

FINANCIAL MATHEMATICS TEAM CHALLENGE

A collection of the four reports from the 2018
Financial Mathematics Team Challenge.

This edition of the FMTC is dedicated to our close
friend and colleague, Prof Coenraad Labuschagne
(16-5-1958 to 10-7-2018)



AIFMRM

AFRICAN INSTITUTE OF FINANCIAL MARKETS AND RISK MANAGEMENT



Preamble

One of the key aims of the FMTC is for South African postgraduate students in Financial and Insurance Mathematics to have the opportunity to focus on a topical, industry-relevant research project, while simultaneously developing links with international students and academics in the field. An allied purpose is to bring a variety of international researchers to South Africa to give them a glimpse of the dynamic environment that is developing at UCT in the African Institute of Financial Markets and Risk Management. The primary goal, however, is for students to learn to work in diverse teams and to be exposed to a healthy dose of fair competition.

The Fifth Financial Mathematics Team Challenge was held from the 26th of June to the 6th of July 2018. The challenge brought together four teams of Masters and PhD students from France, Germany, China, Ireland, South Africa and the UK to pursue intensive research in Financial Mathematics. Each team worked on a distinct research problem over the twelve days. Professional and academic experts from Switzerland, South Africa, and the UK individually mentored the teams; fostering teamwork and providing guidance. As they have in the past, the students applied themselves with remarkable commitment and energy.

This years research included topical projects on (a) South African interest rate dynamics, (b) commitment scheduling for private equity investments, (c) portfolio optimisation under uncertainty, and on (d) the appropriateness of the LFMM model in South African interest rate markets. These were either proposed directly by our industry partners or chosen from areas of current relevance to the finance and insurance industry. In order to prepare the teams, guidance and preliminary reading was given to them a month before the meeting in Cape Town. During the final two days of the challenge, the teams presented their conclusions and solutions in extended seminar talks. The team whose research findings were adjudged to be the best was awarded a floating trophy. Each team wrote a report containing a critical analysis of their research problem and the results that they obtained. This volume contains these four reports, and will be available to future FMTC participants. It may also be of use and inspiration to Masters and PhD students in Financial and Insurance Mathematics.

FMTC V was a great success, so 2019 and FMTC VI is already in the pipeline! It is a great pleasure to see that the FMTC reaches new shores this year. The first edition of the FMTC Brazil takes place at the Fundação Getulio Vargas (FGV) in Rio de Janeiro, 8-18 August 2018.

David Taylor, University of Cape Town

Andrea Macrina, University College London & University of Cape Town

Contents

1. **A. Allan, A. Soane, R. Van Gysen and Y. Wang**
JIBAR Dynamics and Short Dated Caplets
2. **W. Black, B. Mbele, C. Van Jaarsveldt, J. Wiesel**
Commitment Scheduling for Private Equity Investments¹
3. **H. Brannelly, J. Nevin, D. Ngwenza, Y. Wang**
Portfolio Optimisation Under Uncertainty
4. **W. Cresswell, X. Kleisinger-Yu, A. De Savigny, C. Sterley**
An Assessment on the Appropriateness of the use of the LFMM in South Africa

¹Winning team of the fifth Financial Mathematics Team Challenge

JIBAR Dynamics and Short Dated Caplets

TEAM 1

ANDREW ALLAN, University of Oxford
ANDREW SOANE, University of Cape Town
RICHARD VAN GYSEN, University of Cape Town
YIYING WANG, University College London

Supervisor:
JAMES TAYLOR, Private

Contents

1	Introduction	3
1	Assertion	3
2	A review of JIBAR	4
3	Interest rate derivatives	6
2	Observed Irrationality	9
1	Exploratory data analysis	9
2	Theoretical JIBAR behaviour	10
3	Data Analysis	14
1	Methodology	14
1.1	Results	15
2	Time series analysis	16
2.1	JIBAR	17
2.2	Forward rate agreements	19
3	Conclusion	21
4	Option Pricing	23
1	Forward curve construction	23
1.1	Linear vs piecewise interpolation	23
1.2	Rational short-end construction	26
2	Gaussian model	29
3	Log-normal model	32
5	Conclusion and Further Research	36

Chapter 1

Introduction

1 Assertion

This paper reviews current interest rate setting practices and associated derivative pricing in the South African financial markets. It places a special focus on the behaviour of the 3-month Johannesburg Interbank Average Rate (JIBAR) in the lead up to monetary policy committee (MPC) meeting dates. On these dates, the repurchase (Repo) rate is regularly altered by a factor of $\pm 0.25\%$ or $\pm 0.5\%$. There is an intricate relationship between JIBAR and the Repo rate, which is illustrated in Figure 1.1 below. Graph (a) shows an instance where the forward market did not anticipate a jump in JIBAR, which is evident from the sudden change in both JIBAR and the forward rates immediately after the 19th July 2012 MPC meeting. By contrast, graph (b) shows an instance where the forward market anticipated a jump in JIBAR, as indicated by the forward rate jumping well before the MPC meeting on 17th July 2014.

The main assertion of this paper is that on dates when the market anticipates that a Repo rate change will occur, rational market participants should properly account for this future change in their NCD rates, and thus we should expect a more gradual change in JIBAR as the MPC date approaches. Note this assertion is based off a number of assumptions about the tradability of NCDs. These assumptions are not addressed in this paper. We hypothesise that JIBAR has systematic jumps around the MPC dates. This hypothesis will be tested in detail in later sections. The objective of this paper is to present a realistic pricing model for interest rate derivatives that incorporates these irrational jumps.

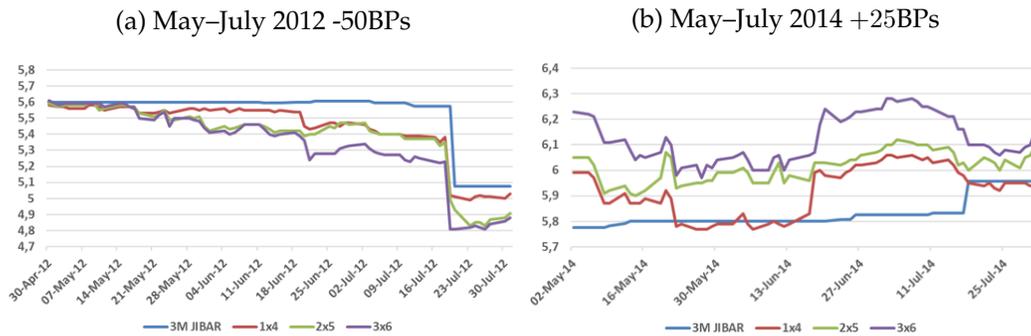


Figure 1.1

The introduction of this paper will analyse the construction of JIBAR and introduce interest rate derivatives in the South African context. Chapter 2 will present evidence of observed irrationality of JIBAR and a formal argument of how JIBAR and forward rates should act rationally. Chapter 3 will further analyse the time series of interest rates and perform a variety of statistical tests to prove our hypothesis. Chapter 4 will then introduce an alternative method of forward curve construction and compare it to the market's current convention. Chapter 5 will then implement an option pricing model that compensates for the jumps in JIBAR.

2 A review of JIBAR

JIBAR is the central focus as a benchmark rate in the South African money market. JIBAR simple yield rates are calculated by the Johannesburg Stock Exchange (JSE). It is calculated by taking the mid-rate of interbank Negotiable Certificates of Deposit (NCDs) randomly from the screens of eight different banks in South Africa. The JSE observes snapshots of the contributors' NCD bid and offer rates from a real-time trading screen randomly between 9:15am and 9:45am each day. The top two and bottom two mid-rates are discarded and an arithmetic mean of the remaining four are calculated, resulting in the JIBAR rate. The JSE reserve the right to use any snapshot figure at their discretion, observe the bid and offer rates, and calculate the mid-rate for each NCD maturity from each bank. NCDs with maturities of 1, 3, 6, 9 and 12-months are sampled to produce a term structure of JIBAR rates. This random rate sampling and mid-point averaging helps to achieve observations of real world trading conditions. See (South African Reserve Bank, 2007), (South African Reserve Bank, 2012) and (South African Reserve Bank, 2018) for further discussion.

The calculation and publishing of JIBAR rates are overseen by the South African Reserve Bank (SARB) which enforces a strict code of conduct. JIBAR underpins R3trn to R4trn worth of transactions in South Africa, with the bulk of these trades

being completed in the inter-bank interest rate derivative market. This motivates its importance as a reference rate. JIBAR is driven primarily by the Repo rate in South Africa which is set by the SARB (Figure 1.2). The Repo rate is the rate at which the SARB lends to banks. Banks have a choice to fund themselves in the overnight market at the Repo rate or through non-bank institutional funding with NCDs. Thus, if these two rates differ greatly, supply and demand forces will adjust the rates towards each other. The SARB sets the Repo rate every two months at monetary policy committee (MPC) meetings in order to control total debt levels and consequently money supply, in order to achieve inflation, GDP and exchange-rate targets.

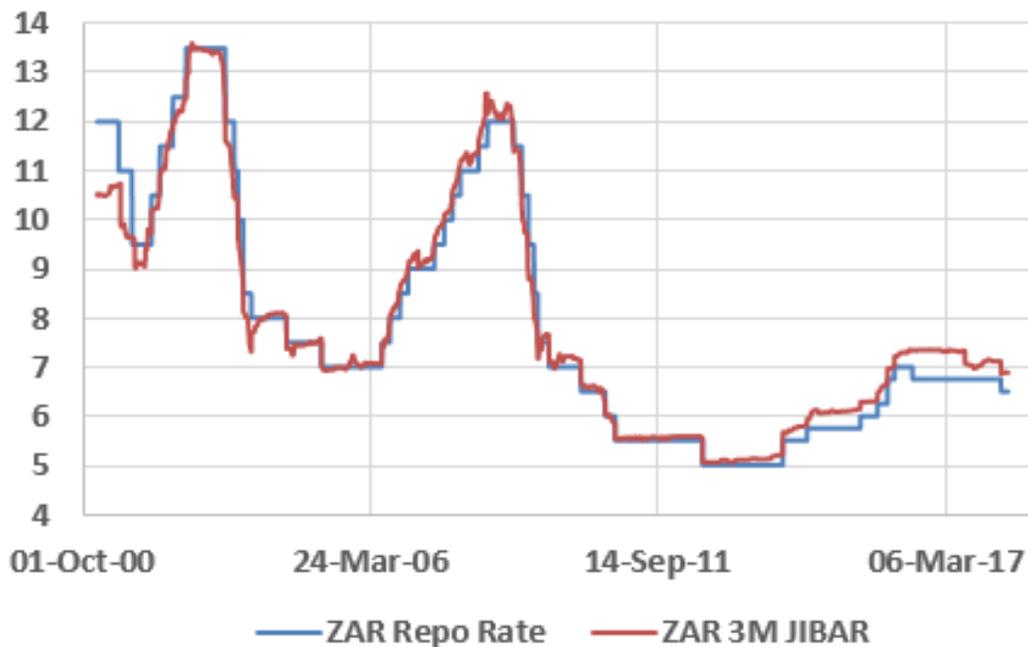


Figure 1.2: Repo VS JIBAR

An international comparative to JIBAR is the London Interbank Offered Rate (LIBOR) which is the most widely used reference rate in the world. 150 LIBOR rates of varying maturities and currencies are determined by the contributions from 15 banks, each estimating the cost of borrowing from other banks. The top and bottom 25% are discarded and the middle rates are then arithmetically averaged to produce the term structure of LIBOR rates.

Controversy has surrounded the LIBOR where manipulation by banks contributing to the LIBOR was clearly demonstrated, and heavy fines were imposed on the

convicted banks. These Banks were convicted of collusion and manipulation of LIBOR to profit from their existing interest rate derivative positions. Indeed, 20 banks were investigated in 2013 for previously manipulating LIBOR rates and were found guilty and punished with severe monetary fines (The Economist, 2013).

In 2012, the SARB reviewed the JIBAR calculation procedure and conducted an investigation into potential manipulation of JIBAR from 2003 to 2012. The report found there were outliers in the rates submitted by contributors for the calculation of JIBAR, but that these did not provide sufficient evidence of deliberate manipulation (South African Reserve Bank, 2012). Nevertheless, the report recommended that JIBAR switch to real-time screen rates instead of submitting estimated rates which are far more susceptible to manipulation via collusion. However, Jager and Parsons (2013) found evidence of the 3-month JIBAR spreads behaving similarly to the spreads observed in the LIBOR rate during its period of manipulation. Furthermore, the 1-month JIBAR was found to have abnormal movements that were not economically justifiable (Jager and Parsons, 2013). This warrants further investigation into the behaviour of JIBAR and its associated interest rate derivatives. This paper will only deal with certain anomalies identified in the next chapter.

3 Interest rate derivatives

JIBAR is referred to as a spot rate, since it is an interest rate which applies from today until some future date t . Let the capitalisation factor be $C(0, t) = (1 + r(0, t) \frac{t}{365})$ which accumulates one unit of currency from current time until time t at rate $r(0, t)$. Similarly, let the discount factor $Z(0, t) = \frac{1}{C(0, t)}$ be the value today of receiving one unit of currency at time t . There exists a relationship between spot rates and forward rates which apply for a future period from t until T where $t < T$. In order to avoid arbitrage, we require $C(0, t) \times C(t, T) = C(0, T)$ where $C(t, T) = (1 + F(0; t, T) \frac{T-t}{365})$ and $F(0; t, T)$ is the forward rate today applying over time period t to T (West, 2009). If this relationship does not hold, then one can profit at maturity with zero initial cost. The forward rate can be rewritten as
$$F(0; t, T) = \frac{365}{T-t} \left(\frac{C(0, T)}{C(0, t)} - 1 \right).$$

A Forward Rate Agreement (FRA) is a derivative instrument where the long party pays a pre-agreed fixed rate K and receives a floating rate rate from a counter party (short party) over a future period from time t to T (West, 2009). Notationally, a FRA that begins in n months time (near date) and maturity in m months time (far date) is referred to as a $n \times m$ FRA. The effective period is the period from time n to m where the underlying rate comes into effect and is deterministic over this period. Common FRAs in South Africa include 1×4 , 2×5 and 3×6 . The value of a long

FRA at any time $s \leq t < T$ can be represented as

$$N(F(s; t, T) - K) \frac{T-t}{365} Z(s, T),$$

per 1 unit of notional. Time t is referred to as the resettlement date, and the future rate and strike rate over period t to T are known at this date. Contracts are typically settled in advance at time t (West, 2009).

FRAs are commonly used to hedge interest rate risk. For example, if an entity wishes to borrow money at a future date and expects floating rates to rise in the near future, then they can enter a long FRA where the near and far dates match the dates of the loan and fix a rate K that will apply to their loan. It is noted that under the principle of no arbitrage the strike rate of FRA will be the forward rate associated with the FRA reset and maturity dates. It is concluded that the forward rates give an expectation of future JIBAR.

The figure below shows a series of 1x4 FRA rates and a series of JIBAR rates. As can be seen, the FRA rates give reasonable estimate for the future JIBAR rates.

