Available online at www.sciencedirect.com
ScienceDirect
The Journal of Finance and Data Science 5 (2019) 127−139
http://www.keaipublishing.com/en/journals/jfds/

Quadratic hedging strategies for private equity fund payment streams

Christian Tausch

AssetMetrix GmbH, Theresienhöhe 13, D-80339 Munich, Germany

Received 23 April 2019; revised 27 May 2019; accepted 27 August 2019
Available online 4 September 2019

Abstract

To better understand the relation between public markets and private equity, we consider quadratic hedging strategies to replicate the typical payment stream pattern associated with private equity funds by traded factors. Our methodology is inspired by the risk-minimization framework developed in financial mathematics and applies the componentwise L_2 Boosting machine learning technique to empirically identify feasible replication strategies. The application to US venture capital fund data further draws on a stability selection procedure to enhance model sparsity. Interestingly a natural connection to the famous Kaplan and Schoar (2005) public market equivalent approach can be established.

Keywords: Private equity fund; Cash flow replication; Public market equivalent; Componentwise boosting; Machine learning

1. Introduction

A fund manager enters their favorite investment boutique and asks: How to replicate private equity fund (PEF) cash flows by public market trading strategies?

Despite the continued popularity of PEFs among institutional investors, with evermore capital committed to these illiquid investment vehicles,[1] the question about liquid alternatives to PEFs remains largely unanswered by the academic literature. Particularly the identification of feasible traded-factor based replication strategies for the cash flow streams associated with PEFs seems appealing from various risk management and performance evaluation perspectives: They can be readily used as (1. replication) easy to execute investment alternative to a PEF commitment, (2. hedging) interim hedging instrument, and (3. benchmarking) advanced public market equivalent methodology for benchmarking purposes. Generally liquid alternatives can be used to temporarily adjust private equity exposure either as under-allocation guidance for over-committed PEF investors or over-allocation guidance for under-committed PEF investors.

E-mail address: christian.tausch@quant-unit.com.

Peer review under responsibility of KeAi.

[1] Typical PEFs exhibit investment horizons of more than ten years and only limited ability to exit a given fund investment before maturity (via a secondary sale).

https://doi.org/10.1016/j.jfds.2019.08.002

Yet some replication concepts exist. The holding based approaches of Madhavan and Sobczyk[1] and Porter and Porter[2] that operate on company level aim at the bottom-up construction of investable (gross of fee) mimicking portfolios; here the identification of suitable portfolio underlyings is facilitated by modern data science technology. Following the fund level risk-adjusted profit concept of Gupta and Nieuwerburgh[3] who form net of fee replicating portfolios for PEFs, a public hedging strategy decomposes PEF cash flows into a systematic traded component and a residual term that is idiosyncratic and private equity specific by construction. Hence the maybe most interesting - though theoretical – application is the extraction of the true private equity 'alpha' by running a (hypothetical) strategy that commits to all PE funds in the universe and simultaneously hedges the public market exposure.

Our fund level and cash flow based methodology relies on the fundamental result established in financial mathematics that in incomplete financial markets the exact replication of contingent claims by means of self-financing strategies is generally not possible. Optimal hedging strategies therefore must either sacrifice the (1) exact replication or (2) self-financing property. Our approach is initially inspired by the risk minimization for payments streams framework introduced by Moller[4] which insists on exact replication. His general payoff stream setting extends the quadratic hedging in incomplete markets literature pioneered by Föllmer and Sondermann[5] and subsequently refined by most notably Martin Schweizer (e.g., Schweizer[6,7]).

To empirically estimate feasible trading strategies, we combine the notion of quadratic hedging strategy known from the risk minimization in incomplete markets literature and the L_2 Boosting machine learning technique.[8] Therefore, we translate the generic risk minimization problem to the new more applied task of searching for the optimal combination of predictors and traded return factors that best describe observed private equity fund net cash flows (and potentially net asset values). Specifically, we employ L_2 Boosting with componentwise linear least squares to sequentially select the best hedge for private equity fund cash flows in a high-dimensional covariate space setting, i.e., there are many potential predictors and traded return factor candidates and we only want to select (the most) significant ones. Here componentwise L_2 Boosting even allows for having more selected covariates than observations. However we have to apply a stability selection procedure to further enhance model sparsity.

The structure of the paper is as follows. Section 2 introduces a private equity adapted quadratic hedging framework inspired by financial mathematics and highlights its connection to the famous Kaplan and Schoar[9] public market equivalent approach. Section 3 explains the componentwise L_2 Boosting algorithm and illustrates a stability selection method to enhance model sparsity. Section 4 demonstrates an empirical application to US venture capital fund data. Section 5 concludes.

2. Quadratic hedging strategies for private equity funds

2.1. Mathematical framework

Private equity fund data typically comes in panel data structure with fund-or-portfolio index $i = 1, \ldots, N$ and (monthly) time index $t = 1,\ldots,T$. Here $\tilde{V}_{i,t}$ and $\widetilde{\mathrm{CF}}_{i,t}$, respectively, denote the (proxy net asset) value and net cash flow corresponding to fund-or-portfolio i and realized at time t. Our hedging strategy, denoted by ξ, draws on a set of J zero-net-investment public market factors $F_{j,t}$ with $j = 1, \ldots, J$ constructed by one long and one complementary short position

$$F_{j,t} = \left(\frac{S_t^{(j+)} - S_{t-1}^{(j+)}}{S_{t-1}^{(j+)}}\right) - \left(\frac{S_t^{(j-)} - S_{t-1}^{(j-)}}{S_{t-1}^{(j-)}}\right)$$

where $S_t^{(j+)}/S_t^{(j-)} \in \mathbb{R}_{\geq 0}$ is the asset value belonging to the jth factor's long/short position, respectively. In analogy to the risk minimization framework, a suitable numeraire $S^* \in \mathbb{R}_{>0}$ is applied to discount all financial variables of interest. We start with the fund proxy valuation

$$V_{i,t} := \frac{\tilde{V}_{i,t}}{S_t^*}$$

and corresponding cumulative cash flow

$$A_{i,t} := \sum_{\tau=1}^{t} \frac{\widetilde{\mathrm{CF}}_{i,\tau}}{S_\tau^*}$$

