Reduced Form Private Equity Fund Asset Modeling

Christian Tausch[a,*], Axel Buchner[b], Georg Schlüchtermann[c,d]

[a] *AssetMetrix GmbH, Theresienhöhe 13, D-80339 Munich, Germany*
[b] *Department of Business and Economics, University of Passau, D-94030 Passau, Germany*
[c] *Faculty of Mathematics, Informatics and Statistics, Ludwig-Maximilians-Universität München, Theresienstraße 39, D-80333 Munich, Germany*
[d] *Department of Mechanical, Automotive and Aeronautical Engineering, University of Applied Sciences Munich, Dachauer Straße 98b, D-80335 Munich, Germany*

Abstract

Risk perception in private equity is notoriously difficult, as the cash flow patterns associated with private capital funds are not well understood on underlying asset level. To account for the incomplete information setting induced by the infrequent and imperfect valuation practice of privately held assets, this paper proposes the first reduced form model tailored for private equity fund investments. Especially their realized exit cash flows are analyzed in a *joint* modeling framework that describes *both* the exit timing *and* exit performance on individual deal level. The corresponding linear parametric models are estimated by means of maximum likelihood for a buy out and venture capital data set and are applied within a Monte Carlo simulation example to emphasize the superiority of our approach in the risk management context.

Keywords: Private equity, Reduced form model, Cash flow risk simulation, Buy out, Venture capital
JEL: G11, G17, G24

Acknowledgements

We thank Markus Rieder, Kevin Gruber, Marcus Pietz, Holger Fink, Stefan Mittnik, Michael Rockinger, and the participants in the CEQURA Junior Research Workshop 2017 and CEQURA Risk Conference 2018. Especially we thank AssetMetrix GmbH for the generous access to data. The views expressed in this paper are those of the authors and do not necessarily reflect those of AssetMetrix.

Declaration of interest

The authors report no declarations of interest.

*I am corresponding author
Email addresses: christian.tausch@quant-unit.com (Christian Tausch), axel.buchner@uni-passau.de (Axel Buchner), georg.schluechtermann@hm.edu (Georg Schlüchtermann)

1. Introduction

Financial assets with a cumulative valuation of \$2.83 trillion are controlled by private equity funds (as of June 2017; Source: Preqin). Fund investors naturally ask (i) what will happen with these assets in the future or, on a small scale, (ii) what is the economic hazard associated with a particular private equity fund stake? Fundamentally (iii) how much cash will be realized and (iv) when? These questions arise because a Private Equity Fund (PEF) is generically constructed as limited partnership with bounded lifetime that is not traded on public markets[1]. The fund manager receives an unfunded upfront commitment from fund investors and then controls the timing of all discretionary investment and divestment cash flows. From the fund investors' perspective these cash flows can be considered as the outcome of exogenous random variables.

This article follows the perception that the illiquid character of PEF investments requires tailored risk measurement methodologies, preferably on underlying asset level to insure the inclusion of company level information, vital in undiversified settings. However, little is known about the dynamic behavior of single private equity fund investments. Especially the interaction between holding period and total return, that ultimately determines the cash flows to investors, are not well grasped yet. The empirical and theoretical private equity literature lacks comprehensive concepts on asset level that jointly describe both (i) the fund manager's endogenous timing of cash flows and (ii) the risk and return of the underlying fund holdings. These highly related aspects are unfortunately often studied in rather isolated and fragmented approaches.

To better understand the exit behavior of private equity fund investments on portfolio company level, we establish the reduced form approach to PEF asset modeling. It is derived in a continuous-time framework by exploiting analogies to credit risk models in incomplete information settings [2]. Since reduced form formulations avoid to model unobservable quantities, our model exclusively describes all cash flow events associated with a given private equity fund by a marked point process. The benchmark approach of Platen [3] can be applied to price these payment streams under the real world measure. Consistent with our reduced form perspective, we analyze the exit dynamics of PEF investments, i.e. the connection between (i) exit timing and (ii) return of underlying fund assets, by two marginal statistical models that are linked by a copula. Additionally the exit performance model is conditioned on exit timing. The exit timing regression is based on a parametric multiplicative hazard rate formulation adapted for time-variant covariates. The return multiple regression employs the so called two-part or hurdle modeling idea to account for the zero-heavy nature of historically observed PEF asset returns.

The access to a proprietary asset level data set of Buy Out (BO) and Venture Capital (VC) fund investments allows the empirical application of our modeling idea. In the first step, the aforementioned exit timing and return multiple regression models are estimated by maximum likelihood for both data sets. Here the asset level granularity permits the inclusion of many covariates that are not available in fund level regressions. This paper

[1]Kaplan and Strömberg [1] describe the nature and economics of PEFs in more detail.

focuses on public market (equity returns and corporate high yield spreads) and timing related (holding period, time to exit, etc.) independent variables. In the second step, the estimated parametric models are applied in a Monte Carlo simulation example that, with its undiversified PEF portfolio setting, reveals the advantages of asset level over fund level cash flow risk simulation approaches.

The article is organized as follows: Section 2 reviews related literature. Section 3 introduces the new reduced form modeling framework for private equity fund assets. Section 4 presents a joint parametric exit timing and exit return multiple model that can be estimated by maximum likelihood. Section 5 reports the empirical results of these regression approaches for a BO and VC data set. Section 6 discusses a risk management application of the model estimates in a Monte Carlo simulation example. Section 7 finally concludes.

2. Related literature

2.1. Empirical private equity analyses

Published empirical analyses on deal level focus *either* on the exit route *or* on the asset performance of private equity investments. By contrast joint empirical analyses of both aspects merely exist as unpublished working papers [4, 5].

The realized exit routes of VC fund investments (e.g. initial public offering, trade sale, liquidation) are analyzed by Giot and Schwienbacher [6] and Félix et al. [7] by means of competing risk models. Jenkinson and Sousa [8] employ a multiplicative hazard model and a trinomial logistic regression for a BO data set. Cumming [9] and Schmidt et al. [10] use multinomial logit models in similar VC and BO studies. Cumming et al. [11] survey the firm level exit performance of governmental and independent VC investments in Europe. In all exit route regressions just static, i.e. time-invariant, covariates are incorporated.

The return and risk of VC companies is studied by Cochrane [12] and Korteweg and Sorensen [13]. They develop sample selection correction methodologies that allow the calculation of the amended return of VC investments from observed financing round valuation data. Both approaches are based on log-normally distributed returns. The value creation of BO firms is examined by Guo et al. [14] and Valkama et al. [15]. Here the return drivers of BO investments are identified in detailed deal level regressions.

2.2. Stochastic private equity models

The first private equity *fund level* model that primarily relies on Gaussian stochastic processes is introduced by de Malherbe [16]. Buchner [17] utilizes a similar framework to distinguish between fund level (i) market, (ii) liquidity, and (iii) cash flow risk. Buchner et al. [18] propose a stochastic model on the typical cash flow dynamics of private equity funds that solely relies on observable cash flow data. Conclusive from a diversified portfolio perspective, these approaches inherently neglect any asset level information.

Further there exist some structural *asset level* models that are designed to address concrete private equity related questions. Bongaerts and Charlier [19] estimate the capital requirements for private equity investments under Basel II. Braun et al. [20] quantify the risk appetite of BO fund managers. Escobar et al. [21] examine the portfolio optimization problem for private equity investors. Dong et al. [22] assess the credit risk associated with a portfolio of private infrastructure projects. Lahmann et al. [23] focus on the stepwise debt reduction associated with BO investments. Each paper deals with private equity specific issues, however the general structural (default) framework applied therein is explicitly developed rather for public debt than private equity.

3. General framework

This paper aims to establish and estimate a suitable stochastic model for the exit cash flows of private equity fund investments. Section 3.1 introduces the generic probability space and describes the information flow to fund investors. In section 3.2 generic stochastic processes are utilized to generally elaborate the distinction between structural and reduced form approach in the private equity context. Section 3.3 sketches PEF pricing.