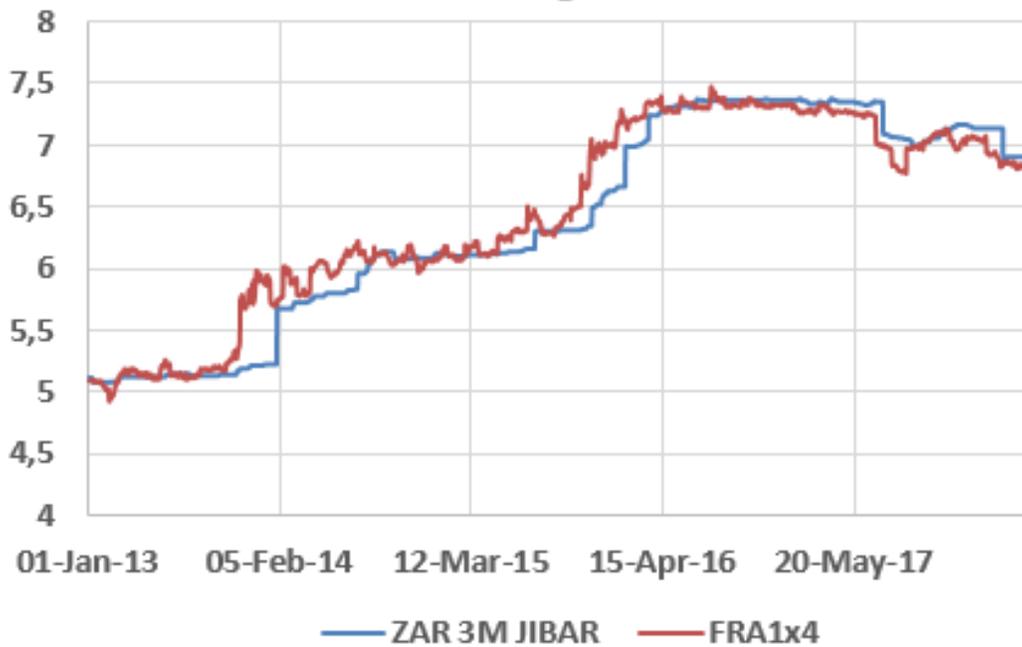


Figure 1.3: JIBAR VS FRA1x4

Caplets and floorlets are options with an interest rate as the underlying asset. A caplet is a call option which gives the holder (long party) the right but not the obligation to receive the underlying interest rate and pay the pre-agreed fixed rate (West, 2007). Similarly, a floorlet is a put option which gives the holder the right

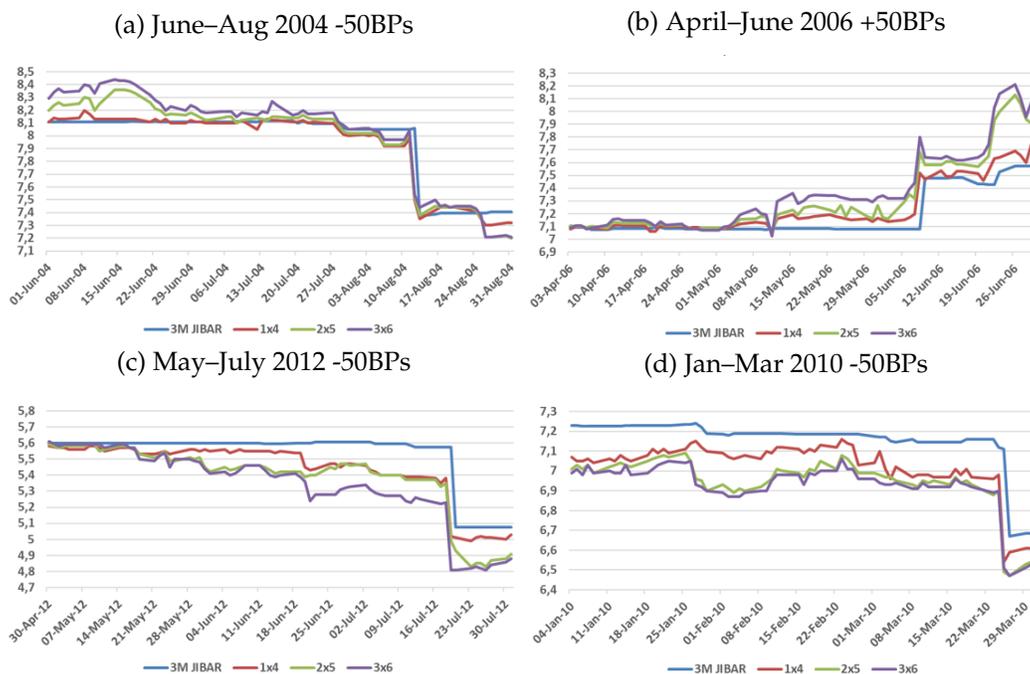
but not the obligation to pay the underlying interest rate and receive the pre-agreed fixed rate. The long party only has upside potential and pays a premium for this optionality. Caplets and floorlets are settled in advance and have payoffs determined at their respective reset date t . Black (1976) provide an analytic pricing framework for caplets and floorlets. A cap is a series of adjacent caplets and a floor is a series of adjacent floorlets.

Chapter 2

Observed Irrationality

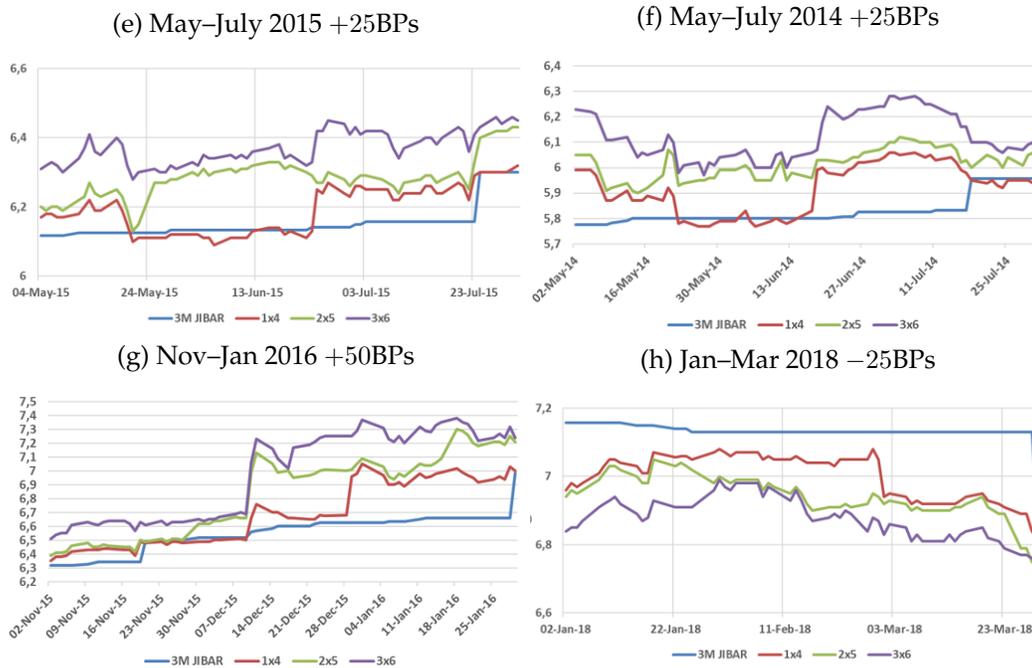
1 Exploratory data analysis

Data of the Repo rate, 3-month JIBAR and 1×4 , 2×5 and 3×6 FRA strike rates were reviewed from 2001 until May 2018. In particular, periods where the Repo rate changed were investigated to explore the relationship between the market rates and the Repo rate.



Graphs (a)–(d) provide visual evidence of the close relationship the FRA rates and JIBAR have when the Repo rate changes. These graphs include large jumps of the

rates when the Repo rate changed unexpectedly. These periods contained MPC dates where the market did not anticipate a Repo rate change. This is evident by the jumps in JIBAR and FRA rates on the MPC dates. Had the repo rate change been anticipated, the FRA rates would have jumped prior to the MPC date as they are an expectation of future JIBAR.



Graphs (e)–(h) show visual evidence of FRA rates moving when their effective periods include an expected Repo rate change, but the 3-month JIBAR rate does not change when it overlaps the same MPC date. This is indicated by the relatively large gap between the market FRA rates and JIBAR until the MPC announcement date. The JIBAR rates appear to only jump according to the Repo rate change on the MPC date. This implies that the market anticipates a repo rate change but JIBAR itself is not accounting for the jumps prior to the date of the jumps. In (g), the rate hike is anticipated nearly 2 months in advance by market FRA rates but JIBAR shows almost no change until the actual announcement. We explore the trades below that could exploit this difference .

2 Theoretical JIBAR behaviour

In a liquid arbitrage-free market one expects that the rates of JIBAR and short-term FRAs should not differ by too great a margin, since their associated contract pe-

riods have a significant overlap. However, in the previous section we identified some irregular behaviour in JIBAR in the lead up to MPC dates. That is, the FRA rates frequently anticipate a change in the Repo rate, and hence in JIBAR, but JIBAR itself does not account for this change before it actually occurs, which causes a discrepancy between these two rates. To see why this may be considered irrational, let us consider the following investment strategy for a bank.

Consider two time periods, $T, \Delta t > 0$, measured in number of days, with $T \leq \Delta t$, such that an MPC date occurs between the times T and Δt . To make this more concrete one may think of T and Δt as being 1 month and 3 months respectively. Suppose that at an initial time (let's say time 0) we observe the Δt -JIBAR rate $J(0, \Delta t)$ and the $T \times (T + \Delta t)$ -FRA rate $F(0, T, T + \Delta t)$.

At time 0 we sell a Negotiable Certificate of Deposit (NCD) with face value N (ZAR), with interest paid at the time of maturity Δt . The interest rate is given by $J(0, \Delta t)$ (indeed, the rates submitted to the JSE for JIBAR calculation are principally based on NCDs). We then immediately invest the full amount received in return for the NCD repeatedly in overnight repurchase agreements (Repo) for the duration T . Denote the initial Repo rate by r_0 , so that at time T the return on our investment is given by $N(1 + r_0 \frac{1}{365})^T$.

When the NCD matures (at time Δt) we pay back $N(1 + J(0, \Delta t) \frac{\Delta t}{365})$. Considering the time value of money, this value corresponds to

$$N \left(1 + J(0, \Delta t) \frac{\Delta t}{365} \right) \left(1 + r_{\Delta t} \frac{1}{365} \right)^T$$

at time $T + \Delta t$, where $r_{\Delta t}$ denotes the Repo rate at time Δt , which may be different to r_0 since an MPC date occurs before time Δt .

Also at the initial time, we enter into a short FRA with notional value $N(1 + r_0 \frac{1}{365})^T$, contract period between the times T and $T + \Delta t$, and with the fixed rate $F(0, T, T + \Delta t)$ while simultaneously investing N at the repo rate until time T , at which point we immediately reinvest these earnings in an NCD from T to Δt with face value $N(1 + r_0 \frac{1}{365})^T$. By doing so we effectively lock in an interest rate of $F(0, T, T + \Delta t)$ over the contract period, so that the return at the final time $T + \Delta t$ is given by

$$N \left(1 + r_0 \frac{1}{365} \right)^T \left(1 + F(0, T, T + \Delta t) \frac{\Delta t}{365} \right).$$

In summary, having started with zero capital, our final portfolio is worth

$$N \left(1 + r_0 \frac{1}{365} \right)^T \left(1 + F(0, T, T + \Delta t) \frac{\Delta t}{365} \right) - N \left(1 + J(0, \Delta t) \frac{\Delta t}{365} \right) \left(1 + r_{\Delta t} \frac{1}{365} \right)^T.$$

Absence of arbitrage then implies that this difference should be equal to zero, so that

$$1 + J(0, \Delta t) \frac{\Delta t}{365} = \left(1 + F(0, T, T + \Delta t) \frac{\Delta t}{365} \right) \frac{\left(1 + r_0 \frac{1}{365} \right)^T}{\left(1 + r_{\Delta t} \frac{1}{365} \right)^T}. \quad (2.1)$$

Of course, the rate $r_{\Delta t}$ is typically unknown at the initial time since an MPC meeting is scheduled to occur before the terminal time, which may result in a change to the Repo rate. However, even in this case the Repo rate will typically only change by 25 or 50 basis points, and such changes are frequently anticipated by the market. As FRAs are liquidly traded, it is reasonable to assume that the FRA rate is correct. Thus, we infer from (2.1) bounds on sensible values of the JIBAR rate $J(0, \Delta t)$.

On the other hand, one can also use (2.1) to provide a rational forecast for JIBAR, by considering the function $T \mapsto F(0, T, T + \Delta t)$, which is naturally interpreted as providing the expected future values of the Δt -JIBAR. This function is shown in Figure 2.2 for the case where $J(0, \Delta t) = 7\%$ with $\Delta t = 3$ months and an MPC date occurs after 50 days.

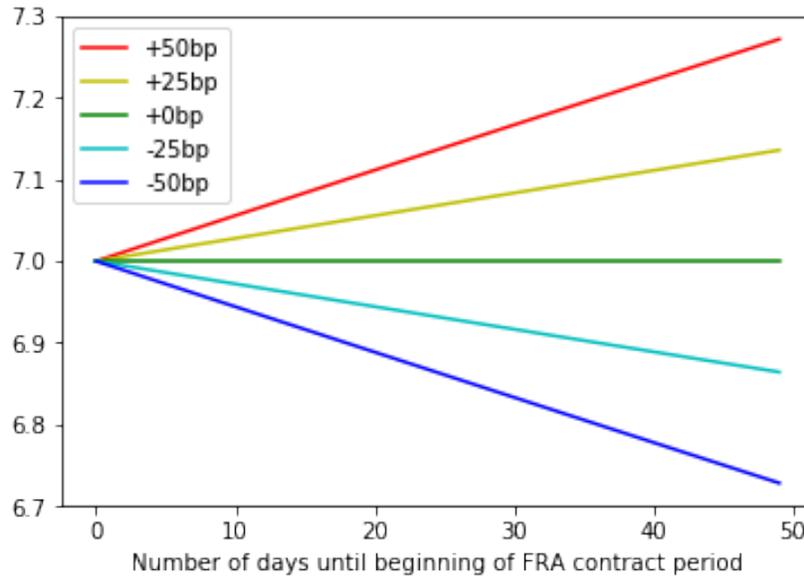


Figure 2.2: Rational forecast for JIBAR in the lead up to an MPC date

We conclude that in anticipation of a change in the Repo rate, JIBAR and forecasts of JIBAR should move gradually in the lead up to this change. In particular, in cases where such a change in the Repo rate is correctly predicted by the market, JIBAR should not jump significantly on MPC dates. As we saw in the previous section, this is consistently in stark contrast with observed behaviour.

Chapter 3

Data Analysis

The aim of this project is to test the hypothesis that there is a systematic jump in JIBAR rates, and hence forward rates, around the monetary policy committee (MPC) meeting dates. It is speculated in previous sections that the forward rate curve underlying JIBAR movements has discontinuities on the short end. As was shown in the previous section, knowing a future change in Repo rates should cause rational market participants to react slowly to the change, and so JIBAR rates should continuously approach the expected change.

1 Methodology

The alternate hypothesis tested below is that market participants are not reacting rationally to expectations of future JIBAR movements. The aim of this section is to use naive statistical tests to determine if there is significant evidence of biases in these rates. The argument is formalised as follows:

Null hypothesis (H_0): There are no systematic jumps in JIBAR movements around the MPC meeting dates.

Alternate hypothesis (H_a): JIBAR rates experience discontinuous jumps at the time of anticipated MPC announcements.

A time series of JIBAR, 1×4 FRA, 2×5 FRA and 3×6 FRA rates from 02/01/2001–17/05/2018 are used for testing. The methods used to test these hypotheses are summarised as follows:

- First, testing the significance of just a jump variable, jumping on MPC announcement dates, in explaining all the variability in the rates. That is, writing R for one of the considered rates, does the existence of the jump term in

$$R_t - R_{t-1} = a + b^+ \mathbb{I}_{t \in \{\text{MPC dates}\} \cap \{\text{The Jump was positive}\}} + b^- \mathbb{I}_{t \in \{\text{MPC dates}\} \cap \{\text{The Jump was negative}\}},$$

provide a significant explanation for the variability in the rate?

