The smooth double Pareto distribution: A model of private equity fund returns

December 2020

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,

henry.lahr@open.ac.uk

Abstract: Whether returns of venture capital and private equity investments exhibit fat tails, particularly in the upper tail, affects how entrepreneurs and investors view the attractiveness of such investments. Using fund performance data, we propose and test a random growth model with a lognormally distributed diffusion process and lognormally distributed initial valuation of funds to explain the observed distribution of funds' residual-value and payout multiples. This model endogenously generates power-law tails in the cross-section. We find that the resulting smooth double Pareto distribution fits the data better than competing lognormal or ordinary double Pareto models. Fat tails are particularly pronounced in seed stage, early stage and generalist venture capital funds and suggest returns with infinite variance. The smooth double Pareto distribution has wide applicability to growth processes with a random initial size.

Keywords: size distribution; income distribution; Pareto law; power-law distribution; fat tails; private equity; venture capital; financial returns

JEL codes: C46; D31; G24; G32; L11; R12

1. Introduction

The purpose of this paper is to introduce the smooth double Pareto distribution, which can describe random growth processes with a random initial size. This distribution has Pareto tails and a central body that smoothly connects the tails. It can be applied to a wide range of phenomena, including people's incomes 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, 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 $\mathrm{P}(W > x) = k / x^{\alpha}$ for some constant *k*. The special case of $\alpha \approx 1$ 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 α is some positive real number. Research has established power laws in firm sizes, city sizes, short-term stock price fluctuations and trading volume (Gabaix, 2009). The tail of the distribution of the productivity of 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 et al., 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 et al., 2009).

Mechanisms known to generate power laws are preferential attachment (Yule, 1925; Simon, 1955), random growth with a reflective barrier (Champernowne, 1953; Gabaix, 1999) or death (Luttmer, 2007; Gabaix, 2009) to stabilise the distribution, self-organized criticality (Bak et al., 1993), optimisation (Mandelbrot, 1953), random observation (Huberman and Adamic 1999; Huberman and Adamic, 2000), and exponential technological progress combined with exponential technology diffusion (Hilbert, 2013). Mitzenmacher (2003), Newman (2005) and Gabaix (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 and Thiel, 2014). Power-law tails would be an attractive feature for entrepreneurs and investors because they can offer extraordinary gains from a modest investment. The continued search for "unicorns", startup companies with a value of greater than $1bn, through seed and early-stage investments could be justified by rational investors even if historical samples provide 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 may 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 tests based on a (right) tail distribution is to find a cut-off point above which the power law is presumed to hold. If part of the distribution cannot be explained by the power law to be tested, this threshold may 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 thus indistinguishable from a power law.

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 in that growth in value for subsequent periods is lognormally distributed (Ang et al., 2018; Korteweg and Nagel, 2016; Driessen et al., 2012; Cochrane, 2005).

The objective of this paper is to model the cross-section of private equity returns in its entirety, taking into account the life cycle of private equity funds and potential power-law behaviour in the tails. By fitting the new smooth double Pareto distribution to private equity data, the assumption of lognormally distributed returns can be tested.

We derive the stationary return distribution from a standard diffusions process that is augmented by a lognormally distributed initial value to match the investment process of private equity funds. 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 Gabaix (2009) and Gabaix et al. (2016). However, this distribution features a discontinuity at the fixed point at which new entities are continuously added that does not seem to agree with observed private equity valuations. 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 and 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 supports in the data.

This paper contributes to the theoretical literature on random growth models (Gabaix et al., 2016, 1999; Toda and Walsh, 2015; Reed, 2003; Reed and Jorgensen, 2001) by deriving the stationary distribution of a random growth process with random initial size and constant death. It has wide applicability to random growth processes in which the initial state can be described by a lognormal distribution, such as personal wealth (Benhabib, 2018), the size distributions of firms (Coad, 2009; Luttmer, 2007; Cabral and Mata, 2003) and settlements (Rozenfeld et al., 2011). The results further contribute to the literature on private equity returns by establishing bounds on the likely processes generating the distribution of returns. We 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 and Nagel, 2016).

Are private equity returns distributed according to a power-law? Probably. But this does not mean, however, 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 growth 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 make the cross-sectional distribution's tails heavier.

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

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 simpler measure that is often used in the private equity industry and which we adopt in this paper is the ratio of distributions by the fund 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. It can be interpreted as the fund's size, normalised by the contribution made by the fund's investors. Similar to 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 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.[1] In practice, a fund's "vintage year" is typically defined as

[1] 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

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.

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.[2] Once all investments have been exited and proceeds distributed to fund investors, the fund is liquidated.

2.2. Characterising the entire distribution of returns

In this section, we model the events in a fund's life as a random growth process with constant death of entities and rebirth at a random location and derive its closed-form stationary distribution. The new element in this process 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

---

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 our model because no market value has been observed yet.

[2] 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.

processes may produce new entities in random rather than fixed locations, the proposed process is more general and may better describe observed datasets compared with existing models. While prior literature describes the tail behaviour of such growth processes, the goal in this section is to describe the entire stationary distribution of this growth process, including its central part.