3.1. Generic probability space

Let $(\Omega, \mathscr{G}, \mathbb{P})$ be a filtered probability space (satisfying the usual hypotheses) with the sample space Ω, $\mathscr{G}$ a σ-algebra of subsets of Ω, and the real world probability measure $\mathbb{P}$. To model the incomplete information setting of a typical private equity fund investor, we introduce the smaller *investor filtration* $(\mathscr{F}_t)_{t\in[0,T^*]}$ with $\mathscr{F}_t \subset \mathscr{G}_t, \forall t \in [0, T^*]$. We use the notation $\mathscr{F} = \mathscr{F}_{T^*}$ and $\mathscr{G} = \mathscr{G}_{T^*}$.

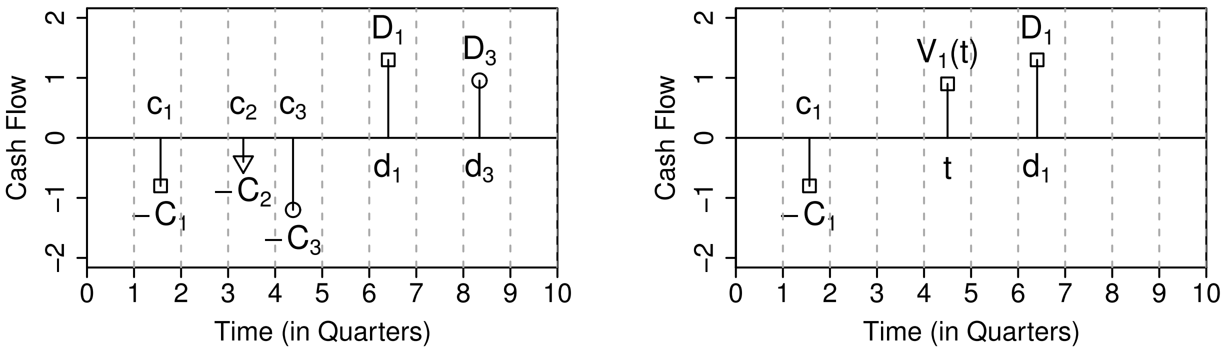
Investors can observe three distinct types of cash flows associated with a generic fund investment. Contribution cash flows $\mathcal{C}$ are spent by fund managers to buy new companies and distribution cash flows $\mathcal{D}$ result when the corresponding companies are sold. Other cash flows $\mathcal{O}$, i.e. mainly fee payments, are ignored in the remainder, since they can be modeled as deterministic functions of $\mathcal{C}$ and $\mathcal{D}$. Funds invest in several companies and we assume for simplicity that each investment $i \in \{1, 2, \ldots, n\}$ is characterized by one single entry event $\mathcal{C}_i = (c_i, C_i)_{i=1,\ldots,n}$ and one single exit event $\mathcal{D}_i = (d_i, D_i)_{i=1,\ldots,n}$ [2]. The timing of investment and divestment is given by c_i and d_i, respectively, and the cash amounts associated with investment and divestment are denoted by C_i and D_i, respectively. As can be seen from figure 1a, it holds $c_i < d_i$, $C_i > 0$, $D_i \geq 0$ for all $i = 1, 2, \ldots, n$.

More specifically, the timing of investment cash flows can be described by the univariate counting process $N^{(c)}$ that is governed by a nonnegative $\mathscr{G}$-progressively measurable intensity process $\lambda = \{\lambda_t, t \geq 0\}$. The exit timing is governed by an n-dimensional point process $\mathbf{N^{(d)}}(t)$, where each $N_i^{(d)}(t) = \mathbf{1}_{\{d_i \leq t\}} \vec{e_i}$ indicates if the ith company's final divestment time d_i has been exceeded before time t; with $\mathbf{1}_{\{\}}$ as indicator function and $\vec{e_i}$ is the ith unit vector. To assure $c_i < d_i$ we construct exit timing by $d_i = c_i + T_i$ where T_i is a positive random variable. Appendix B outlines the $\mathscr{F}$-adapted intensity based model for $N_i^{(d)}(t)$.

Additionally, there exist two auxiliary information processes for each entered investment

$$V_i(t) = \mathring{V}_i(t) + \upsilon_{t,i} \qquad \text{and} \qquad \mathbf{X}_i(t)$$

[2]Appendix A accounts for the more realistic case of multiple investments and divestments per company, needed when the model is empirically applied in equation 7.



(a) Multiple investments

(b) One investment and valuation

Figure 1: (a) Visualization of fund cash flows $(\mathcal{C}, \mathcal{D})$. Three entry, but only two exit events are observed. For the right-censored exit event (d_2, D_2) we know that $d_2 > 10$.
(b) Private Equity Fund Asset Model: A PEF-AM connects $V(t)$ with $\mathcal{D} = (d, D)$ for $t \in [c, d[$. There are two general modeling approaches to this problem: structural and reduced form.

where $V_i(t)$ is the (observable) proxy net asset value process and $\mathring{V}_i(t)$ is the (latent) true net asset value process of the ith asset, both locally integrable and $\mathscr{G}$-progressively measurable. Thus fund mangers repeatedly report *imperfect approximations of the true asset value* that result in untradeable proxy valuations. The valuation error random variable $\upsilon_{t,i}$ is left unspecified for $t \in]c_i, d_i[$ and we further assume

$$V_i(c_i) = C_i \qquad \text{and} \qquad V_i(d_i) = D_i$$

The locally integrable multivariate covariate process $\mathbf{X}_i(t)$ contains e.g. supplementary macro-economic, public market, fund level, or asset specific information.

Fund managers disclose informations regarding $V(t)$ to their fund investors just on a quarterly basis. Due to this reporting practice in the private equity industry, PEF data on asset level is characterized by a longitudinal (or panel) data structure with synchronous interval-censoring. This means we face a ‘doubly censored data structure’ in the sense of Sun [24, 1.3.3], since entry and exit timings are both interval-censored

$$c_i \in \left] c_i^{(L)}, c_i^{(R)} \right] \qquad d_i \in \left] d_i^{(L)}, d_i^{(R)} \right]$$

For the ith asset let $A_{q,i}$ resp. $B_{q,i}$ $(q = 1, \ldots, Q)$ represent the quarter start resp. end dates, i.e.

$$0 \leq A_{1,i} \leq B_{1,i} \leq \ldots \leq A_{Q,i} \leq B_{Q,i} \leq T^*$$

Thus, $c_i^{(L)}$ and $d_i^{(L)}$ are components of the vector of quarter start dates $\mathbf{A}_i = (A_{q,i})_{q=1,\ldots,Q}$ as $c_i^{(R)}$ and $d_i^{(R)}$ are components of the corresponding vector of end dates $\mathbf{B}_i = (B_{q,i})_{q=1,\ldots,Q}$. Hence, we define the underlying censoring and filtering processes as

$$Z_i^{(\text{cens})}(t) := \mathbf{1}_{\left\{t > c_i^{(L)}\right\}} \mathbf{1}_{\left\{t \leq d_i^{(R)}\right\}} \tag{1}$$

$$Z^{(\text{filt})}(t) := \mathbf{1}_{\{t \in (\mathbf{A} \vee \mathbf{B})\}} \tag{2}$$

The *investor filtration* $(\mathscr{F}_t)_{t\in[0,T^*]}$ is thus generated by the self-exciting filtrations of the respective (filtered) stochastic processes

$$\mathscr{F}_t := \mathscr{N}_t^{(c)} \vee \mathscr{N}_t^{(d)} \vee \mathscr{V}_t \vee \mathscr{X}_t \vee \mathscr{Z}_t^{(\text{cens})} \vee \mathscr{Z}_t^{(\text{filt})}$$

with

$$\begin{array}{lcl} \mathscr{N}_t^{(c)} & = & \sigma\left\{N^{(c)}(s)Z^{(\text{filt})}(s) : 0 \leq s \leq t\right\} \\ \mathscr{N}_t^{(d)} & = & \sigma\left\{\mathbf{N}^{(d)}(s)Z^{(\text{filt})}(s) : 0 \leq s \leq t\right\} \\ \mathscr{V}_t & = & \sigma\left\{\mathbf{V}(s)Z^{(\text{filt})}(s) : 0 \leq s \leq t\right\} \\ \mathscr{X}_t & = & \sigma\left\{\mathbf{X}(s) : 0 \leq s \leq t\right\} \\ \mathscr{Z}_t^{(\text{cens})} & = & \sigma\left\{\mathbf{Z}^{(\text{cens})}(s) : 0 \leq s \leq t\right\} \\ \mathscr{Z}_t^{(\text{filt})} & = & \sigma\left\{Z^{(\text{filt})}(s) : 0 \leq s \leq t\right\} \end{array}$$

This means $\mathscr{F}$-information allows us only to observe $N^{(c)}(t)$, $\mathbf{N}^{(d)}(t)$, and $\mathbf{V}(t)$ on a quarterly time grid. Therefore when working under filtration $\mathscr{F}_t$ we just regard times $t \in (\mathbf{A} \vee \mathbf{B})$.