- Second, testing whether there is a significant difference in the mean jump sizes on days with an MPC announcement and days that do not. In other words, does the mean jump size on MPC dates differ from the mean jump size on other dates?

$$\mathbb{E}[R_t - R_{t-1} | \mathbb{I}_{t \in \{\text{MPC dates}\}}] \neq \mathbb{E}[R_t - R_{t-1} | \mathbb{I}_{t \notin \{\text{MPC dates}\}}].$$

1.1 Results

The following figures show the results from the first hypothesis test. The following is a snapshot of the output from R:

(a) Jump Significance in JIBAR Rates

```
Call:
lm(formula = dJibar3m ~ 0 + JibarJumpminus + JibarJumpplus)

Residuals:
    Min     1Q   Median     3Q      Max
-0.860  0.000  0.000  0.000  0.468

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
JibarJumpminus -0.32438    0.01396  -23.23  <2e-16 ***
JibarJumpplus  0.13046    0.01095   11.91  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03949 on 4386 degrees of freedom
Multiple R-squared:  0.1345, Adjusted R-squared:  0.1341
F-statistic: 340.8 on 2 and 4386 DF, p-value: < 2.2e-16
```

(b) Jump Significance in FRA1x4 Rates

```
Call:
lm(formula = dFRA1x4all ~ 0 + FRA1x4Jumpminus + FRA1x4Jumpplus)

Residuals:
    Min     1Q   Median     3Q      Max
 -0.91  -0.01  0.00  0.01  1.00

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
FRA1x4Jumpminus 0.002571    0.017413  0.148  0.88261
FRA1x4Jumpplus  0.090000    0.024626  3.655  0.00026 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06515 on 4492 degrees of freedom
Multiple R-squared:  0.002969, Adjusted R-squared:  0.002525
F-statistic: 6.689 on 2 and 4492 DF, p-value: 0.001257
```

(c) Jump Significance in FRA2x5 Rates

```
Call:
lm(formula = dFRA2x5all ~ 0 + FRA2x5Jumpminus + FRA2x5Jumpplus)

Residuals:
    Min     1Q   Median     3Q      Max
 -0.88  -0.02  0.00  0.02  1.15

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
FRA2x5Jumpminus -0.06013    0.02496  -2.409  0.016048 *
FRA2x5Jumpplus  0.06692    0.01958   3.418  0.000637 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0706 on 4484 degrees of freedom
Multiple R-squared:  0.003884, Adjusted R-squared:  0.003439
F-statistic: 8.741 on 2 and 4484 DF, p-value: 0.0001626
```

(d) Jump Significance in FRA3x6 Rates

```
Call:
lm(formula = dFRA3x6all ~ 0 + FRA3x6Jumpminus + FRA3x6Jumpplus)

Residuals:
    Min     1Q   Median     3Q      Max
 -1.05  -0.02  0.00  0.02  1.85

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
FRA3x6Jumpminus -0.01667    0.02612  -0.638  0.523
FRA3x6Jumpplus  -0.02592    0.02262  -1.146  0.252
---
Residual standard error: 0.07835 on 4491 degrees of freedom
Multiple R-squared:  0.0003829, Adjusted R-squared: -6.223e-05
F-statistic: 0.8602 on 2 and 4491 DF, p-value: 0.4231
```

Figures (a)-(d) show summaries of a linear regression calculated for the model testing the first hypothesis, for the four reference rates. The first observation from these results is the diminishing significance of the jump term going from JIBAR to 1×4 , 2×5 and 3×6 FRA rates. This is in line with our main assertion, as we expect the

discontinuities to be more prevalent at the shorter end of the yield curve. It is, however, surprising that neither the 2×5 nor 3×6 FRA rates are significantly explained by a single jump term. This result suggests that expected Repo rate changes have a slower reaction the further along the yield curve they are observed.

The second hypothesis test, examines whether there is a significant difference in the mean jump size in the four reference rates around MPC and non-MPC dates. The following is a snapshot of the output from R:

<p>(a) Jump Differences in JIBAR</p> <pre> Welch Two Sample t-test data: dJibarjump and dJibartot t = 4.6827, df = 20.007, p-value = 0.0001429 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: 0.1082515 0.2821586 sample estimates: mean of x mean of y 0.204333333 0.009128275 </pre>	<p>(b) Jump Differences in FRA1x4</p> <pre> Welch Two Sample t-test data: dFRA1x4 and dFRA1x4tot t = 3.3254, df = 20.127, p-value = 0.003352 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: 0.02980385 0.13001084 sample estimates: mean of x mean of y 0.13319048 0.05328313 </pre>
<p>(c) Jump Differences in FRA2x5</p> <pre> Welch Two Sample t-test data: dFRA2x5 and dFRA2x5tot t = 2.2472, df = 20.035, p-value = 0.03605 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: 0.008253945 0.221494811 sample estimates: mean of x mean of y 0.1771905 0.0623161 </pre>	<p>(d) Jump Differences in FRA3x6</p> <pre> Welch Two Sample t-test data: dFRA3x6 and dFRA3x6tot t = 1.6536, df = 20.026, p-value = 0.1138 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.03027545 0.26192105 sample estimates: mean of x mean of y 0.18757143 0.07174863 </pre>

These tests conclude (at the 5% significance level) that all but the 3×6 FRA rates show significant jump sizes around the MPC dates compared with the jump sizes on other dates. This naive test gives further evidence of a systematic jump in market rates around the MPC dates. Further, more rigorous tests are needed to finally conclude that these jumps are present.

2 Time series analysis

In this section we demonstrate that a discrete jump term on an MPC date is necessary when modelling the spot rate (JIBAR) and forward rates (FRA) by using time series analysis. The jump term exists due to both expected and unexpected Repo rate movements on MPC dates.

To generate time series models for JIBAR and the various FRA rates, we use data from January 2001 through to May 2018 obtained from the South African Reserve Bank. We concentrate on autoregressive integrated moving average (ARIMA) models and afterwards try to improve them with the addition of a jump term to each model.

2.1 JIBAR

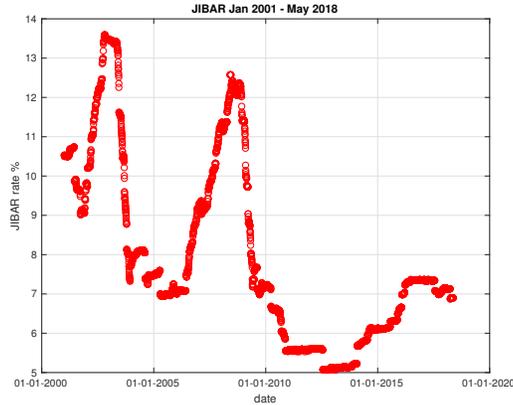
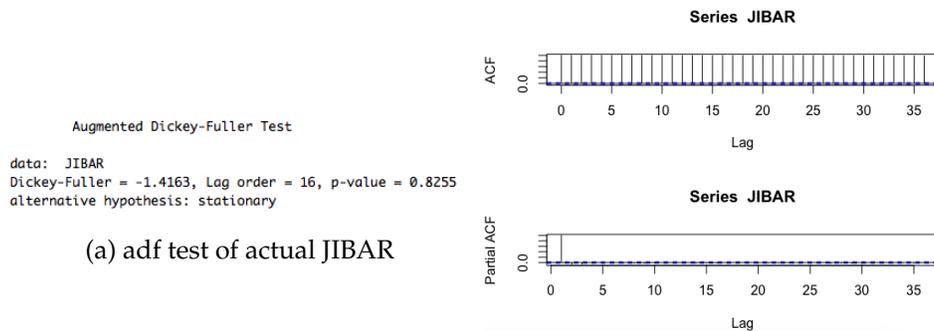


Figure 3.2: Actual JIBAR Jan 2001 – May 2018

JIBAR over a 17.5 year horizon is plotted in Figure 3.2. A visual inspection of the series doesn't show a deterministic, upward or downward pattern, nor any seasonality effect, consistent with the unpredictability of the monetary market. According to the augmented Dickey Fuller test, the series appears to be non-stationary and changes are observed discretely, which can also be illustrated by the autocorrelation function and the partial autocorrelation function. Thus, we infer that it is reasonable to take a difference of the time series. Furthermore, the series may experience co-integration, but the rates do not appear to indicate any mean-reversion.



(a) adf test of actual JIBAR

(b) acf, pacf of actual JIBAR

ARIMA model for JIBAR

The information of the fitted ARIMA model is shown in Figure 3.4.

```

Series: JIBAR
ARIMA(3,1,1)

Coefficients:
      ar1      ar2      ar3      ma1
      1.0550  -0.0228  -0.0410  -0.9657
s.e.  0.0163  0.0216  0.0153  0.0065

sigma^2 estimated as 0.001665:  log likelihood=8053.25
AIC=-16096.5  AICc=-16096.49  BIC=-16064.42

```

Figure 3.4: ARIMA model for JIBAR

The points shown in Figure 3.5 give a good approximation of the JIBAR series over periods of low volatility. However, the model is not well fitted given the outliers shown. By inspection, one can easily notice that there is a much larger jump than the actual curve on the date of a repo rate change. Thus, it would make more sense to add a jump term into the model.

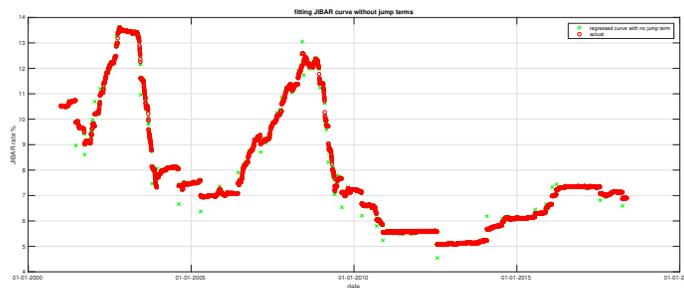
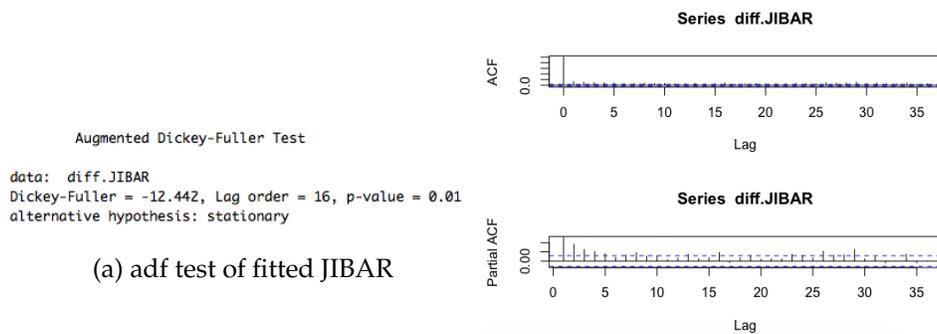


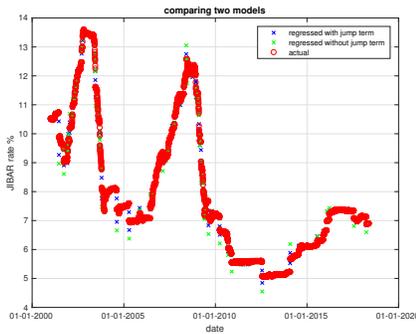
Figure 3.5: Fitted ARIMA curve for JIBAR



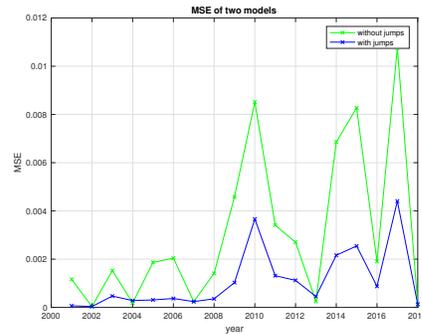
(b) acf, pacf of JIBAR after first differencing

Model improvement

The observations above provide reasonable evidence to include a jump term to the fitted ARIMA model. We take the average of the actual JIBAR jump sizes on MPC dates as the parameter of the indicator function term. As can be seen in graph (a) below, though not perfectly fitted, the mean squared error of each model shows that the model with the jump term (blue crosses) performs better than that without jumps (green crosses), and gives a more precise approximation of the actual JIBAR.



(a) Model Comparison

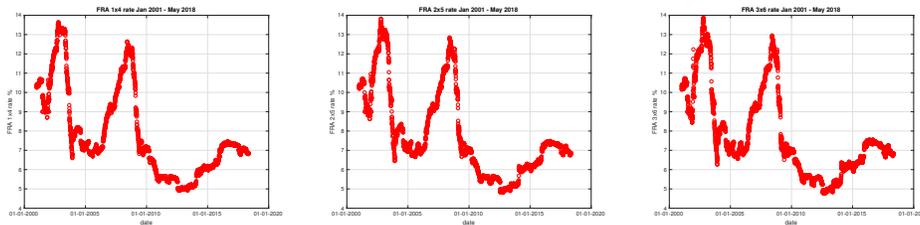


(b) MSE of the two models

The better fit of the ARIMA-with-jumps against that of the original ARIMA model suggests JIBAR is not in fact continuous but rather involves a discrete jump on an MPC date.

2.2 Forward rate agreements

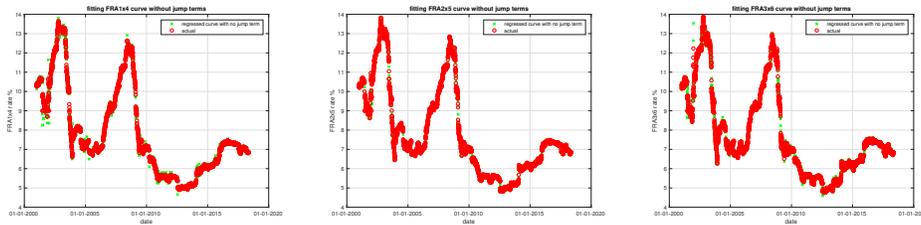
Since the FRA rates are the expectation of future JIBAR rates, which depend on the Repo rates, different term FRAs with same contract period duration have similar moving tracks, as can be seen in the figures below. Thus we can study these three curves in the same way.



ARIMA model for FRAs

Series: FRA1x4 ARIMA(1,1,2)	Series: FRA2x5 ARIMA(3,1,1)	Series: FRA3x6 ARIMA(5,1,5)
Coefficients: ar1 ma1 ma2 0.9896 -1.0882 0.1127 s.e. 0.0038 0.0154 0.0149	Coefficients: ar1 ar2 ar3 ma1 -0.4919 0.8502 0.0829 0.4459 s.e. 0.1538 0.0179 0.0160 0.1538	Coefficients: ar1 ar2 ar3 ar4 ar5 ma1 ma2 ma3 ma4 ma5 0.2363 -0.1782 -0.6223 0.3963 -0.5815 -0.3171 0.2003 0.6046 -0.4122 0.4746 s.e. 0.2109 0.1490 0.1811 0.1868 0.2050 0.2220 0.1449 0.0876 0.1355 0.0989
sigma^2 estimated as 0.004122: log likelihood=6002.37 AIC=-11996.75 AICc=-11996.74 BIC=-11971.08	sigma^2 estimated as 0.004827: log likelihood=5645.97 AIC=-11281.94 AICc=-11281.93 BIC=-11249.86	sigma^2 estimated as 0.005937: log likelihood=5180.75 AIC=-10339.3 AICc=-10339.44 BIC=-10268.91

As was expected, the ARIMA model for the 1×4 , 2×5 and 3×6 FRAs behave well in estimating the actual FRA rates when the actual rates move smoothly as the time passes. Nevertheless, as we previously observed for JIBAR, larger-sized jumps tend to occur on the MPC dates.



Model improvement

So as to minimize the gap between the estimated and actual FRA rates, an upward/downward jump term is added to the calculation of the rates on the MPC dates, and the mean of upward/downward jump size is quoted as the parameter of the indicator term.

Figure 3.8: comparing the two models for FRA 1x4

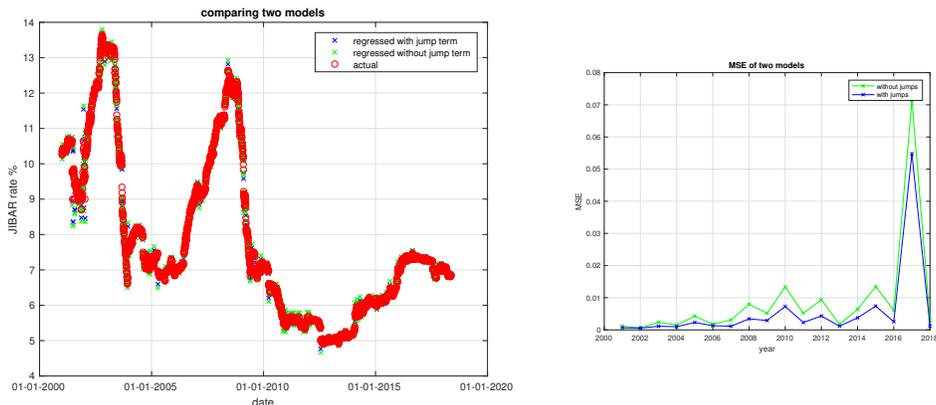


Figure 3.9: comparing the two models for FRA 2x5

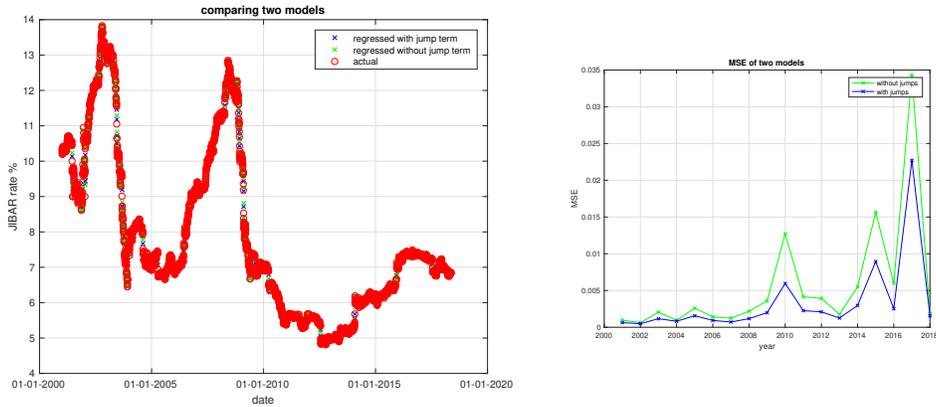
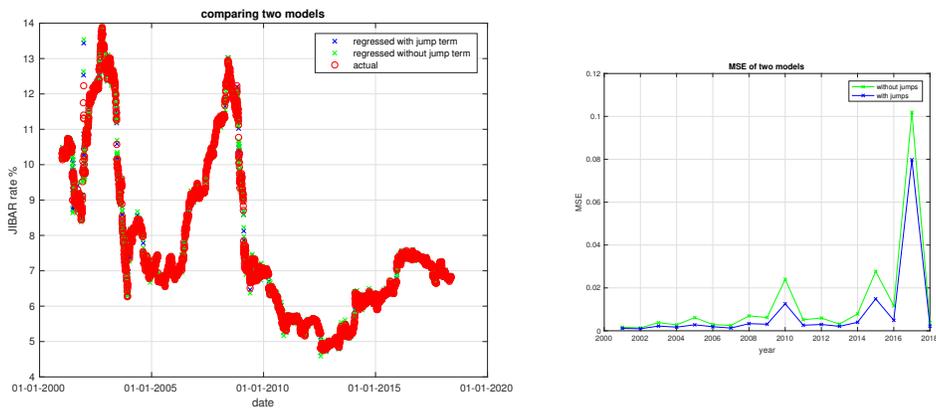


Figure 3.10: comparing the two models for FRA 3x6



Again, although not perfectly fitted, the MSE becomes smaller after the application of the jump terms.

