the fixed point at which new entities are continuously added, which does not agree with observed private equity valuations in the stationary cross-section (see Fig. 1). A process by which new funds are born with a random initial valuation solves this problem. The result is a distribution function that is smooth everywhere.

Results reveal a good fit of the smooth double Pareto distribution to the data. The new distribution accurately describes return multiples of private equity funds across the samples studied in this paper, both in terms of tail behaviour and goodness of fit in the distribution’s central body. It performs either as well as or better than alternative distributions such as the standard double Pareto or lognormal distribution. These findings imply that the random growth process with random initial size is a plausible candidate for the process generating the stationary distribution of private equity multiples. In particular, the commonly used hypothesis of lognormal growth finds support in the data.

Are private equity returns distributed according to a power-law? Probably. But this does not mean that individual fund investments exhibit power-law behaviour. The difference is due to the unequal time periods that funds spend in the growth process. Individual funds are still described by the lognormal process inherent in the diffusion model. A power law arises only because an unequal life time gives funds the chance to follow their lognormal path longer and thus increase the weight of the cross-sectional distribution’s tails.

This paper contributes to the theoretical literature on random growth models (Gabai, 1999; Gabai et al., 2016; Reed, 2001; Reed & Jorgensen, 2003; Toda & Walsh, 2015) by deriving the stationary distribution of a random growth process with random initial size and constant death. This distribution has wide applicability to random growth processes in which the initial state can be described by a lognormal distribution, such as personal wealth (Benhabib & Bisin, 2018; Benhabib, Bisin, & Luo, 2017; Blanchet, Fournier, & Piketty, 2021; Kuhn, Schularick, & Steins, 2020; Levy & Levy, 2003) or income (Atkinson, 2017; Gabai et al., 2016; Jenkins, 2017), the size distributions of firms (Coad, 2010, Coad, 2009; Luttmert, 2007; Cabral & Mata, 2003) and settlements (Hwu, 2012; Rozenfeld, Rybski, Gabai, & Makse, 2011). The modelling of these growth processes may benefit from the full characterisation in

---

1 The purpose of this paper is not to propose a statistical method to estimate the power-law tail of a distribution above or below a cut-off point but to eliminate the need to specify a cut-off point by analytically describing the entire distribution. Because this approach replaces one set of assumptions with another, its suitability will depend on the specific situation, that is, whether the growth process to be modelled includes birth and death of entities.
this paper of the stationary distribution of a process whose starting point is not fixed but random.\textsuperscript{3} Findings in this paper further contribute to the literature on private equity returns by establishing bounds on the likely processes generating the distribution of returns. More specifically, fitting the stationary distribution allows for a test whether innovations in the return process are normally distributed, leading to lognormal growth for individual entities as is often assumed in the literature (Ang et al., 2018; Korteweg & Nagel, 2016).

From a practical viewpoint, the model proposed in this paper completely characterises the return distribution rather than being restricted to its tail properties. This feature allows for a more comprehensive modelling of risk and return properties where sample moments can be unreliable due to unstable or non-existent moments. For example, the distribution’s mean, variance and higher moments (if they exist) can be calculated more precisely from the distribution’s estimated closed-form density function.

2. Random birth and random growth

2.1. Return observations in the private equity investment process

Interim private equity fund returns are usually expressed as the internal rate of return (IRR) implicit in the string of cash flows between the fund and its investors, typically treating the unrealised net asset values as a final cash flow. A slightly simpler measure that is often used in the private equity industry and which I adopt in this paper is the ratio of the fund’s cash distributions to its investors plus the fund’s net asset value, relative to the capital paid in by investors. The model described in this paper aims to account for the empirical distribution of this total-value-to-paid-in (TVPI) ratio at any point in time. This return measure can be interpreted as the fund’s size, normalised by the contributions made by the fund’s investors. Similar to the IRR, this ratio exists from the first observed cash inflow to the fund and can be calculated at any point during the fund’s life time.

At the start of a fund’s life, its first valuation typically occurs when the fund makes its first investment and enters some value for it on its balance sheet. In this paper, a fund’s birth is defined as the moment a value for its portfolio is observed. Capital is drawn down from investors when the fund makes its first capital call to finance this investment.\textsuperscript{4} In practice, a fund’s “vintage year” is typically defined as the year of the first drawdown of capital. The following investment period generally lasts for about three to six years depending on the fund’s strategy. Additional capital may be called down after this initial period to meet any capital needs of existing portfolio companies and to pay expenses of the fund (Blackstone, 2020; Gilligan & Wright, 2020).

During the investment period, additional valuations can be observed whenever the fund manager updates the book value of existing investments or adds new ones to the fund’s portfolio. Valuation events, such as disposals of portfolio companies or additional funding rounds in existing portfolio companies, typically lead to updates of the fund’s net asset value. In the absence of valuation events, the trajectory of underlying portfolio value is not observed, which often creates a time series of seemingly smooth or “stale” interim returns due to fund managers that hesitate to update fund values.

When investments are sold by the fund, it enters its distribution phase, during which investors receive the proceeds from the divestments of portfolio companies as distributions from the fund. This phase often overlaps the investment period as new investments are still being made while the fund crystallises returns from older investments. Investors may thus receive both distributions and capital during this period. Return observations become more frequent as the fund exists its investments and injects additional capital into existing portfolio companies in funding rounds that are typically priced.\textsuperscript{4} Once all investments have

\textsuperscript{3} Processes have been proposed in the literature whose stationary distribution allows for random starting points of new entities (e.g., a process with a reflective boundary as discussed by Gabaix (1999, 2009)). However, studies of these processes are typically limited to the right tail of a stationary distribution or assume that new entities are injected outside the region of interest. The analytical solution introduced in this paper describes the entire stationary distribution of a growth process in which new entities are injected randomly within the region of interest.

\textsuperscript{4} Some funds may not immediately invest the capital raised from investors or may value its investment at cost, both of which will lead to a valuation multiple of 1. As the fund pays for its ongoing expenses, its multiple may drop below 1. This initial period of stationary or decreasing multiples, also called the J-curve for the shape of the multiple curve over time, is not meaningful from the viewpoint of the model developed in this paper because no portfolio value has been observed yet.

\textsuperscript{4} Seed rounds in start-ups are often unpriced injections of convertible debt to avoid the costs associated with a company valuation and contracting between the parties.
been exited and proceeds distributed to fund investors, the fund is liquidated typically after about 7 to 15 years with seed and early-stage VC funds having longer life cycles due to the nature of their investments.

2.2. Characterising the distribution of fund returns

A fund’s life can be modelled as a random growth process with constant death of funds and rebirth at a random location and derive the closed-form stationary distribution of its valuation multiple. The new element in the process proposed here is a (re-)birth location (i.e., initial normalised size) that is not fixed but distributed normally with unconstrained location and scale parameters. Because real-life growth processes may produce new entities in random rather than fixed locations, the proposed process is not limited to private equity and may better describe observed datasets in other domains compared with existing models. While prior literature describes the tail behaviour of such funds having longer life cycles due to the nature of their investments.

models. While prior literature describes the tail behaviour of such observed datasets in other domains compared with existing phenomena if there is randomness in the initial size.