3.2. Structural vs reduced form approach

The credit risk literature distinguishes between structural and reduced form modeling approaches. Jarrow and Protter [25] outline their differences from an information based perspective: 'Structural models assume complete knowledge of a very detailed information set, akin to that held by the firm's managers. [...] In contrast, reduced form models assume knowledge of a less detailed information set, akin to that observed by the market'. As a consequence, reduced form model formulations naturally arise in some incomplete (partial/imperfect/latent/noisy) information settings [2].

We suggest that reduced form models constitute at least a fruitful alternative to existing structural PEF models (cf. section 2.2), since incomplete information settings can be assumed characteristic for private equity investments. Generally reduced form models strive to describe the data but not necessarily the underlying cause and effect phenomenon like structural approaches. The reduced form approach to PEF asset modeling is introduced by the following definitions, where we omit the index 'i' for better reading:

Definition 1. *A* Private Equity Fund Asset Model (PEF-AM) $\mathbf{P}$ *connects the intermediate proxy valuation $V(t)$ with the resulting final divestment cash flow $\mathcal{D}$ for $t \in [c, d[$, i.e.*

$$\mathcal{D} = (d, D) := \mathbf{P}\left(V(t), ...|\, \mathcal{C}\right)$$

where $\mathbf{P}()$ is a two-dimensional function and $\mathcal{C}$ is assumed to be known for each entered investment (cf. figure 1b).

Definition 2. *A* structural *PEF-AM calculates the exit amount by the pathwise Stieltjes integral over either the observable value process V or the unobservable value processes $\mathring{V}$*

$$\begin{array}{lcl} D_{\text{struc}}^{(1)} & := & V(t) + \int_t^d \mathrm{d}V(s) \\ D_{\text{struc}}^{(2)} & := & \mathring{V}(t) + \int_t^d \mathrm{d}\mathring{V}(s) \end{array} \tag{3}$$

and the exit timing is defined as model based first hitting time

$$d_{\text{struc}} := \inf\{t > c : \; V(t) > E(t)\} \wedge \inf\left\{t > c : \; \mathring{V}(t) \leq 0\right\} \wedge T^* \tag{4}$$

with the $\mathscr{G}$-progressively measurable dynamic exit acceptance process $E(t)$.

Definition 3. *A* reduced form *PEF-AM defines the exit amount in terms of the observable proxy value process V and a stochastic multiplier m_D that is modeled conditional on exit timing information*

$$D = m_D\left(t \,|\mathscr{F}_d\right) \cdot V(t) \tag{5}$$

therefore the stochastic multiplier is defined as

$$m_D\left(t\,|\mathscr{F}_d\right) := \frac{D}{V(t)}$$

The exit timing is determined in the reduced form approach by the point process $N^{(d)}$

$$d := \inf\left\{t > c : \; N^{(d)}(t) > 0\right\} \wedge T^* \tag{6}$$

with conditional survival probability $\mathbb{P}\left[d > t\,|\mathscr{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[1 - N^{(d)}(t)\,|\mathscr{F}_t\right]$.

In a realistic incomplete information setting, private equity investors can observe neither $\mathring{V}$ nor E. Fund managers have to base their subjective exit decision E on the proxy value process V, since they are possibly also unable to observe $\mathring{V}$. Then fund mangers may be even surprised by unexpected firm defaults, i.e. when the latent firm value process $\mathring{V}$ hits zero. Due to this difficulties, existing structural PEF-AMs often regard exit timings as exogenously fixed variables, e.g. by using hypothetical fund manager estimates or average durations [19, 21]. From a mathematical viewpoint equations (3, 5, 6) can be calculated with $\mathscr{F}_d$ information, but equation (4) describes a totally inaccessible stopping time under $\mathscr{F}$. If we assume $\mathscr{F}$ to constitute a reasonable filtration for a private equity investment situation, reduced form formulations are to be favored over structural approaches (cf. table 1 for an overview). Consequently, definition 3 is used to model D and d in the remainder.

3.3. Real world pricing in the reduced form approach

If we assume our covariate process $\mathbf{X}(t)$ to contain (a proxy for) the growth optimal numeraire portfolio $S^*(t)$ in the sense of [26, 3], the $\mathscr{F}$-fair company value/price $\dot{V}$ is

$$\dot{V}_i(t) = S^*(t) \cdot \mathbb{E}^{\mathbb{P}}\left[\left.\frac{D_i}{S^*(d_i)}\right|\mathscr{F}_t\right] = S^*(t) V_i(t) \cdot \mathbb{E}^{\mathbb{P}}\left[\left.\frac{m_D\left(t|\mathscr{F}_{d_i}\right)}{S^*(d_i)}\right|\mathscr{F}_t\right] \quad \forall t \in (\mathbf{A} \vee \mathbf{B})$$

In practice the expected value can be estimated by (Monte Carlo) simulation relying on a suitable model for $S^*(t)$ and $\mathcal{D}$.

Using the benchmark approach in combination with the global filtration $\mathscr{G}$ yields

$$\mathring{V}_i(t) = S^*(t) \cdot \mathbb{E}^{\mathbb{P}}\left[\left.\frac{\mathring{V}_i(u)}{S^*(u)}\right|\mathscr{G}_t\right] = S^*(t) \cdot \mathbb{E}^{\mathbb{P}}\left[\left.\frac{D_i}{S^*(d_i)}\right|\mathscr{G}_t\right]$$

for $t < u < d_i$. Especially the $\mathscr{G}$-true value can be just observed at exit, i.e. $\mathring{V}_i(d_i) = D_i$.

Table 1: Distinction between structural and reduced form approach with respect to the stochastic processes that need to be specified for the exit behavior estimation. Unspecified stochastic processes can be perceived as realized data for past times and unknown for future times.

Process	Description	Approach		$\mathscr{F}$-observable
		structural	reduced form	(quarterly)
$X(t)$	regression covariates	known	known	yes
$N^{(c)}(t)$	counts entry events	known	known	yes
$N^{(d)}(t)$	indicates exit events	unspecified	specified	yes
$Z^{(\text{cens})}(t)$	indicates censoring	unspecified	specified	yes
$Z^{(\text{filt})}(t)$	quarterly filtering	unspecified	specified	yes
$m_D\left(t\,\middle\vert\mathscr{F}_d\right)$	multiplier	unspecified	specified	yes
$V(t)$	proxy asset value	specified	unspecified	yes
$\mathring{V}(t)$	true asset value	specified	unspecified	no
$E(t)$	exit decision set	specified	unspecified	no
c	entry timing	via $N^{(c)}(t)$	via $N^{(c)}(t)$	-
C	entry amount	$V(c)$	$V(c)$	-
d	exit timing	equation (4)	equation (6)	-
D	exit amount	equation (3)	equation (5)	-