3 Conclusion

Despite the fact that JIBAR and FRA rates seem to progress smoothly for the majority of the last 17 years, they can be very unpredictable due to their original composition. In forming and executing the proper monetary policy, the SARB inspects and readjusts the Repo rate, which is the determinant of JIBAR, on MPC dates, which typically occur every two months.

The naive tests see the diminishing significance of the jump term as the starting date of the effective period increases, suggesting that expected Repo rate changes have a slower reaction the further along the yield curve they are expected. Moreover, a structural jump appears to exist in market rates around the MPC dates.

The time series analysis provides evidence that including a discrete jump term on an MPC date is necessary when modelling the spot rate and forward rates. By fitting the historical rates, we come to the conclusion that the jump sizes on each MPC date are observed to be much larger than on usual days. Therefore, inclusion of a jump term in the original model gives a more precise approximation.

Chapter 4

Option Pricing

1 Forward curve construction

Having now provided sufficient evidence to prove the existence of irrational jumps in JIBAR, the next step is to propose a new pricing framework to incorporate these jumps. The following section will present an extension of the Black Pricing formula to incorporate the jumps, noting that jump times are known while jump sizes are not. In order to price short-dated interest rate derivatives, an adaptation of the conventional forward curve construction techniques should be developed to incorporate the observed inconsistencies in JIBAR.

In this section, a simple adaptation of the linear interpolation short-end curve construction technique is applied. Additionally, a rudimentary version of a rational short-end curve construction based off previous arguments on the natural progression of JIBAR is implemented.

1.1 Linear vs piecewise interpolation

Forward rate curves and hence zero rate curves are traditionally constructed at the short-end using linear interpolation. We propose the use of a piecewise linear interpolation method that interpolates between forward rates, between MPC dates. This artificially builds a jump in the forward rates on MPC dates.

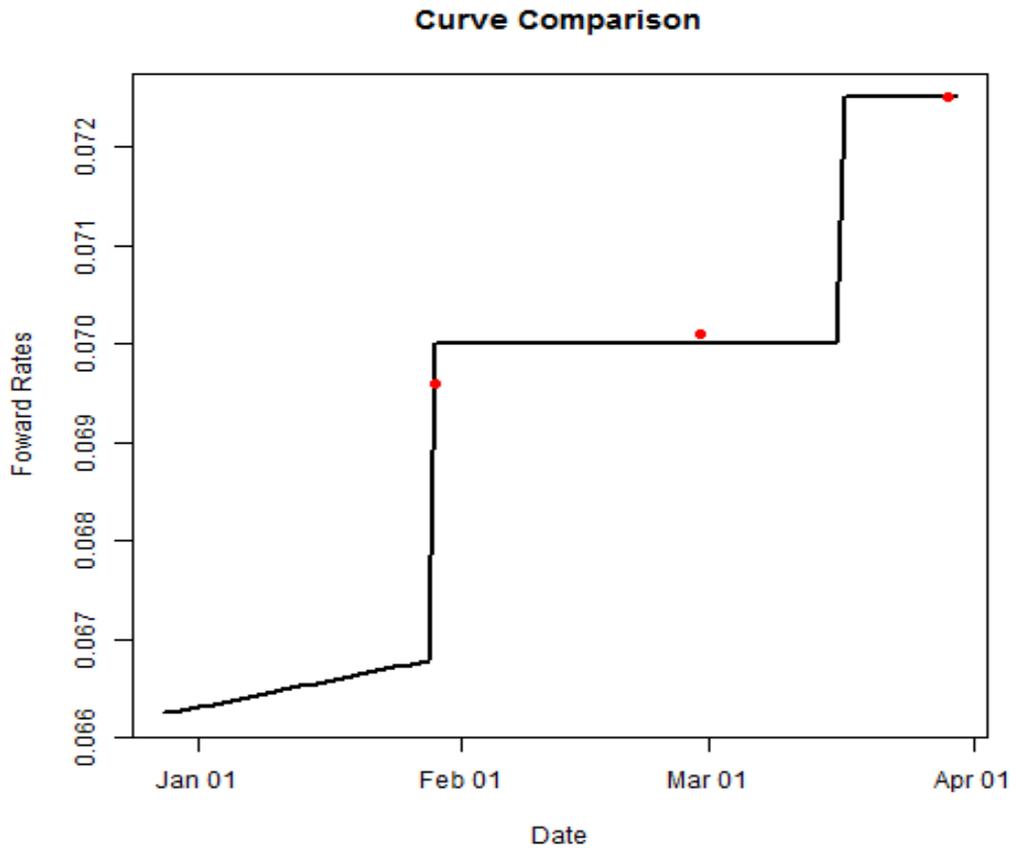


Figure 4.1: Piecewise linear interpolation on 28th December 2015

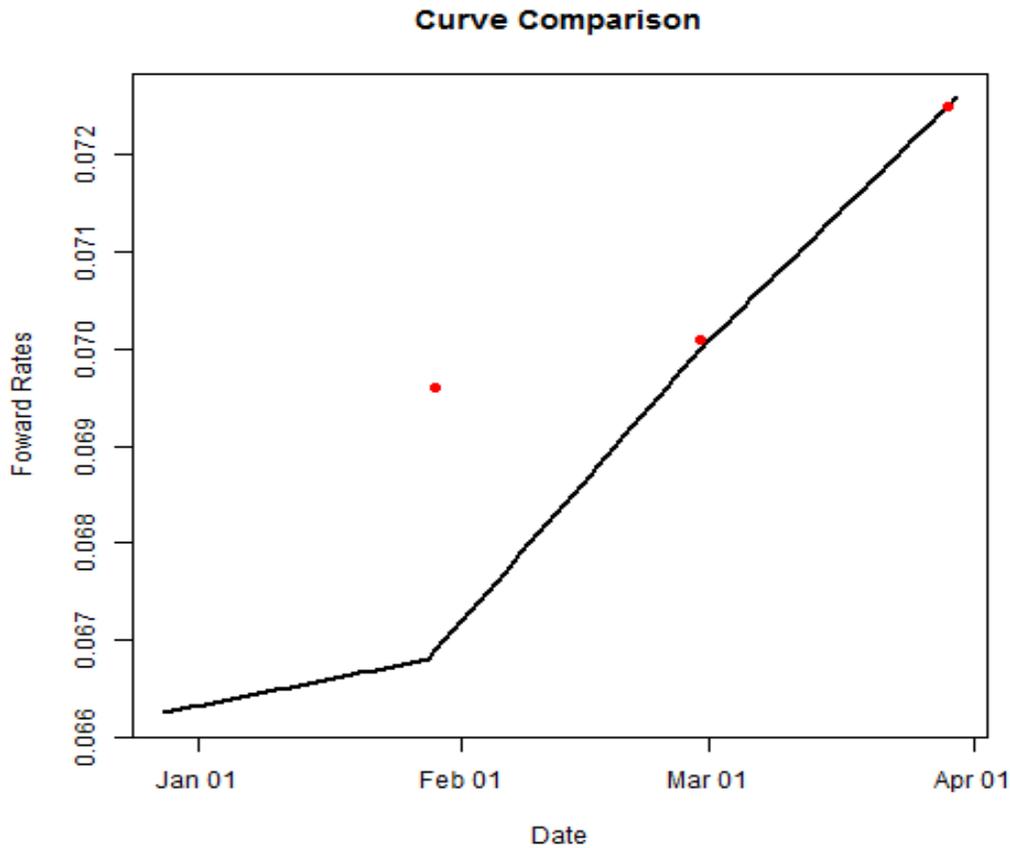


Figure 4.2: Linear interpolation on 28th December 2015

Figure 4.1 shows an implementation of the piecewise linear interpolation alongside the standard linear interpolation in Figure 4.2. These curves represent an approximation of the underlying forward rate curve on 28th December 2015. Imposed on the curves are the actual market 1×4 , 2×5 and 3×6 forward rates on the 29th December 2015, this is to gauge the accuracy of the curves at predicting future forward rates. A change in the Repo rate occurs on 28th of January 2016 so we therefore expect (and observe) that the 1×4 forward rate will have a jump from 28th December to 29th December. By construction, the piecewise linear curve incorporates this jump, whereas the standard linear interpolation does not. We propose that further tests should be conducted to test the accuracy of these two curve construction techniques using the next days observed rates as a proxy for the actual underlying forward rate curve. It is noted that one can only observe the benefit of the piecewise interpolation once a month when the effective date of the forward rates passes over the MPC dates.

1.2 Rational short-end construction

In accordance with Chapter 2, Section 2, we will now build a rudimentary short-end curve construction technique given only observed JIBAR, 1×4 FRA, 2×5 FRA and 3×6 FRA rates. It is presented as follows:

Let T_0 be the date at which the market rates are observed, and let T_i be the date i months from T_0 , for $i = \{1, 2, 3, 4, 5, 6\}$. Moreover, let $\{t_1, t_2, t_3\}$ be the MPC meeting dates in the next six months. It is noted that there are at most three MPC dates in the next six months and as few as no dates.

The rational curve construction technique is based off of a term structure of implied continuously compounded rates $\{r_1, r_2, r_3, r_4\}$ such that:

$$\begin{aligned}
 & 1 + \text{JIBAR} \times \frac{T_3 - T_0}{365} = \\
 & \exp\left(r_1 \frac{t_1 - T_0}{365}\right) \exp\left(r_2 \frac{\min(T_3, t_2) - t_1}{365}\right) \exp\left(r_3 \max\left(\frac{T_3 - t_2}{365}, 0\right)\right) \\
 & 1 + \text{FRA}_{1 \times 4} \times \frac{T_4 - T_1}{365} = \\
 & \exp\left(r_1 \max\left(\frac{t_1 - T_1}{365}, 0\right)\right) \exp\left(r_2 \frac{t_2 - \max(t_1, T_1)}{365}\right) \exp\left(r_3 \frac{T_4 - t_2}{365}\right) \\
 & 1 + \text{FRA}_{2 \times 5} \times \frac{T_5 - T_2}{365} = \\
 & \exp\left(r_2 \frac{t_2 - T_2}{365}\right) \exp\left(r_3 \frac{\min(t_3, T_5) - t_2}{365}\right) \exp\left(r_4 \max\left(\frac{T_5 - t_3}{365}, 0\right)\right) \\
 & 1 + \text{FRA}_{3 \times 6} \times \frac{T_6 - T_3}{365} = \\
 & \exp\left(r_2 \max\left(\frac{t_2 - T_3}{365}, 0\right)\right) \exp\left(r_3 \frac{t_3 - \max(t_2, T_3)}{365}\right) \exp\left(r_4 \frac{T_6 - t_3}{365}\right)
 \end{aligned}$$

Essentially, this algorithm is solving for a term structure of continuously compounded rates that imply the four observed three month market rates. Where the term structure has discontinuities at MPC meeting dates. Once this term structure is found, forward rates applicable on days between T_0 and T_3 can be approximated as an accumulation using the term structure between MPC meeting dates. Figure 4.3 shows the solution to the term structure when three MPC dates are occurring in the next six months.

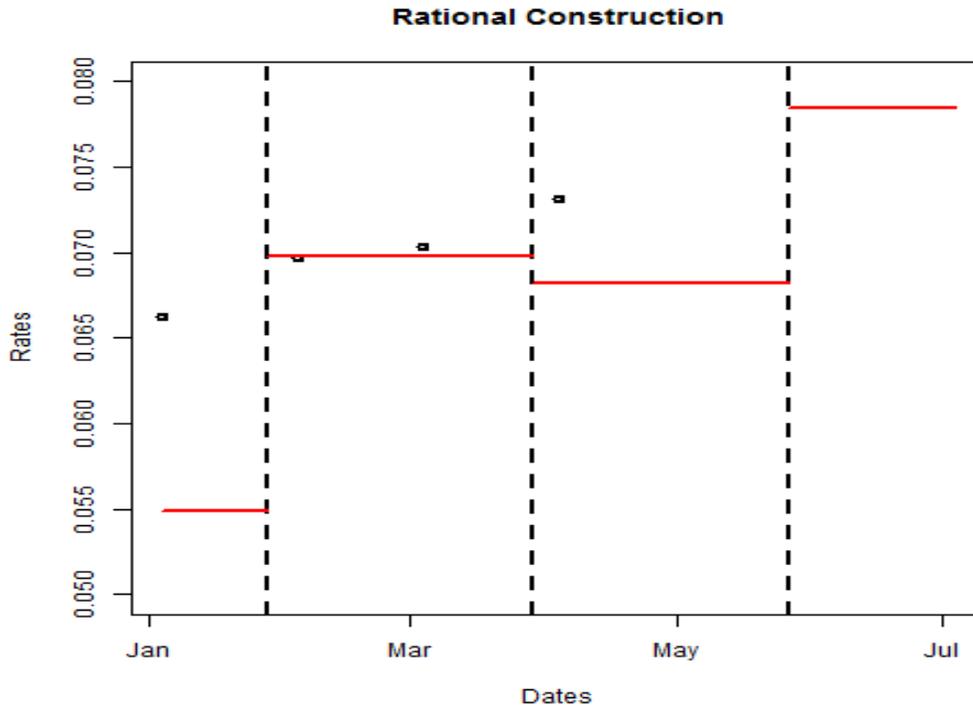


Figure 4.3: Term Structure of continuously compounded rates

It is noted that the system of equations required to solve for the term structure of continuously compounded rates is fully determined when there are three MPC meeting dates in the next three months. The system is not fully determined when there are less than three meeting dates. This then requires an optimisation with respect to some cost function to solve for the four required rates. This is left as an area of further research and implementation.

Figure 4.4 shows an implementation of the rational curve construction technique on the following date with the associated market rates:

Table 4.1: Sample of dates used in the illustration of the rational curve construction.

