MODELING PE RETURNS

TAKAHASHI-ALEXANDER MODEL REVISITED

GUSTAV ERIKSON
MÅNS HOLMBERG

Master Thesis
Stockholm School of Economics
2024

Modeling PE Returns: Takahashi-Alexander Model Revisited

Abstract:

In this study, we revisit the Takahashi-Alexander (TA) model, an influential framework to model private equity cash flows, by incorporating macroeconomic variables to enhance its predictive accuracy. Our data covers 2,671 global buyout funds with vintages between 1990 through 2022. We adapt the TA model to reflect macroeconomic impacts on three crucial parameters: the rate of capital contributions, the NAV growth rate, and the bow factor, which determines the phasing of distributions. By training explanatory models on historical data, we find that our macro-calibrated TA model significantly outperforms benchmarks in predicting future capital contributions. However, the predictive accuracy for net cash flows is less definitive, as modeling capital distributions proves to be more complex. This paper presents an extension to the TA model and provides private equity investors with a potentially more robust tool to manage liquidity risk.

Acknowledgements:

We would like to thank our thesis supervisor Professor David Robinson for his continuous support and guidance.

Keywords:

Private equity, liquidity, financial modeling, market cyclicality, Takahashi-Alexander Model

Authors:

Gustav Erikson (24154)
Måns Holmberg (50659)

Tutors:

David Robinson, Professor at the Fuqua School of Business at Duke University, & Visiting Professor at the Stockholm School of Economics, Department of Entrepreneurship, Innovation and Technology

Master Thesis
Master Program in Finance
Stockholm School of Economics
© Gustav Erikson and Måns Holmberg, 2024

1. Introduction **6**

2. Literature review **9**

2.1 Private equity as an asset class **9**

2.2 The limited partnership fund structure **9**

2.2.1 Private equity cash flows 9

2.2.2 Agency conflicts and governance mechanisms 10

2.2.3 Modeling private equity returns 11

2.3 Takahashi-Alexander model **11**

2.3.1 Parameter estimation 13

2.4 Cyclicality and private equity returns **15**

2.4.1 Implication on TA model parameters 16

3. Data **18**

3.1 Descriptive overview **18**

3.2 Data manipulation **18**

3.2.1 Variable creation 18

3.2.2 Upper bound of J-curve 18

3.2.3 Recallable distributions 19

4. Empirical methodology **20**

4.1 Model definitions **20**

4.1.1 Macroeconomic indicators 21

4.1.2 Stepwise selection technique of explanatory variables 23

4.1.3 Robustness check 23

4.1.4 Validation of model predictive performance 24

4.2 Contribution rate parameter estimation **25**

4.2.1 Variable selection 25

4.2.2 Model specification 27

4.3 Growth rate parameter estimation **28**

4.3.1 Estimating current public returns (first-stage regression) 28

4.3.2 Estimating growth rate parameter (second-stage regression) 29
4.3.3 Model specification 30
4.4 Bow factor parameter estimation 30
4.4.1 Fixed bow factor 31
4.4.2 Optimizing bow factor 31
5. Results 32
5.1 Contribution rate parameter results 32
5.1.1 Model specification 32
5.1.2 Robustness check 33
5.1.3 Validating predictive performance 33
5.2 Growth factor parameter results 36
5.2.1 Estimating current public returns 36
5.2.2 Estimating growth rate parameter 36
5.3 Bow factor parameter results 38
5.3.1 Fixed bow factor 38
5.3.2 Optimized bow factor 41
6. Discussion and conclusions 43
6.1 Rate of contribution 43
6.2 Growth rate 44
6.3 Bow factor 45
6.3.1 Fixed bow factor 45
6.3.2 Optimized bow factor 45
6.4 Further considerations 46
7. References 49
7.1 Published papers 49
7.2 Published literature 50
7.3 Industry reports 51
7.4 Online resources and working papers 51

8. Appendix **53**

Appendix A: Benchmark TA model plot for bow factor **53**

Appendix B: Data descriptives **54**

Appendix C: Robustness check for multicollinearity **55**

Appendix D: Robustness check of regression residuals **56**

Abbreviations

Abbreviation	Definition
PE	Private Equity
LP	Limited Partner
GP	General Partner
AUM	Assets Under Management
IRR	Internal Rate of Return
LTM	Last Twelve Months
TA	Takahashi-Alexander
C	Capital Contribution
CC	Committed Capital
D	Capital Distribution
NC	Net Cash
NAV	Net Asset Value
UC	Uncalled Capital
PIC	Paid In Capital
RC	Rate of Contribution
RD	Rate of Distribution
G	Growth Rate
Y	Yield
L	Lifetime
B	Bow Factor

1. Introduction

Emerging around half a century ago as a niche enterprise, private equity has today risen to become a household name. In 2023, global PE dry powder reached $3.9 trillion, representing nearly 800% growth from $500 billion twenty years ago (Statista, 2024). Simultaneously, the consistent delivery of high returns—averaging 12.5% annually in the period between 2001-2020 compared to 8.7% from the S&P500 in the same period—has driven a higher target allocation among existing investors and an overall expansion of the investor pool (Barath, 2022).

However, private equity cash flows are characterized by uncertainty concerning both their magnitude and timing, leading to significant liquidity risk for investors. Commitments to private equity funds are irrevocable once made, and the General Partners (GPs) can call for the capital at their sole discretion during the investment period. Correspondingly, the timings of capital distributions are conditioned on the exit strategy of the fund. Therefore, keeping allocations at target level is a complex task for investors opting for the asset class (Kieffer et al., 2023). This unpredictability often compels Limited Partners (LPs) to both decrease their overall commitments and target allocation to PE, and to maintain their uncalled capital in low-yield investments, such as treasury bills (Jeet & O'Shea, 2018). Accurately forecasting these cash flows would greatly benefit LPs by enhancing their ability to pace commitments effectively. This, in turn, would facilitate the maximization of returns while adhering to target allocations and ensuring liquidity.

Modeling PE cash flows is albeit a complex exercise and prior research on the area is scarce, with only a handful of exceptions. Historically, investors have often surrendered to rules-of-thumb and general guidelines rather than data analysis. This comes as little surprise, considering the opaque nature of the asset class, characterized by significant information imbalances between LPs and GPs, making it hard for investors to make accurate predictions.

In this report, we present the landmark Takahashi-Alexander (TA) model to model PE cash flows. While this framework still effectively balances predictive power and maintains practical ease of use, we propose an updated framework that includes an additional degree of freedom for the market environment, allowing for more complexity and improved predictability.

We propose three distinct ways in which macroeconomic effects can impact cash flows through the TA model: either through altering the rate at which GPs call for capital (RC), distribute capital (RD), or the underlying growth rate of the NAV (G).

The empirical approach in this study aims to refine TA model-predictions by enhancing the estimation of RC and G using macroeconomic indicators. In this process, we make them time-dependent, potentially changing between periods. In estimating the parameters RC and G, the core method is to fit regression models on training data, with the aim of predicting actual values by incorporating macroeconomic indicators as explanatory variables. The predictive performance of the selected models is evaluated on unseen test data, simulating a real-world scenario, and compared against a benchmark model. The data is sourced from Burgiss and consists of pooled data from 2,671 global buyout funds with vintages 1990-2022.

The results show that the predictions of RC and G are significantly improved, when compared against the benchmark model. The enhanced prediction of RC further translates into an improved prediction of capital contributions, which holds positive implications for liquidity-constrained LPs. However, the predictive performance from net cash flows is less definitive, as modeling capital distributions proves to be more complex due to the fixed and non-linear nature of RD, due its inherent bow factor (B). The benchmark model performs better at predicting capital distributions and net cash flows, assuming a fixed B. This gap narrows when a dynamic estimate of B is used, optimized for each vintage and model.

This thesis contributes to prior research by introducing explanatory variables intended to capture these macroeconomic effects in the TA model. While this has been done to capture the impact on the growth rate parameter (Jeet, 2020), we have not found any previous research that explicitly addresses the impact of macroeconomic conditions on the other model parameters. In fact, industry practitioners often rely on loosely defined benchmarks and general guidelines to define parameter assumptions. Given that the TA model is a deterministic model, it is heavily dependent on parameter assumptions. Our aim is to contribute with a more robust cash flow model tailored for private equity fund investors which is less reliant on general guidelines.

The remainder of this paper is organized as follows. Section 2 gives the reader an introduction to private equity as an asset class, and specifically to why modeling its returns presents a challenge for investors. We highlight how the illiquid nature of its cash flows impacts modeling techniques, and important aspects of the LP-GP relationship. On the topic of modeling cash flows, the TA model is then introduced, by also touching upon

potential macroeconomic contingencies. Furthermore, Section 3 provides a descriptive overview of the data used in this study. Section 4 specifies our empirical approach and Section 5 exhibits our main results. Finally, the results are deliberately discussed in Section 6.

2. Literature review

2.1 Private equity as an asset class

We begin this review by defining Private Equity (PE) following the standard definition, wherein investment firms engage in leveraged buyouts (LBOs), using a relatively small portion of equity, and a relatively large portion of external debt financing (Kaplan & Strömberg, 2009; McGrath & Nerkar, 2023).

The value proposition of PE fundamentally involves the expansion of the equity value upon realization of the investment. The means to achieve this end have been categorized into both operational (e.g. strategic transformation & cost-cutting) governmental (e.g. human capital) and financial (e.g. tax & debt financing) initiatives to create shareholder value (Jensen, 1989; Kaplan & Stromberg, 2009).

2.2 The limited partnership fund structure

Private equity funds are organized as limited partnerships, in which managers from the PE firm serve as GPs, with a mandate to oversee the investment process. Conversely, LPs contribute capital without active participation in the investment process. The LPs are mostly institutional investors, such as sovereign wealth funds, high-net-worth individuals, pension funds, or endowments (Kaplan & Strömberg, 2009).

2.2.1 Private equity cash flows

A generic illustration of net cash flows from the LP point of view is indicated in Figure 1. Typically, LPs commit to invest a fixed amount upon the formation of the fund. This commitment is then translated to one or several cash flows during the lifecycle of the fund, named either “capital contributions” or “capital calls” if these are negative, or “capital distributions” if positive. In the first stage of the fund, net cash flows are negative since capital contributions outweigh distributions. As the fund cycle reaches a mature stage, investments are realized and distributed back to investors, resulting in net positive cash flow. This pattern is known as the “J-curve” phenomenon in private equity funds. Early days within each fund are associated with a decrease in net cash flows, followed by a gradual increase over the lifetime, eventually tending towards a fixed value—exhibiting a “J-shape” (see Figure 1).

However, the exact shape of the J-curve is in practice difficult to predict for investors, as the timing of capital contributions and distributions is governed by how GPs enter or exit holdings during the time horizon (Karatas et al., 2021).

2.2.2 Agency conflicts and governance mechanisms

The fact that general partners endogenously choose when to call and distribute capital is enabled by how the LP-GP relationship is structured. Through the lens of the principal-agent theory, GPs serves as an agent (managing the fund) and LPs (investors) as principals. This agency relationship is specified by the management contract that is typically signed at the inception of the fund. The management contract is an important commitment device for LPs used to align the value-maximizing agenda with the incentives of GPs. For example, the bulk of compensation to GPs hinges on a successful exit, as it generally follows a "2/20/1" schedule[1].

However, several scholars note that agency frictions between GPs and LPs are not completely mitigated by the management contract. These conflicts predominantly occur because GPs expend significant resources in selecting, monitoring, and exiting investments. Thus, they hold privileged information regarding the quality of these investments. For example, this information asymmetry empowers GPs to potentially manipulate the timing of exits (Metrick et al., 2011), management fees (Robinson & Sensoy, 2016), and NAV valuations in connection with fundraising, thus leveraging contractual clauses intended to safeguard the returns of LPs. These effects are augmented by the fact that investing in PE involves financial commitment for up to 15 years, and that LPs have limited governance mechanisms beyond the management contract (Robinson & Sensoy, 2016).

Collectively, these findings illustrate that the LP-GP relationship presents a double-edged sword for the LPs. On one hand, PE funds tend to outperform direct LP investments (Kieffer et al., 2023). On the other hand, the LP-GP information asymmetry can make it harder for LPs to make informed asset allocation decisions. A robust modeling methodology has been argued to potentially alleviate these negative agency conflicts, as LPs can better predict the behavior of GPs before making asset allocation decisions.

[1] 2% management fee on the underlying AUM, 20% carried interest and 1% co-investment from GPs

2.2.3 Modeling private equity returns

Utilizing predictive modeling techniques to forecast cash flows from PE investments has the potential to refine asset allocation decisions made by LPs. That said, modeling PE cash flows presents a complex challenge given the nature of the asset class (Butler et al., 2023; Jeet, 2020).