4. Parametric joint reduced form model

4.1. Marginal MULTIPLE *and* TIMING *distributions*

In accordance with the reduced form approach introduced by definition 3 we define the exit MULTIPLE variable as $Y = D/V(t)$ and the exit TIMING variable as $y = d - t$. Due to the semicontinuous (zero-heavy) nature of the MULTIPLE $Y \in \mathbb{R}_{\geq 0}$ we split the univariate cumulative distribution function (CDF) and probability density function (PDF) into a point mass at zero and an absolutely continuous part

$$\begin{aligned} F_Y(\bar{Y}) &= \mathbb{P}\left[\tfrac{D}{V(t)} \leq \bar{Y}|\mathscr{F}_t\right] &&= \pi_0(\mathbf{X}) + (1-\pi_0)G_Y\left(\bar{Y}|\mathbf{X},\xi_Y\right) \\ f_Y(\bar{Y}) &= \tfrac{\delta}{\delta \bar{Y}}F_Y(\bar{Y}) &&= \pi_0(\mathbf{X})\mathbf{1}_{\{\bar{Y}=0\}} + (1-\pi_0)g_Y\left(\bar{Y}|\mathbf{X},\xi_Y\right) \end{aligned}$$

where the default probability $\pi_0(\mathbf{X})$ is conditioned on some covariates $\mathbf{X}$. G_Y and g_Y represent a continuous CDF and PDF, respectively, that are conditioned on vectors of covariates $\mathbf{X}$ and parameters ξ_Y. Further the TIMING $y \in \mathbb{R}_{>0}$ is specified by the conditional survival model

$$S_y\left(\bar{y}|t\right) = \mathbb{P}\left[d > t + \bar{y}|\mathscr{F}_t\right] = \frac{S_y\left(t+\bar{y}|\mathbf{X},\xi_y\right)}{S_y\left(t|\mathbf{X},\xi_{\mathbf{y}}\right)}$$

where S_y denotes an absolutely continuous survival function with covariate vector $\mathbf{X}$ and parameter vector ξ_y. Due to interval censoring introduced by $\mathscr{F}_t$ the TIMING density function is linearly approximated by rectangle probabilities

$$f_y(\bar{y}|t) = S_y(\bar{y}|t)\, h_y(t+\bar{y}) \approx \frac{S_y(\bar{y}-\Delta|t) - S_y(\bar{y}|t)}{\Delta}$$

where h_y is the corresponding hazard function and $\Delta = 0.25$ for quarterly observations.

4.2. Bivariate copula model

We define the bivariate distribution in terms of the distribution function for Y and survival function for y

$$F_{Yy}(\bar{Y}, \bar{y}|t) = \mathbb{P}\left[\frac{D}{V(t)} \le \bar{Y}, d > t + \bar{y} \,\middle|\, t, \mathscr{F}_t\right]$$

In a similar insurance setting Shi et al. [27] favor a parametric copula model over a straightforward conditional probability decomposition. However since reduced form explicitly models the Multiple conditional on Timing we combine both approaches. First we include Timing in the predictor set $d \in \mathbf{X}$ for both complementary two-part models, π_0 and G_Y, and use $\mathscr{F}_d$-information for all covariates $\mathbf{X}$. Next we assume conditional independence between these two models. Finally, if necessary, we define a parametric copula function $\mathrm{Cop}(u, v)$ for the non-default part of the bivariate distribution, to model residual dependence between Multiple and Timing not captured by the Multiple model covariates $\mathbf{X}$. In this case the CDF is constructed as

$$\begin{array}{lclcl} F_0 & := & F_{Yy}(\bar{Y}, \bar{y}|t, \bar{Y} = 0) & = & 1 - S_y(\bar{y}|t) \\ F_1 & := & F_{Yy}(\bar{Y}, \bar{y}|t, \bar{Y} > 0) & = & \mathrm{Cop}[G_Y(\bar{Y}), S_y(\bar{y}|t)] \\ F_{Yy} & := & F_{Yy}(\bar{Y}, \bar{y}|t, \bar{Y} \ge 0) & = & \mathbf{1}_{\{\bar{Y}=0\}} F_0 \pi_0 + \mathbf{1}_{\{\bar{Y}>0\}} [\pi_0 + (1-\pi_0) F_1] \end{array}$$

and the corresponding PDF is given by

$$\begin{array}{lclcl} f_0 & := & f_{Yy}(\bar{Y}, \bar{y}|t, \bar{Y} = 0) & = & f_y(\bar{y}|t)\pi_0 \\ f_1 & := & f_{Yy}(\bar{Y}, \bar{y}|t, \bar{Y} > 0) & = & g_Y(\bar{Y}) \cdot f_y(\bar{y}|t) \cdot \mathrm{cop}[G_Y(\bar{Y}), S_y(\bar{y}|t)] \\ f_{Yy} & := & f_{Yy}(\bar{Y}, \bar{y}|t, \bar{Y} \ge 0) & = & \mathbf{1}_{\{\bar{Y}=0\}} f_0 + \mathbf{1}_{\{\bar{Y}>0\}} f_1 \end{array}$$

with copula function derivative

$$\mathrm{cop}(u, v) = \frac{\delta^2 \mathrm{Cop}(u, v)}{\delta u \delta v}$$

4.3. Model specification

For the exit Timing regression we apply Cox [28]'s multiplicative hazard modeling idea to specify a parametric Weibull survival function. Our approach allows the integration of exogenous, time-variant variables into the exit Timing regression, which is in contrast to the analyses of Giot and Schwienbacher [6], Félix et al. [7], and Jenkinson and Sousa [8] focusing on internal, time-invariant covariates to examine empirical exit routes in VC and

BO. The survival model construction and associated likelihood function are described in Appendix B.

According to our theoretical framework just one single distribution cash flow D_i and one single contribution cash flow C_i per asset is scheduled. However, in real data sets multiple investment and divestment cash flows can be observed for a given company. To account for this practical consideration, we redefine the MULTIPLE regression's dependent variable as

$$Y_{i,t} = \frac{\check{D}_i(t)}{V_i(t) + \check{C}_i(t)} \geq 0 \tag{7}$$

According to equations (A.2) $\check{C}_i(t)$ resp. $\check{D}_i(t)$ represents the sum of all contribution resp. distribution cash flows occurring after t. The corresponding marginal two-part model consists of a logistic regression for π_0 and a generalized linear model relying on a Gamma distribution for G_Y, which are adopted from the R package GAMLSS introduced by [29].

To detect dependency between both marginal models a 180-degree rotated Joe copula is tested

$$\mathrm{Cop}_{\mathrm{Joe}}(u, v; \theta) \quad = \quad 1 - [u^\theta + v^\theta - u^\theta v^\theta]^{1/\theta}$$

with parameter $\theta \geq 1$.

5. Data & model estimation

5.1. Asset level data set

For the empirical application of the joint regression model we use asset level data (stemming from a fund-of-fund program) split into a BO and a VC subset (cf. table 2). All investments had been entered between 1998-09-30 and 2016-12-31. The underlying companies are clustered into 144 BO and 98 VC funds.

The MULTIPLE regression data set excludes all non exited observations and all companies entered after 2009-12-31 in order to alleviate possible sample selection bias. Allowing exit observations of recently entered companies (quick flips) to enter the Multiple analysis data set, causes biased estimates, if there is a significant relation between exit timing and exit performance. Yet exactly this presumed connection is one of the objects under investigation in this paper.

The empirical distributions of both dependent variables are visualized in figure 2. For both fund types, the maximum holding period is approximately 15 years and at least 10% of the Multiple observations are exactly zero.

5.2. Explanatory variables

In both marginal regression models we focus on (i) public market and (ii) timing related covariates. The common public market variables, shared by both marginal analyses, cover

Table 2: Descriptive statistics of the private equity data sets used for the TIMING and MULTIPLE regressions.

TIMING data set	BO	VC
Censored events (exit date > 2016-12-31)	649	531
Observed events (exit date ≤ 2016-12-31)	1,542	2,536

MULTIPLE data set	BO	VC
Realized exits (entry date ≤ 2009-12-31)	1,231	2,179
Thereof with unique company ID	1,108	1,836

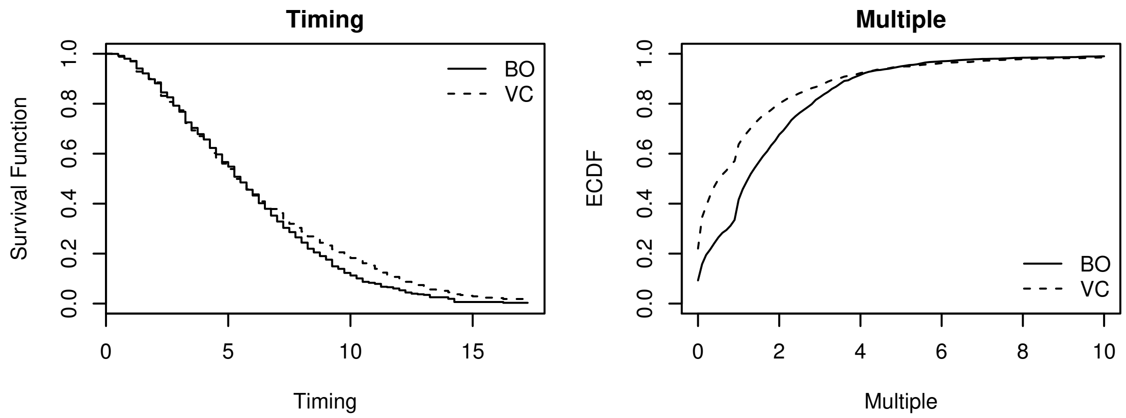


Figure 2: Empirical distribution functions of dependent variables: Kaplan-Meier estimate of the survival function for the TIMING variable in the left plot and empirical cumulative distribution function (ECDF) for the entry-to-exit MULTIPLE variable in the right plot.