Start Date	MPC Date 1	MPC Date 2	MPC Date 2
2016-01-04	2016-01-28	2016-03-29	2016-05-27
JIBAR	FRA1x4	FRA2x5	FRA3x6
6.625%	6.97%	7.03%	7.31%

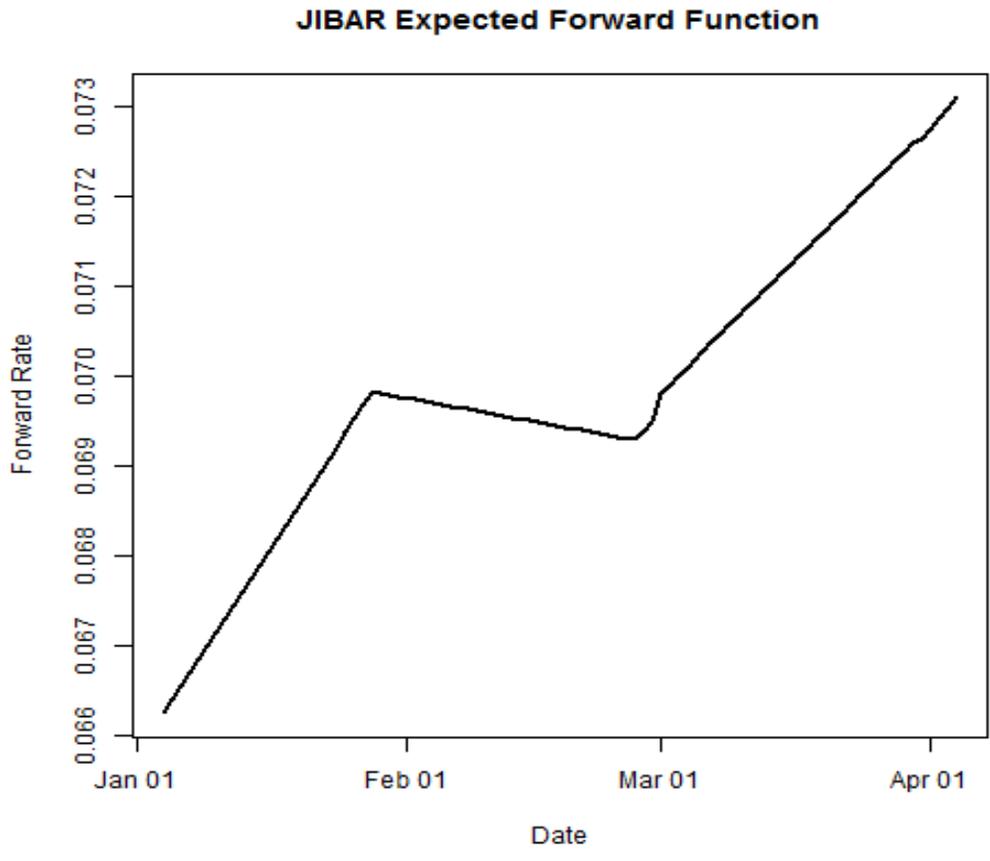


Figure 4.4: Rational Curve Construction

Interestingly, we observe a turning point in the curve 3 months prior to an MPC date. Furthermore, it is noted that irregularities in the smoothness of the curve appear from business day effects causing irregularities in the compounding of the term structure.

2 Gaussian model

Suppose we are working on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where all processes considered are adapted to \mathbb{F} , furthermore we assume \mathbb{F} satisfies the usual conditions. As we have seen, JIBAR appears to exhibit irrational behaviour in the lead up to the dates of MPC meetings. Nevertheless, while this behaviour continues to persist, one would naturally like to be able to price options based on a model that takes it into account. Thus, the purpose of the following sections is to consider pricing options with JIBAR as the underlying, using a stochastic model which allows for jumps on MPC dates.

Let us first consider the Gaussian model, where JIBAR is assumed to satisfy

$$dJ_s = \sigma dW_s + \eta_s dN_s + d\Lambda_s, \quad s \geq t, \quad (4.1)$$

where J denotes the Δt -JIBAR, W is a standard Brownian motion, and $\sigma > 0$ is the constant volatility. Here, η is a process, independent of W , which represents the size of the jumps at MPC dates, and N is a deterministic counting process, which jumps up by 1 at MPC dates. The process Λ is deterministic, and is chosen in order to prescribe the forward condition

$$\mathbb{E}[J_T - F(t, T, T + \Delta t) | \mathcal{F}_t] = 0.$$

As usual, here $F(t, T, T + \Delta t)$ is the $T \times (T + \Delta t)$ -FRA rate which, for notational simplicity, we will sometimes denote by F_T instead of $F(t, T, T + \Delta t)$. The purpose of the previous section was to allow us to calibrate this function for every $T \geq t$.

We anticipate that Λ should take the form

$$d\Lambda_s = -\mathbb{E}[\eta_s] dN_s + \lambda_s ds.$$

Substituting this into (4.1) and integrating, we have

$$J_T = J_t + \sigma(W_T - W_t) + \int_t^T (\eta_s - \mathbb{E}[\eta_s]) dN_s + \int_t^T \lambda_s ds \quad (4.2)$$

Taking an expectation, we obtain

$$F_T = \mathbb{E}[J_T | \mathcal{F}_t] = J_t + \int_t^T \lambda_s ds,$$

and substituting this back into (4.2) gives

$$J_T = \sigma(W_T - W_t) + \int_t^T (\eta_s - \mathbb{E}[\eta_s]) dN_s + F_T.$$

Consider a caplet with underlying three-month JIBAR and strike rate K . For simplicity, we make use of the fact that the variance of the discount factor is approximately of order $(\Delta t)^2$ where as the variance of the payoff is approximately of order Δt , this leads to us assuming a deterministic discount factor for short dated derivatives. Furthermore, we use the current three-month JIBAR rate as a proxy for the appropriate discount rate corresponding to the term of the derivative. That is, we are interested in computing

$$\frac{1}{1 + J_t \frac{T-t}{365}} \mathbb{E}[(J_T - K)^+ | \mathcal{F}_t],$$

Let us write Z for an auxiliary random variable with a standard normal distribution, which is independent of η . We have

$$\mathbb{E}[(J_T - K)^+ | \mathcal{F}_t] = \mathbb{E}[\mathbb{E}[(J_T - K)^+ | \eta] | \mathcal{F}_t],$$

where

$$\begin{aligned} \mathbb{E}[(J_T - K)^+ | \eta] &= \mathbb{E}\left[\left(\sigma(T-t)^{\frac{1}{2}}Z + \int_t^T (\eta_s - \mathbb{E}[\eta_s])dN_s + F_T - K\right)^+\right] \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\sigma(T-t)^{\frac{1}{2}}x + \int_t^T (\eta_s - \mathbb{E}[\eta_s])dN_s + F_T - K\right)^+ dx. \end{aligned}$$

Note that the integrand is positive if and only if

$$x > A := -\frac{1}{\sigma(T-t)^{\frac{1}{2}}} \left(\int_t^T (\eta_s - \mathbb{E}[\eta_s])dN_s + F_T - K\right).$$

Hence,

$$\begin{aligned} \mathbb{E}[(J_T - K)^+ | \eta] &= \int_A^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\sigma(T-t)^{\frac{1}{2}}x + \int_t^T (\eta_s - \mathbb{E}[\eta_s])dN_s + F_T - K\right) dx \\ &= \sigma(T-t)^{\frac{1}{2}} \int_A^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} (x - A) dx \\ &= \sigma(T-t)^{\frac{1}{2}} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{A^2}{2}} - A\Phi(-A)\right), \end{aligned} \tag{4.3}$$

where Φ denotes the cumulative distribution function of a standard normal distribution.

Since η corresponds to jumps in the Repo rate, which can only take discrete values, typically only 50, 25, 0, -25 or -50 basis points, one can simply calculate (4.3) for each possible value of η , and then weight each such value accordingly. Thus, we have derived an explicit formula for the caplet price.

The discounted payoff of a caplet, priced according to the formula derived above, is shown in Figures 4.5 and 4.6 as a function of its maturity time (and strike). The forward rates and the distribution of η here were calibrated according to data from 4th January 2016. In particular, the forward rates were calculated according to the curve construction exhibited in the previous section. MPC meetings occurred on 24 and then on 85 days after this date and, as one would expect, we observe jumps in the price at these dates.

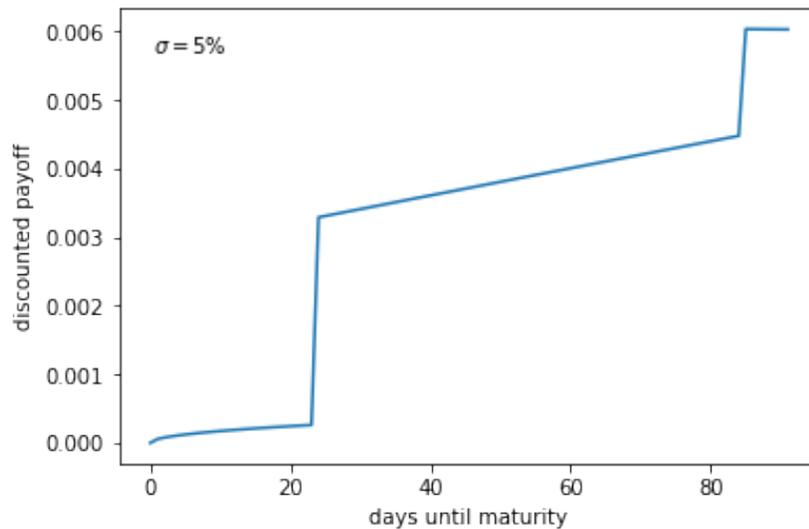


Figure 4.5: Theoretical caplet price under the Gaussian model

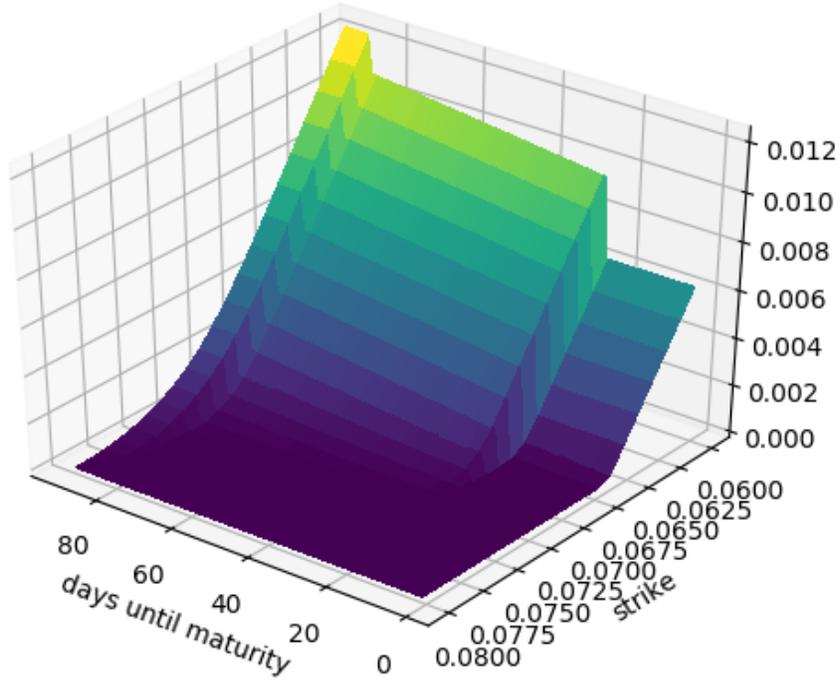


Figure 4.6: Theoretical caplet price against strike and maturity date

3 Log-normal model

Caplets and floorlets have been priced under a log-normal model in a standard setting in Gumbo (2012), where a closed form pricing formula was derived. In the current setting it is not clear that such an explicit formula is attainable. Nevertheless, we shall now investigate a log-normal model for JIBAR from which the distribution at any given time may be obtained from repeated simulations, and we will then use this to price caplets.

Using the same notation as above, let us suppose that JIBAR satisfies

$$dJ_s = J_s \sigma dW_s + \eta dN_s + d\Lambda_s, \quad s \geq t, \quad (4.4)$$

where for simplicity we consider only one MPC date with jump given by the random variable η . As before we take Λ to be of the form

$$d\Lambda_s = -\mathbb{E}[\eta] dN_s + \lambda_s ds,$$

so that

$$dJ_s = J_s \sigma dW_s + (\eta - \mathbb{E}[\eta]) dN_s + \lambda_s ds. \quad (4.5)$$

In the following we shall denote by $\mathcal{E}(X)$ the stochastic exponential of a given process X , i.e. $\mathcal{E}(X)_T = \exp(X_T - \frac{1}{2}\langle X \rangle_T)$.

The solution of the SDE (4.5), up until the first MPC date, is given by

$$J_T = \mathcal{E}(\sigma(W. - W_t))_T \left(J_t + \int_t^T \lambda_s \mathcal{E}(\sigma(W. - W_t))_s^{-1} ds \right), \quad T < \text{MPC}.$$

At the MPC date there is a jump, so that the solution at the MPC date is given by

$$J_{\text{MPC}} = \mathcal{E}(\sigma(W. - W_t))_{\text{MPC}} \left(J_t + \int_t^{\text{MPC}} \lambda_s \mathcal{E}(\sigma(W. - W_t))_s^{-1} ds \right) + \eta - \mathbb{E}[\eta].$$

Solving the equation after the MPC date, using J_{MPC} as the new initial condition, we obtain

$$J_T = \mathcal{E}(\sigma(W. - W_{\text{MPC}}))_T \left(J_{\text{MPC}} + \int_{\text{MPC}}^T \lambda_s \mathcal{E}(\sigma(W. - W_{\text{MPC}}))_s^{-1} ds \right), \quad T > \text{MPC}.$$

Taking expectations in the above we deduce, as in the Gaussian model above, that

$$F_T = \mathbb{E}[J_T | \mathcal{F}_t] = J_t + \int_t^T \lambda_s ds,$$

for all $T \geq t$ (both before and after the MPC date). In particular $dF_s = \lambda_s ds$, so we can actually eliminate λ by rewriting the solution as

$$J_T = \mathcal{E}(\sigma(W. - W_t))_T \left(J_t + \int_t^T \mathcal{E}(\sigma(W. - W_t))_s^{-1} dF_s \right), \quad T < \text{MPC},$$

$$J_{\text{MPC}} = \mathcal{E}(\sigma(W. - W_t))_{\text{MPC}} \left(J_t + \int_t^{\text{MPC}} \mathcal{E}(\sigma(W. - W_t))_s^{-1} dF_s \right) + \eta - \mathbb{E}[\eta].$$

$$J_T = \mathcal{E}(\sigma(W. - W_{\text{MPC}}))_T \left(J_{\text{MPC}} + \int_{\text{MPC}}^T \mathcal{E}(\sigma(W. - W_{\text{MPC}}))_s^{-1} dF_s \right), \quad T > \text{MPC}.$$

To visualise this solution, below we present two histograms resulting from simulations of JIBAR according to the above model. The forward rates used are the same as those in the previous section, which were constructed from the observed rates on 4th January 2016. Figure 4.7 shows the solution the day before the MPC date which occurred 24 days after the initial date, and Figure 4.8 shows the solution on the day after this MPC date. Here we have supposed that η takes the values 0, 25 and 50 basis points, with probabilities 0.5, 0.3 and 0.2 respectively.

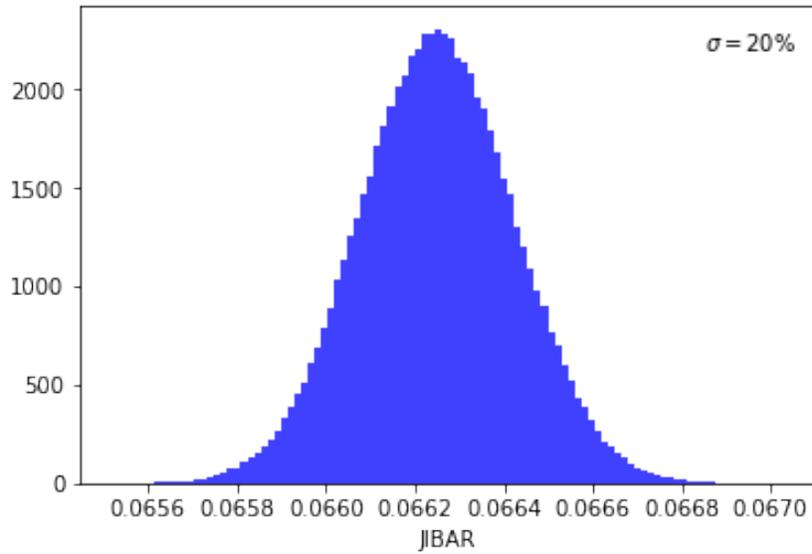


Figure 4.7: Simulations of JIBAR the day before the MPC date

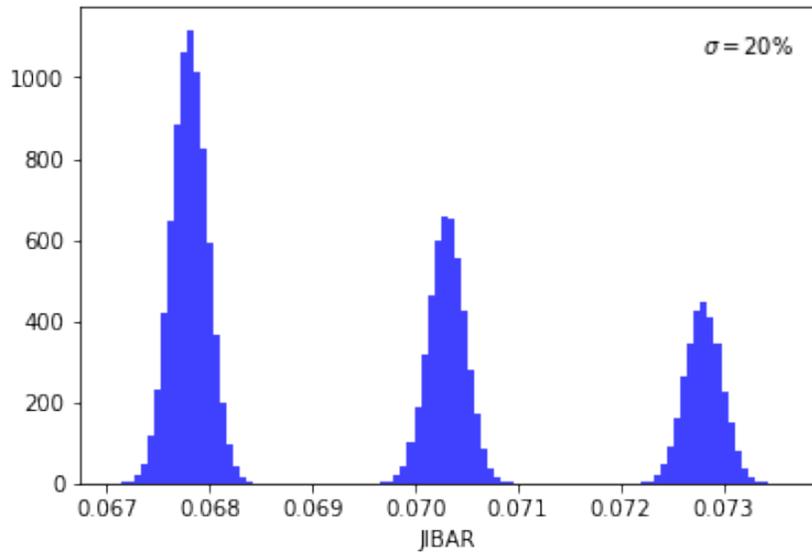


Figure 4.8: Simulations of JIBAR the day after the MPC date

We can also use simulations of the log-normal model to price caplets. In Figure 4.9 we show the discounted payoff of a caplet over the same time period as above (with an MPC date occurring after 24 days), based on simulations of JIBAR.