During the nascent period of the PE industry, LPs mainly made asset allocation decisions based on intuition rather than data analysis (Butler et al., 2023). This dependency on heuristics was partly attributed to agency conflicts (Section 2.2.2), fueled by the obscurity and lack of transparency from GPs about their operations (Johan & Zhang, 2021). Even today, estimating cash flows from PE investments is complicated due to their irregular nature and limited data availability (Ang & Sorensen, 2012).

To this end, Butler et al. (2023) highlight several amplifying mechanisms that underline this difficulty. Firstly, the variability in returns among PE funds and PE firms (Sharma & Kaushal, 2012), is significantly higher than public benchmarks. Secondly, the illiquid nature of the asset class makes projections of capital contributions and distributions difficult, as GPs endogenously choose when to call and distribute capital (Section 2.2.2). Thirdly, solely relying on past performance to extrapolate future returns is problematic, since research indicates that the ability of PE firms to sustain past returns across buyout funds has decreased since the 2000s (Harris et al., 2022). Taken together, these challenges make it difficult for LP investors to make informed asset allocation decisions for PE investment vehicles, since the planning and forecasting of cash flows is overall a complex exercise.

2.3 Takahashi-Alexander model

Recognizing the inherent challenge for investors in relying on general guidelines and best guesses, Takahashi and Alexander (2001) provided a framework for modeling private capital cash flows. The authors cited three main criteria when defining the model, it had to be: 1) simple and sensible on a theoretical basis, 2) flexible to incorporate actual observations during the projection window with varying return and timing scenarios, and 3) useful for a variety of asset types. Today, the TA model is one of the most influential and popular modeling techniques for illiquid assets (Meyer et al., 2013). Apart from demonstrating proficiency in forecasting anticipated cash flows, another reason for its popularity has been its simplicity and ease of use (Jeet, 2020). Nevertheless, the TA model is a non-probabilistic model, where all cash flows are modeled as deterministic (i.e. for a

given set of input parameters the outcome variable only yields one value) (Meyer et al., 2013).

While being simple and easy to implement, this deterministic specification means that accurate parameter inputs are essential for the predictive power of the model (Meyer et al., 2013). In contrast to probabilistic models that account for elements of randomness in outcomes (for reference, see Buchner et al. (2006); Forsell & Furenstam, (2018)), this non-probabilistic approach means that the TA model does little to reflect the volatility of cash flows, which is important when managing liquidity risk. For professionals and researchers relying on the model, this underscores the importance of having accurate projections, which again hinges on accurate input parameters.

The TA model consists of projections of capital contributions, capital distributions, and net asset value. Below, we describe this framework:

Contribution model

The contribution model states that capital contributions in period t are equal to the rate of contribution $(RC_{(t)})$ multiplied by uncalled capital—equivalent to committed capital (CC) minus paid-in-capital $(PIC_{(t-1)})$.

$$C_{(t)} = RC_{(t)}\left(CC - PIC_{(t-1)}\right) \quad (1)$$

$PIC_{(t)}$ in period t is simply the cumulative sum of capital contributions $C_{(t)}$.

Distribution model

The distribution formula is based on both the rate of distribution ($RD_{(t)}$) and the net asset value ($NAV_{(t)}$):

$$D_{(t)} = RD_{(t)}\left[NAV_{(t-1)} \times (1 + G)\right] \quad (2)$$

$RD_{(t)}$ is broken down into two components: yield (Y) and realizations $\left(\frac{t}{L}\right)^B$:

$$RD_{(t)} = Max\left[Y, \left(\frac{t}{L}\right)^B\right] \quad (3)$$

The yield is relevant for income-generating assets such as real estate but is typically set to zero for funds containing assets that do not generate any real dividends. In the latter years of the fund, $RD_{(t)}$ will approach 100% as time (t) passes. The monetary value of D_t will however approach zero as NAV declines in the latter stage of the fund lifecycle. The bow factor (B) is a model-specific parameter that determines the phasing of distributions along the fund lifetime. A higher B implies a slower initial growth in RD and a faster acceleration towards the end of the period—and vice versa. For example, at age ten, a fund with a lifetime of twelve years with $B = 2$, rate of distribution ($RD_{(10)}$) will equal:

$$\left(\frac{10}{12}\right)^2 = 69\%$$

Net Asset Value model

Finally, the NAV model is formed by multiplying NAV at the beginning of the period with the projected growth rate (G), adding capital contributions ($C_{(t)}$) net of capital distributions ($D_{(t)}$):

$$NAV_{(t)} = \left[NAV_{(t-1)} \times (1 + G)\right] + C_{(t)} - D_{(t)} \qquad (4)$$

In this equation, growth rate G equals unrealized IRR—it measures how NAV changes between periods, net the change in cash inflows and outflows between LPs and the fund.

2.3.1 Parameter estimation

As exhibited in the previous section, the TA model relies on input parameters, namely, the RC, B, L, and G. These input parameters are typically estimated with best guesses or historical averages. Jeet & O'Shea (2018) use common industry averages for L and G. B and RC are statistically calibrated using pooled historical data from US buyout funds to minimize error between the model's predicted cash flows and observed cash flows. The results from these estimations are presented in Table 1, along with an illustrative plot in Figure 1 of the resulting TA model output, expressing all predictions as a percentage of CC.

Throughout this paper, we will use the above TA model specification as a benchmark, denoting it as $TA_{Benchmark}$. Further, each parameter within $TA_{Benchmark}$ has a denotation, as specified in Table 1.

Table 1.

Benchmark Takahashi-Alexander ($TA_{Benchmark}$) model parameters

This table lists the input parameter assumptions for the benchmark Takahashi-Alexander model ($TA_{Benchmark}$), using the same inputs parameter values as in Jeet & O'Shea (2018), but converted to quarterly instead of yearly numbers.

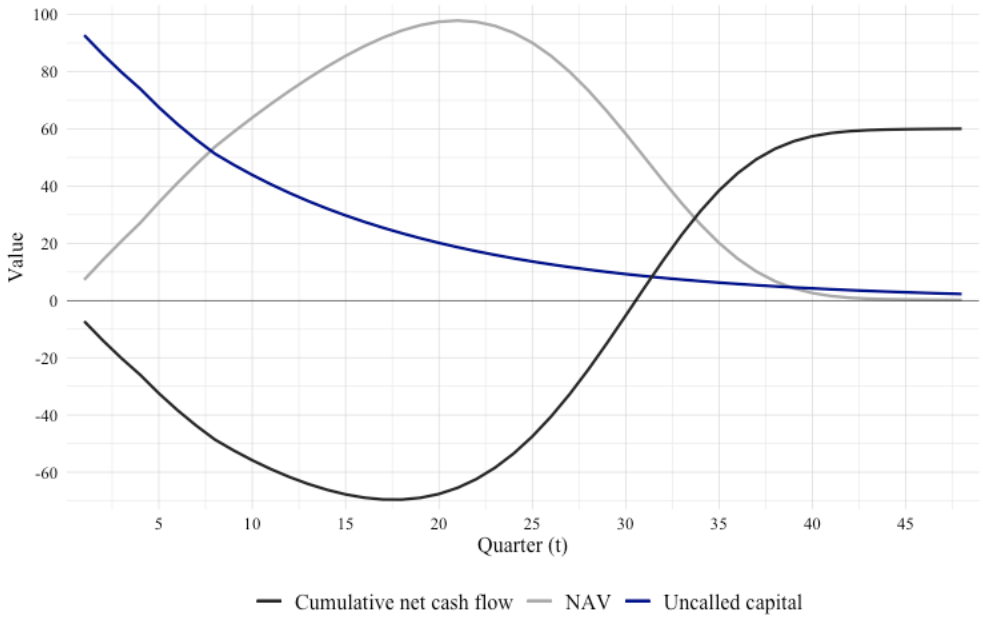
Model	Parameter	Age interval	Value	Denotation
Contribution	RC (quarterly)	$0 \leq \text{Age} \leq 4$	7.3%	$RC_{Benchmark}$[2]
Contribution	RC (quarterly)	$4 \leq \text{Age} \leq 8$	8.8%	$RC_{Benchmark}$
Contribution	RC (quarterly)	$8 \leq \text{Age} \leq \infty$	7.5%	$RC_{Benchmark}$
Distribution	B		4[3]	$B_{Benchmark}$
Distribution	L (quarters)		48	$L_{Benchmark}$
Distribution	G (quarterly)		2.6%	$G_{Benchmark}$
Distribution	Y (quarterly)		0%	$Y_{Benchmark}$

Figure 1.

Benchmark Takahashi-Alexander model plots

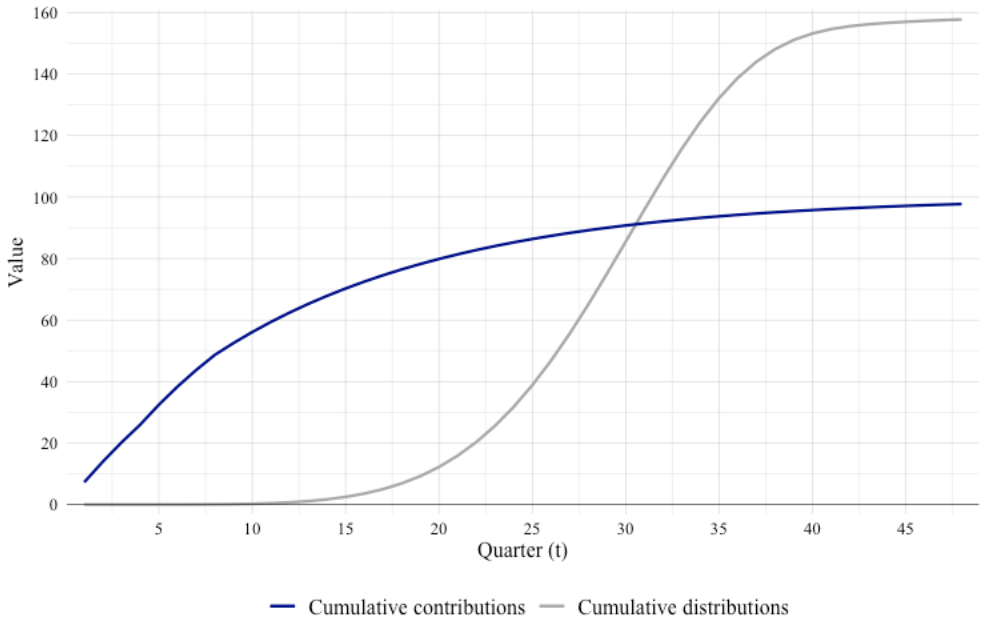
Figure 1a and 1b visually illustrates cash flow predictions from the benchmark TA model specification, using the sample model inputs from Table 1, and assuming a CC of 100.

Figure 1a: NAV, UC, and cumulative NC



[2] Although this parameter has three time-dependent estimates, we $RC_{Benchmark}$ as denotation for all three
[3] In accordance with Jeet (2020), a yearly $B = 2$ can be approximated to $B = 4$ on a quarterly basis

Figure 1b: Cumulative C and D



2.4 Cyclicality and private equity returns

Takahashi and Alexander (2001) constructed their model based on fund-specific (internal) parameters. While this enables periodic updates, to accommodate changes along the lifecycle of a fund, it does not account for external factors that might affect the underlying cash flows, and thus the predictability of the model.

This section expands on how macroeconomic variables can affect the model parameters. To formally address this, consider the following expression for uncalled capital ($UC_{(t)}$):

$$UC_{(t)} = \left(CC - PIC_{(t-1)}\right) = UC_{(t-1)} \times \left(1 - RC_{(t)}\right) \tag{5}$$

$$UC_{(0)} = 100 \tag{6}$$

$$NAV_{(0)} = 0 \tag{7}$$

By substitution, we can now write Equation 4 as:

$$NAV_{(t)} = NAV_{(t-1)} \times \left(1 + G_{(t)}\right) \times \left[1 - \left(\frac{t}{L}\right)^{B}\right] + UC_{(t-1)} \times RC_{(t)} \tag{8}$$

Equation 8 now illustrates the full TA model in one equation. From this equation, we can deduce that macroeconomic factors can potentially affect the TA model through three dynamics. The first two dynamics are found in the first term of Equation 8, which isolates the effect with regards to capital distributions: Firstly, macroeconomic factors can affect the rate at which NAV grows (G). Secondly, capital distributions can be affected by the bow factor (B), which depends on how the GPs of the fund phase distributions (Section 2.2.2). The third and final dynamic is the effect on rate of contributions (RC) which is seen in the second term of Equation 8 and isolates the effect on capital contributions.