It is mathematically convenient to analyse the growth dynamics of fund value in terms of its logarithm, \( \log w \). Suppose that fund values evolve according to a standard diffusion process

\[ dx_t = \mu_G dt + \sigma_G dZ_t, \]

where \( \mu_G \) is the drift parameter and \( \sigma_G \) is the parameter scaling the standard Brownian motion \( Z_t \) (the subscript \( G \) stands for “growth” to distinguish it from the location and scale of the birth process introduced below, and using \( Z \) for the Wiener process to distinguish it from fund value, or wealth, \( w \)). The instantaneous growth of this process, \( dx_t \), thus follows a normal distribution.

We are interested in the stationary distribution \( p(x) \) of this process. As we will show below, the stationary distribution has an exponential tail

\[ P(x > x) \sim C e^{-\alpha x} \]

where \( C \) is a constant and \( \alpha > 0 \) is a simple function of the parameters \( \mu_G \) and \( \sigma_G \) (Gabaix, 2009). Equivalently, the distribution of fund value \( w \) has a Pareto tail,

\[ P(w > w) \sim C w^{-\alpha} \]

The observed right fat tail of the distribution of fund multiples can potentially be described by this power law.

If this growth process was used to describe the stationary distribution of fund multiples, however, the tail distribution must be cut off at some arbitrary value (typically \( x = 0 \)) because using the tail distribution function for one tail will typically not fit the other tail, making the standard random growth process inappropriate to fit the entire stationary distribution of fund multiples. The tail distribution may also be insufficient to describe an empirical distribution in which entities are born at random locations. More importantly, it fails to model the beginning and end of a fund’s life as important determinants of the stationary distribution of fund multiples.

Growth processes of private equity funds can be modelled by explicitly accounting for births and deaths of funds. When combined with the standard diffusion process described above, the forward Kolmogorov equation associated with this process describes the evolution of the distribution defined by the diffusion process over time (Gabaix et al., 2016, Eq. (5)):

\[ \frac{\partial}{\partial t} p(x,t) = - \frac{\partial}{\partial x} \left( \mu_G p(x,t) + \frac{\partial^2}{\partial x^2} \left( \frac{\sigma_G^2}{2} p(x,t) \right) - \delta p(x,t) + \delta \psi(x) \right) \]

Intuitively, the first term on the right-hand side describes the gain in density in locations with a locally decreasing density (i.e., the distribution is shifting in the direction of the drift parameter \( \mu \)), the second term increases the density at locations whose neighbourhood has a higher density and vice versa (similarly to the heat equation in physics), the third term describes a constant loss of entities across the distribution (e.g., firm death) at a rate of \( \delta \), and the last term creates new entities at a random location governed by the distribution \( \psi(x) \).

A contribution of this paper is to solve for the stationary distribution described by equation Eq. (1) by assuming a normal distribution for \( \psi(x) \). While arbitrary functional forms of \( \psi(x) \) typically do not lead to a closed-form solution, a normal distribution admits such a solution while being reasonably general for practical applications. In the remainder of this section, I will derive the stationary distribution and demonstrate its desirable properties as a model for private equity fund valuations.

Substituting the stationary distribution \( p(x,t) = p(x) \) in Eq. (1) gives

\[ 0 = - \mu_G p(x) + \frac{\sigma_G^2}{2} \psi(x) - \delta p(x) + \delta \psi(x) \]

where \( \psi(x) \) is the distribution of new-born entities that replace those that disappear \( (\delta p(x)) \), hence both terms are multiplied by the constant rate of death \( (\delta) \). The prime notation (e.g., \( p'(x) \)) denotes first and second derivatives with respect to \( x \).

Eq. (2) is a nonhomogeneous ordinary differential equation in \( p(x) \), whose solution consists of the solution to the corresponding homogeneous differential equation plus the particular solution for the term not involving \( p(x) \). The homogeneous equation corresponding to Eq. (2) is

\[ 0 = - \mu_G p(x) + \frac{\sigma_G^2}{2} \psi(x) - \delta p(x) \]

The solution to this equation can be found using the guess-and-verify approach. By guessing

\[ p(x) = Ce^{-\alpha x} \]

and substituting in (3) we find (Gabaix et al., 2016, Eq. 3)

\[ a_1 = -\mu_G + \sqrt{\mu_G^2 + 2\delta \sigma_G^2} \]

\[ a_2 = -\mu_G - \sqrt{\mu_G^2 + 2\delta \sigma_G^2} \]

The constants \( a_1 \) and \( a_2 \), when substituted in \( p(x) \), lead to the two equations comprising the steady-state solution of a random growth process in which dead entities are reborn at a constant location, typically \( x = 0 \).

\[ p(x) = \begin{cases} C e^{-\alpha x}, & x > 0 \\ C e^{-\alpha x}, & x < 0 \end{cases} \]

The constant \( C \) is used to normalise the distribution \( p(x) \). There is only a single constant \( C \) in this equation because the stationary distribution must be continuous at \( x = 0 \), and its integral must be finite (i.e., \( p(x) \) approaches zero as \( x \) approaches positive and negative infinity). In the remainder of this paper, assume without loss of generality that \( a_1 > 0 \) and \( a_2 < 0 \).

Eq. (5) is the double Pareto distribution discussed by (Gabaix, 2009) and Gabaix et al. (2016). The cut-off point at which both branches of the function meet can be chosen freely by expressing \( x \) relative to the cut-off point (i.e., replacing \( x \) with \( x - x^* \)). If the double Pareto distribution is fitted by maximum likelihood, this cut-off point can be estimated alongside the parameters \( a_1 \) and \( a_2 \), thereby fitting both the left and right tails of a distribution. However, at the fixed cut-off point, which represents the location at which new entities are injected into the distribution, the distribution is not smooth and may not well describe real-life phenomena if there is randomness in the initial size.

To find the solution of the nonhomogeneous differential Eq. (2) with a random birth location instead of a fixed one, the variation-of-parameters approach\(^5\) can be used to produce a closed-form solution.

First, rewrite Eq. (2) as
\[ p^s(x) - \frac{2\mu_c}{\sigma_c^2} p(x) + \frac{2\delta}{\sigma_c^2} p(x) = -\frac{2\delta}{\sigma_c^2} \psi(x) \]

and define the function \( r(x) = -\frac{2\delta}{\sigma_c^2} \psi(x) \) to simplify the right-hand side.

The two conditions required by the variation-of-parameters approach to solve for \( p(x) \) are
\[ u'(x)p_1(x) + v'(x)p_2(x) = r(x) \quad (6) \]
\[ u'(x)p_1(x) + v'(x)p_2(x) = 0 \quad (7) \]

where \( u \) and \( v \) are unknown functions of \( x \), and \( p_1 \) and \( p_2 \) are the solutions to the homogeneous differential equation in Eq. (5). The general solution to the steady-state Eq. (2) is then given by
\[ p(x) = u(x)p_1(x) + v(x)p_2(x) + Cp_1(x) + Cp_2(x) \quad (8) \]

Substituting (5) and (7) into (6) yields
\[ u' = \frac{r(x)}{(a_1 - a_2)} e^{-\delta x} = -r(x) e^{-\delta x} \quad (9) \]

These expressions must be integrated to find \( u \) and \( v \).