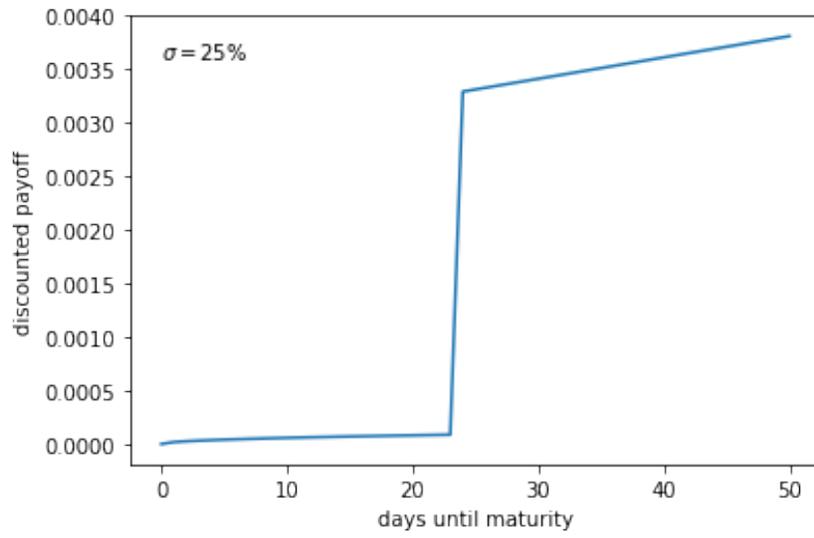


Figure 4.9: Theoretical caplet price under the log-normal model

Chapter 5

Conclusion and Further Research

We asserted that, historically, JIBAR exhibited inconsistencies on monetary policy committee meeting dates when it was clear that the forward market had anticipated a change in Repo. We showed that a rational market participant would price these anticipated changes into JIBAR prior to the Repo change. We used statistical tests to provide evidence that JIBAR and its forward rates were exhibiting these jumps around MPC dates.

Having provided sufficient evidence to support our assertion, we changed our focus to how these jumps can be implemented in an interest rate derivative setting. We implemented a curve stripping and option pricing framework consistent with the observed jumps in JIBAR and show that around Repo rate changes, the option prices will themselves change. The option pricing model implemented in this report extends the standard Gaussian and log normal spot rate model to include jumps at MPC dates. This then allows a jump in caplet prices when their maturity crosses over an MPC date, which is not predicted by classical methods.

Further research should be conducted into the accuracy of the curve stripping techniques. On any particular day, the next day's market rates were used to determine the accuracy of today's predicted forward rate curve. This, as mentioned, is not ideal as the effectiveness of the piecewise linear interpolation can only be observed once a month. A better measure of the accuracy of the predicted forward rate curve should be derived.

Bibliography

- Black, F., 1976. The pricing of commodity contracts. *Journal of financial economics* 3 (1-2), 167–179.
- Gumbo, V., 2012. The SAFEX-JIBAR market models. *Journal of Mathematical Finance* 2 (4), 321–326.
- Jager, P., Parsons, S., 2013. Signs of JIBAR Manipulation? SSRN Electronic Journal.10.2139/ssrn.2224597.
- South African Reserve Bank, 2007. Interest rates and how they work. <https://www.resbank.co.za/AboutUs/Documents/Interest%20rates%20and%20how%20they%20work.pdf>.
- South African Reserve Bank, 2012. A review of the rate-setting process of the Johannesburg Interbank Agreed Rate (Jibar) as an interest rate benchmark. <https://www.resbank.co.za/Lists/News%20and%20Publications/Attachments/5292/JibarReview.pdf>.
- South African Reserve Bank, 2018. Money Market Reference Rates. Jibar: Code of Conduct, Governance Process and Operating Rules. https://www.resbank.co.za/Lists/News%20and%20Publications/Attachments/8225/Jibar%20Code%20of%20Conduct_January%202018.pdf.
- The Economist, 2013. Year of the Lawyer. <https://www.economist.com/finance-and-economics/2013/01/05/year-of-the-lawyer>.
- West, G., 2007. Interest rate derivatives: Lecture notes. Financial Modelling Agency.
- West, G., 2009. South African financial markets. Financial Modelling Agency (September 2009).

Commitment Scheduling for Private Equity Investments

TEAM 2

WILLIAM BLACK, University College London
BANDILE MBELE, University of Cape Town
COLE VAN JAARVELDT, University of Cape Town
JOHANNES WIESEL, Oxford University

Supervisors:

LISA LARSSON, StepStone
MARTIN LARSSON, Eidgenössische Technische Hochschule Zürich



Contents

1	Introduction	4
2	Deterministic Modelling	12
2.1	The Yale Model	12
2.2	The deterministic optimisation problem	15
2.3	Commitment schedule variance penalty	18
2.4	Implementation	19
2.5	Steady state analysis	21
2.6	Sensitivity analysis	24
2.7	Calibration	31
3	Continuous-Time Stochastic Modelling	38
3.1	The de Malherbe Model	39
3.2	The stochastic optimisation problem	43
3.3	Implementation	44
3.4	Sensitivity analysis	46
3.5	Calibration	46
4	Conclusion	51

1 Introduction

The majority of equity and debt trading is conducted in the public market. Public companies (PC's) have to endure external scrutiny, comply with governmental regulations and have to make their financial transactions and compensation of its officers publicly available. When an agent purchases a single share of a PC (that has a listed price), she owns part of the company and this entitles her to certain rights as a shareholder. She has equity which can be sold in a liquid market at any given time for a known price.

What does she own if she invests money in private equity instead? How much is it worth? Who would buy it? Given that private equity is – as the word implies – private, these questions prove quite intricate. Nevertheless investing in private companies is a daily occurrence that has ancient roots. Venture capitalism is undoubtedly a familiar term and is generally understood to be a form of private equity investing in emerging companies with high growth potential. The fact that the company is unproven, the illiquid nature of the equity you own in it and the difficulty in valuing it, all contribute to the risk and associated higher returns of venture capitalism. In this introduction we aim to give an overview of the world of private equity, learn the vocabulary and understand its operating principles.

We begin with the definition of a private equity fund (PEF). It is a legal entity that arises from a contractual agreement between the investors in the fund, referred to as the limited partners or LPs, and the fund manager, referred to as the general partner or GP. A clear distinction must be established between the fund and the fund manager with a single fund manager potentially raising many funds throughout its existence. The companies that require the private equity funding have various courses of action they can pursue in order to come into contact with the fund managers. Often, the fund managers host events where potential candidates can pitch their companies. These candidates are then interrogated by a team of corporate experts that work alongside the fund manager and will inevitably be responsible for the management of the companies that the fund manager decides to purchase.

Should you, as a private investor, wish to acquire interests in private companies, you have three options based on the level of involvement you wish to have in the running of the company and the amount of work you are prepared to carry out:

- You can acquire ownership in a company without the use of intermediaries. This is known as a direct investment. Your capital contribution is then exchanged for a stake in the company in the form of equity, debt or some hybrid variation of the two. This contribution is a tangible contribution, with other intangible contributions being your time and effort in ensuring the success

of the company and, by extension, a desirable return on your initial capital contribution. The upside of this form of ownership is the satisfaction of seeing your time, money, expertise and effort grow into a tangible return when it is realised by selling your stake in the company. The downside is that you could lose your entire initial investment along with all of your effort being for none.

- You can obtain ownership in a private company with no managerial responsibilities. This is where an investor will make use of the PEF. These funds are established for a fixed term during which investors commit money to the fund that is then invested on their behalf by the manager of the fund enabling you to own part of the company without the added responsibility of managing the company. Thus your risk is reduced in exchange for fees charged by the fund for investing on your behalf which translates to a reduced return on your investment. In this report, this approach to investing in private equity will be our primary focus.
- Finally, should you not have the means to invest directly in a fund (most funds have a minimum capital contribution requirement), you could invest in a private equity fund of funds. The contributions are then pooled and invested on your behalf by a fund manager. On an operational level there are no differences between funds of funds and PEFs.

Venture capitalist funding of private enterprises usually occurs after the so called seed funding round. Seed funding is the early stage of funding provided to companies that are not yet able to generate significant cash flows. The seed money provides financing for the company until it generates enough cash flows or is ready for another round of investment. This funding is usually in the form of money from friends and family, "angel" funding and crowdfunding. Angel investors are wealthy, unrelated individuals who provide funding in exchange for a stake in the company. Crowdfunding has become increasingly significant with the advent of the internet and pooled investment platforms such as Kickstarter. After the seed funding round, which can be seen as an initial viability screening stage, venture capitalists start to invest in the company in a series of rounds until the company is ready to go public and the funding gains can be realised.

In addition to the obvious financial incentives that align the interests of the fund manager and the LP's, the agency risk for the LP is further reduced if the GP personally invests in the fund. A general rule of thumb prevalent in the private equity world suggests that the ratio of capital provided by LPs to the ratio provided GPs should be approximately 99:1. When the GP personally has money invested in the fund it is referred to as having skin in the game.

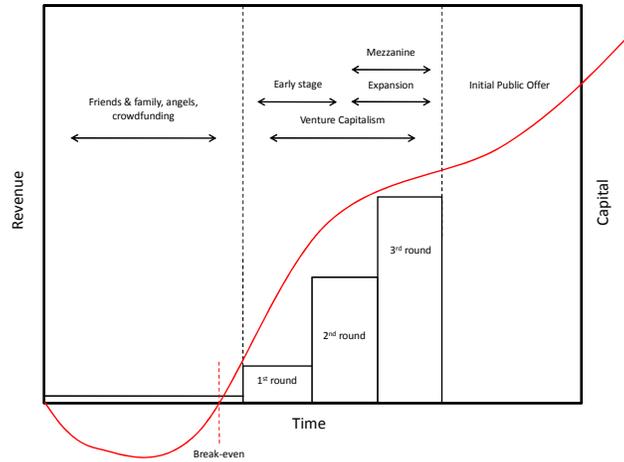


Figure 1.1: Typical financing procedure for startups

We see from Figure 1.1, a startup company is provided with capital in discrete increments that are referred to as rounds of financing. The first round of financing is referred to as Series A, the second round as Series B, and so on. A pre-financing event valuation of the company and a post-financing valuation is associated with each series. This adds the capital provided by the investors to the pre-financing event valuation. A regular occurrence is an inside round of funding that only involves existing investors. The valuation of the company at these subsequent financing events is clearly of vital importance for the GP of the fund. A down round describes a lower pre-financing valuation of a company compared with the previous post-financing valuation of the company. In an ideal situation each subsequent pre-financing valuation is higher than the previous post-financing valuation of the company. With each round of funding the ownership of the company becomes increasingly diluted.

The LPs of a PEF are provided with quarterly written reports that detail any significant events such as new investments, milestones being reached that were outlined in previous reports, new products being launched as well as information on companies that either feature in the portfolio or that the GP hopes to feature in the portfolio in future. The report contains the current valuation of all companies in the portfolio as well as their respective carry costs and possibly the proceeds from the LPs investments. Companies that repeatedly show down rounds should be written down in value or written off entirely. These investments are notorious for draining the already thinly-stretched time and energy of the GP and are referred to in the industry as the “living dead”.

In order to gain investors, a fund will need to sell its vision. This is done by using a so-called private placement memorandum (PPM), in which the strategy of the fund is laid out. A major selling point for the fund is the proprietary deals that they allege to have already secured this is either through the GP having exclusive knowledge of a specific company or through deals in which the GP will supply funding to companies.

To be successful, a fund needs to distribute the PPM to as many potential investors as possible and to subsequently meet with them. The closing or signing of a partnership agreement is the clear sign that a fund is in operation. Despite being called the closing, this does not mean that the fund is closed to new investors. In fact a fund will most likely have several closings throughout its lifetime. Thus the total fund size and number of investors are not known until well into the funds operations. The relationship between investors and fund managers is diverse in itself. In particular, older, established firms with impressive reputations (and lower management fees) commonly attract investment without the need of any self-promotion.

Venture capital funds have been our primary focus thus far as this is the original form of the PEF. The most common form of private equity investing, however, is buyout funds, with venture capitalism being the next most common. In contrast to venture capitalists that invest in emerging companies, the incentivised goal for the GP of a buyout fund is to take established, potentially struggling, companies private in order to economise them, improve efficiency or target a new market segment with the hope of increasing the revenue so they can be sold at a profit. Buyout funds will often purchase all the equity of a company ensuring complete control. The fund manager can then institute any changes to improve the financial benchmarks of the company without any consultation. In contrast, venture capital funds will mostly purchase a non-controlling interest in a company, thus enabling the venture capitalist to diversify risk as much as possible by being able to invest in many companies.

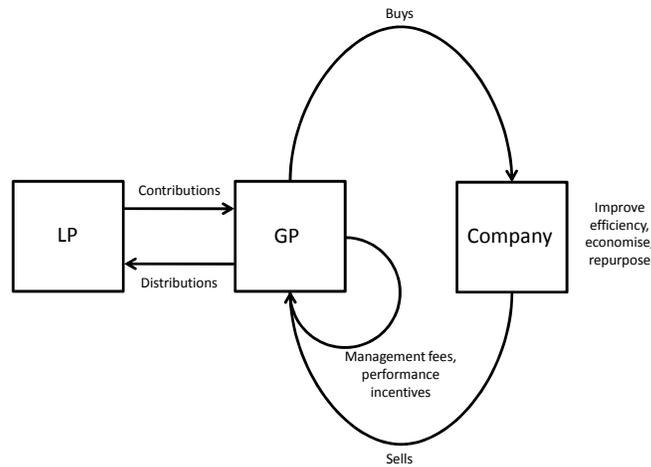


Figure 1.2: Cash flows of a typical private equity fund

Buyout funds and venture capital funds occupy the vast majority of the PEFs in practice. It would, however, be misleading not to state the other types of funds in existence. Other popular PEF strategies are described in Cumming (2009) and here:

- Mezzanine funds: These funds primarily invest in emerging companies in the later stages of their development. They offer a hybrid form of funding, i.e. provide funding in form of debt that can be converted into equity in the case that a company should default on its debt. This procedure makes it an interesting field of study in itself as a company that is unable to satisfy its creditors will more likely have high volatility in its share price. Therefore the contractual agreement on the receivable equity is an intricate task.
- Secondary purchase funds: These focus on the purchase of existing investments in PEFs and are a particularly labour and due diligence intensive style of PEF due to the difficulties that arise in the valuation of private equity and the illiquid nature thereof. When an investor sells her existing stake in private equity she sells not only the investments in companies, but also the total undrawn commitments.
- Direct Investment funds: Also known as foreign direct investment funds. These provide funding to foreign companies in exchange for equity in the company. In addition to the list of risks PEFs face, direct investment funds also face the added risk of adverse movements in the exchange rate. The fund will also have additional expenses in the form of foreign taxes, but it can minimize risk by diversifying across multiple countries and industries.

A buyout fund operates in much the same fashion as a venture capital fund: LPs commit capital to the fund that is then invested by the GP on their behalf. The total capital committed to the fund is the fund size, which is the maximum amount the GP can draw down for management fees and invest in private equity which yield performance incentives for the GP.