The subsequent sections will explore how each of these parameters are connected to established research on macroeconomic factors.

2.4.1 Implication on TA model parameters

2.4.1.1 Growth (G)

The growth factor determines how NAV grows from one period to the next, independently of capital contributions and distributions within that interval. Holding RC, L and B, the IRR of the fund is simply a function of the growth rate parameter (Jeet, 2020). Measures of private equity returns (IRR) have been found to be correlated with broader macroeconomic conditions (Tolkamp, 2007). During periods of economic prosperity, asset valuations generally increase, translating into an elevated growth factor. This heightened state of growth drives up the NAV, boosting the fund's perceived performance. Conversely, economic downturns can depress asset prices, diminishing the growth factor and thus negatively impacting the NAV.

Jeet (2020) and Tolkamp (2007) both proposed extensions of the baseline TA model by not assuming fixed growth rates. This modification enabled the incorporation of market indicators in the TA framework, as these can be used to predict actual growth rates in period t, thus potentially arriving at a parameter input that lies closer to the actual.

2.4.1.2 Rate of contribution (RC)

Within the LP-GP fund structure, investors can only time their commitments: not capital contributions since GPs endogenously choose when to call capital (Section 2.2.2). Robinson & Sensoy (2016) find that this endogenous effect on capital contributions varies with conditions in the public market. Their findings highlight that the magnitude of capital contributions increase in hot capital markets, using price-dividend ratio and yield spread

as market indicators. Specifically, capital contributions follow a procyclical trend with market price-dividend ratios and yield spreads.

No prior research has, to the best of our knowledge, addressed the relationship between macroeconomic factors and the contribution function in the TA model. However, within the model framework, any macroeconomic effect on capital contributions would be seen through the rate of contribution $RC_{(t)}$, as this parameter captures the phasing of capital contributions. Following the conclusions of Robinson & Sensoy (2016), we would then period a higher magnitude of capital contributions in periods of favorable market conditions, and vice versa.

2.4.1.3 Rate of distribution (RD)

GPs do not only choose when to call capital, they also decide when to realize exits and distribute back capital. Robinson & Sensoy (2016) also find that macroeconomic indicators can significantly predict capital distributions, again using proxies for price-dividend ratio and yield spread as explanatory variables.

As in the case of capital contributions, we have not found any no prior research directly incorporating macroeconomic indicators into the distribution model within the TA model. However, theoretically the effect of macroeconomic indicators on the capital distributions would be seen through the bow factor B. For instance, in an economic boom, rate of distributions could increase due to quicker than expected realization of investments. This increase in rate of distribution would imply a smaller optimal value of B. Conversely, prolonged economic contractions might extend the lifecycle of investments, leading to a decrease in rate of distribution and therefore a larger optimal value of B, suggestive of delayed distributions. Appendix A visualizes predicted rates of distribution given different values of B.

3. Data

3.1 Descriptive overview

Our research uses newly available PE cash flow data provided by Burgiss, which is a global provider of data and analytics geared towards LPs who invest in PE funds. Our sample covers 2,671 buyout funds worldwide, across vintage years 1990 through 2022. The dataset contains quarterly end-of-period observations of total capital commitment, number of funds, cumulative distributions and contributions, and net asset value. All columns are aggregated on vintage year level, implying 1) that we have one unique dataset per vintage year, and 2) all fund observations within each vintage year are aggregated.

Burgiss calculates this aggregated value in two steps. First, all fund observations are summarized using a weighted average by the relative size of each fund, as measured by total capital commitment. Then, in the aggregated dataset, each cash flow variable (including NAV) is expressed as a percentage of total capital commitment. Intuitively, this means that within each vintage year, an investor would hold an indexed value of each fund scaled by its relative size on total capital commitment.

3.2 Data manipulation

3.2.1 Variable creation

Given that the cash flow variables in the dataset are all cumulative numbers, we create two new variables: contributions and distributions, which are simply the non-cumulated versions of contributions and distributions. This is needed in order to apply the distribution function and contribution function of the TA model (see Equation 1 and 2).

Finally, we create an age variable, by simply specifying the number of quarters since the beginning of the vintage year. One note on this topic is that in the dataset, the number of funds initially increases during the four quarters before it reaches a fixed value. This is problematic with respect to the age variable since the funds within each row will have slightly different ages. However, given that this difference is four quarters at maximum, and that a fund lifecycle generally is around 50 quarters, the discrepancy is marginal.

3.2.2 Upper bound of J-curve

First, in line with Jeet & O'Shea (2018), we drop observations for each vintage year when the NAV is less than 5% of its fund size, given that the fund age is at least eight years.

Following the J-curve shape of private equity cash flows (see Figure 1), the net change in cash flow measures approaches zero as the fund reaches maturity. Given that the variance in cash flows decreases as the time variable increases, not setting an upper bound can result in an inflated R-squared value in the predictions, as the model continues to account for periods where variability in cash flows is minimal.

3.2.3 Recallable distributions

The accumulated capital contributions over the fund lifetime should by definition not exceed the capital commitment unless an agreement on recallable distributions has been formed. Recallable distributions are distributions that by the discretion of the GP can be recalled as a new capital contribution (Meads et al., 2016). Thus, the initial capital commitment will be understated as a proxy for total invested capital by the LP.

The dataset used in this paper does not strip out the value of recallable distributions as a separate vector. Instead, we can only observe that the PIC for some vintage years exceeds the initial capital commitment. On average, this value is 3% of the initial capital commitment. To deal with this issue, we define the following condition for adjusted capital commitment:

$$CC_{ADJ} = \max(max\ PIC, CC) \tag{9}$$

This states that the adjusted capital commitment in our dataset is always the maximum of capital commitment and maximum paid-in-capital. This ensures that the recallable distributions are added to the capital commitment value.

4. Empirical methodology

As illustrated in Section 2.3, the TA model is by nature deterministic, where the predictive performance is highly dependent on accurate values of its input parameters. A string of working papers such as Jeet (2020) and Jeet & O'Shea (2018) have introduced calibration techniques for estimating some of these parameters.

Meanwhile, Section 2.4 illustrates how macroeconomic factors are potentially linked to the TA model through three of these parameters: the rate of contribution (RC), the growth rate (G), and the bow factor (B). Our empirical approach aims to refine the TA model prediction by enhancing the estimation of RC and G using macroeconomic indicators. In this process, we make them time-dependent, potentially changing between periods. We focus on RC and G to align one input parameter each with the contribution and distribution function. This paper does not directly estimate B, instead, we assume a fixed value as this enables us to predict actual cash flows.

The next sections are organized as follows: Section 4.1 describes the main method of our enhanced parameter estimation, together with the benchmarks used to evaluate its performance and the macroeconomic input variables used. Sections 4.2, 4.3, and 4.4, respectively, define the exact methodology and variables used for the RC, G and B parameter estimations.

4.1 Model definitions

In estimating the parameters RC and G, the core method is to fit a linear regression model which estimates coefficients by minimizing OLS between actual and predicted values in the data. In the models we include macro-level indicators as explanatory variables, to explain variances in each estimated parameter. For clarification, Table 2 highlights the denotation practice we use. TA_{Macro} denotes the TA model with our macro-estimated parameters. All other parameters follow the same assumptions used by $TA_{Benchmark}$, outlined in Table 1. The bow factor parameter estimate is further expanded upon in Section 4.4.

Table 2.

Macro-dependent Takahashi-Alexander (TA_{Macro}) model parameters

This table lists the input parameter assumptions and denotations for the macro-dependent Takahashi-Alexander model (TA_{Macro}).

Model	Parameter	Value	Denotation
Contribution	RC (quarterly)	Time-dependent	$\widehat{RC_t}$
Distribution	B	4[4] (= $B_{Benchmark}$)[5]	
Distribution	L (quarters)	48 (= $L_{Benchmark}$)	
Distribution	G (quarterly)	Time-dependent	$\widehat{G_t}$
Distribution	Y (quarterly)	0% (= $Y_{Benchmark}$)	

4.1.1 Macroeconomic indicators

Prior research, for example Axelson et al. (2009) and Robinson & Sensoy (2016), has shown that PE cash flows exhibit co-movement with business cycle variation. Guided by this finding, our selection of macroeconomic explanatory variables has been focused on those that prior literature has explicitly tested with regards to the correlation with private equity cash flows. Below, we present these variables:

Price-dividend ratio

Robinson & Sensoy (2016) uses the price-dividend (P/D) ratio as the primary proxy for market liquidity to study the relationship with PE cash flows. The price-dividend ratio has been shown to strongly correlate with the market cycle (Fama & French, 1989), together with variables such as IPO and M&A activity (Gompers et al., 2008). Overall, the P/D ratio can be viewed as a measure for all variables that account for fluctuations in investment and exit opportunities (Robinson & Sensoy, 2016). Robinson & Sensoy (2016) find that the direction of the relationship between net cash flows is positive: as the business cycle goes up, PE funds become a liquidity provider. Both capital contributions and capital distributions have a positive relationship with the P/D ratio, but the effect is larger for distributions, leading to a procyclical relationship with net cash flows.

In defining the price-divided ratio, we calculate it by observing the S&P500 Price Index ($PI\ Index$) and the S&P500 Total Return Index ($TR\ Index$). The method of calculation is defined below.

[4] In accordance with Jeet (2020), a yearly B = 2 can be approximated toB = 4 on a quarterly basis
[5] Why a fixed value of B = 4 is used is clarified in section 4.4

$$P/D_{(t)} = \frac{TR\ Index_t}{\mathrm{D}_t} \tag{10}$$

where:

$$D_{(t)} = TR\ Index_{t-1} \times (TR\ Return_t - PI\ Return_t) \tag{11}$$

and:

$$TR\ return_{(t)} = \frac{TR\ Index_t}{(TR\ Index_{t-1})} - 1 \tag{12}$$

$$PI\ return_{(t)} = \frac{PI\ Index_t}{PI\ Index_{t-1}} - 1 \tag{13}$$

FED Balance Sheet

We introduce the FED Balance Sheet ($ million) ($Balance\ Sheet_t$) as an alternative proxy for market liquidity. The rationale for this decision is that the size of the federal balance sheet typically correlates with higher market liquidity due to the infusion of capital into the economy. As market liquidity is a multi-dimensional concept, the FED's balance sheet provides a more macroeconomic perspective of liquidity, rather than an immediate transactional measure as the P/D ratio.

Yield Spread

Robinson & Sensoy (2016) also find that both capital contributions and distributions are significantly negatively correlated with the yield spread. Intuitively, this is natural since the yield spread is a measure of the cost of debt. When the financing cost increases for leveraged buyouts, buy-side activities decrease for the PE fund, lowering the need to call capital. Likewise, the higher financing costs also impact potential buyers in sell-side M&A transactions, lowering potential distributions to LPs. We measure the yield spread ($Yield\ spread_t$) using the spread between a 10-year and a 2-year US treasury bond (constant maturity).

High-yield Spread

We also include a high-yield index ("Bank of America US High Yield Index Spread") as an additional proxy for yield spread, since LBOs are typically highly levered. We denote this variable ($High_yield\ spread_t$).

Public returns

As stated in Section 2.4.1, numerous studies have found that private equity returns, and cash flows, show correlation with public market returns (Tolkamp, 2007). We choose the S&P500 index return as our measure for public returns ($Public\ returns_t$).

4.1.2 Stepwise selection technique of explanatory variables

The method for selecting explanatory variables in our models is done using a stepwise selection technique for both RC and G, in line with Wagner & Shimshak (2007). In a stepwise procedure, explanatory variables are added or removed based on a predetermined selection criterion, namely statistical significance. This procedure can start with a model that includes all potentially significant variables and then uses backward elimination to remove non-significant variables. Alternatively, forward elimination can be used when starting with no variables and adding significant ones in each iteration.

In this paper, backward elimination is used, where the removal criterion is based on the highest p-value. Backward elimination is typically preferable when several predictors are suspected to be relevant in the specification. Given that we use several different macroeconomic indicators with different lags, all of which could reasonably be significantly related to the outcome variable, we have chosen to use the backward elimination approach.

There are both pros and cons when opting for this approach. On the one hand, the exclusion of variables is to some extent arbitrary since the selection depends on what order variables are removed from the model. However, it serves as a simplistic approach to fit regression models with many potential explanatory variables without large computational effort. Additionally, there is a risk of overfitting, resulting in a model that predicts poorly on unseen data. We mitigate this risk by adopting a cross-validation technique, detailed further in Section 4.1.4.1.

4.1.3 Robustness check

To check the robustness of our regression models, we begin with evaluating multicollinearity among explanatory variables by calculating the variance inflation factor

(VIF) for each variable. A VIF above ten suggests significant multicollinearity, potentially compromising the stability and reliability of the regression coefficients.