It is mathematically convenient to analyse the growth dynamics of fund value in terms of its logarithm, $x_{it} = \log w_{it}$. Suppose that fund values evolve according to a standard diffusion process

$$dx_{it} = \mu_G dt + \sigma_G dZ_{it},$$

where μ_G is the drift parameter and σ_G is the parameter scaling the standard Brownian motion Z_{it} (the subscript *G* stands for "growth" to distinguish it from the location and scale of the birth process described below, and using *Z* for the Wiener process to distinguish it from fund value, or wealth, *W*).

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

$$P\left(x_{it} > x\right) \sim Ce^{-ax},$$

where *C* is a constant and $a > 0$ is a simple function of the parameters μ_G and σ_G (Gabaix, 2009). Equivalently, the distribution of fund value (*w*) has a Pareto tail,

$$P\left(w_{it} > w\right) \sim Cw^{-a}.$$

The observed right fat tail of the distribution of fund values can potentially be described by this power law. However, this distribution must be cut off at some arbitrary value (typically $x = 0$) and may be insufficient to describe an empirical distribution in which entities are born at random locations.

The forward Kolmogorov equation describes the evolution of the distribution defined by the diffusion process over time:

$$\frac{\partial}{\partial t} p(x,t) = -\frac{\partial}{\partial x}\left(\mu_G p(x,t)\right) + \frac{\partial}{\partial x^2}\left(\frac{\sigma_G^2}{2} p(x,t)\right) - \delta p(x,t) + \delta\psi(x) \qquad (1)$$

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 μ), 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 is a constant loss of entities across the distribution (e.g., firm death), and the last term creates new entities at a random location governed by the distribution $\psi(x)$.

Substituting the stationary distribution $p(x,t) = p(x)$ in eq. (1) gives

$$0 = -\mu_G p'(x) + \frac{\sigma_G^2}{2} p''(x) - \delta p(x) + \delta \psi(x) \tag{2}$$

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 (δ). The prime notation (e.g., $p'(x)$) denotes first and second derivatives with respect to *x*.

Equation (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

$$0 = -\mu_G p'(x) + \frac{\sigma_G^2}{2} p''(x) - \delta p(x)\,. \tag{3}$$

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

$$p(x) = Ce^{-ax}$$

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

$$a_1 = \frac{-\mu_G + \sqrt{\mu_G^2 + 2\delta\sigma_G^2}}{\sigma_G^2},\; a_2 = \frac{-\mu_G - \sqrt{\mu_G^2 + 2\delta\sigma_G^2}}{\sigma_G^2}\,. \tag{4}$$

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 $x = 0$.

$$p(x) = \begin{cases} Ce^{-a_1 x}, x \geq 0 \\ Ce^{-a_2 x}, x < 0 \end{cases} \quad (5)$$

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 $a_2 < 0$.

Equation (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 adjusted 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 tail 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 equation with a random birth location instead of a fixed one, the variation-of-parameters approach can be used to produce a closed-form solution. First, rewrite eq. (2) as

$$p''(x) - \frac{2\mu_G}{\sigma_G^2} p'(x) + \frac{2\delta}{\sigma_G^2} p(x) = -\frac{2\delta}{\sigma_G^2} \psi(x)$$

and define the function $r(x) = -2\delta / \sigma_G^2 \psi(x)$ 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 equation (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

$$u' = r(x)\frac{e^{a_1 x}}{(a_1 - a_2)C}, \; v' = -r(x)\frac{e^{a_2 x}}{(a_1 - a_2)C}. \tag{9}$$

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 μ_B and standard deviation σ_B, and generalise the location of rebirth by substituting $x - \mu_B$ for x in eq. (5), such that

$$u = \int u'dx = \int r(x)\frac{e^{a_1(x-\mu_B)}}{(a_1 - a_2)C}dx = \int \frac{2\delta}{\sigma_G^2}\frac{e^{a_1(x-\mu_B)}}{(a_1 - a_2)C}\frac{1}{\sigma_B\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu_B}{\sigma_B}\right)^2\right)dx$$

$$= \frac{\delta}{\sigma_G^2(a_1 - a_2)C}\exp\left(a_1^2\sigma_B^2/2\right)erf\left(\frac{a_1\sigma_B^2 + \mu_B - x}{\sigma_B\sqrt{2}}\right) + const_1, \tag{10}$$

where $erf(\cdot)$ is the error function and $const_1$ is the constant of integration. Likewise,

$$v = \int v'dx = -\frac{\delta}{\sigma_G^2(a_1 - a_2)C}\exp\left(a_2^2\sigma_B^2/2\right)erf\left(\frac{a_2\sigma_B^2 + \mu_B - x}{\sigma_B\sqrt{2}}\right) + const_2. \tag{11}$$

We obtain the general solution by substituting these functions into eq. (8):

$$\begin{aligned} p(x) = &\frac{\delta}{\sigma_G^2(a_1 - a_2)}\exp\left(a_1(a_1\sigma_B^2/2 + \mu_B - x)\right)erf\left(\frac{a_1\sigma_B^2 + \mu_B - x}{\sigma_B\sqrt{2}}\right) + c_1\exp(-a_1(x-\mu_B)) \\ &-\frac{\delta}{\sigma_G^2(a_1 - a_2)}\exp\left(a_2(a_2\sigma_B^2/2 + \mu_B - x)\right)erf\left(\frac{a_2\sigma_B^2 + \mu_B - x}{\sigma_B\sqrt{2}}\right) + c_2\exp(-a_2(x-\mu_B)) \end{aligned} \tag{12}$$