high yield spreads[3] and public equity performance[4]. Additionally in the MULTIPLE regression Fama and French [30] factors were tested, but only the Conservative-Minus-Aggressive (CMA) investment factor turns out to be significant. The TIMING regression is obviously not capable of incorporating holding period $(t-c)$ and time to exit $(d-t)$ as independent variables, since they are directly derived from the dependent variable. Instead the fund age at entry date is used as the only time related TIMING covariate. The asset level Residual Value to Paid In (RVPI), defined as value proxy to cumulative contributions ratio (R/V), must be regarded as time-variant *internal* variable [31, section 6.3.2] and therefore is excluded from the set of possible TIMING predictors.

5.3. Estimation procedure

5.3.1. Timing

In the TIMING regression all exited and right-censored observations are included. We explicitly keep multiple entries for a company with numerous fund investors, since the entry and exit TIMING is endogenously determined by each fund manager. Thus the estimation procedure for the marginal TIMING model, i.e. maximizing equation (B.3), is relatively straightforward. However it is computationally more intensive than a parametric multiplicative hazard rate regression with only time-invariant covariates since it involves the numerical integration (stepwise approximation) over the hazard rate function in equation (B.4). The numerical maximum likelihood optimization for our Cox Weibull model is performed in R by the function optimx(... , method= "nlminb") from the optimx package.

5.3.2. Multiple

The estimation procedure for the marginal MULTIPLE model is more intricate, since we have to account for (i) the longitudinal data structure and (ii) economically negligible observations. Here we propose a one-per-company resampling scheme to resolve the within company autocorrelation of the dependent variable characteristic for panel data. In our iterative procedure, in each step exactly one observation per company identifier is randomly chosen to enter the likelihood optimization. This also resolves the issue when multiple funds invest in the same company. However, observations with an RVPI of less than 10% are totally excluded within the resampling algorithm, since they can be regarded economically irrelevant in our opinion. We assume that these observations add more noise than information, as fund managers seem to come up with sloppy firm valuation proxies in these situations. Certainly this solution is just one possible RVPI related weighting method and could be replaced by more elaborate approaches, e.g. weighting the likelihood function by the RVPI in the optimization procedure. The choice of a resampling based estimation method is clearly associated with high computational costs, but on the other hand enables the application of

[3] We use the BofA Merrill Lynch US High Yield Option-Adjusted Spread (https://fred.stlouisfed.org/series/BAMLH0A0HYM2).

[4] In the TIMING regression we incorporate monthly Public Equity Returns (i.e. $\frac{\text{Index}_{t+1}}{\text{Index}_t} - 1$) and in the MULTIPLE regression we use same-horizon Public Equity Multiples (i.e. $\frac{\text{Index}_d}{\text{Index}_t}$) both calculated with the MSCI World Total Return Index in US Dollar.

a simple, yet informative marginal MULTIPLE model and simultaneously provides resampling based standard error estimates (free of charge). The generalized linear models for π_0 and G_Y are separately estimated by the functions gamlss(..., family = BI(mu.link = logit)) and gamlss(..., family = GA)), resp., from the R package GAMLSS [29].

5.3.3. Copula

For the copula estimation we use the inference functions for margins approach of Joe and Xu [32]. This means, first the marginal models are estimated from separately maximized univariate likelihoods, where survival models for TIMING can incorporate non exited investments in contrast to MULTIPLE models. In a second step, the dependence parameter's significance is examined. The 180-degree rotated Joe copula derivative is obtained by the function BiCopPDF(..., family = 16) from the R package VineCopula.

5.4. Parameter estimates

5.4.1. Timing

The coefficient estimates associated with the multiplicative hazard rate model from Appendix B explain the impact of (i) public equity returns, (ii) high yield spreads, and (iii) the fund age at entry on the exit TIMING of individual fund investments (cf. table 3).

Favorable public market conditions, i.e. high public equity returns and low high yield spreads, and a high fund age at entry date result in faster exit TIMINGS. The corresponding Akaike Information Criterion (AIC)[5] values indicate that the relative quality of TIMING models with covariates is superior to a Weibull distribution model without covariates for both BO and VC data sets, since models with minimum AIC are to be preferred. For the BO subset the minimum AIC model (a) contains all three covariates, but for the VC subset the minimum AIC model (b) contains just public equity returns and the fund age at entry as covariates.

In summary the public market related estimates indicate, that high yield spreads posses more predictive power for the BO set and public equity returns posses more predictive power for the VC set. The fund age at entry effect is highly significant for both fund types, however the magnitude of this effect is stronger for the BO set.

5.4.2. Multiple

The coefficient estimates obtained from the two-part model introduced in section 4 explain the impact of (i) public market, (ii) private (proxy) valuation, and (iii) exit timing related covariates on the exit MULTIPLE of individual fund investments (cf. table 4).

Favorable public market conditions, i.e. now a high public equity multiple and a low high yield spread, lead to high MULTIPLE estimates in both sub models, since the signs of the public equity multiple coefficients are positive and the signs of the high yield spread coefficients are negative for π_0 and $\mu(G_Y)$, where π_0 denotes the probability of default,

[5]The AIC is calculated as $AIC = 2k - 2\ln\left(\hat{L}\right)$ where k gives the number of parameters used in a given regression and $\hat{L}$ denotes the maximized likelihood value.

Table 3: Parameter and coefficient estimates of the Timing regression. Three distinct set of covariates (a - c) are tested for each fund type. The standard errors (in parentheses) are obtained from the corresponding Hessian matrix.

Timing estimates	Buy Out			Venture Capital		
Variables	(a)	(b)	(c)	(a)	(b)	(c)
Public equity return	1.892	2.478	-	5.201	5.169	-
	(1.084)	(0.959)	-	(0.756)	(0.775)	-
High yield spread	-3.071	-	-3.588	0.651	-	-0.43
	(1.076)	-	(1.051)	(0.674)	-	(0.67)
Fund age (at entry)	0.086	0.088	0.085	0.036	0.036	0.037
	(0.014)	(0.014)	(0.014)	(0.01)	(0.01)	(0.01)
Scale (Weibull)	6.44	7.221	6.261	7.554	7.355	6.977
	(0.308)	(0.19)	(0.285)	(0.269)	(0.165)	(0.24)
Shape (Weibull)	1.651	1.654	1.652	1.474	1.474	1.482
	(0.034)	(0.034)	(0.034)	(0.269)	(0.165)	(0.24)
AIC (including covariates)	12,598	12,604	12,599	21,058	21,057	21,103
AIC (without covariates)	12,647	12,647	12,647	21,113	21,113	21,113

i.e. a zero multiple, and $\mu(G_Y)$ represents the Gamma distribution mean. The significant negative CMA factor coefficients indicate that this public portfolio could be potentially used to partially hedge both BO and VC exposure. Unsurprisingly, low company valuation proxies (in relation to the initial investment amount) increase the probability of default. Time related covariates in the two-part model just influence the variance of G_Y, i.e. $\text{Var}(G_Y) = \sigma(G_Y)^2 \cdot \mu(G_Y)^2$. Lower past holding periods decrease $\sigma(G_Y)$, whereas high future time-to-exits increase $\sigma(G_Y)$.

The AIC values for the zero and continuous part of the Multiple model indicate that the relative quality of the full covariate regressions are superior to their corresponding intercept only equivalents for both BO and VC data sets.