Suppose that the birth location (i.e., initial size) of entities is normally distributed with mean \( \mu_b \) and standard deviation \( \sigma_b \), and generalise the location of rebirth by substituting \( x - \mu_b \) for \( x \) in Eq. (5), such that
\[ u = \int u(x) \, dx = \int \left( \frac{2\delta}{\sigma_c^2} e^{(\delta - \sigma_b^2)} \right) \frac{1}{(a_1 - a_2)^2} \exp \left( -\frac{1}{2} \frac{(x - \mu_b)}{\sigma_b} \right) \, dx \]
\[ = \frac{\delta}{\sigma_c^2(a_1 - a_2)} \exp \left( a_1^2 \sigma_b^2 / 2 \right) \exp \left( \frac{a_1 \sigma_b^2 + \mu_b - x}{\sigma_b \sqrt{2}} \right) + \text{const}_1, \quad (10) \]

where \( \text{erf}(x) \) is the error function and \( \text{const}_1 \) is the constant of integration. Likewise,
\[ v = \int v(x) \, dx = -\frac{\delta}{\sigma_c^2(a_1 - a_2)} \exp \left( a_1^2 \sigma_b^2 / 2 \right) \exp \left( \frac{a_1 \sigma_b^2 + \mu_b - x}{\sigma_b \sqrt{2}} \right) + \text{const}_2. \quad (11) \]

We obtain the general solution by substituting these functions into Eq. (8):
\[ p(x) = \frac{\delta}{\sigma_c^2(a_1 - a_2)} \exp \left( a_1^2 \sigma_b^2 / 2 + \mu_b - x \right) \text{erf} \left( \frac{a_1 \sigma_b^2 + \mu_b - x}{\sigma_b \sqrt{2}} \right) + c_1 \exp \left( -a_1(x - \mu_b) \right) \]
\[ -\frac{\delta}{\sigma_c^2(a_1 - a_2)} \exp \left( a_1^2 \sigma_b^2 / 2 + \mu_b - x \right) \text{erf} \left( \frac{a_1 \sigma_b^2 + \mu_b - x}{\sigma_b \sqrt{2}} \right) + c_2 \exp \left( -a_2(x - \mu_b) \right) \quad (12) \]

Notice that the single constant \( C \) in the model with constant rebirth has been replaced with two constants, \( c_1 \) and \( c_2 \). These constants must be chosen appropriately for \( p(x) \) to be a well-behaved probability distribution. If we require that the probability of large positive or negative observations tends to zero,
\[ \lim_{x \to \pm \infty} p(x) = 0 \]

we find that
\[ c_1 = -\frac{\delta}{\sigma_c^2(a_1 - a_2)} \exp \left( a_1^2 \sigma_b^2 / 2 \right) \]
\[ c_2 = -\frac{\delta}{\sigma_c^2(a_1 - a_2)} \exp \left( a_2^2 \sigma_b^2 / 2 \right), \]

assuming again that \( a_1 > 0 \) and \( a_2 < 0 \).

The general stationary solution to the random growth process with normally distributed rebirth location is then given by substituting these constants into Eq. (12) and simplifying:
\[ p(x) = \frac{\delta}{\sigma_c^2(a_1 - a_2)} \exp \left( a_1^2 \sigma_b^2 / 2 + \mu_b - x \right) \left( \text{erf} \left( \frac{a_1 \sigma_b^2 + \mu_b - x}{\sigma_b \sqrt{2}} \right) - 1 \right) \]
\[ -\frac{\delta}{\sigma_c^2(a_1 - a_2)} \exp \left( a_2^2 \sigma_b^2 / 2 + \mu_b - x \right) \left( \text{erf} \left( \frac{a_2 \sigma_b^2 + \mu_b - x}{\sigma_b \sqrt{2}} \right) + 1 \right). \quad (13) \]

Eq. (13) is a smooth double Pareto distribution: It consists of a Pareto distribution for each tail with a smooth central part that connects both tails. For large positive values of \( x \) in the right tail, the first term approaches \( A_1 \exp(-a_1 x) \) for some constant \( A_1 \), while the second term approaches zero. Conversely, for large negative values of \( x \) in the left tail, the second term approaches \( A_2 \exp(-a_2 x) \) for some constant \( A_2 \), and the first term approaches zero. This shows that the logarithm of private equity fund multiples has exponential tails, which corresponds to the multiplex itself having Pareto tails. In the central part of the distribution, both Pareto distributions are smoothly mixed according to weights governed by the two error functions.

Contrary to the process with a constant rebirth location described by Gabaix (2009), this distribution does not require the splicing together of two separate functions at the point of rebirth. As expected from a (re-) birth distribution that is differentiable everywhere in combination with a standard diffusion process, the stationary distribution is smooth everywhere, too. One can show that this distribution converges to the double Pareto distribution with constant rebirth if the variation of this location (\( \sigma_b \)) tends towards zero and thus includes the standard double Pareto distribution described by Gabaix (2009) as a limiting case.

2.3. Estimating the stationary distribution

While Eq. (13) provides a closed-form solution for the stationary distribution of the growth process, estimation of its five parameters requires them to be uniquely identified. This is not the case for the parameters governing the growth component of the process: \( \mu_G \), \( \sigma_G \) and \( \delta \). These parameters are only identified up to scale.

To see this, consider the pre-factor in Eq. (13),
\[ \frac{\delta}{\sigma_c^2(a_1 - a_2)} = \frac{\delta}{2 \sqrt{\mu_G^2 + 2 \sigma_G^2}} \]
which has been simplified by plugging in the expressions for \( a_1 \) and \( a_2 \). If we now scale the death rate \( \delta \) and the drift parameter \( \mu_G \) by a constant \( t \) and the diffusion parameter \( \sigma_G \) by \( \sqrt{t} \), these terms cancel in the pre-factor of the stationary distribution,
The terms $a_1$ and $a_2$ are similarly unaffected by this scaling scheme. This scaling behaviour of the stationary distribution is expected because units in which to measure the growth process have not been defined. Drift and diffusion are typically expressed per unit time and grow linearly and as the square root of time, respectively, if the measurement interval is lengthened. Therefore, many combinations of the parameters $\mu_G$, $\sigma_G$ and $\delta$ result in the same stationary distribution. If the growth rate (and its standard deviation) is increased, for example, the stationary distribution is not affected if entities die at some higher rate.

For estimation purposes, the three parameters $\mu_G$, $\sigma_G$ and $\delta$, which describe the drift and diffusion of the growth process and the death rate of funds, represent only two degrees of freedom. If we let $t \equiv 1/\sigma_G^2$, Eq. (13) can be estimated by maximum likelihood using

$$
\frac{t \delta}{2 \sqrt{(\mu_G)^2 + 2(\delta)(\sqrt{1/\sigma_G^2})^2}} - \frac{t \delta}{2 \sqrt{\mu_G^2 + 2\delta^2\sigma_G^2}} - \frac{\delta}{2 \sqrt{\mu_G^2 + 2\delta^2\sigma_G^2}}
$$

where $a_1 = -\beta_{\mu_G} + \sqrt{\beta_{\mu_G}^2 + 2\beta_{\sigma_G}}$, $a_2 = -\beta_{\mu_G} - \sqrt{\beta_{\mu_G}^2 + 2\beta_{\sigma_G}}$.

The death rate ($\delta$) and growth rate ($\mu_G$) are only identified up to scale through the parameters $\beta_{\mu_G} \equiv \delta/\sigma_G^2$ and $\beta_{\sigma_G} \equiv \mu_G/\sigma_G^2$. The location ($\mu_G$) and scale ($\sigma_G$) of the birth process are always identified because the (re)birth distribution does not depend on time.