We then analyze four different diagnostic plots. Firstly, the residuals versus fitted values are plotted, to detect non-linearity, heteroscedasticity, and outliers. Secondly, we plot standard normal quantiles versus actual quantiles (normal q-q plot) to assess the assumption of normality among residuals, which is important to validate confidence intervals and hypothesis tests. Thirdly, we plot the square root of standardized residuals versus fitted values (scale-location plot) to further confirm homoscedasticity, ensuring consistent variance of residuals across predictor levels. Lastly, we plot residuals versus leverage to identify influential cases that could skew model estimation. If any issues are identified when analyzing the aforementioned diagnostic plots, we implement necessary actions suited to address these issues (Wooldridge, 2018).

4.1.4 Validation of model predictive performance

4.1.4.1 Data split

To mitigate the risk overfitting our regression models and thereby modeling randomness, we adopt a cross-validation technique. This is done by splitting the data into two subsets. The first subset (training set) is used to train the models using stepwise selection, representing vintage years 1990-2009. The second subset (test set) is used to test how the models perform on unseen data, simulating a real-world scenario. This subset represents the remaining vintage years 2010-2022.

4.1.4.2 Statistical measure

Throughout our analysis, we use the mean-squared error (MSE) measure on the test set to compare predictive power across models, defined as:

$$\mathrm{MSE} = \frac{\sum(y_t - \hat{y}_t)^2}{n} \tag{14}$$

In other words, the MSE measures the average squared distance (residual) between the observed and predicted values. Values closer to zero are therefore always better. Another feasible measure is the mean absolute error (MAE), which simply calculates the average of all absolute residual values. However, unlike the MAE, the MSE squares the residuals. Due to this, large errors (outliers), will have a relatively larger effect on MSE than the

MAE (Wooldridge, 2018). As a performance metric, we primarily choose the MSE because it may be specifically desirable to penalize large outliers due to the deterministic nature of the TA model. Nevertheless, we will examine MAE as an additional metric to compare prediction accuracy in absolute terms.

4.1.4.3 Benchmark to evaluate performance

In Section 2.3.1, we presented the parameter estimates used by Jeet & O'Shea (2018) ($TA_{Benchmark}$). In our analysis, we will use $TA_{Benchmark}$ as a benchmark to evaluate the performance of TA_{Macro}.

4.2 Contribution rate parameter estimation

The dependent variable in our model is the actual RC. We observe this by dividing observed capital contributions with the opening balance of uncalled capital, defined as:

$$RC_t = \frac{C_t}{CC - PIC_{t-1}} \tag{15}$$

We filter for observations where $CC - PIC_{t-1} > 1$ and $Age < L$. The reason is that for observations where either the uncalled capital is low or where the fund is mature, the observed rate of contribution has no significant impact on contribution cash flows.

4.2.1 Variable selection

In this section, we list the explanatory variables used to try and explain the variation in actual RC. The full definitions of macroeconomic variables are found under Section 4.1.1 and the resulting regression model is specified under Section 5.

Price-dividend ratio (t-1)

First, we include the P/D ratio as an explanatory variable. Given that capital contributions follow a procyclical pattern with market liquidity (Section 4.1.1), we expect a positive coefficient with regard to RC.

FED Balance Sheet (t-1)

Given that FED balance sheet is an alternative proxy for market liquidity (Section 4.1.1), we would expect that our coefficient has a positive direction with regards to the rate of contribution, as in the case for the P/D ratio.

Yield spread (t-1)

We predict a negative coefficient following Robinson & Sensoy (2016), with the rationale that when the financing costs increase for leveraged buyouts, buy-side activities decrease for the PE fund, lowering the need to call capital.

High-yield spread (t-1)

Following the above reasoning, we also predict a negative coefficient for the high-yield spread variable.

Maturity

Robinson & Sensoy (2016) find that cash flow variation can also be explained by the fund's age, not only by macroeconomic factors. To capture this lifecycle effect, we include age divided by fund lifetime as an explanatory variable and denote this term $Maturity_t$:

$$Maturity_t = \frac{Age_t}{Lifetime_t} \tag{16}$$

Investment period (dummy variable)

Another way to analyze the age effects on RC, is to divide age into phases. Looking at Figure 2, the data suggests a noticeable change in RC level at age 5. To investigate this phenomenon further, we categorize the lifecycle of investments into two distinct periods based on the age of the fund:

- *Investment Period 1 ("Initial Investment Phase"):* Defined as the period when the fund age is 5 years or less. This phase is characterized by higher capital contributions as the fund actively pursues investment opportunities.
- *Investment Period 0 ("Mature Phase"):* Defined as the period when the fund age is above 5 years. This phase typically exhibits a decline in capital contributions, indicating that the initial period of intensive investment has passed, and the fund has moved into a more mature stage.

This binary variable, $Investment_Period$ is used to capture the differential impact of the investment's lifecycle stage on the RC, where $Investment_Period$ is set to 1 for the initial investment phase and 0 for the mature phase.

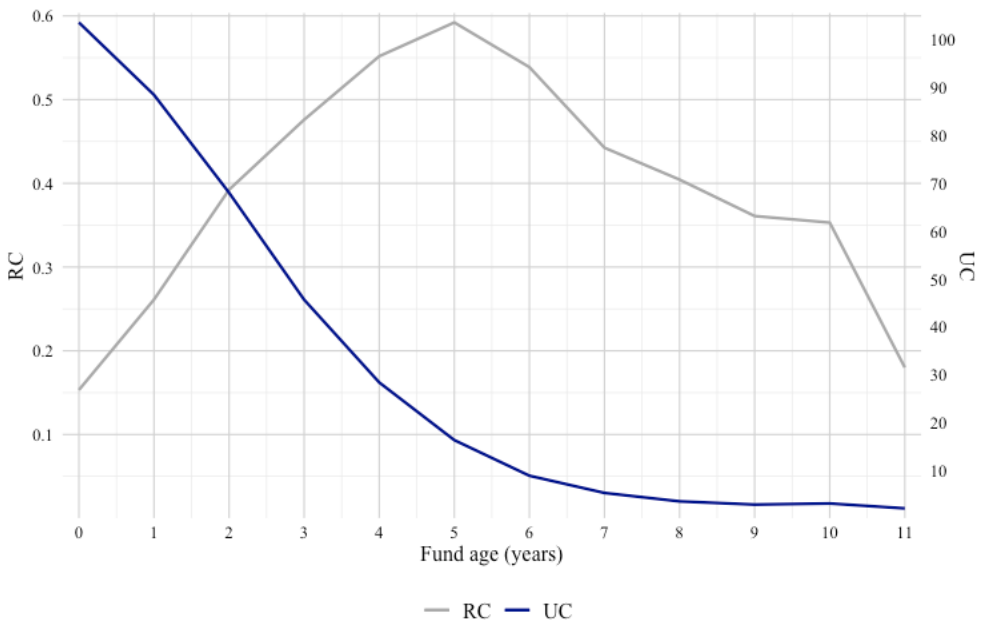
Uncalled capital (t-1)

The lifecycle effect on RC can also be captured by the level of uncalled capital (UC) the previous year. The level of uncalled capital (portion of committed but not yet called capital) reflects how much "dry powder" the GP has available to invest. Uncalled capital is directly observed in our data and its definition is found in Equation 5. Looking at Figure 2, it is reasonable to assume that there exists a relationship between RC and UC.

Figure 2.

Average RC and UC by age

Figure 2 plots average RC (across all vintages) and UC, by fund age.



4.2.2 Model specification

Our dataset contains quarterly observations by vintage year (Section 3.1). When estimating the regression, we combine all vintage-level cash flows into a new dataframe and calculate actual RC within each vintage. We do this on a yearly level, summarizing the data by year, to smooth out the volatility of RC since the quarterly observations are deemed too volatile. Intuitively, predicting contributions on an annual basis is also more in line with reality, as decision-making typically extends beyond a single quarter.

As described in Section 4.1.2, we use a stepwise selection technique on the training set to define our explanatory model. The explanatory variables we start with are found in Section 4.2.1. The resulting regression model is found in Section 5.

The validation of the model's predictive performance is done by comparing the MSE of $\widehat{RC_t}$ and $RC_{Benchmark}$ on the test set. The final validation will be to compare MSE of implied predicted contributions.

4.3 Growth rate parameter estimation

As described in Section 2.4.1.1, another way macroeconomic conditions enters the TA model is through the quarterly return on the NAV. Industry practitioners use historical IRRs to estimate this return, which in the TA model is proxied by G and is constant across the lifecycle. Jeet (2020) uses a regression analysis with several lags of public returns to examine how past economic conditions impact current growth rates, suggesting that historical data can predict future returns.

Following Jeet's methodology, we also regress the quarterly returns of NAV against lagged public returns to understand how previous economic shifts influence fund returns. By doing so, we expect to capture the delayed effects of economic changes on private equity returns, rather than simply assuming a fixed parameter value of G.

Jeet (2020) uses four lags of public returns to explain G, including lag 0, which is the same time period as the predicted outcome variable. This is problematic when predicting $G_{(t)}$ since lag 0 of public returns is unobserved in period t. Therefore, we begin by predicting lag 0 of public returns as a first-stage regression, before using it to predict $G_{(t)}$. Specifically, we use the lagged P/D ratios to explain lag 0 of public return in the same period as the predicted outcome for the growth rate.

In our second-stage regression to we use this estimated value as an explanatory variable to explain $G_{(t)}$. The following sections define the exact variables used in this specification.

4.3.1 Estimating current public returns (first-stage regression)

In this section, we list the explanatory variables used when forming the starting model of the stepwise selection. The full definitions are found under Section 4.1.1, and the resulting regression model is specified under Section 5.

4.3.1.1 Variable selection

Dependent variable

Public returns

In the first-stage regression, we use the S&P500 total return as the dependent variable and define it as Public Return ($Public\ Return_t$).

Explanatory variables

Price-dividend ratios (t-1, t-2, t-3, t-4, LTM)

In the first stage, we include four lags of P/D ratios and an LTM value as explanatory variables. We expect a positive coefficient with regard to public returns given that the P/D ratio shows strong correlation with the market cycle (Section 4.1.1).

4.3.2 Estimating growth rate parameter (second-stage regression)

In this section, we list the explanatory variables used when forming the starting model of the stepwise selection. The full definitions are found under Section 4.1.1 and the resulting regression model is specified under Section 5.

4.3.2.1 Variable selection

Dependent variable

Growth rate

In the second-stage regression, our dependent variable is the growth rate ($G_{(t)}$). The growth rate is defined using the NAV function in the TA model (Equation 4). Solving for the growth rate we get the following definition:

$$G_{(t)} = \frac{NAV_{(t)} - C_{(t)} - D_{(t)}}{NAV_{(t-1)}} - 1 \qquad (17)$$

Explanatory variables

Public returns (t) (estimated)

As stated in the section above, the estimated public returns ($\widehat{Public\ Return_t}$) from the first-stage regression becomes an explanatory variable in the second-stage regression. As explained in Section 4.1.1, private equity returns have been found to correlate with broader macroeconomic conditions. During periods of economic prosperity, asset

valuations generally increase, translating into an elevated growth factor. We therefore expect a positive coefficient of $\widehat{Public\ Return_t}$.

Public returns (t-1, t-2, t-3, t-4)

Additionally, we include four lags of public returns, in line with Jeet (2020).

FED balance sheet (t-1)

We predict that this coefficient has a positive direction given the positive correlation between market liquidity and fund returns (Fama & French, 1989; Robinson & Sensoy, 2016).

Yield spread (t-1)

We predict a negative coefficient following Robinson & Sensoy (2016), given that high interest rates are negatively associated with PE returns (G).

High-yield spread (t-1)

We predict a negative coefficient using the same argument as above.

4.3.3 Model specification

As described in Section 4.1.2, we use a stepwise selection technique on the training set to define our explanatory model. The explanatory variables we start with are found in 4.3.1 for the first-stage regression, and 4.3.2 for the second-stage regression. The resulting regression models are found under Section 5.

The validation of the model's predictive performance is done by comparing the MSE of $\widehat{G_t}$ and $G_{Benchmark}$ on the test set.

4.4 Bow factor parameter estimation

As deducted in Equation 8, the third and final parameter, after G and RC, which macroeconomic factors potentially could impact the TA model through is B. Although our core purpose of this paper is to improve the estimation of RC and G, it would still be interesting to address the bow factor for two reasons.