5.4.3. Copula

The 180-degree rotated Joe copula parameter estimates are 1.135 (0.018) for the BO set and 1.101 (0.015) for the VC set. This means Multiple and Timing can be regarded almost conditional independent, which would be the case for parameter estimate $\theta = 1$.

6. Monte Carlo model on portfolio level

Our reduced form exit dynamics approach is suited for straightforward and computationally inexpensive cash flow simulations, which help to genuinely understand the risk of a given static PEF portfolio, since the underlying framework is capable of processing detailed asset level information. Here portfolio aggregation relies on conditional independence, i.e.

Table 4: Parameter and coefficient estimates of the two-part Multiple regression, where the default probability π_0 is estimated by a logit model and the Gamma distribution G_Y is specified by parameters μ, σ. The resampling procedure is iterated 1,000 times for each fund type. The estimates are mean and standard deviation (in parentheses) that are naturally estimated within the resampling based methodology.

Multiple estimates	Buy Out			Venture Capital		
Covariates	$1-\pi_0$	$\mu(G_Y)$	$\sigma(G_Y)$	$1-\pi_0$	$\mu(G_Y)$	$\sigma(G_Y)$
Intercept	1.720	0.674	-0.089	1.295	0.596	0.384
	(0.156)	(0.066)	(0.034)	(0.106)	(0.101)	(0.019)
Holding period	0.105	-	-0.032	0.133	-	-0.029
	(0.034)	-	(0.013)	(0.023)	-	(0.010)
Time to exit	-	-	0.086	-	-	0.048
	-	-	(0.009)	-	-	(0.005)
RVPI - 1	0.790	-	-	0.341	-	-
	(0.186)	-	-	(0.074)	-	-
Public equity multiple - 1	1.305	0.499	-	0.869	0.524	-
	(0.167)	(0.107)	-	(0.098)	(0.158)	-
High yield spread	-5.260	-3.101	-	-6.264	-3.668	-
	(2.113)	(0.993)	-	(1.403)	(1.787)	-
CMA factor multiple - 1	-0.596	-0.676	-	-0.545	-0.700	-
	(0.344)	(0.193)	-	(0.162)	(0.239)	-
Link function	logit	log	log	logit	log	log
Mean AIC (including covariates)	893	2,743	2,743	1,883	2,392	2,392
Mean AIC (intercept only)	938	2,885	2,885	1,959	2,499	2,499

dependence between companies is solely introduced by common covariates $\mathbf{X}$. Then liquidity and cash flow risk at the α-level for a given fixed horizon z can be conveniently assessed by a portfolio level Cash-Flow-at-Risk (CFaR) measure

$$\begin{aligned}\mathrm{CFaR}_{\alpha,z}(\mathrm{PF}) &= \inf\{Y_t \geq 0 : F_{\mathrm{PF}}(Y_t, z) > \alpha\} \\ F_{\mathrm{PF}}(Y_t, z) &= \mathbb{P}\left[Y_t \leq \sum_{i=1}^{n} w_{it} Y_{it} \mathbf{1}_{\{y_{it} \leq z\}} \mid \mathscr{F}_t\right]\end{aligned}$$

with portfolio weights w_{it} by means of Monte Carlo simulation. Monte Carlo simulation formally constitutes the application of the Multiple conditional on Timing model specified in section 4 with the parameter estimates from section 5 for each portfolio company. Specifically our simulation procedure relies on inverse transform sampling, i.e.

$$\begin{aligned}\widetilde{Y} &= G_Y^{-1}\left(\widetilde{U}_{G_Y} \middle| \xi_Y, \widetilde{\mathbf{X}}, \widetilde{y}\right) \mathbf{1}_{\{\widetilde{U}_\pi > \pi_0(\widetilde{\mathbf{X}})\}} + 0 \\ \widetilde{y} &= S_y^{-1}\left(S_y(t) \cdot \widetilde{U}_y \middle| \xi_y, \widetilde{\mathbf{X}}\right)\end{aligned}$$

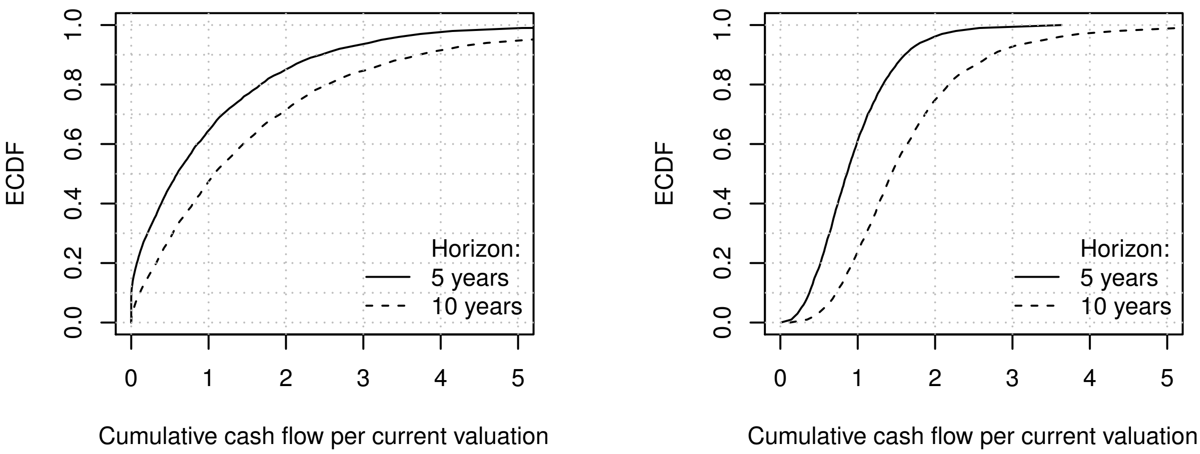
where the tilde symbolizes the simulated nature of a given random variable. Possible future covariate process paths $\widetilde{\mathbf{X}}$ can be held fixed or re-simulated in each iteration step. The default random variable $\widetilde{U}_\pi$ is distributed uniformly i.i.d. and the bivariate uniforms $\left(\widetilde{U}_y, \widetilde{U}_{G_Y}\right)$ are generated from the 180-degree rotated Joe copula $\mathrm{Cop}_{\mathrm{Joe}}$.

Figure 3 shows a Monte Carlo example designed to emphasize the benefit of our (asset level) reduced form approach in the risk management context. Here the cash flow profiles of two VC funds with 3 and 15 companies in their portfolio are compared. The study uses historical public market observations up to 2016-12-31, afterwards *one* hypothetical future path is generated by sampling random permutations from the empirical data. Thus the public market scenario is fixed throughout all iterations. Further all company ages are set to five years and all RVPIs (i.e. V/C) are set to one for both funds, in order to focus on the number of companies in the respective funds. Each run involves 5,000 iterations.

Basically the simulation result suggests that our model can capture large diversification effects on asset level. Typical fund level models on the other hand cannot distinguish funds from the same vintage year at all. Hence, the risk management for small PEF portfolio can substantially benefit from the application of our asset level model instead of crude fund level approaches.

7. Conclusion

Since private equity fund stakes cannot be traded on organized secondary markets, sophisticated cash flow projections are an expedient means to improve the risk understanding of these vehicles. This paper generally contributes a marked point process based framework to describe the inherent cash flow dynamics of private equity funds on single asset level. Specifically the divestment behavior of PEFs is studied in a reduced form approach, that naturally formulates the exit Multiple of a given fund investment conditionally on its exit Timing. With regard to the associated parametric regressions, the asset level granularity enables the inclusion of many insightful covariates, that are otherwise infeasible in pure fund



(a) 3 company VC fund

(b) 15 company VC fund

Figure 3: Monte Carlo simulation example to compare the cash flows associated with generic 3 and 15 company VC funds.

level approaches due to identifiability issues. In our empirical analysis, public market and time related predictors significantly affect both the deal level exit Timing and Multiple of BO and VC investments. Realistic cash flow scenarios for a given static PEF portfolio can be ultimately obtained by Monte Carlo simulations that draw on these model estimates.

The intrinsic risk of undiversified PEFs (with just a few company holdings) may be commonly underestimated by models with implicit or explicit diversification assumptions. Our asset level approach, on the other hand, is capable of reproducing the high default probabilities empirically observed for single BO or VC investments. Innocent fund investors may likely benefit from this enhanced risk perception, while confident fund managers may hardly consent to these estimates in their subjective risk assessments.