The cumulative gain function associated with a given hedging strategy is given by

$$G_t := \sum_{\tau=1}^{t} \sum_{j=1}^{J} \xi_{j,\tau} \frac{F_{j,\tau} - \lambda}{S_\tau^*} \tag{1}$$

relying on a hedging strategy ξ of the following linear functional form

$$\xi_{j,\tau} = \sum_{k=1}^{K} \beta_{j,k} \cdot P_{\tau-1}^{(k)} \cdot \tilde{V}_{\tau-1}$$

where P_τ is a K-dimensional vector of potential predictors and $\beta \in \mathbb{R}^{J \times K}$ are the linear coefficients to be estimated. The constant $\lambda \in \mathbb{R}_{\geq 0}$ represents proportional transaction cost which are assumed to be time- and factor-independent. Instead of using reported net asset values we could employ an autocorrelation model to adjust them for stale pricing. Next, the replication target of our hedging strategy is simply

$$Y_{i,t} = V_{i,t} + A_{i,t} \tag{2}$$

On this basis several quadratic loss function specifications for a given fund-or-portfolio i are possible, e.g.,

$$\begin{aligned} L_i^{(A)} &= \frac{1}{T_i} \sum\nolimits_{t=1}^{T_i} (Y_{i,t} - G_{i,t})^2 \\ L_i^{(B)} &= \frac{1}{T_i} \sum\nolimits_{t=1}^{T_i} [(Y_{i,T_i} - G_{i,T_i}) - (Y_{i,t} - G_{i,t})]^2 \end{aligned}$$

where T_i denotes the ith fund's terminal date. However we favor the squared final hedging error loss function

$$L_i = (Y_{i,T_i} - G_{i,T_i})^2 \tag{3}$$

since just at time T_i the target function solely includes the cash flow term we ultimately want to replicate (as there is no remaining net asset value, i.e., $Y_{i,T_i} = A_{i,T_i}$ and $V_{T_i} = 0$). W.l.o.g. this paper assumes the fund's terminal date T_i to be 15 years after inception in all cases. To generally obtain the optimal hedging gain function G^* that best approximates Y, we have to account for the random process nature of $\tilde{V}, \widetilde{\mathrm{CF}}, S, P$ when minimizing the expected value of the loss function

$$G^* = \underset{G}{\operatorname{argmin}} \mathbb{E}[L]$$

In practice we have to resort on the empirical loss function estimate and determine the optimal β coefficient vector according to

$$\beta^* = \underset{\beta}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} L_i(\beta)$$

Example (US VC 1992): Let us estimate a ex-post (Equation (3) minimizing) replication strategy for US Venture Capital (VC) funds of a given historical vintage, i.e., $N = 1$. The MSCI World Index is chosen as numeraire and the ‘NASDAQ Composite minus MSCI World’ excess return as single traded factor. To keep it simple, we only allow for univariate ‘constant’ trading strategies of functional form (with zero transaction cost $\lambda = 0$)

$$G_t^{(\mathrm{Example})} = \sum_{\tau=1}^{t} \beta \cdot 1 \cdot V_{\tau-1} \cdot \frac{F_\tau^{(\mathrm{NASDAQ-MSCI.World})} - 0}{S_\tau^{(\mathrm{MSCI.World})}}$$

The ex-post optimal strategy for a historical payment stream can be straightforwardly obtained by numerically minimizing the empirical risk function. Exemplarily, the result for US VC funds of vintage 1992 from the Preqin cash flow data set (described in section 4.1) is visualized in Fig. 1. For this vintage the optimal coefficient is $\widehat{\beta} = 2.747399$ and (due the ex-post design) we obtain an almost perfect hedge of numeraire denominated final cash flows.

2.2. Connection to Kaplan and Schoar[9]

Let us take a stochastic discount factor (SDF) perspective on our methodology from section 2.1 and start with the historical background of public market equivalent (PME) approaches. Long and Nickels[10] were the first to propose a method to compare private investment returns to the performance of public markets by exactly matching all fund in- and outflows and investing the remaining capital in a public index (known as index comparison method). In their introduction they adopt the six appropriate benchmark characteristics defined by Bailey[11]; accordingly, a valid benchmark shall be: (1) unambiguous, (2) investable, (3) measurable, (4) appropriate, (5) reflective of current investment opinions, and (6) specified in advance. Particularly the appropriateness of a given public market index as well as the need for additional risk-adjustments remain much debated questions until now.[12,13]

In recent years, the attention of PME approaches tailored for private equity shifted towards performance evaluation by means of SDF valuation models as proposed by Farnsworth et al[14] for mutual funds and by Li et al[15] for hedge funds. In the PEF context, the methodology is best described by the generalized public market equivalent measure introduced by Korteweg and Nagel.[16]

$$\mathrm{GPME}_{i,T_i} := \sum_{\tau=1}^{T_i} \Psi_\tau \cdot \widetilde{\mathrm{CF}}_{i,\tau}$$