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

$$c_1 = -\frac{\delta}{\sigma_G^2(a_1 - a_2)}\exp\left(a_1^2\sigma_B^2/2\right)$$

$$c_2 = -\frac{\delta}{\sigma_G^2(a_1 - a_2)} \exp\left(a_2^2 \sigma_B^2 / 2\right),$$

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:

$$\begin{aligned} p(x) = {} & \frac{\delta}{\sigma_G^2(a_1 - a_2)} \exp\left(a_1(a_1\sigma_B^2 / 2 + \mu_B - x)\right)\left(erf\left(\frac{a_1\sigma_B^2 + \mu_B - x}{\sigma_B\sqrt{2}}\right) - 1\right) \\ & - \frac{\delta}{\sigma_G^2(a_1 - a_2)} \exp\left(a_2(a_2\sigma_B^2 / 2 + \mu_B - x)\right)\left(erf\left(\frac{a_2\sigma_B^2 + \mu_B - x}{\sigma_B\sqrt{2}}\right) + 1\right). \end{aligned} \tag{13}$$

Equation (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 values of *x*, the first term approaches $A\exp(-B_1 x)$ while the second term approaches $A\exp(-B_2 x)$ for some constants *A*, B_1 and B_2. This shows that the logarithm of private equity fund multiples has exponential tails, which corresponds to the VC multiple itself having Pareto tails. In the central part of the distribution, both Pareto distributions are mixed with 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 (σ_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 equation (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: μ_G, σ_G and δ. These parameters are only identified up to scale.

To see this, consider the pre-factor in equation (13),

$$\frac{\delta}{\sigma_G^2(a_1 - a_2)} = \frac{\delta}{2\sqrt{\mu_G^2 + 2\delta\sigma_G^2}},$$

which has been simplified by plugging in the expressions for a_1 and a_2. If we now scale the death rate δ and the drift parameter μ_G by a constant t and the diffusion parameter σ_G by $\sqrt{t}$, these terms cancel in the pre-factor of the stationary distribution,

$$\frac{t\delta}{2\sqrt{(t\mu_G)^2 + 2(t\delta)\left(\sqrt{t}\sigma_G\right)^2}} = \frac{t\delta}{2\sqrt{t^2\mu_G^2 + 2t^2\delta\sigma_G^2}} = \frac{\delta}{2\sqrt{\mu_G^2 + 2\delta\sigma_G^2}}.$$

Similarly, the terms a_1 and a_2 are not affected by this scaling scheme. This scaling behaviour of the stationary distribution is expected because units have not been defined in the growth process. 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. Many combinations of the parameters μ_G, σ_G and δ 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 parameters μ_G, σ_G and δ represent only two degrees of freedom. If we let $t \equiv 1/\sigma_G^2$, equation (13) can be estimated by maximum likelihood using

$$\begin{aligned} p(x) = {} & \frac{\beta_\delta}{2\sqrt{\beta_{\mu_G}^2 + 2\beta_\delta}} \exp\left(a_1(a_1\sigma_B^2/2 + \mu_B - x)\right)\left(erf\left(\frac{a_1\sigma_B^2 + \mu_B - x}{\sigma_B\sqrt{2}}\right) - 1\right) \\ & - \frac{\beta_\delta}{2\sqrt{\beta_{\mu_G}^2 + 2\beta_\delta}} \exp\left(a_2(a_2\sigma_B^2/2 + \mu_B - x)\right)\left(erf\left(\frac{a_2\sigma_B^2 + \mu_B - x}{\sigma_B\sqrt{2}}\right) + 1\right), \end{aligned} \tag{14}$$

where

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

Because the death rate (δ) and growth rate (μ_G) are only identified up to scale through the parameters $\beta_\delta \equiv \delta/\sigma_G^2$ and $\beta_{\mu_G} \equiv \mu_G/\sigma_G^2$.

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 (König et al., 2017), repeated ties in networks (Kogut et al., 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 3,332 funds from Preqin's private equity performance database that were alive (i.e., not liquidated) in April 2020. Because 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 = 2,486). To filter out funds that have not made their first investment yet, funds must have called at least 10% of the capital committed by investors.

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

$$\log\left(TVPI_i\right) = \log\left(\frac{NAV_i + \sum_t D_{it}}{\sum_t C_{it}}\right),$$

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 distributions by the fund to its investors and NAV_i 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 settlement. Because it is composed of cash distributions and net asset values, it contains both realised and unrealised gains. For the purposes of this paper, "valuation multiple", "net multiple" and "return multiple" refer to this ratio. The average net multiple we observe 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.

We further observe 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 equation (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, we 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, we use a likelihood ratio test following Clauset et al. (2009), who adopt the methodology of Vuong (1989).

3.2. Results

The double Pareto distribution with random birth (i.e., smooth double Pareto distribution) produces a good fit to the data. Figure 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 where the standard double Pareto distribution has difficulties fitting the data. A normal distribution, 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.