In a next step it is natural to generally compare the reduced form and structural approach to PEF asset modeling. In our opinion the considerations of section 3.2 indicate the superiority of reduced form formulations in incomplete information settings like private equity. Here structural models always involve specifications of unobservable quantities like true firm valuations or concepts hard to quantify at all like the hypothetical exit decision process.

On a large scale, this paper takes a novel look on the cash flow stream that is expected to be realized over the next years from the \$2.83 trillion assets held by private equity funds (as of June 2017; Source: Preqin). Perspectively, our marked point process framework can be also used to model the contribution cash flows resulting form undrawn private equity fund commitments, as there are \$1.09 trillion of dry powder held by PEFs (as of March 2018; Source: Preqin). Conveniently, even a comprehensive \$3.92 trillion model can draw on our net cash flow process from equation (A.1).

Appendix A. More general marked point process model

Now we relax the restrictive assumption of just one investment and divestment cash flow per given company. Inspired by the marked point process (MPP) construction of Biagini and Zhang [33], we describe the ith company cash flows by a $\mathscr{G}$-adapted marked point process $(\tau_{i,k}, \Psi_{i,k})_{k\in\mathbb{N}_{>0}}$ with two-dimensional marks $\Psi_{i,k} = (\mathrm{A}_{i,k}, \mathrm{L}_{i,k})$. Here $\tau_{i,k} \in \mathbb{R}_{>0}$ gives the timing of the kth cash flow associated with the ith company, $\mathrm{A}_{i,k} \in \mathbb{R}_{\geq 0}$ denotes the corresponding cash amount, and $\mathrm{L}_{i,k} \in \{1, 2, 3\}$ labels (1) contribution, (2) distribution, and (3) other events.

The general net cash flow process on fund level is then defined as

$$NCF_{\text{Fund}}(t) := \sum_{i=1}^{N^{(c)}(t)} \sum_{k=1}^{\infty} \mathbf{1}_{\left\{\tau_{i,k}>t\right\}} A_{i,k} \tag{A.1}$$

Again the benchmark approach, sketched in section 3.3, can be applied to value (the payment stream associated with) the current fund stake by means of the growth optimal portfolio S^*

$$\dot{V}_{\text{Fund}}(t) = S^*(t) \cdot \mathbb{E}^{\mathbb{P}} \left[\sum_{i=1}^{N^{(c)}(t)} \sum_{k=1}^{\infty} \mathbf{1}_{\left\{\tau_{i,k}>t\right\}} \frac{A_{i,k}}{S^*(\tau_{i,k})} \middle| \mathscr{G}_t \right]$$

The simple cash flow model introduced in section 3.1 can be pathwise defined in terms of the general MPP for $t \in [c_i, d_i]$

$$\begin{array}{lll}
\check{c}_i & := & \inf\left[\tau_{i,k} : \Psi_{i,k} \in \mathcal{B}\left(\mathbb{R}_{\geq 0}, \{1\}\right), \ \ k \in \mathbb{N}_{>0}\right] \wedge T^* \\
\check{d}_i & := & \sup\left[\tau_{i,k} : \Psi_{i,k} \in \mathcal{B}\left(\mathbb{R}_{\geq 0}, \{2\}\right), \ \ k \in \mathbb{N}_{>0}\right] \wedge T^* \\
\check{o}_i & := & \sup\left[\tau_{i,k} : \Psi_{i,k} \in \mathcal{B}\left(\mathbb{R}_{\geq 0}, \{3\}\right), \ \ k \in \mathbb{N}_{>0}\right] \wedge T^* \\
\check{C}_i(t) & := & \sum_{k=1}^{\infty} \mathbf{1}_{\left\{\tau_{i,k}>t\right\}} \mathbf{1}_{\left\{\Psi_{i,k}\in\mathcal{B}\left(\mathbb{R}_{\geq 0},\{1\}\right)\right\}} A_{i,k} \cdot (-1) \\
\check{D}_i(t) & := & \sum_{k=1}^{\infty} \mathbf{1}_{\left\{\tau_{i,k}>t\right\}} \mathbf{1}_{\left\{\Psi_{i,k}\in\mathcal{B}\left(\mathbb{R}_{\geq 0},\{2\}\right)\right\}} A_{i,k} \\
\check{O}_i(t) & := & \sum_{k=1}^{\infty} \mathbf{1}_{\left\{\tau_{i,k}>t\right\}} \mathbf{1}_{\left\{\Psi_{i,k}\in\mathcal{B}\left(\mathbb{R}_{\geq 0},\{3\}\right)\right\}} A_{i,k}
\end{array} \tag{A.2}$$

which yields a reasonable conservative timing of cash flows suitable for risk management purposes, if we assume equations (A.2) to be $\mathscr{F}_t$-measurable for all $t \in (\mathbf{A} \vee \mathbf{B})$.

Appendix B. Timing: Parametric multiplicative hazard rate model

The Timing variable of interest is time to exit $y = d - t$. However, in agreement with conventional survival analysis practice, the entire survival function $S(t) = \mathbb{P}[T > t \,|\mathscr{F}_t]$ is estimated for each asset. This means, the total lifetime (holding period) of each company, i.e. $T = d - c$, is used as dependent variable in the Timing model to base the analysis on a common time scale.

Since $N_{i,j}^{(d)}(t)$ is not $\mathscr{F}_t$-observable for $t \notin (\mathbf{A} \vee \mathbf{B})$, we define the lagged exit timing process $N_{i,j}^{(d^*)}(t)$ by the lagged event time which corresponds to the right interval boundary

$$d_{i,j}^* := \inf\left\{t > 0 : N_{i,j}^{(d)}(t) > 0 \middle| \mathscr{F}_t\right\} = d_{i,j}^{(R)} \tag{B.1}$$

Let this point process $N_{i,j}^{(d^*)}(t)$ model the exit TIMING $T_{i,j}$ of the ith company belonging to the jth fund in line with the intensity based definition of Bremaud [34, II.3]

$$\mathbb{E}\left[\int_0^\infty dN_{i,j}^{(d^*)}(u)\bigg|\,\mathscr{F}_t\right] = \mathbb{E}\left[\int_0^\infty h_{i,j}\left(u\middle|\,\mathbf{X}(u)\right)du\bigg|\,\mathscr{F}_t\right]$$

The associated random intensity process is characterized according to [35]

$$h_{i,j}\left(t\,|\mathbf{X}(t)\right) = h_0(t)\,Z_{i,j}^{(\text{cens})}(t)\exp\left(\beta'\mathbf{X}(t)\right)$$

with the $\mathscr{F}_t$-predictable censoring process $Z_{i,j}^{(\text{cens})}(t)$ from equation (1) and the $\mathscr{F}_t$-predictable covariate process $\mathbf{X}(t)$ both defined in section 3.1. Further, we assume Weibull distributed exit TIMINGS, i.e. a parametric model for the base hazard function $h_0(t)$. The corresponding Weibull cumulative hazard function is given by a simple closed form expression

$$H_0^{(\text{wb})}\left(t\,|\xi_y\right) = \int_0^t h_0^{(\text{wb})}\left(u\,|\xi_y\right)du = \left(\frac{t}{\text{scale}_{\text{wb}}}\right)^{\text{shape}_{\text{wb}}}$$

with parameter vector $\xi_y = (\text{shape}_{\text{wb}}, \text{scale}_{\text{wb}})$.