3. The distribution of private equity returns

Private equity fund returns may exhibit power law tails for a number of reasons. Explanations may be found in mechanisms based on the interplay of imitation and innovation in firms’ pursuit of productivity growth (Konig, Lorenz, & Zilibotti, 2016), repeated ties in networks (Kogut, Urso, & Walker, 2007) or innovation-driven multiplicative growth (Ghiglino, 2012). This section tests whether the stationary distribution of returns can be explained by a simple random walk with a lognormally distributed initial state as described in the previous section.

3.1. Sample and method

The main sample consists of all 3332 funds from Preqin’s private equity performance database that were alive (i.e., not liquidated) in April 2020. Because the most recent performance information for December 2019 is available only for a minority of funds ($N = 846$), the sample also includes funds whose most recent performance observation is December 2018 ($N = 2486$). Samples from different points in time can be combined if the stationary assumption for the distribution holds. By definition of the stationary distribution, two subsamples will have the same distribution. Each fund is observed exactly once. To filter out funds that have not made their first investment yet and thus have not begun their life in the sense of the model described in the previous section, funds must have called at least 10% of the capital committed by investors. This threshold serves as a proxy for funds having made an initial investment, as underlying individual investments are not observed in this dataset.

The variable of interest is the natural logarithm of a private equity fund’s net multiple, or total-value-to-paid-in multiple,

$$
\log(TVPI_t) = \log \left( \frac{\text{NAV}_t + \sum_{i} D_{it}}{\sum_{i} C_{it}} \right)
$$

where $C_{it}$ is the contribution by investors to fund $i$ in period $t$ (i.e., capital called by the fund), $D_{it}$ are the cash distributions by the fund to its investors, and $\text{NAV}_t$ is the net asset value of the fund’s unrealised portfolio. Sums are taken over a fund’s entire history. In other contexts, this ratio can also be thought of as the normalised size of a firm or city, or wealth. Because it is composed of cash distributions and net asset values, it contains both realised and unrealised gains. For the purposes of this paper, “return”, “valuation multiple”, “net multiple” and “return multiple” refer to this ratio. The observed average net multiple is 1.60 (median 1.44) with a range from a minimum of 0.06 to a maximum of 89.99. In logarithmic terms, both the mean and median are equal to 0.36.

Other observed fund characteristics include the fund’s age, stage focus and geographic region. Funds are 7.6 years old on average (median 7, max. 24, std. dev. 5.02). Buyout funds represent the most common stage focus ($N = 1084$), followed by venture capital ($N = 828$) and funds of funds ($N = 626$). Other types include growth-stage funds ($N = 274$) and funds focussing on secondary transactions ($N = 166$). Due to the history of both the private equity industry and Preqin, the most frequent geographic location is North America ($N = 2065$), followed by Europe ($N = 769$) and Asia ($N = 305$), the remainder of funds investing in other regions or across multiple regions.

The goodness of fit of a smooth double Pareto distribution estimated through Eq. (14) can be assessed against the empirical distribution, as well as against other candidate distributions such as the standard double Pareto distribution and a lognormal distribution. Following the literature, I employ a Kolmogorov-Smirnov test for the hypothesis that the observed data could have been generated by the fitted distribution. In other words, this test ensures that a fitted distribution describes the data sufficiently well. To assess the relative goodness of fit for competing distributions, I use a likelihood ratio test following Clauset et al. (2009), who adopt the methodology of Vuong (1989).

3.2. Results

The smooth double Pareto distribution (i.e., a double Pareto distribution with random birth) produces a good fit to the data. Fig. 1 shows a histogram of the full sample alongside three fitted candidate distributions. As expected, the double Pareto distribution with random birth produces a smooth central part while the standard double Pareto distribution has difficulties fitting the data. A normal distribution of fund multiples, which corresponds to a lognormally distributed value multiple, can be excluded as a plausible candidate by visual inspection.

A small spike with a region of higher-than-expected frequency can be seen around zero, corresponding to funds valued at a net multiple of 1. These seem to be mainly funds that have not realised any investments yet. If we require funds to have distributed gains from realised investments to enter the sample, this elevated frequency shrinks and aligns better with the fitted distribution.

Model statistics in Table 1 indicate a good fit of the smooth double Pareto distribution, which achieves the best fit among the three distributions tested. It fits the baseline sample significantly better than the standard double Pareto distribution. This better fit is also more plausible for theoretical reasons, as there is no theoretical justification for entities to appear at a precise location (e.g., 0.372 as estimated in model 1). Using different proxies for whether a fund has made some investments yields similar results. Alternative sample inclusion criteria consistently show the best fit for the smooth double Pareto distribution. The condition of having returned some cash to investors is the tighter constraint in the sense that requiring any positive distribution markedly reduces the sample size, whereas adding a constraint that at least 10% of the capital
The death rate reported at the bottom of the table is the maximum-likelihood estimate of the proportion of funds leaving the sample each year. Model 1 includes all of the fund.

This figure shows log-log plots of the logarithm of the cumulative distribution (left) and complementary cumulative distribution (right) as a function of the logarithm of the fund’s (log) net multiple.

Table 1

Parameter estimates for random-birth distribution and double Pareto distribution.

<table>
<thead>
<tr>
<th></th>
<th>(1) Baseline sample (called capital ≥10%)</th>
<th>(2) Realised &gt;0</th>
<th>(3) Realised &gt;0 &amp; called capital ≥10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef. S.E.</td>
<td>Coef. S.E.</td>
<td>Coef. S.E.</td>
</tr>
<tr>
<td>Panel A: Random birth location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drift ((\beta_0))</td>
<td>-0.032 (0.083)</td>
<td>-0.074 (0.077)</td>
<td>-0.070 (0.076)</td>
</tr>
<tr>
<td>Death rate ((\delta))</td>
<td>6.571 (0.326)</td>
<td>6.227 (0.280)</td>
<td>6.223 (0.279)</td>
</tr>
<tr>
<td>Birth mean ((\mu_0))</td>
<td>0.362 (0.011)</td>
<td>0.415 (0.010)</td>
<td>0.415 (0.010)</td>
</tr>
<tr>
<td>Birth Std. Dev. ((\sigma_0))</td>
<td>0.129 (0.014)</td>
<td>0.072 (0.017)</td>
<td>0.070 (0.017)</td>
</tr>
<tr>
<td>Right tail exp. ((a_1))</td>
<td>3.657 (0.120)</td>
<td>3.604 (0.113)</td>
<td>3.599 (0.113)</td>
</tr>
<tr>
<td>Left tail exp. ((a_2))</td>
<td>-3.593 (0.125)</td>
<td>-3.456 (0.108)</td>
<td>-3.459 (0.107)</td>
</tr>
<tr>
<td>Observations</td>
<td>3332</td>
<td>2848</td>
<td>2836</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>-1612.65</td>
<td>-1307.09</td>
<td>-1299.34</td>
</tr>
<tr>
<td>Panel B: Double Pareto</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right tail exp. ((a_1))</td>
<td>3.418 (0.086)</td>
<td>3.556 (0.095)</td>
<td>3.552 (0.095)</td>
</tr>
<tr>
<td>Left tail exp. ((a_2))</td>
<td>-3.252 (0.080)</td>
<td>-3.312 (0.086)</td>
<td>-3.322 (0.086)</td>
</tr>
<tr>
<td>Birth location ((s^*))</td>
<td>-0.372 (0.008)</td>
<td>-0.424 (0.008)</td>
<td>-0.424 (0.007)</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>-1630.19</td>
<td>-1311.68</td>
<td>-1303.63</td>
</tr>
<tr>
<td>Panel C: Fit tests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-value KS-test RB</td>
<td>0.0907</td>
<td>0.2111</td>
<td>0.2084</td>
</tr>
<tr>
<td>P-value KS-test DP</td>
<td>0.0382</td>
<td>0.1256</td>
<td>0.1281</td>
</tr>
<tr>
<td>P-value KS-test LN</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>P-value RB vs DP</td>
<td>0.0046</td>
<td>0.1044</td>
<td>0.1153</td>
</tr>
<tr>
<td>P-value DP vs LN</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Death rate / year</td>
<td>0.1308</td>
<td>0.1163</td>
<td>0.1158</td>
</tr>
</tbody>
</table>