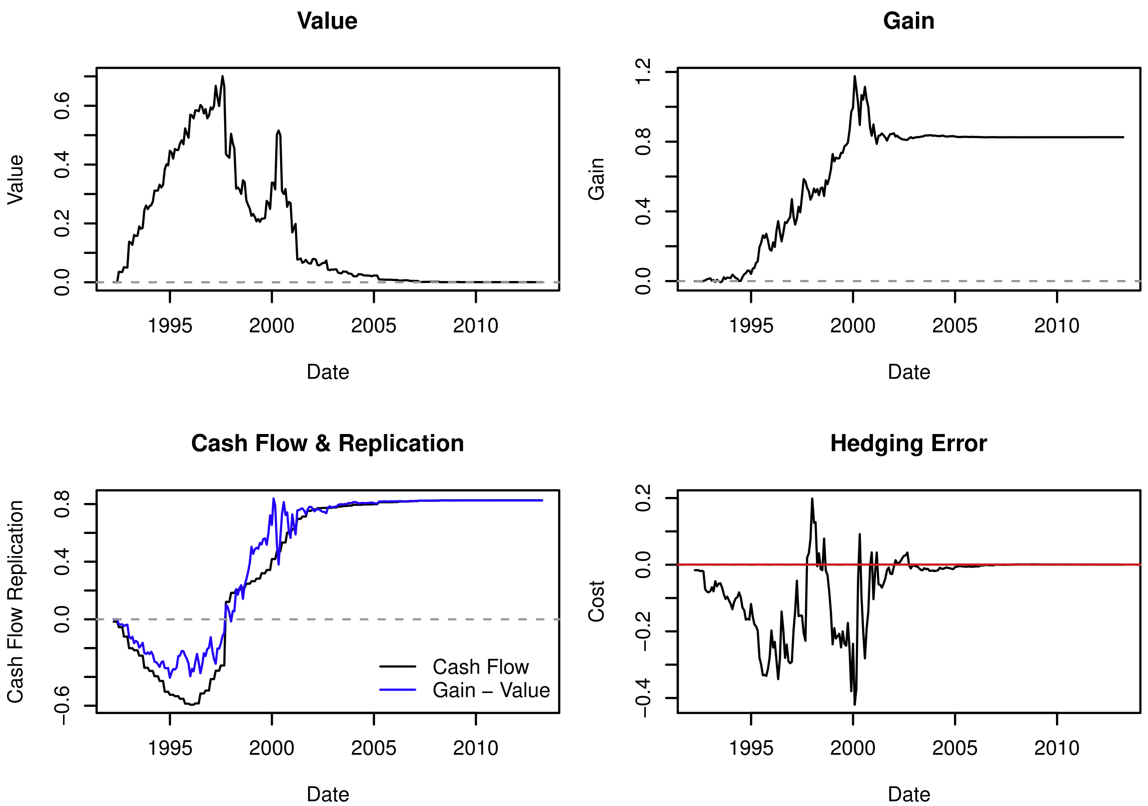

Fig. 1. Simple ex-post replication strategy for US VC funds of vintage 1992. Here the variables introduced in section 2.1 are used as follows: Value is V, Gain is G, Cash Flow is A, and Hedging Error is calculated as $A - (G - V) = Y - G$.

where Ψ denotes a generic SDF (i.e., generally a positive random variable) and $\widetilde{\mathrm{CF}}_i$ is the vector of fund cash flows as previously defined. To apply a given parametric SDF, the parameters associated with Ψ are usually estimated beforehand within a Generalized Method of Moments procedure by pricing predictable dynamic trading strategies that draw on so-called primitive public market securities.[14] Linear pricing models like Driessen et al,[17] which are intrinsically investable by their linear construction, suffer from the well-known issue of permitting arbitrage opportunities by assessing negative prices to strictly non-negative payoffs when strongly negative factor returns occur in combination with high factor loadings. In other words, these linear factor models generally cannot be used to form an always positive numeraire portfolio in the sense of Long[18] to discount PEF cash flows like in the seminal Kaplan and Schoar[9] public market equivalent approach.

Numeraire denomination of cash flows, an idea inherent to every risk-minimization framework,[6] can be conveniently incorporated into the GPME framework and yields a methodology almost identical to the Kaplan and Schoar[9] public market equivalent approach, except that here the difference of discounted fund in- and outflows is taken instead of the ratio.[16]

$$\mathrm{GPME}_{i,T_i}^{(\mathrm{KS05})} := A_{i,T_i} = \sum_{\tau=1}^{T_i} \frac{1}{S_\tau^*} \cdot \widetilde{\mathrm{CF}}_{i,\tau}$$

This means the loss function from Equation (3) can be re-written as

$$L_i = \left(\mathrm{GPME}_{i,T_i}^{(\mathrm{KS05})} - G_{i,T_i} \right)^2$$

Moreover we can also use the gain function from Equation (1) to define an analog to the classical Kaplan and Schoar[9] public market equivalent ratio

$$\mathrm{PME}_{i,t}^{(\mathrm{KS05})} := \frac{A_{i,t}^{(+)} + V_{i,t}}{-A_{i,t}^{(-)}} = 1 + \frac{G_{i,t}}{-A_{i,t}^{(-)}} + \varepsilon$$

where $A_{i,t}^{(+)}$ and $-A_{i,t}^{(-)}$ denote numeraire-denominated cumulative distributions and contributions, respectively, and ε is the error term between the traditional measure (on the left-hand side) and our new ratio approach (on the right).

To conclude, our framework suggested in section 2.1 is the Kaplan and Schoar[9] related special case of the following general SDF-based quadratic loss function

$$L_i^{(\mathrm{SDF})} = \left\{ \left[\sum_{\tau=1}^{T_i} \Psi_\tau \cdot \widetilde{\mathrm{CF}}_{i,\tau} \right] - \left[\sum_{\tau=1}^{T_i} \sum_{j=1}^{J} \Psi_\tau \cdot \xi_{j,\tau} \left(F_{j,\tau} - \lambda \right) \right] \right\}^2$$

which is the squared difference between Ψ-discounted observed cash flows $\widetilde{\mathrm{CF}}$ and the Ψ-discounted payoff associated with the linear replication strategy. The identification of the most suitable SDF model (generally or in the private equity context) is certainly an interesting topic, although not tackled by this paper. Our aim is to develop investable and liquid investment alternatives to long-term private equity fund commitments. With this special emphasis on the ability to actually invest, we rather focus on replication cash flows than the SDF applied.