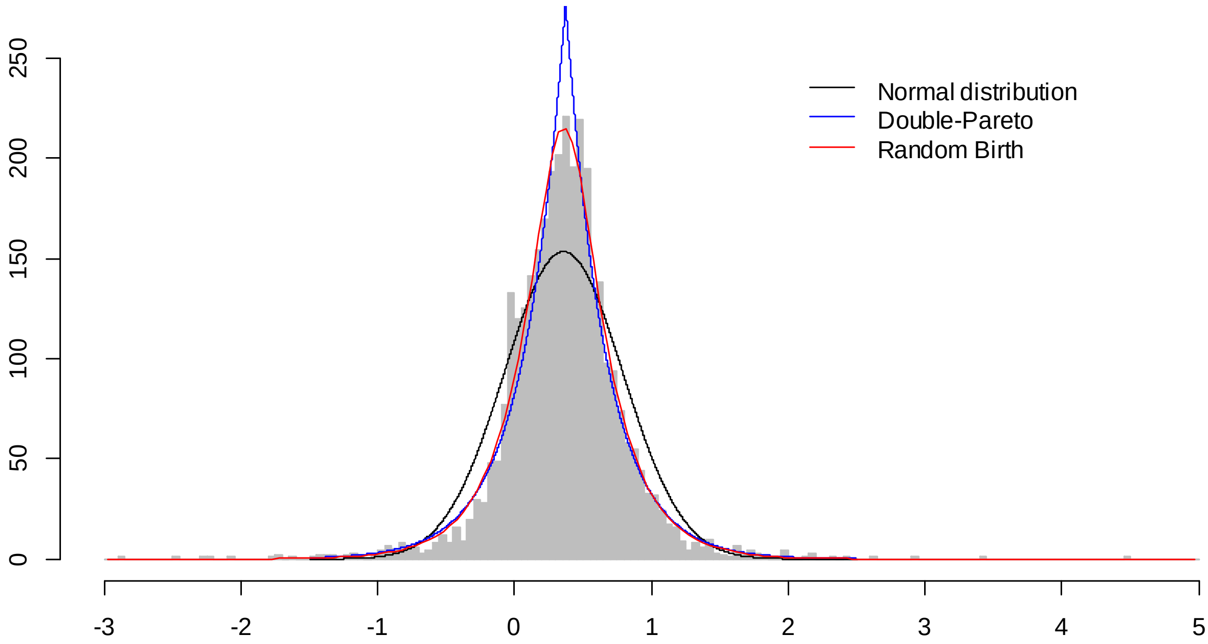

Figure 1. Frequency distribution of net multiple (natural logarithm)

The logarithm of the fund's net multiple is shown as a histogram with a bin width of 0.05. The dataset for this diagram contains all funds that have called at least 10% of their capital. The "random birth" distribution is the smooth double Pareto distribution described in the previous section.

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 reason, 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 has been called does not reduce the sample size much further. Because model fit is nearly identical, we focus our 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 negligible 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 $\beta_{\mu_G} \equiv \mu_G / \sigma_G^2$ from the previous

Table 1. Parameter estimates for random-birth distribution and double Pareto distribution

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). 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 greater than 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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)		(2)		(3)	
	Baseline sample (called capital $\geq$ 10%)		Realised > 0		Realised > 0 & called capital $\geq$ 10%	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Panel A: Random birth location						
Drift (β_{μ_G})	-0.032	(0.083)	-0.074	(0.077)	-0.070	(0.076)
Death rate (β_δ)	6.571	(0.326) ***	6.227	(0.280) ***	6.223	(0.279) ***
Birth mean (μ_B)	0.362	(0.011) ***	0.415	(0.010) ***	0.415	(0.010) ***
Birth Std. Dev. (σ_B)	0.129	(0.014) ***	0.072	(0.017) ***	0.070	(0.017) ***
Right tail exp. (a_1)	3.657	(0.120) ***	3.604	(0.113) ***	3.599	(0.113) ***
Left tail exp. (a_2)	-3.593	(0.125) ***	-3.456	(0.108) ***	-3.459	(0.107) ***
Observations	3332		2848		2836	
Log-Likelihood	-1612.65		-1307.09		-1299.34	
Panel B: Double Pareto						
Right tail exp. (a_1)	3.418	(0.086) ***	3.556	(0.095) ***	3.552	(0.095) ***
Left tail exp. (a_2)	-3.252	(0.080) ***	-3.312	(0.086) ***	-3.322	(0.086) ***
Birth location (x^*)	0.372	(0.008) ***	0.424	(0.008) ***	0.424	(0.007) ***
Log-Likelihood	-1630.19		-1311.68		-1303.63	
Panel C: Fit tests						
P-value KS-test RB	0.0907		0.2111		0.2084	
P-value KS-test DP	0.0382		0.1256		0.1281	
P-value KS-test LN	< 0.0001		< 0.0001		< 0.001	
P-value RB vs DP	0.0046		0.1044		0.1153	
P-value DP vs LN	< 0.0001		< 0.0001		< 0.0001	
Death rate / year	0.1308		0.1163		0.1158	

section, the mean drift is $\mu_G = \beta_{\mu_G}\sigma_G^2 = -0.032 \cdot 0.2^2 = 0.00128$ per year. Similarly, the death rate annualised using this annual standard deviation is $\mu_G = \beta_{\mu_G}\sigma_G^2 = 6.571 \cdot 0.2^2 = 0.26284$. 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 the death rate. If we use a rate of 13.08% per year, as shown in Table 1, it is possible deduce the mean growth rate and diffusion rate of the underlying process. From the previous section, we use $\beta_\delta \equiv \delta / \sigma_G^2$ and $\beta_{\mu_G} \equiv \mu_G / \sigma_G^2$, which gives $\sigma_G^2 = \delta / \beta_\delta = 0.1308 / 6.571 = 0.0199$ and thus $\sigma_G = 0.1411$. This result can also be obtained from estimating equation (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 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 equation (4), we have