Firstly, estimating the bow factor gives the opportunity to make the rate of distribution dependent on the behavior of the GPs. A manager estimating the model can through the bow factor control when the underlying NAV will be distributed based on these expectations. As seen in Appendix A, a higher bow factor drives a slower initial growth

in RD and a faster acceleration towards the end of the lifecycle, and vice versa. Secondly, a value of B is needed in order to apply the TA models, thus analyzing how our adjusted parameter values of RC and G are de facto impacting net cash flows.

It is beyond the scope of this paper to examine a new method of estimating B. However, we can still obtain a B with conventional methods. In all predictions we use benchmark L = 48 quarters as referenced in Table 1.

4.4.1 Fixed bow factor

To compare the forecasted distributions and net cash flows, we run both TA_{Macro} and $TA_{Benchmark}$, with a fixed B of 4[6] across all vintages. This aligns with the approach taken by Jeet (2020) and maintains consistency between the models. Subsequently, we derive MSE and MAE values for comparison of predictive performance on distributions and net cash flows.

4.4.2 Optimizing bow factor

Additionally, we will calculate the optimal bow factor for each vintage and model (TA_{Macro} and $TA_{Benchmark}$). This is done with an optimization function, which applies Brent's method to find the parameter value of B that minimizes the sum squared residuals of capital distributions—henceforth defined as B_{OPT}. This exercise allows us to compare and analyze what would happen to the prediction performance of TA_{Macro}, given that we would have known the B_{OPT} estimate for each vintage. It also enables comparison of patterns in the phasing of distributions between vintages.

[6] In accordance with Jeet (2020), a yearly B = 2 can be approximated to B = 4 on a quarterly basis

5. Results

In this section, we present the main results of our study. Section 5.1 provides the results from the contribution rate estimation, 5.2 the growth rate estimation, and finally 5.3 the bow factor estimation.

5.1 Contribution rate parameter results

5.1.1 Model specification

The resulting regression model from the stepwise selection process is given below (Model 1).

$$RC_t = \beta_0 + \beta_1 \times Yield\ Spread_{t-1} + \beta_2 \times Investment_period_1_t + \beta_3 \times UC_{t-1} + \varepsilon_t \qquad \text{Model (1)}$$

Table 3 presents the results from Model 1. It is based on 210 observations and explains approximately 30% of the variance in RC_t. The coefficient $Yield\ Spread_{t-1}$ has a value of -0.0293 and is statistically significant, indicating that a higher yield spread in the preceding period is associated with a decrease in RC_t. This supports the overall conclusion by Robinson & Sensoy (2016) that debt financing conditions correlates negatively with capital contributions. $Investment_period_1_t$ has a coefficient value of 0.2479 and is statistically significant. This supports our initial prediction that RC_t is higher within the investment period. The coefficient value of UC_{t-1} is -0.0043 and statistically significant, indicating that high levels of dry powder are associated with a decrease in RC_t.

Table 3.

Contribution rate regression results

Table 3 lists the regression results for Model 1 with RC as dependent variable.

	Dependent variable:
	RC actual
	Model (1)
Yield Spread (t-1)	-0.0293*
	(0.0133)
Investment Period 1	0.2479***
	(0.0353)
Uncalled Capital (t-1)	-0.0043***
	(0.0005)
Number of observations:	210
R-squared:	0.30
Adjusted R-squared:	0.29
Note:	*p<0.01; ** p<0.001; ***p<0.0001

5.1.2 Robustness check

Appendix C lists the VIF for the multicollinearity check in Model 1. All variables are well below ten, implying that there is no indication of multicollinearity among the explanatory variables. Furthermore, looking at Appendix D.1, the diagnostic plots provide further insights into the robustness of the regression model. The residuals-fitted plot displays no clear patterns, suggesting a good model fit with no apparent non-linearity or heteroscedasticity. The normal q-q plot closely follows the theoretical line, indicating that the residuals are normally distributed. The scale-location plot shows some spread, but no distinct patterns, further supporting the assumption of homoscedasticity. Lastly, the residuals-leverage plot identifies only one point slightly outside the Cook's distance line, which is deemed not numerous and significant enough to warrant further actions.

5.1.3 Validating predictive performance

Model 1 was then run on the test set and was evaluated by comparing the performance of $\widehat{RC_t}$ and $RC_{Benchmark}$ in predicting actual RC_t and C_t.

The results show that $\widehat{RC_t}$ had an MSE of 0.0305, while $RC_{Benchmark}$ had an MSE of 0.0829. When using the rates of distribution and predicting actual C_t, MSE from $\widehat{RC_t}$ implied C_t was 19.53, compared to 73.40 from $RC_{Benchmark}$. These results indicate a significant increase in predictive performance on capital contributions with Model 1. The full results are presented in Table 5.

Figure 3.

Cumulative contributions for TA_{Macro} and $TA_{Benchmark}$ vs. actual, by vintage

Plots only include a selection of vintages and are provided for illustrative purposes only.

Figure 3a: TA_{Macro}

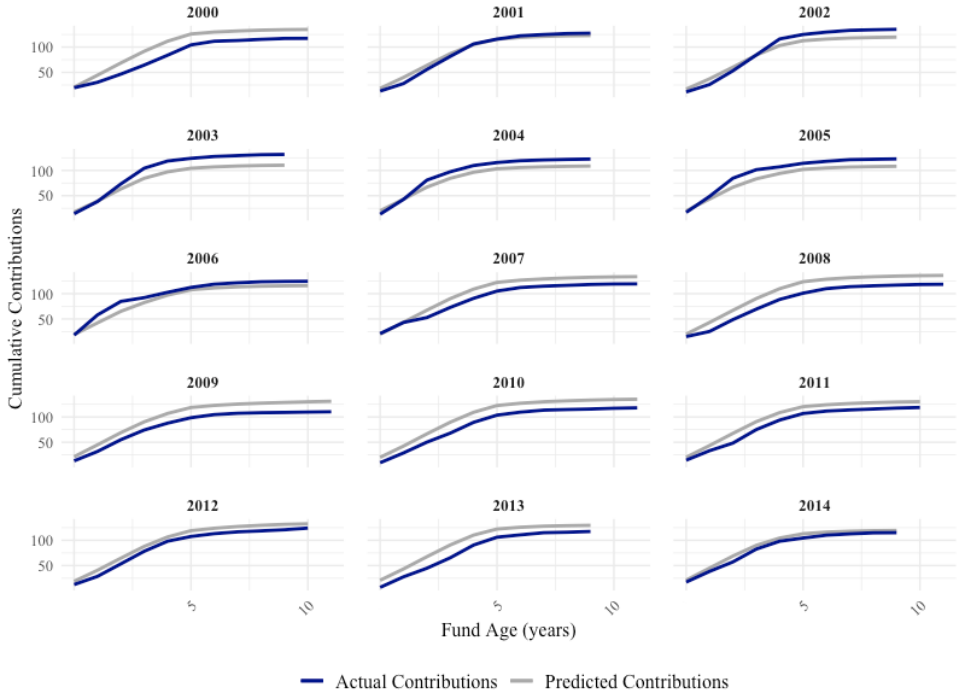
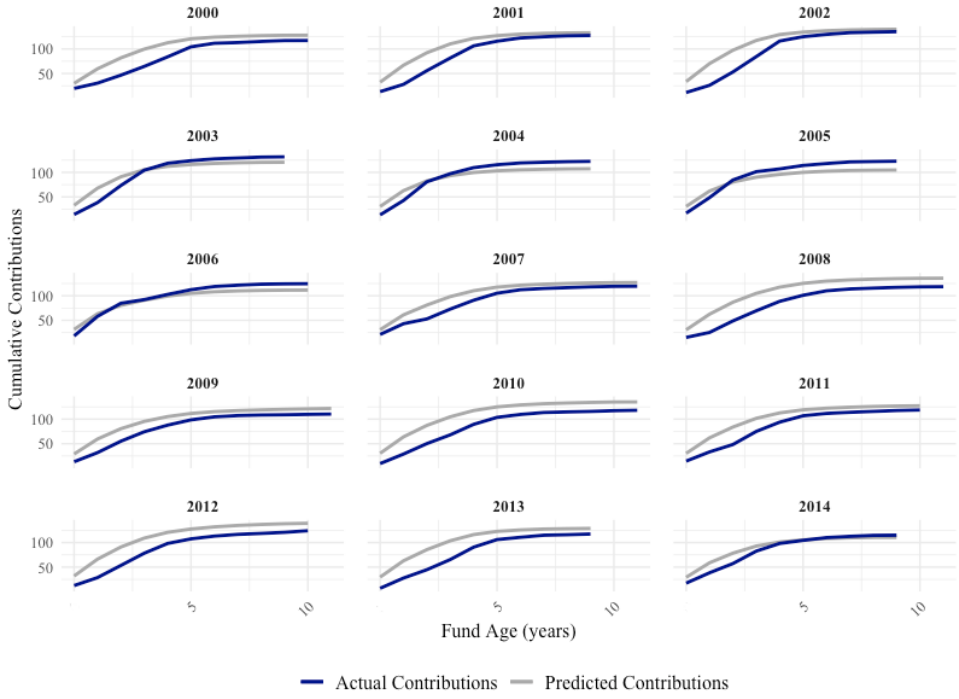


Figure 3b: $TA_{Benchmark}$



5.2 Growth factor parameter results

5.2.1 Estimating current public returns

5.2.1.1 Model specification

Our formal expression for the first-stage regression (Model 2) includes P/D_{t-1} and P/D_{LTM} to predict the $Public\ returns_t$.

$$Public\ returns_t = \beta_0 + \beta_1 \times P/D_{t-1} + \beta_2 \times P/D_{LTM} + \varepsilon_t \qquad \text{Model (2)}$$

Table 4 shows that Model 2 can explain approximately 64 percent of the variation in $Public\ returns_t$ for 78 observations, as indicated by the R-squared value.

The coefficient of P/D_{LTM} is significant with a coefficient value of 0.0141, which is in line with our expectations and previous research (Fama & French, 1989; Robinson & Sensoy, 2016). However, the coefficient of P/D_{t-1} is instead -0.0035 and statistically significant. This is not in line with our expectations. It may indicate a more complex relationship than anticipated between the timings of P/D ratios and public returns.

5.2.2 Estimating growth rate parameter

5.2.2.1 Model specification

In the second-stage regression, we include $\widehat{Public\ returns_t}$ and $High\ yield\ spread_{t-1}$ as explanatory variables to predict G_t.

$$G_t = \beta_0 + \beta_1 \times \widehat{Public\ returns_t} + \beta_2 \times High\ yield\ spread_{t-1} + \varepsilon_t \qquad \text{Model (3)}$$

Table 4 shows that $\widehat{Public\ returns_t}$ is statistically significant and has a coefficient value of 0.3723. Further, $High\ yield\ spread_{t-1}$ is statistically significant with a coefficient value of -0.0083. The direction of the relationship for both variables is in line with our expectations and previous research. Finally, the number of observations in Model 3 (1068) is significantly higher than in Model 2. This is because G_t observations from different vintage years overlap, leading to multiple values per quarter.

Table 4.

Growth rate parameter results

This table lists the regression results for Model 2 with public return ($Public\ returns_t$) as dependent variable (first-stage regression) and for Model 3 with growth rate (G_t) as dependent variable (second-stage regression).

	Dependent variable:	*Dependent variable:*
	Public return	G actual
	Model 2 (1)	Model 3 (2)
P/D (t-1)	-0.0035*** (0.0003)	
P/D (LTM)	0.0141*** (0.0012)	
$\widehat{\text{Public return}}\ (t)$		0.3723*** (0.035)
High yield spread (t-1)		-0.0083*** (0.0008)
Number of observations:	78	1068
R-squared:	0.65	0.30
Adjusted R-squared:	0.64	0.30
Note:	*p<0.01; ** p<0.001; ***p<0.0001	

5.2.2.2 Robustness check

All VIF factors are well below ten, implying no indication of multicollinearity among the explanatory variables (Appendix C). Looking at Appendix D.2, the diagnostic plots provide further insights into the residuals and thus the robustness of the regression model. The residuals-fitted plot displays a fairly even distribution of residuals suggesting a robust model fit with no significant evidence of non-linearity or heteroscedasticity. While there is a slight increase in scatter across certain ranges of fitted values, it is confined to a small number of ranges, and is minimal enough to not require any further actions. The scale-location plot displays a consistent spread across different levels of fitted values, further supporting homoscedasticity. Lastly, the residuals-leverage plot identifies only one point slightly outside the Cook's distance line, indicating no significantly influential outliers.

5.2.2.3 Validating predictive performance

Model 3 was then run on the test set and was evaluated by comparing the performance of $\widehat{G_t}$ and $G_{Benchmark}$ in predicting actual G_t. As shown in Table 5, $\widehat{G_t}$ has an MSE of 0.0030, compared to 0.0048 for $G_{Benchmark}$. These findings indicate that our model (Model 3) more accurately predicted future G_t throughout the test period

Table 5.