This table shows fit statistics and parameter estimates for the stationary distribution of a random growth process with a fixed rebirth location (double Pareto, DP) and with a random birth location (RB, or smooth double Pareto). The variable being modelled is the natural logarithm of private equity funds’ valuation multiples as defined in the text. Fit statistics are also provided for a lognormal distribution. P-values are shown for Kolmogorov-Smirnov tests of the null hypothesis that the observed sample is generated by the candidate distribution. Likelihood ratio tests show results for the hypothesis that the first distribution mentioned does not fit better than the second one mentioned.

Comparisons of smooth double Pareto distributions against lognormal distributions (LN) yield the same qualitative results as comparison of standard double Pareto distributions against lognormal distributions and are thus not shown. Drift and death rate in Panel A are only determined up to scale. The death rate reported at the bottom of the table is the maximum-likelihood estimate of the proportion of funds leaving the sample each year. Model 1 includes all funds with capital called ≥10% of their committed capital, model 2 requires funds to have distributed some cash to investors, and model 3 combines both criteria.

Standard errors are shown in parentheses. Standard errors for tail exponents in Panel A are computed using the Delta method. Significance levels: *** p < 0.01, ** p < 0.05, * p < 0.1.

has been called does not reduce the sample size much further. Because model fit is nearly identical, I focus my analysis on the larger sample.

Surprisingly, the distribution is dominated by the diffusion element of the growth process, as well as the randomness of its birth location. The estimated growth component is insignificant at –0.032. To understand the economic size of this estimate, recall that coefficients are identified only up to scale. If we assume an annual standard deviation of 20% and use \(\mu_0 = \mu G/\sigma G^2\) from the previous section, the mean drift is \(\mu_\delta = \mu G\cdot \sigma G^2 = -0.032 \cdot 0.2^2 = -0.00128\) or –0.128% per year, which is not significantly different from zero. Similarly, the death rate annualised using this annual standard deviation is \(\delta = \mu_\delta/\sigma_\delta = 6.571 \cdot 0.2^2 = 0.26284\). When funds are born, they appear around a logarithmic multiple of 0.362 with a standard deviation of 0.129. These results show that funds achieve exceptionally high (and exceptionally low) multiples through the diffusion component of the growth process, not by drifting to the upper (or lower) tail.
If the rate at which entities die per time period is known, this parameter can be used to constrain the remaining parameters in the full density function (12). Because the age of private equity funds is available from Preqin, its distribution can be used to estimate their rate of death. Using a rate of 13.08% per year from Table 1, it is possible to deduce the mean growth rate and diffusion rate of the underlying process. As described in the previous section, \( \beta_a \equiv \delta/\sigma_G^2 \) and \( \beta_p \equiv \mu_G/\sigma_G^2 \), which gives \( \sigma_G^2 = \delta/\beta_a = 0.1308/6.571 = 0.0199 \) and thus \( \sigma_G = 0.1411 \). This result can also be obtained by estimating Eq. (12) directly while constraining the death rate to 0.1308. The advantage of estimating the 5-parameter equation is that standard errors are available from the estimation procedure and do not need to be calculated ex-post (e.g., via the delta method).

Tail exponents for the entire private equity asset class are in the range 3.2–3.6. The standard double Pareto distribution produces a right-tail exponent of 3.4 and a left-tail exponent of 3.2. These values are similar to those obtained from a smooth double Pareto distribution. Inserting the parameter estimates from Model 1 in Table 1 into Eq. (4), we have

\[
a_1 = -\beta_p + \sqrt{\beta_p^2 + 2\beta_p},
a_2 = -\beta_p - \sqrt{\beta_p^2 + 2\beta_p}
\]

\[
a_1 = 0.032 + \sqrt{0.032^2 + 2 \times 6.571} = 3.6573,
a_2 = 0.032 - \sqrt{0.032^2 + 2 \times 6.571} = -3.5933
\]

The similarity to the standard double Pareto case is expected because both tails are ultimately dominated by the double-Pareto component of the distribution. The random-birth component in the central part of the distribution is lognormal and rapidly decreases in the tails. This similarity in the tails can also be seen in Fig. 2, which shows both tails of the full sample of private equity valuation multiples. In the left tail, the graph again shows a region just below zero with fewer than the expected number of observations.

Extreme observations in both tails are found some distance from their estimated location, while the central body of the distribution fits almost perfectly as shown in Fig. 1 and Fig. 2. These outliers can be an indication that observations are not generated by the same underlying process. Inspection of individual observations in the tails reveals that most extreme observations are venture capital funds. All funds with a log multiple >2 are venture capital or growth-stage funds. Inspection of outliers in the dataset, as well as a manual additional internet search, does not produce any reasons to suspect that outliers may be spurious.

### 3.2.1. Distribution by fund strategy

The accuracy of fitted distributions can be improved by separately fitting subsamples for different fund strategies. The business model of venture capital funds, which typically invest in seed-stage, early-stage and pre-IPO companies, can be expected to generate returns with greater variability than buyout funds, growth funds or funds of funds. Evidence suggests that later-stage investments are less risky (Ang et al., 2018; Cochrane, 2005; Driessen et al., 2012). Returns of a fund of funds, however, may converge to venture returns in the tails if venture funds are part of its portfolio due to the fact that the heavier tails of venture capital returns will eventually dominate other portfolio investments in both tails.

Tail plots for venture capital multiples in Fig. 3 show a better fit if venture capital multiples are analysed separately. Value multiples of non-VC funds produce a better fit, too, if analysed separately as shown in Fig. 4. This finding supports the view of the main sample as a mixture of at least two separate growth processes.

Tail exponents are much smaller for venture capital funds compared to non-VC funds. As shown in Figs. 3 and 4, the fitted line is much steeper in the tails for VC funds. The exact results are shown in Table 2. Tail exponents in VC funds are 2.2 and 2.5 in the right and left tail, respectively (2.3 and 2.6 for the fitted smooth double Pareto distribution). The distribution of non-VC funds, by contrast, has tail exponents between 3.8 and 4.0 (standard double Pareto) or 4.7 and 4.8 (smooth double Pareto).

These differences in tail behaviour between VC and non-VC funds can have important implications for entrepreneurs’ and investors’ decision making. The variance of a power-law distribution is infinite for an exponent of \( a < 3 \) in the right tail. Similarly, the third moment is infinite for \( a < 4 \). In general, any moment \( k \) of a power law exists only if \( k < a - 1 \). This implies that the mean, variance, skewness and possibly kurtosis are finite for non-VC funds. But skewness and kurtosis and most likely – and most importantly – the variance of VC returns are infinite in the stationary distribution. Tail exponents for venture capital are not much bigger than 2 in absolute value as shown in Table 2, indicating value multiples with infinite variance. In addition, the right tail in particular has a third moment that may be unstable and slow to converge to its population value (if it exists) in any finite sample.