$$a_1 = -\beta_{\mu_G} + \sqrt{\beta_{\mu_G}^2 + 2\beta_\delta},\ a_2 = -\beta_{\mu_G} - \sqrt{\beta_{\mu_G}^2 + 2\beta_\delta}$$

$$a_1 = 0.032 + \sqrt{0.032^2 + 2 \cdot 6.571} = 3.6573,\ a_2 = 0.032 - \sqrt{0.032^2 + 2 \cdot 6.571} = -3.5933\,.$$

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 Figure 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 Figure 1 and Figure 2. This can be an indication that observations are not generated by the same underlying process but come from separate distributions. Inspection of individual observations in the tails reveals that most extreme observations are venture capital funds. All funds with a log multiple greater than 2 are venture capital or growth-stage funds.

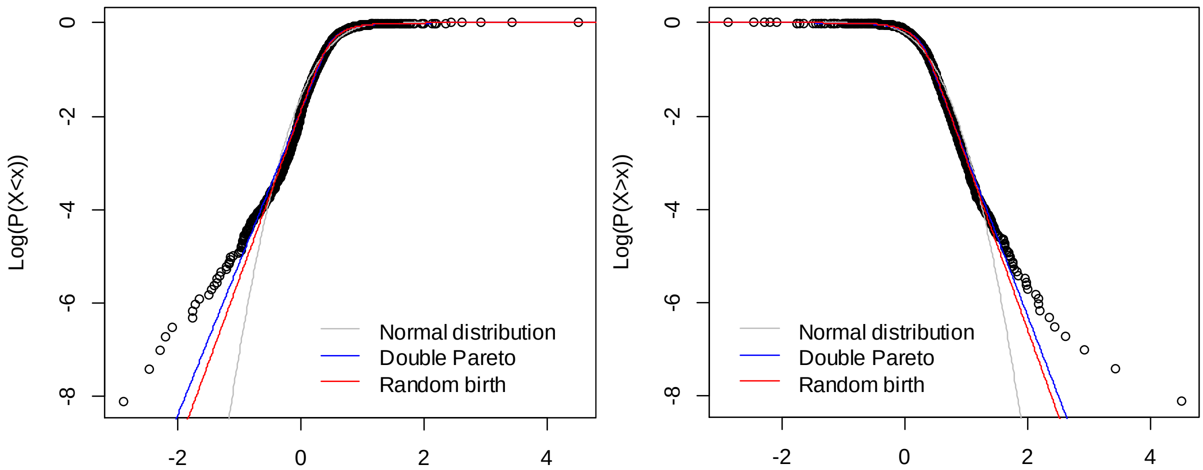

Figure 2. Tail distribution of net multiple for full (baseline) sample

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.

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

Differences in tail exponents for venture capital funds compared to non-VC funds are much smaller. As shown in Figures 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 (double Pareto) or 4.7 and 4.8 (smooth double Pareto).

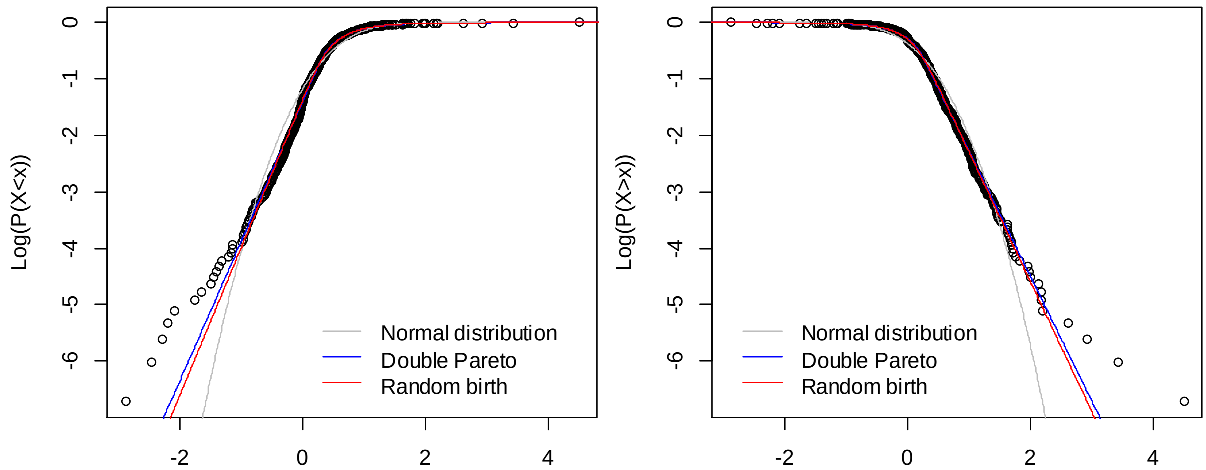

Figure 3. Tail distribution of net multiple for venture capital funds

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.