3. Componentwise L_2 boosting

Now let us estimate general (ex-ante) replication strategies that can be employed for the hedging and replication of future cash flows. Due to the quadratic loss function specification in Equation (3) we use the machine learning technique Componentwise L_2 Boosting (CLB) for that purpose.[8] It is an iterative procedure that combines model estimation and model selection; in every step just the coefficient estimate with maximal marginal univariate explanatory power is updated. The high dimensional variable selection feature is especially important, since our strategy draws on $J \times K$ combinations of J excess return factor candidates and K potential predictors. In contrast to many other statistical learning routines the CLB coefficient estimates possess straightforward interpretability comparable to that of simple linear regressions. Similar componentwise boosting approaches are applied for predictor selection from a large feasible set in an econometric context by Bai and Ng[19] and Mittnik et al.[20]

3.1. Base procedure

We define our L_2 Boosting algorithm with componentwise linear least squares akin to Bühlmann[8], Section 2 of Bühlmann and start with the base procedure specification that selects the univariate factor-predictor combination with maximal explanatory power $\widehat{S} = (\widehat{j,k})$

$$\widehat{S} = \underset{j,k}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \left(u_i - \widehat{g}_i^{(j,k)} \right)^2 \tag{4}$$

with pseudo-response variable u_i and optimal univariate gain function analog to Equation (1)

$$\widehat{g}_i^{(j,k)} = \sum_{\tau=1}^{T_i} \widehat{\beta}_{j,k} \bullet P_{\tau-1}^{(k)} \bullet \tilde{V}_{i,\tau-1} \bullet \frac{F_{j,\tau} - \lambda}{S_\tau^*} \tag{5}$$

In turn, Equation (5) draws on results obtained by the following univariate minimization which estimates the optimal univariate β coefficient for a given factor-predictor combination j, k.

$$\widehat{\beta}_{j,k} = \underset{\beta_{j,k}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \left(u_i - g_i^{(j,k)} \right)^2$$

with

$$g_i^{(j,k)} = \sum_{\tau=1}^{T_i} \beta_{j,k} \bullet P_{\tau-1}^{(k)} \bullet \tilde{V}_{i,\tau-1} \bullet \frac{F_{j,\tau} - \lambda}{S_\tau^*}$$

For a given pseudo-response variable u_i the base procedure initially estimates optimal linear coefficients $\widehat{\beta}_{j,k}$ for all $J \times K$ combinations, which is the computationally intensive part. Based on $\widehat{\beta}_{j,k}$ it then selects the optimal excess return factor and predictor combination $\widehat{S} = (\widehat{j,k})$ that reduces the residual sum of squares most. The univariate function selected and estimated by this base procedure is denoted by

$$\widehat{g}^{(\widehat{S})}(u) = \widehat{g}^{(\widehat{S})}(u, \ldots)$$

where for simplicity the pseudo-response vector u is used as single function argument.

3.2. Algorithm: componentwise L_2boosting

Step 1 (initialization). At the initial step, start with the no hedge situation where all β coefficients are set to zero. A simple application of the base procedure yields the first function estimate

$$\widehat{f}^{(0)}(\bullet) = \widehat{g}^{(\widehat{S})}(Y)$$

where the replication target from Equation (2) is used as pseudo-response variable $u = Y$. Set $m = 0$.

Step 2 (update). Apply the base function to the new residuals and update the previous function estimate according to

$$\widehat{f}^{(m+1)}(\bullet) = \widehat{f}^{(m)}(\bullet) + \nu \bullet \widehat{g}^{(\widehat{S})}\left(u^{(m)}\right)$$

with new residual vector $u^{(m)} = y - \widehat{f}^{(m)}(\bullet)$ and step size $0 < \nu \leq 1$.

Step 3 (iteration). Increase the iteration index m by one and repeat step 2 until a stopping iteration M is achieved.

3.3. Enhancing sparsity by stability selection

The high-dimensional variable selection literature proposes various techniques to further enhance the sparsity of componentwise boosting. These include e.g., re- and sub-sampling methods like stability selection,[21] twin-boosting,[22] sparse boosting,[23] and general sparse boosting.[24] However the standardization of predictor-factor pairs, needed for e.g., twin-boosting, is not perfectly straightforward in our panel data case.

For the detection of false positives (i.e., irrelevant predictor-factor pairs selected by the CLB procedure) we apply an ensemble method similar to stability selection which runs CLB with several distinct, but reasonable numeraire choices S^* and subsequently analyzes all selected predictor-factor combinations. Consequently, only predictor-factor pairs consistently selected (and also exhibiting the same sign) for many numeraire indices are included in the final CLB run. Additionally, the incorporation of proportional transaction costs λ in Equation (1) serves as regularization-like penalty term. Due to the real world interpretation of λ, we again can vary λ values within a reasonable range (e.g., between $0\% - 5\%$) and analyze the effect on the number of selected predictor-factor pairs. To summarize in technical terms, we generate numeraire S^* and transaction cost λ related (i.e., non-randomized/stochastic) stability paths to identify the most stable predictor-factor combinations.

3.4. Summary of comprehensive procedure

The comprehensive procedure introduced in section 3 can be condensed into four general steps. We start our enumeration with the most basic operation, i.e., moving from the inner to the outer loop:

1. To elect which single factor-predictor pair to update, we apply the base procedure described in section 3.1 and iterate over $j \in J$ factors and $k \in K$ predictors.
2. To estimate our multivariate linear model, we apply the componentwise L_2 Boosting algorithm described in section 3.2 and iterate over the number of boosting iterations $m \in M$.
3. To determine the optimal stopping iteration M, we apply Leave-One-Vintage-Out cross validation and iterate over the number of vintages years in the data set N. This means we perform componentwise L_2 Boosting N times with data missing for a distinct vintage year in each run; the out-of-sample hedging error is then always calculated for the held out vintage and finally averaged over the N runs.
4. To identify the most stable factor-predictor pairs, we apply the stability selection procedure described in section 3.3 and iterate over various numeraire S^* and trading cost λ choices.

Once this procedure is completed, we re-run the first three steps but now just using the stable factor-predictor subset (identified in step 4.) to estimate our final multivariate model.