MSE comparison

This table lists the MSE for each estimated model tested on the test dataset.

Parameter	**Model 1**$(\widehat{RC_t})$	$RC_{Benchmark}$	**Model 3** $(\widehat{G_t})$	$G_{Benchmark}$
RC	0.0305	0.0829		
C	19.53	73.40		
G			0.0030	0.0048

5.3 Bow factor parameter results

5.3.1 Fixed bow factor

Next, we run the full TA model using a fixed bow factor, thus allowing us to predict distributions. As seen in Table 6, MSE of predicted distributions per vintage from TA_{Macro} (average = 15.61) is consistently higher than from $TA_{Benchmark}$ (average = 12.32). The difference in MSE is particularly noticeable between vintage years 2011-2016. TA_{Macro} predicts distributions more accurately for 14 out of 32 vintage years.

Table 6 also presents the MSE on net cash flows, where TA_{Macro} (average = 18.29) which still has a higher average value than $TA_{Benchmark}$ (average = 14.68). However, the difference is significantly lower when comparing the MAE on net cash flows, which is less sensitive to outliers than MSE. For illustrative purposes, we present plots for the predicted and actual net cash flows for a selection of vintages. The plots of a selection of vintages for TA_{Macro} are found in Figure 5a and for $TA_{Benchmark}$ in Figure 5b.

.

Table 6.

Cash flow predictive performance with fixed bow factor

This table lists the MSE of predicted distributions from $TA_{Benchmark}$ and TA_{Macro}.

		MSE D		**MSE D**		**MAE NC**	
Vintage year	B_{Fixed}	TA_{Macro}	$TA_{Benchmark}$	TA_{Macro}	$TA_{Benchmark}$	TA_{Macro}	$TA_{Benchmark}$
1990	4	34.32	32.78	41.20	42.10	3.84	3.78
1991	4	32.24	31.94	40.45	39.17	4.60	4.22
1992	4	43.02	34.24	66.08	53.10	4.83	4.13
1993	4	28.67	24.79	31.36	27.72	4.46	4.17
1994	4	19.60	20.89	19.22	19.43	2.92	2.90
1995	4	11.46	13.78	10.93	12.78	2.71	2.84
1996	4	13.12	16.76	13.60	14.46	2.86	2.81
1997	4	3.28	5.34	4.40	6.15	1.62	1.84
1998	4	3.92	4.34	6.04	6.03	1.78	1.77
1999	4	7.06	5.56	10.15	9.12	2.28	2.09
2000	4	11.99	9.48	12.51	10.76	2.69	2.48
2001	4	23.77	25.08	23.87	23.73	3.55	3.63
2002	4	21.14	25.64	20.68	24.05	3.38	3.72
2003	4	15.88	20.02	16.02	19.47	3.13	3.47
2004	4	10.14	13.48	12.37	13.62	2.59	2.95
2005	4	4.27	6.33	4.74	6.59	1.82	1.90
2006	4	5.43	6.43	8.08	8.02	2.31	2.28
2007	4	7.28	5.98	8.81	7.13	2.37	2.22
2008	4	8.86	5.71	10.13	8.43	2.31	2.24
2009	4	10.11	4.81	11.38	5.91	2.37	1.83
2010	4	15.01	7.40	16.78	11.01	3.18	2.65
2011	4	19.73	8.93	21.22	10.28	3.36	2.57
2012	4	21.60	9.31	22.80	12.03	3.16	2.54
2013	4	21.32	6.70	25.67	12.01	3.77	2.46
2014	4	27.52	7.34	30.04	7.54	4.13	2.13
2015	4	31.43	11.47	34.30	14.34	3.78	2.77
2016	4	25.19	12.29	26.56	13.64	3.51	2.75
2017	4	14.21	10.11	16.07	11.75	2.90	2.62
2018	4	5.63	5.03	5.78	4.30	1.87	1.58
2019	4	1.76	1.79	3.46	5.41	1.44	1.67
2020	4	0.40	0.45	6.60	8.13	1.91	2.01
2021	4	0.12	0.13	4.01	1.51	1.64	0.96
Average		**15.61**	**12.32**	**18.29**	**14.68**	**2.91**	**2.62**

Figure 5.

Cumulative net cash flow for TA_{Macro} and $TA_{Benchmark}$ vs. actual, by vintage

With $B_{Fixed} = 4$. Provided for illustrative purposes only.

Figure 5a: TA_{Macro}

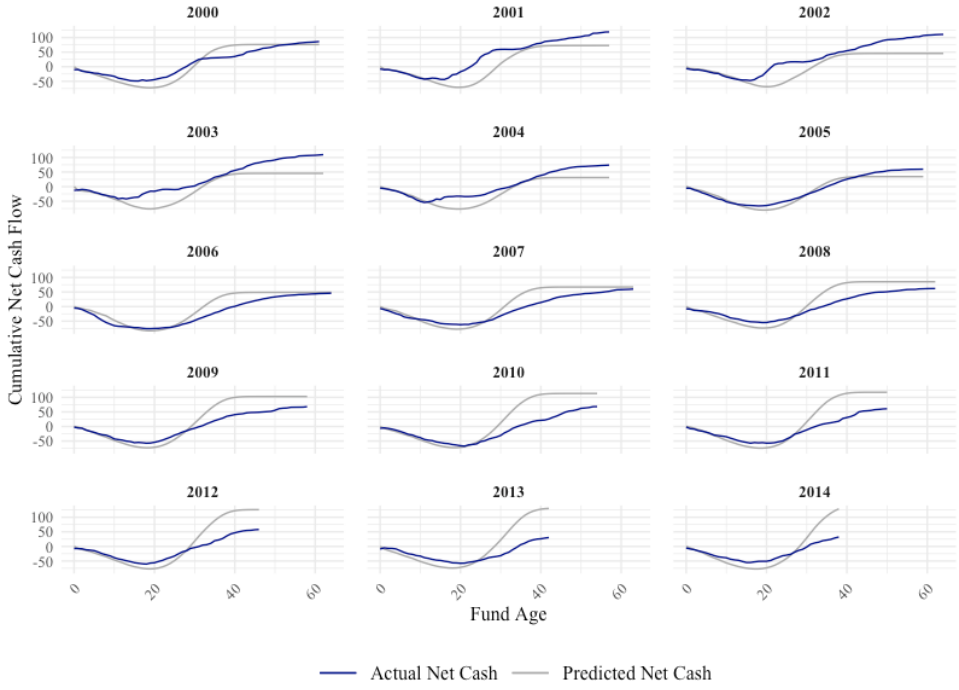
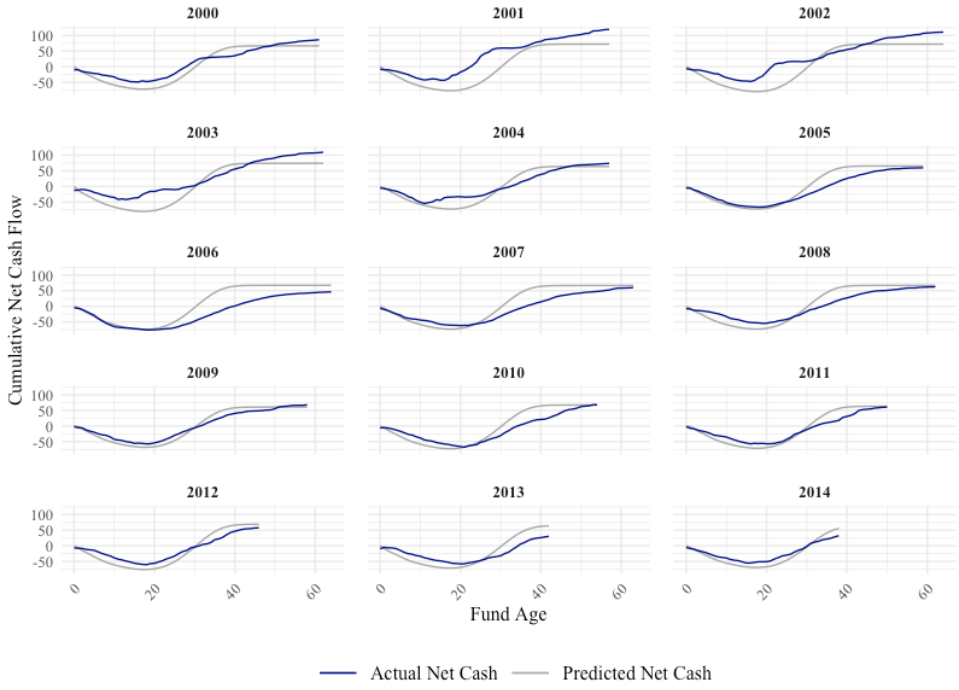


Figure 5b: $TA_{Benchmark}$



5.3.2 Optimized bow factor

As seen in Table 7, when optimizing B per vintage, the MSE of predicted distributions per vintage year from TA_{Macro} (average = 9.62) is still higher than from $TA_{Benchmark}$ (average = 7.80). The difference in MSE is however smaller. TA_{Macro} predicts distributions more accurately for 12 out of 32 vintage years. Looking at net cash flow, the performance is, as expected, more even: TA_{Macro} has an average MSE of 12.83 and $TA_{Benchmark}$ 12.49. For MAE, TA_{Macro} has a value of 2.52 and $TA_{Benchmark}$ 2.51.

Table 7.
Cash flow predictive performance with optimal bow factor

MSE and MAE for net cash flows from $TA_{Benchmark}$ and TA_{Macro}.

	TA_{Macro}				$TA_{Benchmark}$			
Vintage year	B_{OPT}	MSE D	MSE NC	MAE NC	B_{OPT}	MSE D	MSE NC	MAE NC
1990	3.95	34.32	41.20	3.83	5.05	31.55	39.90	3.78
1991	3.48	29.93	38.44	4.17	3.68	31.34	39.05	4.06
1992	1.72	18.70	43.37	3.36	1.97	19.16	43.84	3.51
1993	2.35	11.49	15.12	2.93	2.54	12.18	18.23	3.13
1994	2.30	9.78	10.48	2.30	2.39	10.25	11.83	2.36
1995	4.77	10.60	10.30	2.61	1.80	9.44	14.87	3.04
1996	1.70	7.49	7.02	2.13	1.62	6.24	7.72	2.32
1997	3.96	3.28	4.40	1.62	4.03	5.34	6.12	1.84
1998	3.28	2.61	4.47	1.64	3.37	3.48	5.46	1.76
1999	3.02	2.70	6.54	1.89	3.22	3.18	8.09	2.00
2000	2.65	3.90	6.32	1.97	2.88	3.66	7.70	2.20
2001	2.14	7.17	8.55	2.13	2.37	8.04	11.52	2.65
2002	1.61	7.60	9.44	2.42	1.97	8.39	13.64	3.12
2003	1.04	9.91	12.55	2.84	1.48	9.28	16.16	3.21
2004	3.71	10.10	12.19	2.61	1.37	8.91	14.52	3.20
2005	4.01	4.27	4.74	1.82	3.82	6.28	6.62	1.94
2006	3.68	5.22	7.79	2.33	3.81	6.38	8.07	2.32
2007	3.23	6.16	8.09	2.34	3.55	5.62	7.47	2.30
2008	2.88	6.43	8.32	2.33	3.25	4.65	8.76	2.45
2009	2.95	5.98	7.75	2.32	3.17	3.34	5.65	1.95
2010	3.24	11.90	14.33	2.95	3.71	7.15	11.35	2.67
2011	2.80	11.61	13.78	2.85	3.23	7.22	10.06	2.49
2012	2.55	11.35	12.41	2.93	3.05	6.93	11.67	2.76
2013	2.84	16.22	23.17	3.99	3.97	6.70	12.08	2.48
2014	2.64	11.13	14.58	3.07	3.11	4.21	6.00	1.96
2015	2.61	10.49	14.86	2.96	2.96	4.61	10.69	2.69
2016	6.65	17.38	18.14	3.05	2.48	4.99	9.19	2.65
2017	4.91	12.47	13.36	2.67	2.08	4.34	12.07	2.86
2018	4.32	5.45	5.22	1.80	3.57	4.82	4.87	1.71
2019	3.85	1.74	3.54	1.48	3.40	1.56	6.21	1.98
2020	3.37	0.28	6.40	1.86	3.09	0.25	8.67	2.11
2021	3.19	0.09	3.77	1.56	3.00	0.09	1.67	0.98
Average (std.dev)	**3.17 (1.10)**	**9.62**	**12.83**	**2.52**	**2.97 (0.84)**	**7.80**	**12.49**	**2.51**

6. Discussion and conclusions

In this section, we discuss the implications of our main results. The discussion is structured as follows: Section 6.1 focuses on the rate of contribution, Section 6.2 the growth rate, and Section 6.3 the bow factor. Finally, we expand on further considerations in Section 6.4.