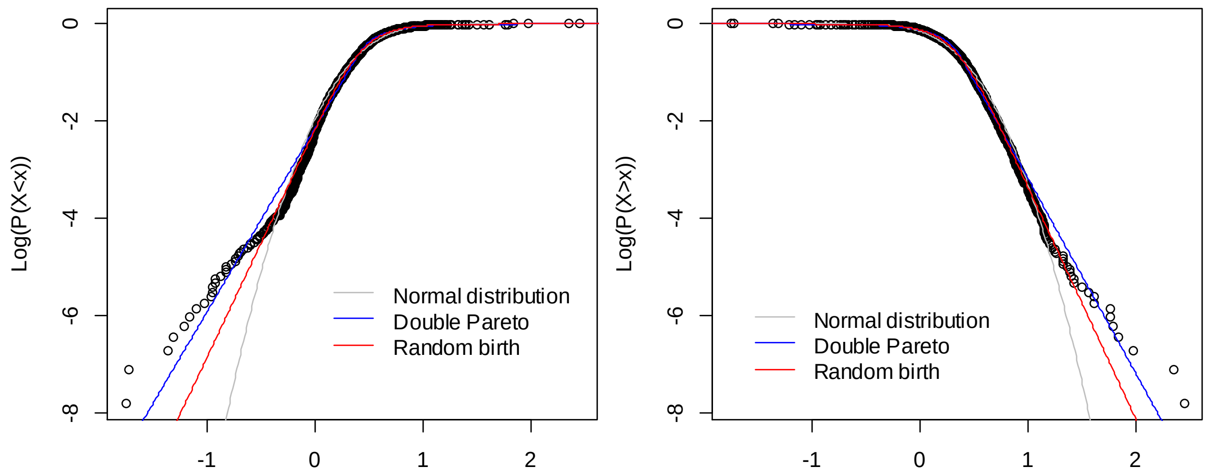

Figure 4. Tail distribution of net multiple for non-VC funds

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.

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 $\alpha < 3$ in the right tail. Similarly, the third moment is infinite for $\alpha < 4$. In general, any moment k of a power law exists only if $k < \alpha - 1$. 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. 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.

This 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 will only be realised after a long waiting time, since the 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 cross-section and a 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 a long time.

Table 2. Venture capital vs buyout funds

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 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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)		(2)		(3)	
	VC		Non-VC		Buyout	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Panel A: Random birth location						
Drift (β_{μ_G})	0.158	(0.105)	-0.042	(0.153)	-0.496	(0.338)
Death rate (β_δ)	3.007	(0.252) ***	11.306	(0.840) ***	11.898	(1.741) ***
Birth mean (μ_B)	0.255	(0.028) ***	0.378	(0.013) ***	0.420	(0.026) ***
Birth Std. Dev. (σ_B)	0.135	(0.034) ***	0.165	(0.013) ***	0.241	(0.019) ***
Right tail exp. (a_1)	2.300	(0.130) ***	4.797	(0.238) ***	5.400	(0.618) ***
Left tail exp. (a_2)	-2.615	(0.163) ***	-4.714	(0.230) ***	-4.407	(0.350) ***
Observations	828		2504		1084	
Log-Likelihood	-696.55		-790.57		-475.13	
Panel B: Double Pareto						
Right tail exp. (a_1)	2.216	(0.096) ***	4.036	(0.121) ***	3.685	(0.153) ***
Left tail exp. (a_2)	-2.474	(0.112) ***	-3.794	(0.110) ***	-3.146	(0.122) ***
Birth location (x^*)	0.260	(0.011) ***	0.390	(0.008) ***	0.425	(0.009) ***
Log-Likelihood	-698.78		-824.64		-510.71	
Panel C: Fit tests						
P-value KS-test RB	0.4430		0.7135		0.6186	
P-value KS-test DP	0.4260		0.0045		0.0028	
P-value KS-test LN	< 0.0001		0.0003		0.1650	
P-value RB vs DP	0.2703		< 0.0001		< 0.0001	
P-value DP vs LN	< 0.0001		0.0058		0.3947	
Death rate / year	0.1216		0.1342		0.1351	

There is some evidence that VC funds rely more on continuous growth than buyout funds. While the drift rate is insignificant in both cases, its difference between model 1 and model 3 in Table 2 is significant at $p < 0.1$. In line with these fund types' business models, this finding suggests 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 compares fitted smooth double Pareto distributions across subsamples of the baseline sample by fund strategy. These separate models can be combined into a single model for the full sample by describing the four identified parameters in equation (14) as linear functions of fund strategies and estimating the parameters by maximum likelihood as before. As a further benefit of this approach, additional covariates can be added to the linear part of the model.

This leads to four equations,

$$\beta_{\mu_G} = \mathbf{X}\boldsymbol{\gamma}_1,\ \beta_\delta = \mathbf{X}\boldsymbol{\gamma}_2,\ \sigma_B = \mathbf{X}\boldsymbol{\gamma}_3,\ \mu_B = \mathbf{X}\boldsymbol{\gamma}_4,$$

where **X** is a matrix of covariates including a constant and $\gamma_{1..4}$ are the vectors of hyperparameters analogous to the parameters estimated for subsamples in Table 2. Subscripts indicating individual funds have been omitted. In our 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 σ_B may become zero or negative during the numerical optimisation is a concern, σ_B in equation (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. $\sigma_B = 0$).

Table 3. Fund strategies as covariates

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=3,332. Standard errors are shown in parentheses. Significance levels: *** p < 0.01, ** p < 0.05, * p < 0.1.