4. Empirical application to venture capital funds

4.1. Preqin US venture capital fund data

For the empirical application of our methodology we draw on the Preqin (net of fee) cash flow data set as of August 2018, which is frequently used in the private equity literature (cf. Korteweg and Nagel[16], Ang et al[25], Gupta and Nieuwerburgh[3]). We restrict our example to (liquidated and non-liquidated) US Venture Capital funds of vintages $1986 - 2000$ and specifically include Preqin Category Types: Balanced, Early Stage, Early Stage: Seed, Early Stage: Start-up, Expansion/Late Stage, Venture (General), Venture Debt. Comparable to Driessen et al[17] we pool all fund cash flows and valuations stemming from one vintage and thus form vintage year portfolios which are normalized by the aggregated vintage year commitment; this means index $i = 1, \ldots, N$ now denotes vintage years.[2]

[2] Ang et al[27] illustrate efficiency issues associated with portfolio formation when testing asset pricing models using cross-sectional data which potentially can also apply to our panel data.

4.2. Public market factors and predictors

The public market factors F include the NASDAQ.Numeraire (NASDAQ Composite minus numeraire index excess return) factor and the five Fama and French[26] factors: Mkt-RF, SMB, HML, RMW, CMA. The predictors P cover an intercept term (i.e., a vector of ones), one-month-lagged fund net asset value, fund age, and the following (also one-month-lagged) economic time-series sourced from fred.stlouisfed.org: TB3MS (3-Month Treasury Bill: Secondary Market Rate), T10Y3M (10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity), TEDRATE (TED Spread), USSLIND (Leading Index for the United States), VXOCLS (CBOE S&P 100 Volatility Index), CFNAIDIFF (Chicago Fed National Activity Index: Diffusion Index), BAA10Y (Moody's Seasoned Baa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity). In total we have six traded factors and ten predictor variables.

4.3. Componentwise L_2 boosting results

We use step size $\nu = 0.3$, the maximal number of iterations is 200, and we apply Leave-One-Vintage-Out cross validation to determine the optimal stopping iteration M. As described in section 3.3 we re-run the whole procedure for several numeraire portfolio and trading cost λ choices to identify the most stable factor-predictor pairs. In total we perform 12 runs with the MSCI World and MSCI North America as numeraire candidates $S^* \in \{\text{World}, \text{NoAm}\}$ and six yearly trading cost choices $\lambda^* \in \{0\%, 1\%, 2\%, 3\%, 4\%, 5\%\}$ with $\lambda^* = 12 \cdot \lambda$.

Let us exemplarily analyze the $\lambda^* = 2\%$ case in greater detail before turning to the stability selection analysis. With $\lambda^* = 2\%$ the optimal stopping iteration is $M_1 = 108$ for the MSCI World and $M_2 = 13$ for MSCI North America (cf. Fig. 2). The corresponding coefficient estimates are depicted in Table 1 and we see that the MSCI North America model with 8 selected factor-predictor pairs is sparser than the MSCI World model with 14 factor-predictor pairs. In Table 2 the in-sample performance in terms of final replication error is analyzed. Here the Observed columns exhibit the most prominent discrepancies due to the different numeraire portfolios used for discounting all observed cash flows. As the sum of hedging errors (i.e., $\sum_i (Y_{T_i} - G_{T_i})$) are positive for both numeraires, these replication strategies are not able to outperform empirical venture capital cash flows in the sample period.

The stability selection result is summarized in the Selected column of Table 1. Here we see that the following factor-predictor pairs CMA_T10Y3M.lag1, HML_Age, and NASDAQ.Numeraire_CFNAIDIFF.lag1 are chosen in 9 out of 12 componentwise L_2 Boosting runs. Further, the pair NASDAQ.Numeraire_TEDRATE.lag1 is selected in 8 out of 12 runs. Finally, we re-run componentwise L_2 Boosting using only these four variables with $\lambda^* = 2\%$ and $S^* =$ World to obtain our sparse model estimate. The corresponding coefficient estimates for each of the M iterations are depicted in Fig. 3. The replication error by vintage year outlined in Table 3 exhibits only minor differences between the sparse $M = 20$ and $M = 200$ models. This indicates that entirely minimizing the loss function for the sparse model may not be necessary at all. However leave one vintage out cross validation for the sparse model run reveals no overfitting for $M = 200$, which suggests that our stability selection procedure turns out to be successful. The

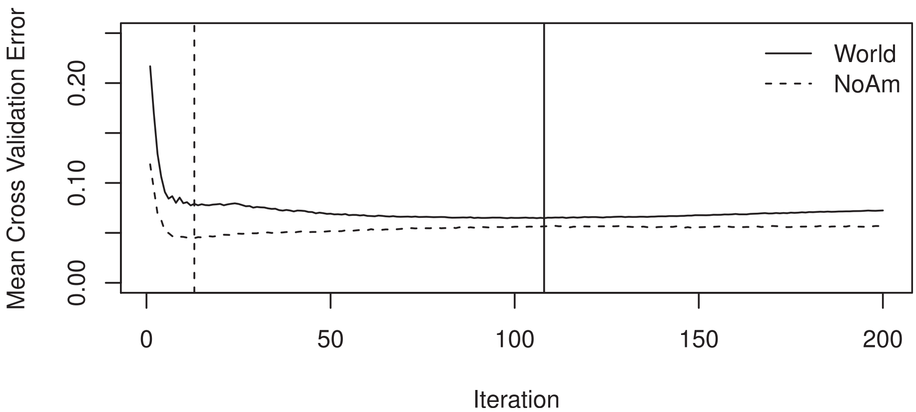

Fig. 2. Leave-One-Vintage-Out cross validation results for λ = 2%.

Table 1
Selected factor-predictor pairs and corresponding coefficient estimates for componentwise L_2 Boosting with $\lambda = 2\%$.