The finding of power-law behaviour in fund multiples and infinite variance in the VC subsample is in agreement with the power-law returns at the investment level suspected by Prencipe (2017). It is worth noting that he studies a distribution that is neither the stationary distribution nor the distribution of instantaneous growth but the distribution of investment returns over the entire duration of the

---

Fig. 3. Tail distribution of net multiple for venture capital funds.

This figure shows log-log plots of the logarithm of the cumulative distribution (left) and complementary cumulative distribution (right) as a function of the logarithm of the fund’s (log) net multiple. The sample contains only venture capital funds.
This figure shows log-log plots of the logarithm of the cumulative distribution (left) and complementary cumulative distribution (right) as a function of the logarithm of the fund’s (log) net multiple. The sample consists of all funds in the full sample that are not venture capital funds.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>(1) VC</th>
<th></th>
<th>(2) Non-VC</th>
<th></th>
<th>(3) Buyout</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>S.E.</td>
<td>Coef.</td>
<td>S.E.</td>
<td>Coef.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Panel A: Random birth location (RB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drift ($\mu$)</td>
<td>0.158</td>
<td>(0.105)</td>
<td><strong>-0.042</strong></td>
<td>(0.153)</td>
<td><strong>-0.496</strong></td>
<td>(0.338)</td>
</tr>
<tr>
<td>Birth Std. Dev. ($\sigma$)</td>
<td>0.255</td>
<td>(0.028)</td>
<td><strong>0.378</strong></td>
<td>(0.013)</td>
<td><strong>0.420</strong></td>
<td>(0.026)</td>
</tr>
<tr>
<td>Right tail exp. ($\alpha_1$)</td>
<td>2.300</td>
<td>(0.130)</td>
<td><strong>4.797</strong></td>
<td>(0.238)</td>
<td><strong>5.400</strong></td>
<td>(0.618)</td>
</tr>
<tr>
<td>Left tail exp. ($\alpha_2$)</td>
<td><strong>-2.615</strong></td>
<td>(0.163)</td>
<td><strong>-4.714</strong></td>
<td>(0.230)</td>
<td><strong>-4.407</strong></td>
<td>(0.350)</td>
</tr>
<tr>
<td>Observations</td>
<td>828</td>
<td></td>
<td>2504</td>
<td></td>
<td>1084</td>
<td></td>
</tr>
<tr>
<td>Log Likelihood</td>
<td><strong>-696.55</strong></td>
<td></td>
<td><strong>-790.57</strong></td>
<td></td>
<td><strong>-475.13</strong></td>
<td></td>
</tr>
<tr>
<td>Panel B: Double Pareto (DP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right tail exp. ($\alpha_1$)</td>
<td><strong>2.216</strong></td>
<td>(0.096)</td>
<td><strong>4.036</strong></td>
<td>(0.121)</td>
<td><strong>3.685</strong></td>
<td>(0.153)</td>
</tr>
<tr>
<td>Left tail exp. ($\alpha_2$)</td>
<td><strong>-2.474</strong></td>
<td>(0.112)</td>
<td><strong>-3.794</strong></td>
<td>(0.110)</td>
<td><strong>-3.146</strong></td>
<td>(0.122)</td>
</tr>
<tr>
<td>Birth location ($x^*$)</td>
<td><strong>0.260</strong></td>
<td>(0.011)</td>
<td><strong>0.390</strong></td>
<td>(0.008)</td>
<td><strong>0.425</strong></td>
<td>(0.009)</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td><strong>-698.78</strong></td>
<td></td>
<td><strong>-824.64</strong></td>
<td></td>
<td><strong>-510.71</strong></td>
<td></td>
</tr>
<tr>
<td>Panel C: Fit tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-value KS-test RB</td>
<td>0.4430</td>
<td></td>
<td>0.7135</td>
<td></td>
<td>0.6186</td>
<td></td>
</tr>
<tr>
<td>P-value KS-test DP</td>
<td>0.4260</td>
<td></td>
<td>0.0045</td>
<td></td>
<td>0.0028</td>
<td></td>
</tr>
<tr>
<td>P-value KS-test LN</td>
<td>&lt; 0.0001</td>
<td></td>
<td>0.0003</td>
<td></td>
<td>0.1650</td>
<td></td>
</tr>
<tr>
<td>P-value RB vs DP</td>
<td>0.2703</td>
<td></td>
<td>&lt; 0.0001</td>
<td></td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>P-value DP vs LN</td>
<td>&lt; 0.0001</td>
<td></td>
<td>0.0058</td>
<td></td>
<td>0.3947</td>
<td></td>
</tr>
<tr>
<td>Death rate / year</td>
<td>0.1216</td>
<td></td>
<td>0.1342</td>
<td></td>
<td>0.1351</td>
<td></td>
</tr>
</tbody>
</table>

This table shows fit statistics and parameter estimates for the stationary distribution of a random growth process with fixed rebirth location (double Pareto) and with random birth location (smooth double Pareto). The variable being modelled is the natural logarithm of private equity funds’ valuation multiples as defined in the text.

The main sample is split by fund strategy: all venture capital funds in model 1, non-VC funds in model 2 and buyout funds in model 3. P-values are shown for Kolmogorov-Smirnov tests of the null hypothesis that the observed sample is generated by the candidate distribution. Likelihood ratio tests show results for the hypothesis that the first distribution mentioned does not fit better than the second one mentioned. Comparisons of smooth double Pareto distributions against lognormal distributions (LN) yield the same qualitative results as comparison of double Pareto distributions against lognormal distributions and are thus not shown. The death rate reported at the bottom of the table is the maximum-likelihood estimate of the proportion of funds leaving the sample each year. Standard errors are shown in parentheses. Standard errors for tail exponents in Panel A are computed using the Delta method. Significance levels: ***, p < 0.01, **, p < 0.05, *, p < 0.1.

### Notes

- Investment from initial investment to exit through sale or write-off. Because returns in his sample have been observed over intervals of varying length, however, the resulting distribution is similar to the stationary distribution. If a process is in its stationary state, sampling investments only at the moment of exit still produces the stationary distribution if investments are exited (i.e., “die” in the terminology of this paper) at random. It is thus plausible that power-law returns at the fund level are generated by similar power-law behaviour at the firm level due to the aggregation properties of power-law distributions (Jessen & Mikosch, 2006).

- The feature of infinite variance and unstable skewness may attract founders and investors who see potentially infinite gains as a chance to find a unicorn in the distribution. Mathematically, the same applies to the left tail of the distribution, which in theory may balance the right tail if the likelihood of large gains is equal to the likelihood for near-total losses. However, this cancellation of large gains and losses may only be a concern if all wealth is concentrated in the portfolio, which may be nearly wiped out after having found a unicorn or, conversely, make an extraordinary gain after having nearly lost all value. If additional wealth is available outside the portfolio (e.g., through a separate portfolio, via
family or friends, or through an entrepreneur’s human capital), then entrepreneurs or investors may be able to replenish their portfolio using this external wealth to enable a new bet. From an entrepreneur’s viewpoint, for example, frequently the main investment is their time to work for the start-up, which can be redeployed into a new project should the start-up not turn into a unicorn.