	(1) Fully parameterised			(2) Tail parameters only			(3) Intercepts only		
	Coef.	S.E.		Coef.	S.E.		Coef.	S.E.	
Equation for Drift (β_{μ_G})									
Intercept	0.119	(0.174)		-0.063	(0.103)		-0.032	(0.083)	
VC (general)	-0.387	(0.243)		-0.242	(0.134)	*			
Late stage / growth	0.334	(0.365)		-0.204	(0.164)				
Buyout	-1.028	(0.410)	**	0.124	(0.144)				
Fund of funds	-1.367	(0.454)	***	0.489	(0.351)				
Other	-0.234	(0.360)		0.032	(0.331)				
Not North America	0.613	(0.1679	***	0.074	(0.100)				
Equation for Death rate (β_δ)									
Intercept	2.876	(0.399)	***	2.599	(0.333)	***	6.571	(0.326)	***
VC (general)	0.863	(0.584)		0.962	(0.461)	**			
Late stage / growth	4.150	(1.129)	***	3.127	(0.713)	***			
Buyout	10.186	(1.958)	***	5.650	(0.716)	***			
Fund of funds	18.842	(3.135)	***	23.346	(3.524)	***			
Other	10.091	(1.805)	***	16.655	(2.842)	***			
Not North America	-0.576	(0.472)		-0.436	(0.396)				
Equation for Birth mean (μ_B)									
Intercept	0.232	(0.047)	***	0.366	(0.012)	***	0.362	(0.011)	***
VC (general)	0.094	(0.064)							
Late stage / growth	0.010	(0.061)							
Buyout	0.230	(0.053)	***						
Fund of funds	0.235	(0.050)	***						
Other	0.151	(0.051)	***						
Not North America	-0.081	(0.016)	***						
Equation for Birth standard deviation (σ_B)									
Intercept	0.140	(0.043)	***	0.164	(0.011)	***	0.129	(0.014)	***
VC (general)	0.049	(0.071)							
Late stage / growth	0.060	(0.056)							
Buyout	0.117	(0.046)	**						
Fund of funds	-0.005	(0.046)							
Other	-0.063	(0.055)							
Not North America	-0.025	(0.017)							
Log-Likelihood	-1365.99			-1420.63			-1612.65		

Results in Table 3 show that the drift rate standardised by the standard deviation of the growth process (β_{μ_G}) 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 (β_δ). The smallest tail exponents are found in seed and early-stage funds with a left tail exponent of −2.52 ($-\beta_{\mu_G} + \sqrt{\beta_{\mu_G}^2 + 2\beta_\delta} =$ $-0.119 + \sqrt{0.119^2 + 2 \cdot 2.876} = -2.52$) and a right tail exponent of 2.283. Valuation multiples of funds of funds have the lightest 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 β_δ), 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 in particular 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 Figure 1. This density drop in the left tail is more pronounced in the VC sample shown in Figure 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). 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 Figure 5, is much less pronounced for other types of funds.

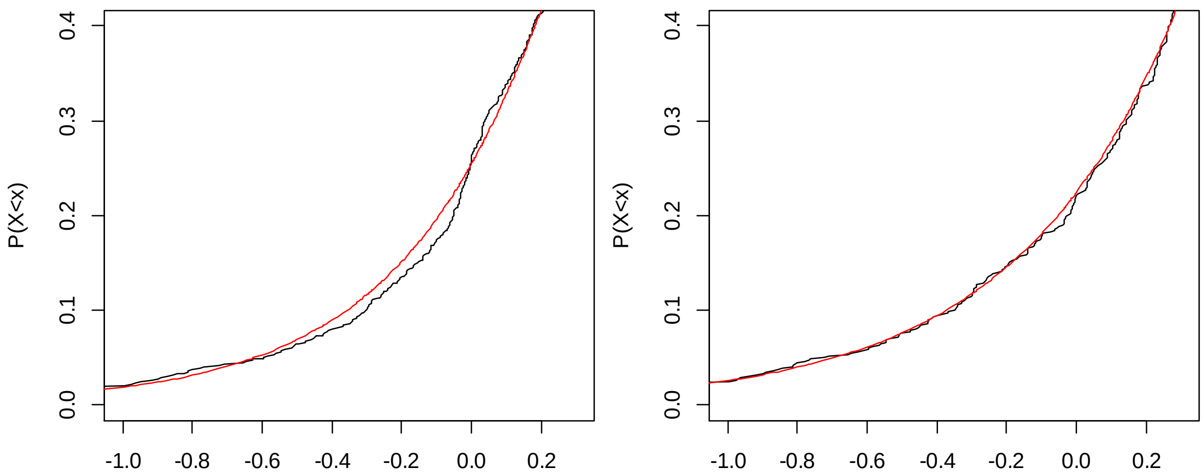

Figure 5. Distribution of all VC funds vs VC funds with distributions

The left-hand graph shows the empirical distribution and a fitted smooth double Pareto distribution for the full sample, corresponding to Figure 1. The right-hand graph shows the same sample after removing funds that have not made any distributions to investors.

When testing whether the spike near zero can be explained by funds without investments, another proxy for investment activity can offer some insights. The right-hand graph in Figure 5 shows the distribution of value multiples for VC funds with the added constraint that funds must have made a positive 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.