		$S^* = \text{World}$ $\lambda = 2\%$ $M_1 = 108$	$S^* = \text{NoAm}$ $\lambda = 2\%$ $M_2 = 13$	
Factor	Predictor	Coef	Coef	Selected
CMA	T10Y3M.lag1	−0.54	−0.09	9/12
HML	Age	0.23	0.06	9/12
NASDAQ.Numeraire	CFNAIDIFF.lag1	1.80		9/12
NASDAQ.Numeraire	TEDRATE.lag1	1.93	0.89	8/12
HML	TEDRATE.lag1	−0.72		7/12
NASDAQ.Numeraire	USSLIND.lag1	0.06	0.87	7/12
RMW	USSLIND.lag1	−0.24	−0.17	7/12
SMB	NAV.lag1	0.27	0.09	7/12
CMA	NAV.lag1	0.18	0.08	6/12
CMA	TEDRATE.lag1	−0.21		6/12
NASDAQ.Numeraire	Age	−0.03		6/12
CMA	Age		0.33	5/12
Mkt-RF	NAV.lag1	−0.36		5/12
SMB	TB3MS.lag1	−0.07		4/12
Mkt-RF	T10Y3M.lag1	0.04		3/12

Table 2
Replication error by vintage year for $\lambda = 2\%$. Here the variables introduced in section 2.1 are used as follows: Observed is Y_{T_i}, Replication is G_{T_i}, and Error is calculated as $(Y_{T_i} - G_{T_i})$.

Vintage	$S^* = \text{World}, M = 108, \lambda = 2\%$			$S^* = \text{NoAm}, M = 13, \lambda = 2\%$		
	Observed	Replication	Error	Observed	Replication	Error
1986	0.09	0.30	−0.21	−0.10	0.00	−0.10
1987	0.29	0.16	0.13	0.01	−0.14	0.15
1988	0.44	0.20	0.24	0.13	−0.02	0.15
1989	0.52	0.25	0.27	0.23	0.05	0.17
1990	0.18	0.39	−0.21	0.01	0.27	−0.26
1991	0.49	0.45	0.03	0.29	0.31	−0.02
1992	0.83	0.76	0.06	0.53	0.42	0.10
1993	0.93	0.79	0.14	0.58	0.48	0.09
1994	1.67	1.73	−0.06	1.01	0.91	0.10
1995	1.21	1.33	−0.11	0.80	0.82	−0.02
1996	1.20	1.13	0.07	0.91	0.83	0.08
1997	0.55	0.39	0.16	0.47	0.39	0.08
1998	0.39	0.31	0.08	0.40	0.35	0.05
1999	−0.32	−0.13	−0.19	−0.30	−0.01	−0.28
2000	−0.43	−0.40	−0.03	−0.39	−0.39	−0.01
Sum	8.03	7.67	0.37	4.57	4.30	0.28

corresponding ex-ante sparse hedging result for vintage 1992 is depicted in Fig. 4 (compare to the simple ex-post hedge from Fig. 1). As the sum of hedging errors (i.e., $\sum_i (Y_{T_i} - G_{T_i})$) are positive for both $M = 20$ and $M = 200$ in Table 3, our sparse replication strategy is not able to outperform empirical venture capital cash flows in the sample period.

4.4. Comparison to simpler hedging approaches

To examine the necessity and performance of our rather sophisticated and computational intensive hedging methodology, we compare it to two similar but simpler replication approaches. The commonality arises since these

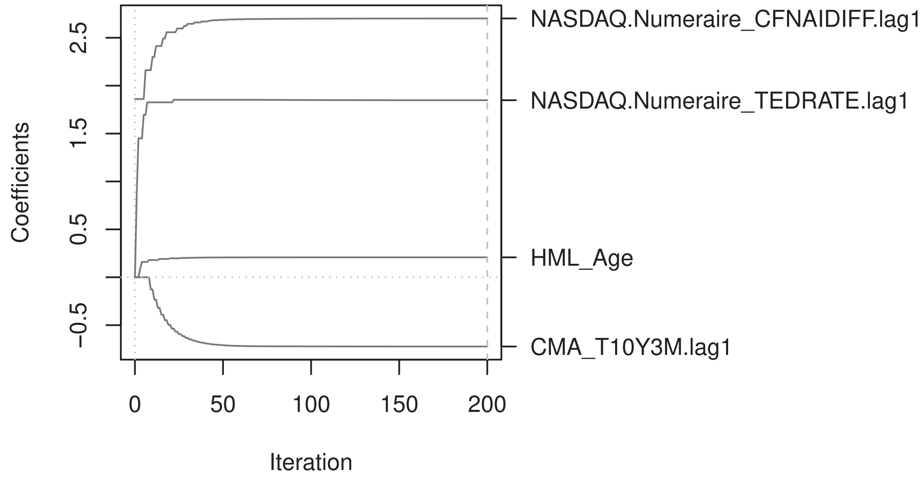

Fig. 3. Sparse model coefficients based on stability selection for $\lambda^* = 2\%$ and $S^* =$ World.

Table 3
Replication error by vintage year for sparse model with $M \in \{200, 20\}$, $\lambda^* = 2\%$, and $S^* =$ World. Here the variables introduced in section 2.1 are used as follows: Observed is Y_{T_i}, Replication is G_{T_i}, and Error is calculated as $(Y_{T_i} - G_{T_i})$.

Vintage	# Funds	Observed	$M = 200$		$M = 20$	
			Replication	Error	Replication	Error
1986	7	0.09	0.28	−0.19	0.29	−0.20
1987	4	0.29	0.07	0.23	0.08	0.21
1988	3	0.44	0.13	0.32	0.14	0.31
1989	5	0.52	0.28	0.24	0.28	0.24
1990	7	0.18	0.45	−0.27	0.46	−0.27
1991	2	0.49	0.58	−0.09	0.58	−0.10
1992	9	0.83	0.92	−0.09	0.88	−0.06
1993	9	0.93	0.79	0.14	0.77	0.16
1994	11	1.67	1.62	0.05	1.53	0.13
1995	15	1.21	1.29	−0.08	1.26	−0.04
1996	14	1.20	1.08	0.12	1.06	0.14
1997	20	0.55	0.38	0.17	0.40	0.15
1998	31	0.39	0.39	0.00	0.41	−0.02
1999	37	−0.32	−0.08	−0.24	−0.03	−0.29
2000	73	−0.43	−0.61	0.18	−0.56	0.12
Sum	247	8.03	7.57	0.46	7.56	0.47

models also build on (i) a linear public market factor model, (ii) a quadratic loss function, and (iii) the Kaplan and Schoar[9] framework.