The management fee is usually around 2% of the total fund size. For larger funds with well-established reputations, this figure could drop to 1.5%, with emerging smaller funds charging up to 2.5% according to Kocis et al. (2009). In most funds the management fee will drop as the fund matures and gains are realised. The performance incentive is known as the "carry" and derives its name from an old nautical term for a captain's compensation. The carry cannot start until a certain percentage return on the LPs investment is realised, known as the hurdle rate. Once this minimum return is met, most of the gains are returned to the GP until a contractually agreed upon ratio of 20% – 30% is achieved for the profit-sharing (Kocis et al., 2009). Nevertheless the GP is still incentivised to achieve the best return: If a fund performs poorly, the carry previously received by the GP has to be returned to the LP through a process known as "clawback".

The GP draws the capital over the investment period of the fund, typically around 5 years according to de Malherbe (2004). The "starting year" of a fund is most commonly referred to as the vintage of a fund. Interestingly, most funds that started in 2006 were fully invested within two years. When the global financial crisis struck and markets experienced a recession, these funds were sitting with deals struck at a far higher price than 2009 valuations. Thus investors were left with holdings in private equity for a significant period of time until market conditions stabilised and a reasonable return could be achieved through the sale of the equity. In other vintages, the investment happens more gradually with a few companies being bought each year. In the aftermath of the recent recession, investors become more cautious and fund managers draw capital at a much slower rate.

Potential investors who lack the knowledge and resources to invest in a PEF directly, can invest in a fund of funds. A fund of funds operates almost identically to a fund, except that a fund of funds invests primarily in PEFs as opposed to companies. Furthermore it may add an additional layer of management fees and performance incentives, but has many advantages such as the diversification of an investors portfolio across location, asset class, strategy and industry, all for a single commitment. It is uniquely situated as the bridge between LPs and GPs. Like an LP, the fund of funds does a due diligence on the managers of the funds, assesses potential investments and invests according to specified strategies. Most of the capital used by a fund of funds is sourced from third parties in events almost identical to those hosted by GPs of PEFs. The fund of funds also receives manage-

ment fees and performance incentives from the PEFs in which they have primarily invested.

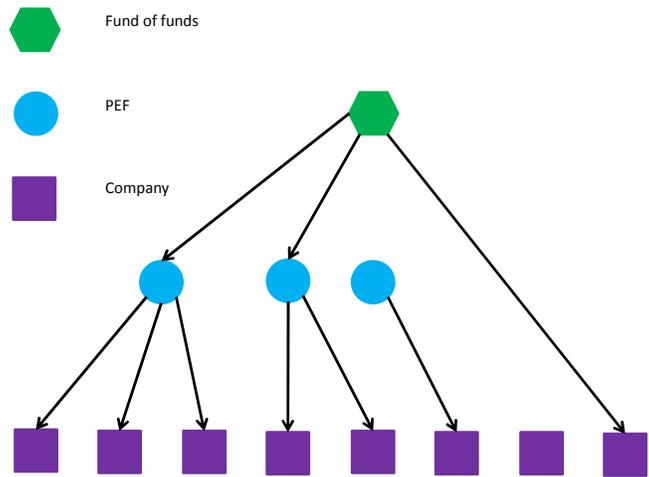


Figure 1.3: Structure of private equity market

In conclusion, investment in private equity can have many different facets. No matter if individual fund or a funds of funds, these entities all perform due diligence on potential investments in the hope of making a profit. Examples of investors include endowments, private foundations and pension funds.

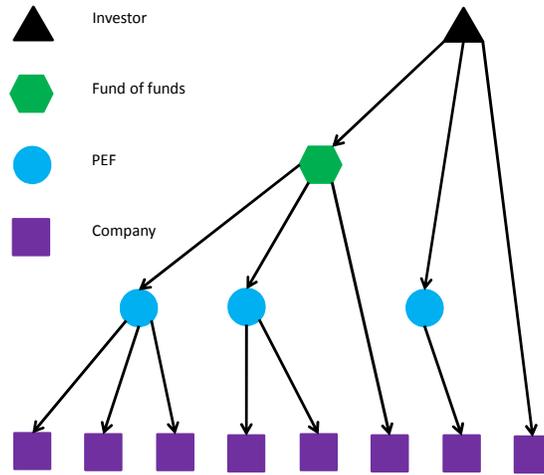


Figure 1.4: Any entity can take on the role of an investor

Private equity has become an increasingly significant portion of institutional investors' portfolios owing to the attractive returns of PEFs in recent decades. Upon committing to a PEF the LP and GP contractually agree upon (among other elements) the total amount committed, the investment period and the legal maturity of the fund. During the investment period, the committed funds can be fully drawn, but often this does not occur with capital being drawn sporadically. The timing and size of the distributions are also unknown, but must occur before the legal maturity of the fund. The potentially erratic timing and size of the draw-downs and distributions coupled with the unforeseeable changes in the value of the fund's investments make it difficult to model the future value of the fund's interests.

Despite the high level of uncertainty, investors target a percentage allocation to private equity which translates into an asset value and therefore some form of model must be employed. The model should furthermore attempt to take future capital commitments as well as future distributions into account in order to provide liquidity to investors. Due to the growing allocation of funds to private equity, the inability of simple rules of thumb previously employed becomes apparent. The amount of capital that is involved gives a strong incentive for more sophisticated models. In the remainder of this report we set out in detail what such a model might look like.

2 Deterministic Modelling

2.1 The Yale Model

Takahashi and Alexander (2002) propose a deterministic model for the dynamics of private equity funds. It was originally developed in response to Yale University's need to model the impact of investing in private equity and more specifically in many different vintages and various types of funds in its overall portfolio. The model has become the industry standard, in this report we often refer to it as the Yale Model and use it as a benchmark for our more sophisticated stochastic model developed in Section 3

Private equity funds have very different fund dynamics compared with other funds like mutual or hedge funds, which invest in public markets. When an investor commits an amount of capital to the fund, this capital is not immediately drawn. This only happens when the GP calls for a contribution, which is usually the case when a suitable company has been found to invest in, and will often be less than the initial capital commitment. Similarly, once an investment in a company has paid off, the GP will distribute the return back to the investors. Over the lifetime of a fund there are usually multiple such contribution calls and distributions.

Importantly, the contributions and distributions happen at random times. Any capital that has been committed to a fund, but has not been called upon does not contribute to the portfolio's private equity exposure. Consequently the dynamics of these contribution calls and distributions must be modelled, if an investor wants to assess the performance of a fund and its future cash flows.

Takahashi and Alexander (2002) create a model that is fairly simple and general. It primarily models the dynamics of contributions c , distributions d , and net asset value of a fund nav . Changing the inputs, it is then possible to assess the impact on these three quantities. Furthermore the model can be easily updated to correspond to actual capital commitment and asset values.

For most funds the contributions are concentrated in the early parts of the funds life, as the GP searches for opportunities to invest the LPs' commitments. After an initial flurry of activity, the contribution rate slows down in subsequent years as the fund begins to shift toward managing and selling the investments. The contribution at time t_i is given by

$$c(t_i) = rc(t_i) \left[\text{Commitment} - \sum_{j=0}^{i-1} c(t_j) \right], \quad (1)$$

where $rc(t_i)$ is the so-called contribution rate, which determines how much of the

remaining committed funds are called upon at time t_i . A standard rate of contribution schedule would be: 25% at time t_1 , 33.3% at time t_2 , and 50% for all subsequent times. We employ this simple schedule in Subsection 2.4.

The distribution and net asset value (nav) of the fund are interlinked in the model. For a standard fund, the distributions tend to be more heavily weighted towards the end of the fund's life, as investments come to maturity and capital is returned to investors. The distribution at time t_i is modeled by

$$d(t_i) = rd(t_i)(nav(t_{i-1})(1 + g^{yale})), \quad (2)$$

where g^{yale} is the growth rate of the fund. The rate of distributions rd is given by

$$rd(t_i) = \max(y, (t_i/L)^b). \quad (3)$$

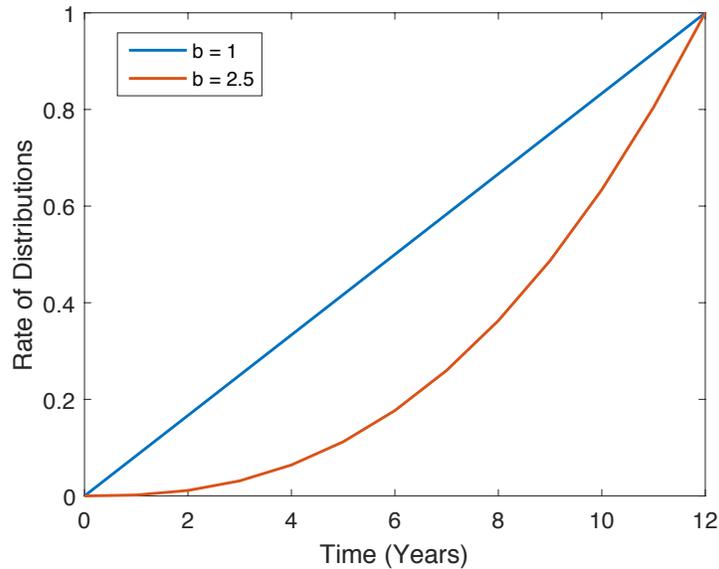


Figure 2.1: The impact of varying bow factor on the rate of distribution

This equation contains three constants, namely the yield y , the lifetime of the fund L and the bow factor b . The yield constant y is only applicable in income generating funds, such as real estate. For buyout funds, which we primarily study in this report, the yield is normally set to zero. The bow factor controls the speed at which the distribution rate changes over time. Figure 2.1 shows how varying bow factors change the rate of distribution. We note that if the lifetime of the fund is reached (i.e. $t_i = L$), $rd(L)$ is equal to one, so all the remaining assets of the fund are distributed to the LPs.

Finally, the net asset value is calculated via

$$\text{nav}(t_i) = [\text{nav}(t_{i-1})(1 + g^{\text{yale}})] + c(t_i) - d(t_i). \quad (4)$$

These equations together with the six inputs rc , Commitment, L , b , g^{yale} , y deterministically model the dynamics of the contributions, distributions and net asset values. A sample plot for the corresponding dynamics is shown in Figure 2.2

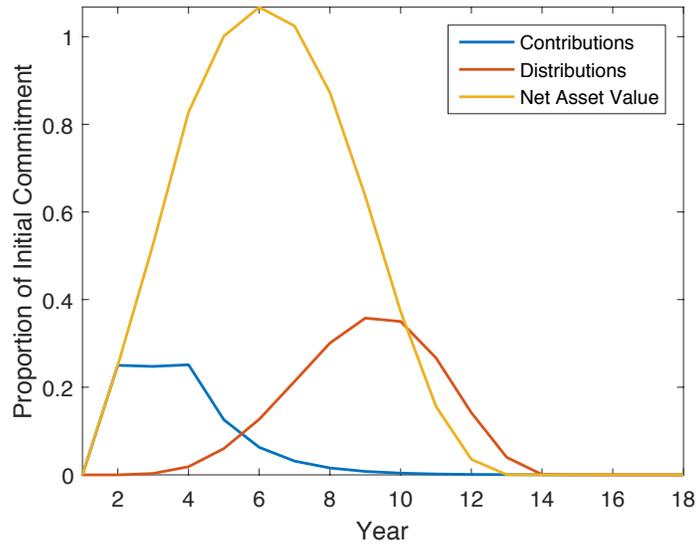


Figure 2.2: Dynamics of the contributions, distributions and net asset values for an example fund using the Yale model. The parameters used are: $L = 14$, $b = 2.5$, $g^{\text{yale}} = 0.13$, $y = 0$.

Let us highlight some key features of the dynamics in Figure 2.2: The contributions are heavily concentrated in the early years of the fund, while distributions primarily occur towards the end of life of the fund. The peak of the net asset value is slightly above the initial commitment value due to the growth factor g^{yale} . Furthermore the sum of capital contributions in this model usually never equals the capital commitment. This is quite realistic as we have discussed in Section 1, since many funds are unable to draw down all the capital due to the illiquid nature of the underlying investments.

Given actual fund data the input parameters of this model can then be calibrated, which makes predictions of future dynamics possible. We elaborate on this procedure in Section 2.7.

2.2 The deterministic optimisation problem

Suppose an investor presently holds a portfolio of value p_0 that consists of one type of asset, called the public asset, that grows at a constant rate $g > 0$. If no external funds are inserted or withdrawn, and no investments in other assets are made, the portfolio value at time t , $p(t)$ is given by $p(t) = p_0 e^{gt}$.

As she observes higher returns in private equity (PE), the investor wishes to shift part of the portfolio into this market. On the grounds of risk management and investment objectives she desires to hold a constant target ratio $\pi \in (0, 1)$ of her total exposure in PE. She thus has to find an optimal commitment schedule $\{a(t_j)\}$, such that the amount $v(t_i)$ invested in PE quickly approaches $\pi p(t_i)$ and stabilizes about it, where $p(t_i)$ is the total fund value at time t_i .

In order to model the problem of finding the optimal commitment schedule $a(t_j)_{j \in \mathbb{N}}$ we make the following simplifying assumptions:

- The portfolio is self-financing: no external funds are inserted or withdrawn. The equations for the portfolio dynamics below reflect this assumption.
- For every t_j there is exactly one PE fund starting at time t_j (vintage), and $a(t_j)$ is the amount committed to the fund at this time. In total there are $m \in \mathbb{N}$ different PE funds.
- The commitment schedule $\{a(t_j)\}_{j=1, \dots, m}$ is deterministic.

Furthermore we establish the following notation:

- $\text{nav}(s, t)$ denotes the net asset value of the fund of vintage s at time t . This implies that $\text{nav}(s, t) = 0$ for all $t \leq s$.
- $c(s, t)$ and $d(s, t)$ are the contributions and distributions at time t , which are transferred to/paid out of the fund of vintage s . In particular $c(s, t) = d(s, t) = 0$ for $t \leq s$.
- $R(s, t)$ is the rate of return over $[t - \Delta t, t)$ on the assets held by the fund of vintage s .

We express all of the above quantities per committed dollar. For clarity we often use vector/matrix notation in the rest of this report. For example we write

$$\text{nav} = (\text{nav}_{j,i})_{j=1, \dots, m}^{i=1, \dots, n} = \text{nav}(t_j, t_i)_{j=1, \dots, m}^{i=1, \dots, n} \in \mathbb{R}^{m \times n}$$

on the following pages. Furthermore the matrix product of $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{m \times k}$ is denoted by $A \cdot B$. If not explicitly stated otherwise we assume that all vectors v are row vectors, i.e. $v \in \mathbb{R}^n = \mathbb{R}^{1 \times n}$.