The survival function for a parametric Cox model with time-variant covariates is calculated by integrating over the multiplicative intensity process

$$S_{\text{Cox}}(t) = \exp\left(-\int_0^t h_0\left(u\,|\xi_y\right)Z_{i,j}^{(\text{cens})}(u)\exp\left(\beta'\mathbf{X}(u)\right)du\right) \tag{B.2}$$

To account for the lagged definition of $N_{i,j}^{(d^*)}(t)$ in equation (B.1), we apply the following full likelihood approach to estimate the parametric Cox model with (i) time-variant covariates, (ii) an interval-censored data structure and (iii) possible final (i.e. nonrandom) right-censoring

$$L\left(\beta,\xi_y\,|\mathbf{T},\mathbf{X}\right) = \prod_{j=1}^{J}\left[\prod_{i=1}^{n_j}\left\{\left[S_{\text{Cox}}\left(T_{i,j}^{(L)}\right) - S_{\text{Cox}}\left(T_{i,j}^{(R)}\right)\right]^{(1-\mathcal{R}_{i,j})}\left[S_{\text{Cox}}\left(T_{i,j}^{(R)}\right)\right]^{\mathcal{R}_{i,j}}\right\}\right] \tag{B.3}$$

where the left and right boundaries of the exit TIMING interval are given by

$$T_{i,j}^{(L)} = d_{i,j}^{(L)} - c_{i,j}^{(L)} \qquad \text{and} \qquad T_{i,j}^{(R)} = \min\left(f_{i,j}, d_{i,j}^{(R)}\right) - c_{i,j}^{(L)}$$

and the final observation time for a given asset is denoted by $f_{i,j}$ with corresponding right censoring indicator

$$\mathcal{R}_{i,j} = \mathbf{1}_{\left\{f_{i,j} < d_{i,j}^{(R)}\right\}}$$

This allows us to construct the adjusted likelihood for non-informative right-censoring [36, section 5.1.2]. Here n_j gives the number of investments for the jth fund and $j = 1, 2, \ldots, J$ counts the number of distinct funds in the data set.

Further we assume a step function for the dynamic covariate process $\mathbf{X}(t)$ since we have to integrate over the history of this process in the course of calculating the survival function. In the maximum likelihood estimation procedure for the Cox model with a Weibull hazard rate function we compute a quarterly time-discrete approximation of equation (B.2), i.e.

$$\begin{aligned} S_{\text{Cox}}^{(\text{wb})}(t) &= \exp\left(- \sum_{q:t_q \leq t} \exp\left(\beta' \mathbf{X}(t_q)\right) \int_{t_{q-1}}^{t_q} Z_{i,j}^{(\text{cens})}(u) h_0^{(\text{wb})}(u \,|\xi_y)\, du \right) = \\ &= \exp\left(- \sum_{q:t_q \leq t} \exp\left(\beta' \mathbf{X}(t_q)\right) Z_{i,j}^{(\text{cens})}(t_q) \left[H_0^{(\text{wb})}(t_q \,|\xi_y) - H_0^{(\text{wb})}(t_{q-1} \,|\xi_y) \right] \right) \end{aligned} \tag{B.4}$$

with $t_{q-1} \in \mathbf{A}$ and $t_q \in \mathbf{B}$ where $q = 1, 2, \ldots, Q$. The covariate information $\mathbf{X}(t_q)$ is here assumed to be relevant for the time in between t_{q-1} and t_q.

References

[1] S. N. Kaplan, P. Strömberg, Leveraged buyouts and private equity, Journal of Economic Perspectives 23 (1) (2009) 121–146.

[2] K. Giesecke, Default and information, Journal of Economic Dynamics and Control 30 (2006) 2281–2303.

[3] E. Platen, A benchmark approach to finance, Mathematical Finance 16 (1) (2006) 131–151.

[4] S. R. Das, M. Jagannathan, A. Sarin, The private equity discount: An empirical examination of the exit of venture backed companies, unpublished working paper, 2002.

[5] A. Ljungqvist, M. Richardson, The investment behavior of private equity fund managers, URL `https://archive.nyu.edu/jspui/bitstream/2451/27031/2/S-FI-03-11.pdf`, unpublished working paper, 2003.

[6] P. Giot, A. Schwienbacher, Ipos, trade sales and liquidations: Modelling venture capital exits using survival analysis, Journal of Banking and Finance 31 (2007) 679–702.

[7] E. G. S. Félix, C. P. Pires, M. A. Gulamhussen, The exit decision in the European venture capital market, Quantitative Finance 14 (6) (2014) 1115–1130.

[8] T. Jenkinson, M. Sousa, What determines the exit decision for leveraged buyouts?, Journal of Banking and Finance 59 (2015) 399–408.

[9] D. Cumming, Contracts and exits in venture capital finance, Review of Financial Studies 21 (5) (2008) 1947–1982.

[10] D. M. Schmidt, S. Steffen, F. Szabó, Exit strategies of buyout investments: An empirical analysis, Journal of Alternative Investments 12 (4) (2010) 58–84.

[11] D. Cumming, L. Grilli, S. Murtinu, Governmental and independent venture capital investments in Europe: A firm-level performance analysis, Journal of Corporate Finance 42 (2014) 439–459.

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

[13] A. Korteweg, M. Sorensen, Risk and return characteristics of venture capital-backed entrepreneurial companies, Review of Financial Studies 23 (10) (2010) 3738–3772.

[14] S. Guo, E. S. Hotchkiss, W. Song, Do buyouts (still) create value?, Journal of Finance 66 (2) (2011) 479–517.

[15] P. Valkama, M. Maula, E. Nikoskelainen, M. Wright, Drivers of holding period firm-level returns in private equity-backed buyouts, Journal of Banking and Finance 37 (7) (2013) 2378–2391.

[16] E. de Malherbe, Modeling private equity funds and private equity collateralised fund obligations, International Journal of Theoretical and Applied Finance 7 (3) (2004) 193–230.

[17] A. Buchner, Risk management for private equity funds, Journal of Risk 19 (6) (2017) 1–32.

[18] A. Buchner, C. Kaserer, N. Wagner, Modeling the cash flow dynamics of private equity funds: Theory and empirical evidence, Journal of Alternative Investments 13 (1) (2010) 41–54.
[19] D. Bongaerts, E. Charlier, Private equity and regulatory capital, Journal of Banking and Finance 33 (2009) 1211–1220.
[20] R. Braun, N. Engel, P. Hieber, R. Zagst, The risk appetite of private equity sponsors, Journal of Empirical Finance 18 (5) (2011) 815–832.
[21] M. Escobar, P. Hieber, M. Scherer, L. Seco, Portfolio optimization in a multidimensional structural-default model with a focus on private equity, Journal of Private Equity 15 (1) (2011) 26–35.
[22] F. Dong, N. Chiara, N. Kokkaew, A. Xu, Copula-based portfolio credit risk assessment in infrastructure project financing, Journal of Private Equity 15 (2) (2012) 31–30.
[23] A. Lahmann, M. Schreiter, B. Schwetzler, Modeling dynamic redemption and default risk for lbo evaluation: A boundary crossing approach, unpublished working paper, 2016.
[24] J. Sun, The Statistical Analysis of Interval-censored Failure Time Data, Springer, New York, 2007.
[25] R. A. Jarrow, P. Protter, Structural versus reduced-form models: A new information-based perspective, Journal of Investment Management 2 (2) (2004) 1–10.
[26] J. B. Long, The numeraire portfolio, Journal of Financial Economics 26 (1990) 29–69.
[27] P. Shi, X. Feng, A. Ivantsova, Dependent frequency severity modeling of insurance claims, Insurance: Mathematics and Economics 64 (2015) 417–428.
[28] D. R. Cox, Regression models and life-tables, Journal of the Royal Statistical Society 34 (2) (1972) 187–220.
[29] R. A. Rigby, D. M. Stasinopoulos, Generalized additive models for location, scale and shape,(with discussion), Applied Statistics 54 (2005) 507–554.
[30] E. F. Fama, K. R. French, A five-factor asset pricing model, Journal of Fin. Econ. 116 (1) (2015) 1–22.
[31] J. D. Kalbfleisch, R. L. Prentice, The Statistical Analysis of Failure Time Data, John Wiley & Sons, Hoboken, NJ, 2002.
[32] H. Joe, J. J. Xu, The estimation method of inference functions for margins for multivariate models, Tech. Rep. 166, Department of Statistics, University of British Colombia, 1996.
[33] F. Biagini, Y. Zhang, A unified modeling framework for life and non-life insurance, URL `https://arxiv.org/pdf/1802.07741.pdf`, unpublished working paper, 2018.
[34] P. Bremaud, Point Processes and Queues. Martingale Dynamics, Springer, New York, 1981.
[35] P. Andersen, R. D. Gill, Cox's regression model for counting processes: A large sample study, Annals of Statistics 10 (4) (1982) 1100–1120.
[36] O. O. Aalen, O. Borgan, H. K. Gjessing, Survival and Event History Analysis, Springer, New York, 2008.