Owning an option on an underlying asset with a potentially infinite volatility is more valuable than one with a finite volatility. It is important to keep in mind, however, that any value may only be realised after a long waiting time, since fund valuations are still governed by the diffusion process outlined above. In other words, if a VC fund is observed for a fixed length of time, its distribution will be lognormal and its return variance will be finite. This seeming contradiction between power laws in the (stationary) cross-section and a more benign lognormal distribution in the time series sense may be at the heart of the debate whether it is worthwhile investing in venture capital. A conclusion from our findings may be that unicorns can be found but only if entrepreneurs or investors are willing to wait.

This finding is in agreement with the generative process described by Huberman and Adamic (1999, 2000), in which power law tails are growing for a random amount of time when it is observed due to the maximum likelihood as before. As a further benefit of this approach, findings suggest that most systematic value creation in buyout funds occurs when investments are made, whereas VC funds add value continuously throughout the fund’s life.

The smooth double Pareto distribution is the only one that fits all three subsamples of private equity strategies in Table 2. In Kolmogorov-Smirnov goodness-of-fit tests, the distribution with random birth also performs significantly better than the double Pareto distribution with fixed birth location in the non-VC and buyout samples. These findings show that a random growth process with random birth size can accurately describe valuation multiples in the private equity asset class.

### 3.2.2. Fund strategy as a covariate

The previous section uses subsamples of the baseline sample to compare fitted smooth double Pareto distributions across fund strategies. These separate models can be combined into a single model for the full sample by describing the four identified parameters in Eq. (14) as linear functions of fund strategies and estimating the parameters by maximum likelihood as before. As a further benefit of this approach,

#### Table 3

<table>
<thead>
<tr>
<th>Fund strategies as covariates.</th>
<th>(1) Fully parameterised</th>
<th>(2) Tail parameters only</th>
<th>(3) Intercepts only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>S.E.</td>
<td>Coef.</td>
</tr>
<tr>
<td>Equation for Birth mean ($\mu_0$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.232 (0.047)</td>
<td>***</td>
<td>0.366 (0.012)</td>
</tr>
<tr>
<td>VC (general)</td>
<td>0.094 (0.064)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late stage / growth</td>
<td>0.010 (0.061)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buyout</td>
<td>0.230 (0.053)</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>Fund of funds</td>
<td>0.235 (0.050)</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.151 (0.051)</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>Not North America</td>
<td>−0.081 (0.016)</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>Equation for Birth standard deviation ($\sigma_0$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.140 (0.043)</td>
<td>***</td>
<td>0.164 (0.011)</td>
</tr>
<tr>
<td>VC (general)</td>
<td>0.049 (0.071)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late stage / growth</td>
<td>0.060 (0.056)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buyout</td>
<td>0.117 (0.046)</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Fund of funds</td>
<td>−0.005 (0.046)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>−0.063 (0.055)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not North America</td>
<td>−0.025 (0.017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>−1365.99</td>
<td>−1420.63</td>
<td>−1612.65</td>
</tr>
</tbody>
</table>

This table shows models fitting valuation multiples for the baseline sample. Model 1 treats all four identified parameters of the smooth double Pareto distribution as a function of fund strategy (VC (general), Late stage / growth, buyout, fund of funds, other) and fund location. The omitted baseline fund stage category is Seed/early stage. Model 2 keeps birth location and birth standard deviation fixed and models the two parameters that determine the tails as functions of fund strategy and location. Model 3 is the baseline model from Table 1 for comparison. $N = 3332$. Standard errors are shown in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 

...
additional covariates can be added to the linear part of the model.

This leads to four equations,
\[ \beta_{\mu_0} = X_{T1}, \beta_{\mu_1} = X_{T2}, \sigma_\delta = X_{T3}, \mu_\delta = X_{T4}, \]

where \( X \) is a matrix of covariates including a constant and \( \gamma_{1,0,4} \) are the vectors of hyperparameters analogous to the parameters estimated for subsamples in Table 2. Subscripts indicating individual funds have been omitted. In the present case, the matrix of covariates includes dummy variables for fund strategy (Late stage / growth, Buyout, Fund of funds, Other; Seed/early stage as the omitted baseline category) and a dummy variable indicating whether a fund is located outside of North America. Covariates need not be the same for each of the parameters.

If the possibility that the standard deviation \( \sigma_\delta \) may become zero or negative during the numerical optimisation is a concern, \( \sigma_\delta \) in Eq. (14) may be reparametrized to avoid this scenario. This may be useful if it is suspected that the distribution is not smooth around the birth location but follows a standard double Pareto shape without variation in birth location (i.e., \( \sigma_\delta = 0 \)).

Results in Table 3 show that the drift rate standardised by the standard deviation of the growth process (\( \beta_{\mu_0} \)) is large of earlier stages in the portfolio companies’ life cycle. Buyout funds and funds of funds have negative growth rates, while growth rates for venture capital funds are indistinguishable from zero. Interestingly, North American funds grow faster than funds located elsewhere. Funds that invest in later stages also start their life with a higher valuation multiple, which suggests that a large proportion of any value creation in funds of funds and buyout funds occurs at the start of their life (i.e., when they initially invest), while venture capital funds create value more continuously.

As expected from results in Table 2, funds that invest in later stages exhibit more stable tail behaviour as indicated by the standardised death rate (\( \beta_{\mu_1} \)). The smallest tail exponents are found in seed and early-stage funds with a left tail exponent of \(-2.52 = - \beta_{\mu_0} - \sqrt{\beta_{\mu_0}^2 + 2\beta_{\mu_1} = -0.119 - \sqrt{0.119^2 + 2.2876} = -2.52} \) and a right tail exponent of 2.283. Valuation multiples of funds of funds have the tightest tail with exponents of \(-5.46 \) on the left and 7.96 on the right. The left tail exponent for generalist VC funds is almost identical at \(-2.48 \), while the right tail exponent of 3.02 is just at the boundary between a defined and undefined variance of the underlying valuation multiple. This result shows that usual OLS estimates of VC returns based on means and variances may not lead to econometrically stable results if returns of seed and early-stage funds in particular are sensitive to extreme observations in the tails. Estimation within a single model allows easy comparison of coefficients and shows that standardised death rates significantly drive the tail exponents across fund strategies. Intuitively, the large variation of growth rates in venture capital funds dominates their death rate and causes fat-tailed valuation multiples (via a small \( \beta_{\mu_1} \), while for other funds a relatively higher standardised death rate keeps valuation multiples closer to their birth value.

3.2.3. Shifted observations in the left tail?

One feature stands out in the left tail of the distribution of logarithmic valuation multiples: observations seem to be missing in a small region to the left of zero. This can be seen as the gap between the estimated density and the histogram in Fig. 1. This density drop in the left tail is more pronounced in the VC sample shown in Fig. 4 than in the full sample. At the same time, there is a surplus in a small region around a value multiple of zero.

Missing observations in one region and a surplus in another can be explained by funds deliberately shifting their valuations slightly to appear more profitable in a region that matters to investors (e.g., break-even). This behaviour would be similar to earnings management to beat benchmarks (Degeorge, Patel, & Zeckhauser, 1995). However, a simpler explanation involves funds not having acknowledged any return on investment on their balance sheet yet. These funds may not have made an investment yet or have booked their investments at cost due to the uncertain nature of VC investments. The s-shaped empirical distribution around zero, shown for venture capital funds in Fig. 5 in the left-hand panel, is much less pronounced for other types of funds.