6.1 Rate of contribution

Our results indicate that TA_{Macro}, which incorporates yield spread, investment period (dummy variable), and uncalled capital, can better predict the actual yearly RC than $TA_{Benchmark}$. The TA model, which focuses on modeling cash flows by subtracting contributions from distributions, demonstrates that RC impacts the volume of contributions. However, the impact is not clear, as the relative impact of RC on contributions differs along the lifetime of the fund. Therefore, we also test the cash flow implications, which show that our model predicts actual contributions more precisely than the benchmark, reducing the MSE by more than two-thirds.

The highly significant explanatory variables uncalled capital and investment period, present the primary reason as to why TA_{Macro} performs better at predicting capital contributions. This indicates that GP-timing of capital contributions is more closely associated with certain stages of the fund, rather than with certain conditions in the market. However, the yield spread is still significant: it has a p-value of less than 0.03 on the training set, adds about two percentage points to a total R-squared of around 30%, and removing it increases the MSE on test data by approximately two units. From a macroeconomic perspective, we hypothesize that the negative coefficient on yield spread is linked with increased financing costs for leveraged buyouts, decreasing the appetite for deals and by extension contributions.

The prospect of more reliable predictions of capital contributions has significant implications for both LPs and GPs—who endogenously determine when to call for capital. As we have seen, these terms represent a double-edged sword for LPs. On the one hand, the flexibility in the timing of contributions is a key driver for PE funds in outperforming public markets since the IRR only begins to accrue once the GP calls capital. If instead IRR calculations were based on the total committed capital from day one of the vintage, achieving comparable IRR figures would be significantly more difficult, since this phasing of contributions would not be possible. On the other hand, the

uncertainty in the timing of capital contributions can pose a significant challenge, particularly for LPs who face ongoing liquidity demands. To maintain sufficient liquidity and target allocations, these investors might reduce the overall exposure to PE and hold funds in low-yield, liquid "temporary transit" assets, pending to be called (see Section 2). Meyer et al. (2013) argue that this illiquidity and unpredictability is a natural part of the limited partnership, where investors earn an illiquidity premium in return.

More accurately predicting capital contributions mitigates this unpredictability, while still enabling LPs to leverage the upside of the illiquidity premium. There might also be potential second-order effects from the GP side as well, since less liquidity-constrained LPs would imply higher target allocations into the asset class. All else equal, this would reasonably make it easier for GPs to raise capital from this investor pool.

6.2 Growth rate

Performing a two-stage regression with lagged public returns and high yield spread as explanatory variables, we explain approximately 30% of the variance in G during the training period. TA_{Macro} predictions also achieve a significantly lower MSE on the test set.

As mentioned, lagged public returns served as a significant explanatory variable, aligning with Jeet (2020). Interestingly, we also found that the high yield spread (t-1) is a strong predictor of G, with a p-value of less than 2e-16. From a macroeconomic perspective, LBOs are typically highly levered transactions, which supports the high yield spread's predictive strength for G. However, it raises the question of why the high yield spread is a more robust predictor for G, while the treasury yield spread appears to predict RC more accurately. This could possibly be linked to the treasury yield spread's connection to the risk-free rate, though this relationship is not fully understood.

Enhancing the prediction of G may offer potential benefits for LPs. However, it is important to understand the complexities involved. More accurate predictions may hold value in themselves for PE investors, but these predictions are in fact unrealized IRRs. The reliability and usefulness of these unrealized IRRs for LPs have been problematized in recent papers (e.g., Huthner, 2023), since the rolling valuation of private assets is in fact difficult and should only be seen as indicative. More important is the second-order implication on cash flows of our enhanced G estimation, as the objective of the TA model is to predict cash flows. Nevertheless, the effect on actual cash flows is more ambiguous than in the case of RC. As seen in Equation 8, capital distributions are parameterized in a

more complex way than contributions, including both G and a non-linear bow factor effect. We will elaborate on this in the next section.

6.3 Bow factor

6.3.1 Fixed bow factor

To gauge how our prediction of G affects predicted distributions, we first set B to equal four and run the full TA model. However, $TA_{Benchmark}$ outperformed TA_{Macro} in predicting actual distributions, as evidenced by a lower MSE. We also compared MSE on net cash flows, with the rationale that what matters for LPs is the net effect on cash flows, rather than isolated impacts on distributions and contributions. This exercise showed that the benchmark model still exhibited a slightly better MSE, raising questions if the MSE's sensitivity to outliers could be penalizing a more volatile macro-dependent model. When comparing MAE, results still favor $TA_{Benchmark}$, albeit less significantly.

When comparing between vintages, it is evident that the predictive accuracy fluctuates in both models across different market conditions. Notably, later vintages (2011-2016) tend to show significant increase in the disparity between TA_{Macro} and $TA_{Benchmark}$ MSEs, suggesting that the impact of macroeconomic adjustments may vary more under certain economic conditions or stages of fund maturity.

This comparison underscores the sensitivity of the TA model's predictive performance to the input parameters. While the fixed bow factor ensures consistency in the model's phasing of distributions, the variations in RC and G parameters significantly influence the model outcomes. The relatively higher MSE from the TA_{Macro} model suggests that while the inclusion of macro- and time-dependent RC and G parameters may provide a more dynamic and realistic framework for modeling certain fund outcomes, it also introduces more variability and potential for error in the model's final predictions.

6.3.2 Optimized bow factor

Fixing B at four might not optimally reflect the dynamics captured by the macro- and time-dependent nature of TA_{Macro}. Specifically, the impact on distributions from RC and G suggests that a fixed B may not adequately adjust to the variability and specific circumstances of different vintages and market conditions.

Therefore, we explored whether a more flexible, optimized B could potentially enhance the performance of the TA_{Macro} model relative to the benchmark. By optimizing B for

each vintage and model, we aimed to discover if knowing the optimal B in advance could have favored TA_{Macro}, in terms of predicting net cash flows.

Interestingly, when B was optimized, the difference in MSE between the models lessened, although $TA_{Benchmark}$ still maintained a slight edge. This reduced difference suggests that an optimized B does refine our models' forecasting ability, but not sufficiently to surpass the benchmark model under all conditions. The MAE compared similarly between the models, however, TA_{Macro} exhibited a higher incidence of outliers, which negatively impacted its MSE. This trend underscores the volatile nature of TA_{Macro}.

The higher variability in B_{OPT} in TA_{Macro} (std.dev = 1.10) compared to B_{OPT} in $TA_{Benchmark}$ (std.dev = 0.84), suggests that distribution rates in TA_{Macro} need to be more dynamic to mimic actuals. Further, the increased performance of TA_{Macro} when introducing a more dynamic B between vintages could also suggest that dynamic adjustments within each vintage may be beneficial. This would imply making it time-dependent, like $\widehat{RC_t}$ and $\widehat{G_t}$. It could also imply including a macro-dependency, for example by predicting larger amounts of distributions when IPO markets are hot. This analysis goes beyond the scope of this paper.

The relatively high variability observed across vintages in B_{OPT} for TA_{Macro} may also indicate that macroeconomic conditions at the time of a fund's inception govern how capital distributions are portioned throughout the fund's lifetime, which could be of interest in future research.

6.4 Further considerations

The time-dependency of our parameter prediction of RC and G presents both pros and cons for the use case of our model. Benefits include more accurate calibration of RC and G across time. Further, as mentioned in Section 2.3, the TA model is deterministic meaning that it only outputs one single prediction given the same input. Including probabilistic parameter estimations, as in TA_{Macro} , has an advantage in that it can output a confidence interval in the predicted parameters. On the downside, its time-dependency does not make long-term simulations possible since t-1 observations of explanatory variables are constantly needed to keep TA_{Macro} up to date. How often cash flow projections should be run is in fact a common discussion point among financial institutions (Meyer et al., 2013), where an additional consequence of increased frequency is more effort in keeping all inputs updated.

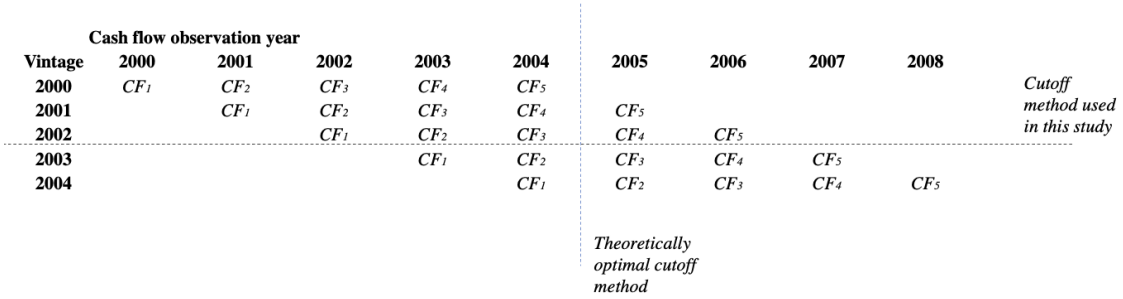
Another important consideration is that using vintage 2009 as a cutoff between training and test datasets may exclude meaningful variation in economic conditions, as TA_{Macro} is not trained on vintages post-2009. This may be influential, given that the 2010s saw a shift in economic conditions, with drastically increased market liquidity and historically low interest rates (FRED, 2024). However, there is a tradeoff between capturing variance in macroeconomic conditions and having enough observations to mitigate the risk of overfitting by validating on a sufficiently large test set. Selecting vintage 2009 as a cutoff lead to a 70-30 split between our training and test datasets when predicting RC and G, which is deemed sufficient by Wooldridge (2018). The question then becomes if we are training the model on meaningful data. Intuitively, training a model on outdated observations may result in a model that does not reflect relevant market dynamics. Fundamentally, this issue arises because private equity is still a relatively new asset class that has not yet experienced many market cycles.

This dichotomy between capturing variation in the prediction period and ensuring model robustness also applies to our internal variables that capture lifecycle effects. Concluding how specific windows across the lifecycle affect cashflows thus also needs variability in observations in these lifecycle windows. This can be problematic given that a fund lifecycle is on average 12 years, and our test period starts with the vintage year 2009. Thus, we can observe the full lifecycle for the first two vintages—later vintages will inevitably not include later-stage observations in their lifecycle, since they will lie in the future. Another issue is that there exist observations from training set-vintages that are dated post-2009. For example, vintage 2005 includes observations until 2017. In theory, one could argue that we should cut at a specific date within each vintage year (see example in Figure 6). This would allow for no overlaps and a more equal spread of fund age observations. However, this is not possible since the TA model requires to be run on a full vintage. Future research may include some kind of sensitivity analysis concerning how the training and test periods are defined.

Figure 6.

Example cutoff between training and test datasets

Scenario with cash flow observations from five vintages, each with a lifetime of five years, for simplicity.



It is also relevant to discuss the generalizability of our results. Our dataset covers a pooled set of global buyout funds (see Section 3). Intuitively, this means that within each vintage, an investor would hold an indexed value of each fund that is scaled by its relative size on total capital commitment. In practice, investors typically invest in a selected set of funds in a selected set of geographies, which incurs additional fixed effects that are not included in our model. Future research may validate the generalizability of our results by estimating TA_{Macro} on datasets covering various selection criteria of funds. Thus, evaluating the extent to which our results hold across various investment contexts.