A disadvantage of using investor distributions as a proxy 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 not 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.

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 incorporate the mechanism whereby funds remain at zero and then jump to the first meaningful valuation as a jump process.

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 part that enables a better fit compared to alternative distributions.

The underlying growth process produces Pareto tails in the cross-section, while individual entities exhibit lognormal growth. This finding supports the assumption of lognormal growth between sparse price observations in studies of private equity fund pricing. The smooth double Pareto distribution's accurate fit further 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.

Tests using subsamples 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 deceptively 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.

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. 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 jump to a more economically meaningful multiple when making or revaluing their first investment.

References

Ang, A., Chen, B., Goetzmann, W. N., and Phalippou, L. (2018). Estimating private equity returns from limited partner cash flows. *Journal of Finance* 73(4), 1751–1783.

Bak, P., Chen, K., Scheinkman, J., and Woodford, M. (1993). Aggregate fluctuations from independent sectoral shocks: Self-organized criticality in a model of production and inventory dynamics. *Ricerche Economiche* 47(1), 3–30.

Benhabib, J., and Bisin, A. (2018). Skewed Wealth Distributions: Theory and Empirics. *Journal of Economic Literature* 56(4), 1261–1291.

Cabral, L. M. B., and Mata, J. (2003). On the Evolution of the Firm Size Distribution: Facts and Theory. *American Economic Review* 93(4), 1075–1090.

Coad, A. (2009). *The growth of firms: A survey of theories and empirical evidence*. Cheltenham: Edward Elgar.

Champernowne, D. (1953). A Model of Income Distribution. *Economic Journal* 63(250), 318–351.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review* 51(4), 661–703.

Cochrane, J. H. (2005). The risk and return of venture capital, *Journal of Financial Economics* 75(1), 3–52.

Crawford, G. C., Aguinis, H., Lichtenstein, B., Davidsson, P., and McKelvey, B. (2015). Power law distributions in entrepreneurship: Implications for theory and research. *Journal of Business Venturing* 30(5), 696–713.

Driessen, J., Lin, T.-C., and Phalippou, L. (2012). A new method to estimate risk and return of nontraded assets from cash flows: The case of private equity funds. *Journal of Financial and Quantitative Analysis* 47(3), 511–535.

Gabaix, X. (1999). Zipf 's Law for Cities: An explanation. *Quarterly Journal of Economics* 114(3), 739–767.

Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics* 1, 255–294, Available at https://doi.org/10.1146/annurev.economics.050708.142940.

Gabaix, X., Lasry, J.-M., Lions, P.-L., Moll, B., and Qu, Z. (2016). The dynamics of inequality. *Econometrica* 84(6), 2071–2111.

Hilbert, M. (2013). Scale-free power-laws as interaction between progress and diffusion. *Complexity* 19(4), 56–65.

Huberman, B. A., and Adamic, L. A. (1999). Evolutionary dynamics of the World Wide Web. *Nature* 401 (6749), 131–131.

Huberman, B. A., and Adamic, L. A. (2000). The nature of markets in the World Wide Web. *Quarterly Journal of Economic Commerce* 1, 5–12.

Kogut, B., Urso, P., and Walker G. (2007). Emergent properties of a new financial market: American venture capital syndication, 1960–2005. *Management Science* 53(7), 1181–1198.

König, M. D., Lorenz, J., and Zilibotti, F. (2016) Innovation vs. imitation and the evolution of productivity distributions. *Theoretical Economics* 11(3), 1053–1102.

Korteweg, A., and Nagel, S. (2016). Risk-adjusting the returns to venture capital. *Journal of Finance* 71(3), 1437–1470.

Mandelbrot, B. (1953). An informational theory of the statistical structure of languages. In: W. Jackson (Ed.), *Communication Theory*, London: Butterworths, 486–502.

Masters, B., and Thiel, P. (2014). *Zero to one: Notes on startups, or how to build the future*. Virgin Digital.

Mitzenmacher, M. (2003). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics* 1(2), 226–251.

Newman, M. E. J., (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46(5), 323–351.

Pareto, V. (1896). *Cours d'Economie Politique*. Geneva: Droz.

Prencipe, D. (2017). The European venture capital landscape: an EIF perspective – Volume III: Liquidity events and returns of EIF-backed VC investments. European Investment Fund Working Paper 2017/41, Available at: https://www.eif.org/news_centre/publications/eif_wp_41.pdf.

Reed, W. J. (2001). The Pareto, Zipf and other power laws. *Economics Letters* 74(1), 15–19.

Reed, W. J., and Jorgensen, M. (2003). The Double Pareto-Lognormal Distribution—A New Parametric Model for Size Distributions. *Communications in Statistics - Theory and Methods* 33(8), 1733–1753.

Rozenfeld, H. D., Rybski, D., Gabaix, X., and Makse, H. A. (2011). The Area and Population of Cities: New Insights from a Different Perspective on Cities. *American Economic Review* 101(5), 2205–2225.

Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika* 42(3/4), 425–440.

Toda, A. A., and Walsh, K. (2015). The Double Power Law in Consumption and Implications for Testing Euler Equations. *Journal of Political Economy* 123(5), 1177–1200.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2), 307–333.

Yule, G. (1925). A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London B* 213(402–410), 21–87.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, Massachusetts: Addison-Wesley.