When testing whether the spike near zero can be explained by funds without investments, a proxy for investment activity can offer some insights. The right-hand graph in Fig. 5 shows the distribution of value multiples for VC funds with the added constraint that funds must have made a positive cash distribution to their investors. Under the assumption that funds do not distribute cash to investors without having first realised an investment, this sample constraint accurately captures valuation activity that has an effect on funds’ value multiples. When fitting a smooth double Pareto distribution to this sample, the apparent effect of shifted valuations disappears.

A disadvantage of using investor distributions as a proxy for valuation events in the fund is its failure to capture funds that have not exited any investments yet. Moreover, it cannot identify funds that have successfully exited portfolio companies but chose to wait before distributing anything to investors. Defining a fund’s birth through this proxy thus fails to capture economically meaningful activity of importance to investors. Similar to the usual definition of the vintage year as the year of the first drawdown, the birth of a fund should coincide with the first purchase or generation of assets capable of exhibiting returns over time.

![Fig. 5. Distribution of all VC funds vs VC funds with distributions.](image-url)

The left-hand graph shows the empirical distribution and a fitted smooth double Pareto distribution for the full sample, corresponding to Fig. 1. The right-hand graph shows the same sample after removing funds that have not made any distributions to investors.
However, if it was possible to accurately identify the first investment of a fund and to define the fund’s birth in this way, this may create a sample of funds in which some funds do not update their valuations despite having made an investment, causing a spike in the distribution of (log) value multiples around zero. There is a trade-off between capturing all meaningful economic activity and observing a distribution that does not include “stale” multiples of funds that have made investments but chose to keep them at cost. An alternative would be to formally model the mechanism by which funds remain at zero and then jump to the first meaningful valuation as a jump process.

3.3. Is fund growth lognormally distributed?

The good fit of the smooth double Pareto distribution to the fund multiple data lends support to the hypothesis that the stationary distribution of fund multiples is produced by the drift-diffusion process assumed in this paper. This indirect test of the data-generating process is justified by the slow-moving nature of funds’ valuation multiples, which partly consist of unrealised gains in portfolio companies held by the fund. Because of the valuation practices of fund managers, published net asset values may not always reflect underlying portfolio values in a timely manner but may be smoothed over time. In this case, the underlying distribution of fund growth can only be observed over longer time periods.

But can the assumption of lognormally distributed growth be tested more directly? If the initial distribution of valuation multiples is lognormal, and underlying growth is lognormal too, then the distribution of valuation multiples of funds with the same age should be lognormal. This is a property of the normal distribution, which is stable in the sense that a linear combination of normally distributed random variables produces another normal distribution. It should therefore be possible to test slices of the fund sample with a given age for normality. More generally, any return over a fixed period should be lognormally distributed, not only returns since the inception of the fund.

The Preqin fund dataset used in this paper can be decomposed into subsamples of funds that have the same age. The total sample for this test includes all fund-year observations in the Preqin dataset – as opposed to funds that were alive in early 2020, which are used in the main analyses above – with the constraint that funds have called at least 10% of their capital to select funds with meaningful economic activity.

To test for normality of log-growth, I explore subsample distributions using quantile-quantile plots against the normal distribution and use Shapiro-Wilk tests, as well as Kolmogorov-Smirnov tests. Because of the potential smoothing of net asset values, I first run tests with one-year growth and then test growth over a ten-year horizon. The latter sample should be less prone to the distortive effects of smoothing. There are between 1551 and 2242 funds in the subsamples of one-year fund growth. For tests of ten-year growth, subsamples contain between 270 (VC subsample) and 772 funds (all funds).

Results for logarithmic one-year multiple growth shown in Fig. 6 indicate a poor fit of the normal distribution. Distributions for all four age strata analysed (i.e., funds that are 0, 3, 7, 12 years old) show heavy tails compared with the normal distribution, which are particularly pronounced for newly established funds. Normality is rejected by
Shapiro-Wilk and Kolmogorov-Smirnov tests at \( p < 0.001 \) in all subsamples. This finding can be explained by short-term fund growth being significantly non-normal or underlying growth being opaque over short horizons.

If underlying (i.e., unobserved or latent) growth is normal, then its observed distribution should become more normal when measured over longer periods. Fig. 7 shows QQ plots similar to those in Fig. 6, but using a ten-year horizon instead of one-year growth. The distribution for newly born funds is still significantly non-normal. However, growth in middle-aged funds looks normal in the upper tail. This finding is true for both venture capital and buyout funds. The left tail remains non-normal, and thus normality is again rejected in all subsamples at \( p < 0.001 \). Similarly pronounced tails that appear to fit a Laplace (i.e., double-exponential) distribution have been observed for one-year growth of firms’ sales, which become thinner when a seven-year horizon is used (Bottazzi & Secchi, 2006, p. 240).

In conclusion, these findings suggest that smoothing of fund net asset values impedes direct normality tests of fund growth at short horizons, but growth may be normal in the upper tail. Overall, departures from normality are not severe enough to reject the diffusion process as the data-generating process in the main analysis in Section 2.2, but exploration of age-based subsamples reveals departures from normality particularly at the beginning of a fund’s life and in the lower tail. It may be possible to model the growth of funds more accurately throughout their life cycle with a jump process for young funds in addition to the standard drift-diffusion process used in this paper. Jumps may, for example, be caused by sparse valuation events in newly established funds at the beginning of their investment period, while mature funds have more frequent valuation events, leading to a smoother and thus more normally distributed growth trajectory. The pronounced left tail calls for a more in-depth theoretical and empirical investigation, as it seems to be a stable feature across subsamples.

4. Conclusion

Findings in this paper show that the stationary distribution of a random growth process with random initial value fits the empirical distribution of private equity valuation multiples. This new distribution – a smooth double Pareto distribution – features a smooth central body that can fit the entire distribution where alternative distributions can only be applied to the tails. It enables the description of growth processes in a range of domains outside private equity that feature a random initial size or value.

The underlying growth process produces Pareto tails in the stationary cross-section, while individual entities exhibit lognormal growth in the time series. This finding supports the assumption of lognormal growth between sparse price observations that is typically found in studies of private equity fund pricing. Additional tests of fund growth for funds of the same age show that observed growth is not lognormal for short time horizons but becomes more lognormal for longer horizons. The smooth double Pareto distribution’s accurate fit thus narrows the bounds on the likely processes generating private equity returns. From a practical viewpoint, its capability to model the entire distribution of valuation multiples rather than being restricted to the tails facilitates...
incorporation into risk management frameworks. Subsample tests of venture capital and buyout funds suggest that the overall distribution of fund returns may be composed of several underlying distributions with distinct tail exponents and birth processes. Venture capital funds in particular show relatively small tail exponents corresponding to infinite variance, which may be a deceptive appealing characteristic for investors. However, fat tails in the cross-section can be explained without having to invoke a fat-tailed growth process for individual entities. A growth process with random death and random birth – and thus random observation of individuals – is sufficient to generate the power law.

An established and growing literature on size distributions in various fields of research might benefit from the increased accuracy afforded by the smooth double Pareto distribution to processes in which initial size is not fixed or constrained to a region below a reflective boundary. Further research may also aim to formally incorporate the dynamics of funds whose first technical observation is a cash pool with a valuation multiple of 1, which then jumps to a more economically meaningful valuation multiple when making or revaluing their first investment.

Data availability

The authors do not have permission to share data.

Acknowledgements

I am grateful for comments from Xavier Gabaix, Rodion Skovoroda, Ze Chen, conference participants at the 2021 European Financial Management conference and the 2021 British Accounting and Finance Association meeting, as well as seminar participants at The Open University. All errors and omissions are my own.

References