7. References

7.1 Published papers

Axelson, U., Strömberg, P., M., & S., Weisbach. (2009). Why Are Buyouts Levered? The Financial Structure of Private Equity Funds. The journal of finance 64(4), 1549-1582. https://doi.org/10.1111/j.1540-6261.2009.01473.x

Ang, A., & Sorensen, M. (2012). Risks, Returns, and Optimal Holdings of Private Equity: A Survey of Existing Approaches. *The quarterly journal of finance 2(3), 1250011*. https://doi.org/10.1142/S2010139212500115

Beutler, D., Billias, A., Holt, S., Lerner, J., & Seet, T. (2023). Takahashi–Alexander Revisited: Modeling Private Equity Portfolio Outcomes Using Historical Simulations. Journal of portfolio management 49(7), 144-158. https://doi.org/10.3905/jpm.2023.1.496

Fama, E., & French, K. (1989). Business conditions and expected returns on stocks and bonds. *Journal of financial economics 25, 23-49.* https://doi.org/10.1016/0304-405X(89)90095-0

Gompers, P., Kovner, A., Lerner, J., & Scharfstein, D. (2008). Venture capital investment cycles: The impact of public markets. *Journal of Financial Economics 87(1), 1-23.* https://doi.org/10.1016/j.jfineco.2006.12.002

Harris, R., S., Jenkinson, T., Kaplan, S., N., Stucke R. (2023). Has persistence persisted in private equity? Evidence from buyout and venture capital funds. *Journal of corporate finance 81, 102361*. https://doi.org/10.1016/j.jcorpfin.2023.102361

Huther, N. (2023). Do private equity managers raise funds on (sur)real returns? Evidence from deal-level data. *Journal of financial and quantitative analysis 58(7), 2959-2992*. https://doi.org/ 10.1017/S0022109022000990

Jenkinson, T., & Sousa, M. (2015). What determines the exit decision for leveraged buyouts? *Journal of banking & finance 59(1), 399-408.* https://doi.org/10.1016/j.jbankfin.2015.06.007

Johan, S., & Zhang, M. (2021). Information Asymmetries in Private Equity: Reporting Frequency, Endowments, and Governance. *Journal of business ethics 174(1), 199-220.* https://doi.org/10.1007/s10551-020-04558-6

Kaplan, S., N., & Strömberg, P. (2009). Leveraged Buyouts and Private Equity. *The journal of economic perspectives 23(1), 121-146.* https://doi.org/10.1257/jep.23.1.121

Kieffer, E., Meyer, T., Gloukoviezoff, G., Lucius, H., & Bouvry, P. (2023). Learning private equity recommitment strategies for institutional investors. *Frontiers in Artificial Intelligence, 6(1), 1014317-1014317.* https://doi.org/10.3389/frai.2023.1014317

McGrath, P., & Nerkar, A. (2023). Private Equity: Antecedents, Outcomes, Mediators, and Moderators. *Journal of management 49(1), 158-195.* https://doi.org/10.1177/01492063221127907

Robinson, D., T., & B., A., Sensoy. (2016). Cyclicality, performance measurement, and cash flow liquidity in private equity. Journal of financial economics, 122(3), 521-543. https://doi.org/10.1016/j.jfineco.2016.09.008

Sharma, M., & Kaushal, S. (2012). The S&P Private Equity Index versus the Private Equity Total Return Index: A Volatility Comparison. *The journal of private equity, 15(2), 51-52.* https://doi.org/10.3905/jpe.2012.15.2.051

Takahashi, D., & Seth, A. (2002). Illiquid Alternative Asset Fund Modeling. *The Journal of Portfolio Management 28(2), 90–100.* https://doi.org/10.3905/jpm.2002.319836

7.2 Published literature

Wooldridge, J., M. (2018). Introductory econometrics: a modern approach. 7th Edition, Cengage, USA

Meyer, T., Cornelius, C., Diller, C., & Guennoc , D. (2013). *Mastering Illiquidity: Risk Management for Portfolios of Limited Partnership Funds.* 1st Edition, John Wiley & Sons Inc, USA

7.3 Industry reports

Barath, A. (2022). Private vs Public Markets: Why investors are flocking to private equity in search of consistently higher returns. BITE Alternative Investments. Viewed March 2024: https://www.biteinvestments.com/wp-content/uploads/2022/05/Private-vs-Public-Markets-Why-investors-are-flocking-to-private-equity-in-search-of-consistently-higher-returns.pdf

7.4 Online resources and working papers

Buchner, A., Kaserer, C., & Wagner, N. (2006). Stochastic modeling of private equity: an equilibrium based approach to fund valuation. *Working Paper, Technische Universität München.* https://www.econstor.eu/bitstream/10419/48440/1/577822632.pdf

Forsell, J., & Furenstam, E. (2018). Cash Flow Simulation in Private Equity: evaluation and comparison of two models. *Master thesis, Umeå University.* https://umu.diva-portal.org/smash/get/diva2:1216322/FULLTEXT01.pdf

FRED. (2024). Total FED (US) assets less items from consolidation. Viewed April 2024: *https://fred.stlouisfed.org/series/WALCL*

Jeet, V. (2020). Building and maintaining a desired exposure to private markets. PGIM IAS-November. https://cdn.pficdn.com/cms/pgim4/sites/default/files/2020-11/IAS-Building-and-Maintaining-a-Desired-Exposure-to-Private-Markets-Nov2020.pdf

Jeet, V. (2020). Modeling Private Investment Cash Flows with Market-Sensitive Periodic Growth. *PGIM IAS-October. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3775008*

Jeet, V., & O'Shea L. (2018). Modeling Cash Flows for Private Capital Funds. *Working Paper, Burgiss Applied Research.* https://static1.squarespace.com/static/5a4e4beb51a5842676483cc5/t/5ae764022b6a28e5da74da67/1525113954551/Modeling+Cash+Flows+For+Private+Capital+Funds

Jensen, M., C. (1989). Eclipse of a public corporation. *Harvard business review* 67(5), *61*. https://doi.org/10.2139/ssrn.146149

Jäckel, C. (2023). *The Takahashi and Alexander model*. Viewed February 2024: https://www.christoph-jaeckel.com/posts/2023-02-05-the-takahashi-and-alexander-model/

Karatas, T., Klinkert, F., & Hirsa, A. (2021). Supervised Neural Networks for Illiquid Alternative Asset Cash Flow Forecasting. *IDEAS Working Paper Series from RePEc.* https://arxiv.org/pdf/2108.02853.pdf

Meads, C., Morandi, N., & Carnelli, A. (2016). Cash management strategies for private equity investors. Pantheon. Viewed April 2024: https://caia.org/sites/default/files/AIAR_Q1_2016_05_MeadsMorandiCamelli_CashManagementStrats.pdf

Metrick, A., Choi, W., W., & Yasuda, A. (2011). A Model of Private Equity Fund Compensation. *Working paper, National Bureau of Economic Research.* https://doi.org/10.3386/w17568

Shen, J., Teng, M. (2021). The rebalancing conundrum: private equity valuations and market dislocations. *PGIM IAS-December*. https://cdn.pficdn.com/cms1/pgim4/sites/default/files/IAS-The-Rebalancing-Conundrum-Private-Equity-Valuations-and-Market-Dislocations-122021-ADA.pdf

Statistia. (2021). *Value of private equity dry powder globally 2003-2023.* Viewed February 2024: https://www-statista-com.ez.hhs.se/statistics/513838/value-of-private-equity-dry-powder/

Tolkamp, C. (2007). Predicting Private Equity Performance: The development of a private equity performance-forecasting model for AEGON Asset Management. *Master's Thesis, University of Twente*. https://essay.utwente.nl/771/1/scriptie_Tolkamp.pdf

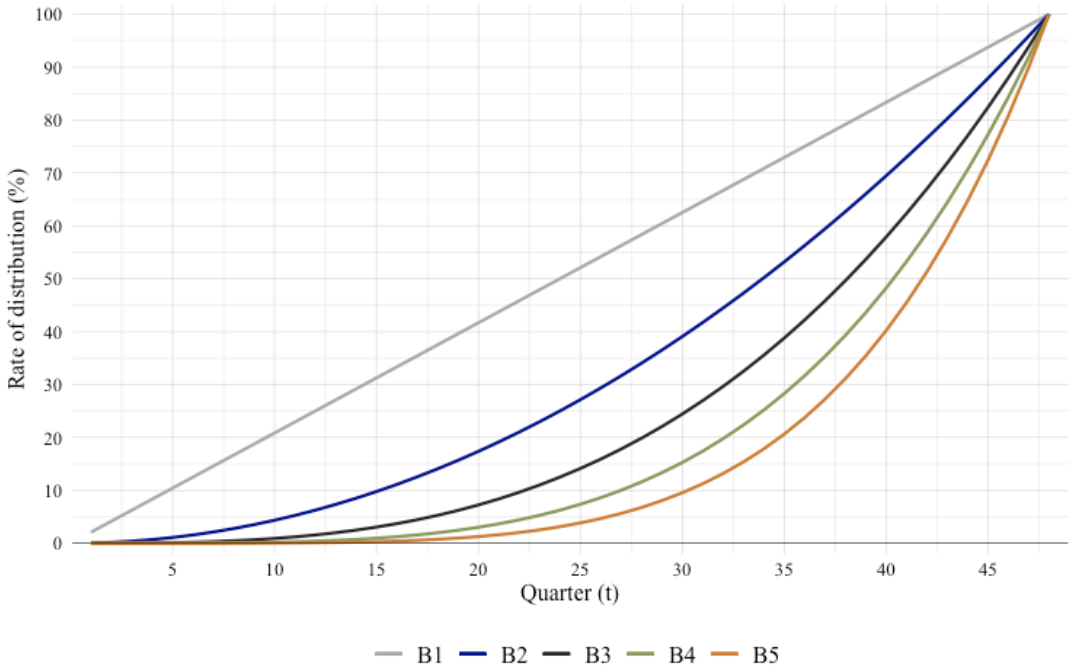
8. Appendix

Appendix A: Benchmark TA model plot for bow factor

Appendix A.

Benchmark TA model distribution rates for different bow factors

This graph plots the rate of distribution, using the sample model inputs from Table 1, but for different assumptions on the bow factor.



Appendix B: Data descriptives

Appendix B.

Summary of data descriptives

Summary of the PE data provided by Burgiss.

Vintage year	Number of funds	Capitalization (MUSD)[7]	Average IRR (%)[8]
1990	9	2,787	13.32
1991	7	1,906	28.25
1992	12	4,649	31.43
1993	14	6,149	23.45
1994	30	10,892	24.09
1995	31	18,466	11.23
1996	28	8,556	15.34
1997	50	39,861	8.23
1998	63	55,563	7.24
1999	50	43,360	9.32
2000	80	78,647	16.70
2001	46	42,244	27.03
2002	44	33,488	23.09
2003	44	34,433	21.77
2004	71	53,997	14.91
2005	96	113,473	9.99
2006	128	227,801	6.67
2007	137	216,689	9.62
2008	122	191,962	12.25
2009	43	46,730	13.44
2010	48	34,972	12.98
2011	84	114,199	14.12
2012	87	100,157	16.19
2013	80	115,585	14.07
2014	121	146,639	18.02
2015	105	127,111	18.09
2016	139	182,973	18.45
2017	116	163,069	19.65
2018	138	263,394	19.26
2019	166	321,623	18.49
2020	138	257,923	15.65
2021	192	436,530	11.36
2022	152	344,205	7.98
Total	**2,671**	**3,840,033**	

[7] Total capital commitment
[8] Represents the average IRR across all funds within each vintage

Appendix C: Robustness check for multicollinearity

Appendix C.

Variance Inflation Factors (VIF)

All values are below ten, indicating no multicollinearity between the explanatory variables.

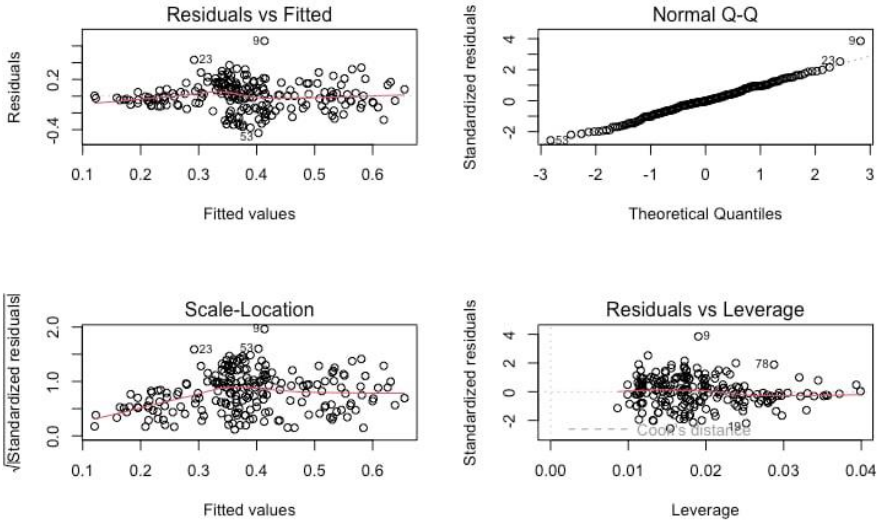
Explanatory variable	**Model 1 ($\widehat{RC}_t$)**	**Model 3 ($\widehat{G}_t$)**
Yield spread (t-1)	1.01	
Investment period 1	2.17	
Uncalled capital (t-1)	2.16	
Public return (t) (estimated)		1.41
High yield spread (t-1)		1.41

Appendix D: Robustness check of regression residuals

Appendix D.

Robustness check of regression residuals

Appendix D.1: Model 1 (CR)



Appendix D.2: Model 3 (G)