First, we use the base procedure described in section 3.1 to estimate the linear one factor model known from the example in section 2.1. This means we just select factor NASDAQ.Numeraire, the vector of ones as predictor, step size $\nu = 1$ and just run one boosting iteration $M = 1$. The result is the static One Factor Boosting hedge, where we refrain from dynamic predictors. Second, we employ the Driessen et al[17] methodology to estimate a linear one factor model relying on the same NASDAQ.Numeraire factor. The cumulative discounted replication cash flow is calculated in both instances as

$$G_t^{(\text{Static})} = \sum_{\tau=1}^{t} \frac{V_{\tau-1}}{S_\tau^*} \bullet \beta \bullet \left(F_\tau^{(\text{NASDAQ}-\text{Numeraire})} - \lambda\right)$$

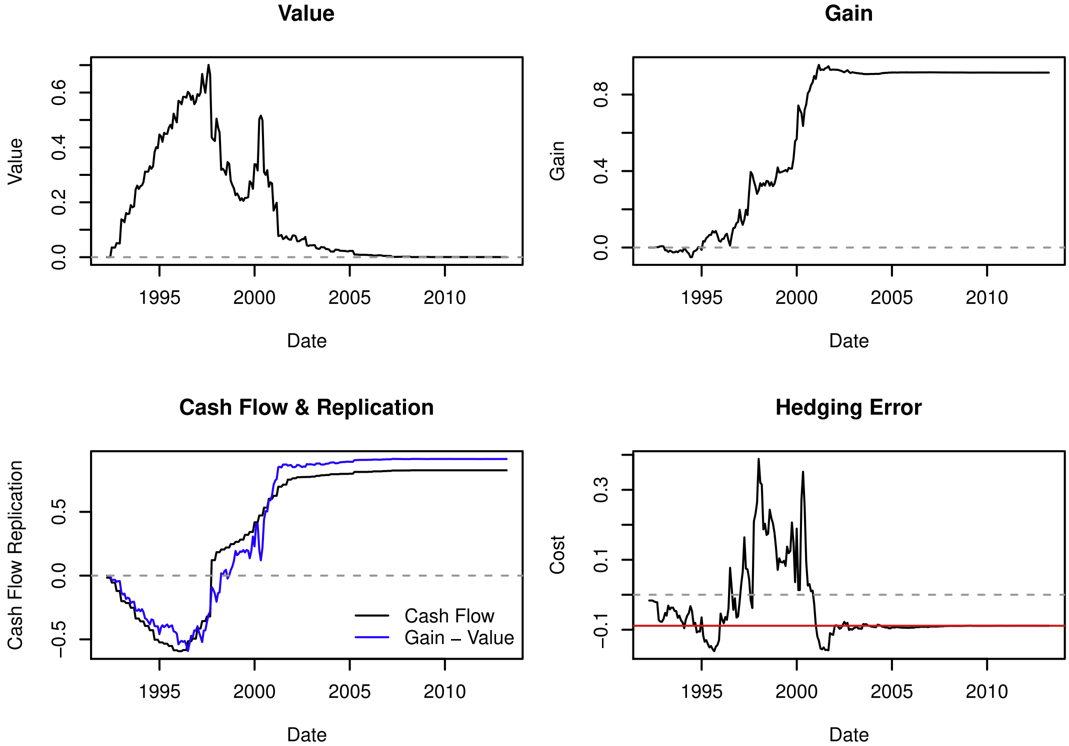

Fig. 4. Sparse model replication strategy for US VC funds of vintage 1992. Here the variables introduced in section 2.1 are used as follows: Value is V, Gain is G, Cash Flow is A, and Hedging Error is calculated as $A-(G-V)=Y-G$.

Table 4
Replication error by vintage year for simpler alternative models. The one-factor one-step boosting approach and a one-factor model constructed by the Driessen et al[17] methodology (with $\lambda^* = 2\%$ and $S^* =$ World in both instances).

Vintage	Observed	One Factor Boosting		Driessen et al[17]	
		Replication	Error	Replication	Error
1986	0.07	0.37	−0.30	0.23	−0.16
1987	0.28	0.14	0.14	0.09	0.19
1988	0.42	0.25	0.16	0.16	0.26
1989	0.51	0.27	0.24	0.17	0.34
1990	0.16	0.50	−0.35	0.32	−0.16
1991	0.49	0.65	−0.17	0.41	0.08
1992	0.82	0.62	0.20	0.39	0.43
1993	0.92	0.52	0.40	0.32	0.60
1994	1.66	0.07	1.59	0.05	1.61
1995	1.20	0.51	0.69	0.32	0.88
1996	1.18	0.24	0.94	0.15	1.03
1997	0.52	−0.21	0.74	−0.13	0.66
1998	0.38	−0.59	0.96	−0.37	0.75
1999	−0.35	−0.50	0.16	−0.32	−0.03
2000	−0.49	−0.18	−0.31	−0.11	−0.38
Sum	7.76	2.67	5.09	1.67	6.09

To facilitate a meaningful comparison to the stability selection model analyzed in Table 3, the models are estimated on the same VC data set, MSCI World is used as numeraire, and yearly trading cost are again $\lambda^* = 2\%$. The results in Table 4 reveal the largest absolute hedging errors for the Driessen et al[17] model, which is of course not explicitly designed to form zero net investment portfolios. Still the One Factor Boosting model exhibits considerably larger errors than the models selected by stability selection. So given that the MSCI World index as numeraire and 2% yearly trading costs are reasonable choices, the numeraire-discounted cash flows observed in our Venture Capital data set can be remarkably preciser replicated by the dynamic trading strategies obtained by our stability selection procedure than by the static linear one factor models presented in this section.

5. Conclusion

This paper proposes the combination of a net cash flow based quadratic loss function and componentwise L_2 Boosting as suitable method to estimate public market trading strategies that replicate private equity fund cash flows. The method can be both employed for the replication and hedging (i.e., a long and short position respectively) of PEF payment streams, since the quadratic loss criterion equally penalizes trading gains and losses. In our opinion, a stability selection like proceeding is the most convenient (non-stochastic) way to enhance model sparsity and at the same time model robustness; thus to obtain more practical trading strategies involving fewer variables. In addition to actual trading considerations, our proposed procedure identifies the most important public factors and - even more interestingly - macro predictors that affect discounted private equity fund cash flows. Since our approach discounts all financial quantities by a common numeraire portfolio, the method perfectly fits into the influential Kaplan and Schoar[9] public market equivalent framework. However, as in Driessen et al[17] estimation exclusively resorts on realized fund level data, i.e., the most recent vintage years data - as well as any company level information - is excluded by design. Moreover, notoriously controversial (lagged) fund net asset values serve as weighting factors in the gain function from Equation (1).

A fund manager leaves their favorite investment boutique and thinks: Is there any machine learning technique that produces sparse results in a more straightforward fashion? And is it possible to find similar trading strategies that do not rely on intermediate private equity fund information at all? Or how about a method that can incorporate unrealized fund data into estimation?

Conflict of interest

The author reports no conflict of interest. The author alone is responsible for the content and writing of the paper.

Acknowledgements

I thank Axel Buchner, Stefan Mittnik, Marcus Pietz, and all participants at the LMU financial econometrics research seminar WS18/19 and DZH St. Patrick's Day event for helpful comments and discussions.

References

1. Madhavan A, Sobczyk A. Fine-tuning private equity replication using textual analysis. *J Financ Data Science*. 2019;1(1):111−121.
2. Porter SC, Porter S. Introducing objective benchmark-based attribution in private equity. *J Financ Data Science*. 2019;1(1):130−140.
3. Gupta A, Nieuwerburgh SV. *Valuing Private Equity Investments Strip by Strip*; 2018. Unpublished Working Paper https://www.aeaweb.org/conference/2019/preliminary/paper/YsGkk6BR.
4. Moller T. Risk-minimizing hedging strategies for insurance payment processes. *Financ Stoch*. 2001;5(4):419−446.
5. Föllmer H, Sondermann D. Hedging of non-redundant contingent claims. In: Hildenbrand W, Mas-Colell A, eds. *Contributions to Mathematical Economics*. 1986:205−223 [chapter 12], [North-Holland].
6. Schweizer M. A guided tour through quadratic hedging approaches. In: Jouini E, Cvitanic J, Musiela M, eds. *Option Pricing, Interest Rates and Risk Management, Chapter 15*. Cambridge University Press; 2001:538−574.
7. Schweizer M. Local risk-minimization for multidimensional assets and payment streams. In: Stettner L, ed. *Advances in Mathematics of Finance, volume 83 of Banach Center Publications*. Polish Academy of Sciences; 2008:213−229.
8. Bühlmann P. Boosting for high-dimensional linear models. *Ann Stat*. 2006;34(2):559−583.
9. Kaplan S, Schoar A. Private equity performance: returns, persistence, and capital flows. *J Financ*. 2005;60(4):1791−1823.
10. Long AM, Nickels CJ. *A Private Investment Benchmark*. The University of Texas System; 1996.
11. Bailey JV. Are manager universes acceptable performance benchmarks? *J Portfolio Manag*. 1992;18(3):9−13.
12. Sorensen M, Jagannathan R. The public market equivalent and private equity performance. *Financ Anal J*. 2015;71(4):43−50.
13. Korteweg A. *Risk Adjustment in Private Equity Returns*; 2018. Unpublished Working Paper https://doi.org/10.2139/ssrn.3293984.
14. Farnsworth H, Ferson W, Jackson D, Todd S. Performance evaluation with stochastic discount factors. *J Bus*. 2002;75(3):473−503.

15. Li H, Xu Y, Zhang X. Hedge funds performance evaluation under the stochastic discount factor framework. *J Financ Quant Anal*. 2016;51(1):231−257.
16. Korteweg A, Nagel S. Risk-adjusting the returns to venture capital. *J Financ*. 2016;71(3):1437−1470.
17. Driessen J, Lin T-C, Phalippou L. A new method to estimate risk and return of nontraded assets from cash flows: the case of private equity. *J Financ Quant Anal*. 2012;47(3):511−535.
18. Long JB. The numeraire portfolio. *J Financ Econ*. 1990;26:29−69.
19. Bai J, Ng S. Boosting diffusion indices. *J Appl Econom*. 2009;24(4):607−629.
20. Mittnik S, Robinzonov N, Spindler M. Stock market volatility: identifying major drivers and the nature of their impact. *J Bank Financ*. 2015;58:1−14.
21. Meinshausen N, Bühlmann P. Stability selection. *J R Stat Soc*. 2010;72(4):417−473.
22. Bühlmann P, Hothorn T. Twin boosting: improved feature selection and prediction. *Stat Comput*. 2010;20(2):119−138.
23. Bühlmann P, Yu B. Sparse boosting. *J Mach Learn Res*. 2006;7:1001−1024.
24. Zhao J. General sparse boosting: improving feature selection of l2 boosting by correlation-based penalty family. *Commun Stat Simulat Comput*. 2015;44(6):1612−1640.
25. Ang A, Chen B, Goetzmann WN, Phalippou L. Estimating private equity returns from limited partner cash flows. *J Financ*. 2018;73(4):1751−1783.
26. Fama EF, French KR. A five-factor asset pricing model. *J Financ Econ*. 2015;116(1):1−22.
27. Ang A, Liu J, Schwarz K. Using stocks or portfolios in tests of factor models. *J Financ Quant Anal*; 2019. https://doi.org/10.1017/S0022109019000255.