Under the above simplifying assumptions the net asset value (per committed dollar) of the private equity fund of vintage t_j evolves as

$$\begin{aligned} \text{nav}(t_j, t_i) &= \text{nav}(t_j, t_{i-1})e^{R(t_j, t_i)\Delta t} + c(t_j, t_i) - d(t_j, t_i) \\ &= \sum_{k=1}^{i-j} e^{R(t_j, t_{j+k+1}) \cdots R(t_j, t_i)\Delta t} (c(t_j, t_{j+k}) - d(t_j, t_{j+k})), \quad t_i \geq t_j. \end{aligned}$$

Given a commitment schedule $(a(t_j))_{j \in \{1, \dots, m\}}$ the resulting PE investment is

$$v(t_i) = \sum_{j=1}^i a(t_j) \text{nav}(t_j, t_i) = a \cdot \text{nav}.$$

Similarly the total net cash flow at time t_i from the fund investments is given by

$$\text{ncf}(t_i) = \sum_{j=1}^i a(t_j) (d(t_j, t_i) - c(t_j, t_i)) = a \cdot (d - c).$$

Given the fund investment $v(t_i)$ the holding in the public asset is $p(t_i) - v(t_i)$ which grows at rate g . Due to the self financing condition the net cash flows are transferred to the investor's holdings in the public asset at every point in time. This evolves as

$$p(t_i) - v(t_i) = (p(t_{i-1}) - v(t_{i-1}))e^{g\Delta t} + \text{ncf}(t_i)$$

Our objective is to minimize the distance between πp and v over time. Choosing the ℓ_2 -distance, our task is thus to solve the optimization problem

$$\text{minimise: } (v - \pi p)^T \cdot (v - \pi p) = \sum_{i=1}^n (v(t_i) - \pi p(t_i))^2 \quad (5)$$

$$\text{such that } a_j \geq 0 \quad \text{for all } j = 1, \dots, m.$$

In order to find an expression for the vector p in terms of a we define the following expressions: We write

$$E := \begin{pmatrix} 1 & e^{g\Delta t} & \dots & e^{ng\Delta t} \\ 0 & 1 & \dots & e^{(n-1)g\Delta t} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{n \times n},$$

$$k := (e^{g\Delta t}, e^{2g\Delta t}, \dots, e^{ng\Delta t}) \in \mathbb{R}^n$$

and

$$C := \text{nav} + (d - c) \cdot E \in \mathbb{R}^{m \times n}.$$

Lemma 2.1. We can rewrite p as

$$\begin{aligned} p &= a \cdot \text{nav} + (p(t_0) - v(t_0)) \cdot k + a \cdot (d - c) \cdot E \\ &= a \cdot C + (p(t_0) - v(t_0)) \cdot k. \end{aligned}$$

Proof. Define $h(t_i) := p(t_i) - v(t_i)$. Then

$$h(t_i) = h(t_{i-1})e^{g\Delta t} + \text{ncf}(t_i)$$

and in particular

$$\begin{aligned} h(t_n) &= h(t_{n-1})e^{g\Delta t} + \text{ncf}(t_n) \\ &= [h(t_{n-2})e^{g\Delta t} + \text{ncf}(t_{n-1})]e^{g\Delta t} + \text{ncf}(t_n) \\ &= h(t_0)e^{ng\Delta t} + \sum_{i=1}^n e^{(n-i)g\Delta t} \text{ncf}(t_i). \end{aligned}$$

Thus

$$p(t_n) = v(t_n) + (p(t_0) - v(t_0))e^{ng\Delta t} + \sum_{i=1}^n e^{(n-i)g\Delta t} (a \cdot (d - c))_i$$

and the claim follows. \square

Proposition 2.2. The minimization problem (5) can be written as a quadratic programming problem

$$\begin{aligned} \text{minimize: } & \frac{1}{2} a \cdot H \cdot a^T + a \cdot f + \text{const.} \\ \text{such that: } & a_j \geq 0 \text{ for all } j \in \{1, \dots, m\}, \end{aligned}$$

where

$$\begin{aligned} H &:= 2(\text{nav} \cdot \text{nav}^T + 2\pi \text{nav} \cdot C^T - \pi^2 C \cdot C^T), \\ f &:= 2\pi(p(t_0) - v(t_0))k^T + \pi^2(p(t_0) - v(t_0))C \cdot k^T. \end{aligned}$$

Proof. Note that

$$\begin{aligned} 2\pi a \cdot \text{nav} \cdot p^T &= 2\pi a \cdot (\text{nav} \cdot C^T \cdot a^T + (p(t_0) - v(t_0))k^T), \\ \pi^2 p \cdot p^T &= \pi^2 (a \cdot C \cdot C^T \cdot a^T + 2(p(t_0) - v(t_0))a \cdot C \cdot k^T \\ &\quad + (p(t_0) - v(t_0))^2 k \cdot k^T). \end{aligned}$$

Thus

$$\begin{aligned} (v - \pi p)^T \cdot (v - \pi p) &= a \cdot \text{nav} \cdot \text{nav}^T \cdot a^T - 2\pi a \cdot \text{nav} \cdot p^T + \pi^2 p \cdot p^T \\ &= a \cdot (\text{nav} \cdot \text{nav}^T + 2\pi \text{nav} \cdot C^T - \pi^2 C \cdot C^T) \cdot a^T \\ &\quad + a \cdot (2\pi(p(t_0) - v(t_0))k^T + \pi^2(p(t_0) - v(t_0))C \cdot k^T) \\ &\quad + \pi^2(p(t_0) - v(t_0))^2 k \cdot k^T, \end{aligned}$$

which concludes the proof. \square

2.3 Commitment schedule variance penalty

The solution $a(t_j)$ to the optimization problem (5) yields a commitment schedule that varies greatly as a function of time. To reduce the fluctuation of $a(t_j)$ and provide the fund manager with a commitment schedule, which is easier to implement, we introduce a variance penalty on the second differences of $a(t_j)$. This produces a smoothed and regular commitment schedule as a function of time and is achieved by modifying the matrix H as follows: Let us introduce the difference matrix

$$D = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

We now add an additional penalisation term in the optimisation problem (5): Indeed, we add the product of a parameter β and the empirical estimator of the variance of the second differences of a given by

$$\widehat{\text{Var}}(aD^2) := \frac{1}{m-2} \sum_{i=1}^{m-2} (aD^2)_i - \frac{1}{(m-2)^2} \sum_{i=1}^{m-2} \sum_{j=1}^{m-2} (aD^2)_i (D^2 a)_j$$

to obtain a smoother commitment schedule $a(t_j)$. This yields the optimization problem

$$\begin{aligned} \text{minimize: } & \frac{1}{2} a \cdot H^* \cdot a^T + a \cdot f + \text{const.} \\ \text{such that: } & a_j \geq 0 \text{ for all } j \in \{1, \dots, m\}, \end{aligned}$$

where the penalty factor β is non-negative,

$$H^* := 2(\text{nav} \cdot \text{nav}^T + 2\pi \text{nav} \cdot C^T - \pi^2 C \cdot C^T + \beta D^2 M D^{2T})$$

and

$$M := \begin{pmatrix} \frac{1}{m-2} - \frac{1}{(m-2)^2} & -\frac{1}{(m-2)^2} & \cdots & -\frac{1}{(m-2)^2} \\ -\frac{1}{(m-2)^2} & \frac{1}{m-2} - \frac{1}{(m-2)^2} & \cdots & -\frac{1}{(m-2)^2} \\ \vdots & \vdots & \cdots & \vdots \\ -\frac{1}{(m-2)^2} & -\frac{1}{(m-2)^2} & \cdots & \frac{1}{m-2} - \frac{1}{(m-2)^2} \end{pmatrix}.$$

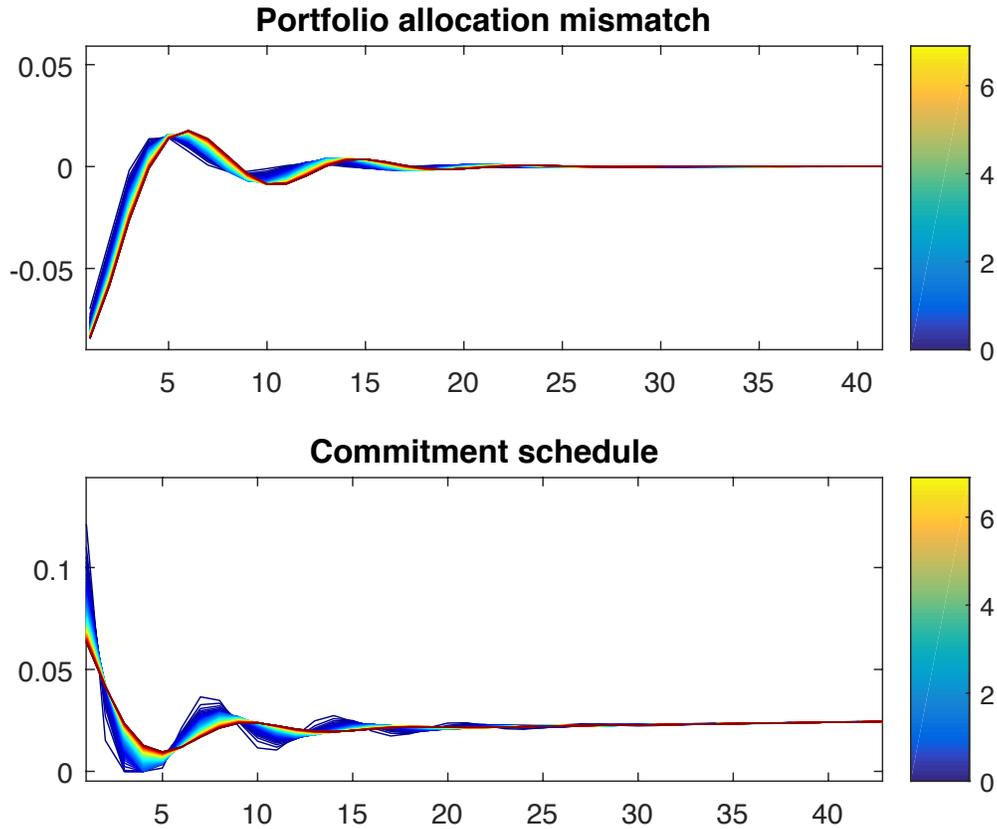


Figure 2.3: Dynamics of varying the weighting between correct exposure to PEF and variance reduction for the commitment schedule.

Figure 2.3 shows how the difference between actual and target PEF, and commitment schedule a changes as the weighting between these factors in the optimiser is varied. A higher weight indicates more bias towards variance reduction. This can be seen in both graphs, as the weight increases, the variance of the commitment schedule is reduced, while the difference in the actual and target exposure takes longer to converge.

2.4 Implementation

We implement the deterministic Yale model specified in Subsection 2.2 in MATLAB and use the Optimization Toolbox “Quadprog” to solve (5). As a proof of concept we set $m = n = 100$ and generate shift-invariant contributions and distributions of the different funds according to the Yale model with parameters $b = 2.5$, $g^{yale} = 0.13$, $y = 0$, $L = 48$, $p(t_0) = 1$, $rc(t_1) = 0.25$, $rc(t_2) = 1/3$, $rc(t_i) = 0.5$ for all

$j \in \{3, \dots, n\}$. We then optimize under the assumption that $g = 0.01$ and $\pi = 0.1$. Figure 2.4 shows the solution of the optimisation problem in terms of optimal exposure to PEF given by $\pi p(t_i)$ and actual exposure to PEF given by $a(t_i)v(t_i)$. After a short ramp period of under 10 quarters, the actual exposure converges quickly to the target exposure. This rapid convergence is mostly due to the homogeneous structure of the underlying funds. Figure 2.5 shows the corresponding optimal commitment schedule $a(t_i)$.

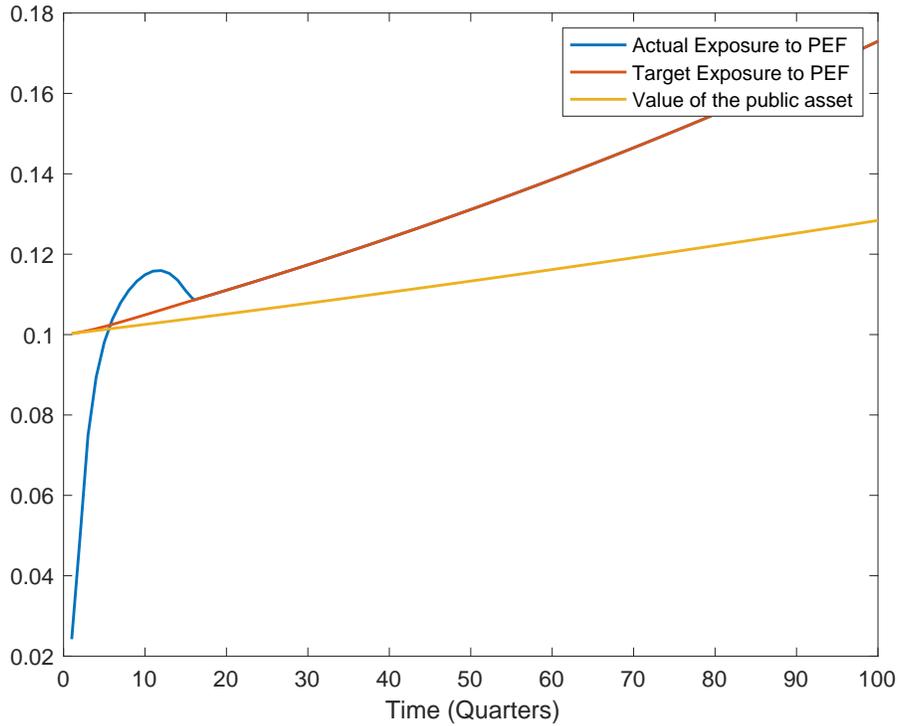


Figure 2.4: Target exposure vs Actual exposure to PEF in the Yale model for 100 funds with parameters $b = 2.5$, $g^{yale} = 0.13$, $y = 0$, $L = 48$, $p(t_0) = 1$, $rc(t_1) = 0.25$, $rc(t_2) = 1/3$, $rc(t_i) = 0.5$ for all $j \in \{3, \dots, n\}$. Growth of the public asset is plotted as a comparison.

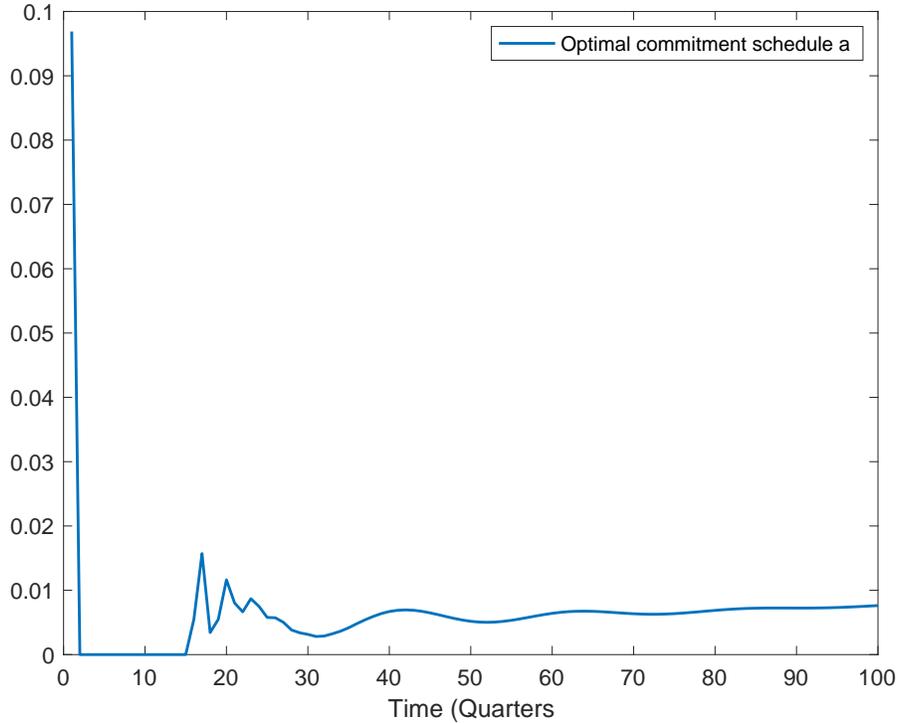


Figure 2.5: Optimal commitment schedule in the Yale model for 100 funds with parameters $b = 2.5$, $g^{yale} = 0.13$, $y = 0$, $L = 48$, $p(t_0) = 1$, $rc(t_1) = 0.25$, $rc(t_2) = 1/3$, $rc(t_i) = 0.5$ for all $j \in \{3, \dots, n\}$.

2.5 Steady state analysis

In this section we investigate the longterm behaviour of the portfolio value $p(t_i)$ given the optimal commitment schedule $a(t_j)_{j \in \mathbb{N}}$. Instead of formally arguing by use of limits we opt for the more intuitive approach of equating quantities for large t_i . In particular we assume that a steady state will be reached in which $p(t_i)$, the total portfolio at time t_i , will grow at a fixed rate r^* so that $p(t_{i+1}) = p(t_i)e^{r^* \Delta t}$ for large t_i . Furthermore we make the following simplifying assumptions:

- PE has a constant rate of return $R(t_j, t_i) = r^{PE}$.
- The deterministic contributions and distributions are stationary, i.e., they depend only on the elapsed time since the fund has started: $c(t_j, t_i) = c(t_i - t_j)$ and $d(t_j, t_i) = d(t_i - t_j)$.

The public portion of the portfolio will grow at constant rate $g > 0$, i.e., $(1 -$

$\pi)p(t_{i+1}) = (1 - \pi)p(t_i)e^{g\Delta t}$. The private portion of the portfolio will grow at constant rate $r^{PE} > 0$. Setting

$$[c(t_{i-k} - t_j) - d(t_{i-k} - t_j)] =: \kappa(t_{i-k} - t_j)$$

we have:

$$\begin{aligned} v_{t_{i+1}} &= \sum_{j=1}^{i+1} a(t_j) \text{nav}(t_j, t_i) = \sum_{j=1}^{i+1} a(t_j) [\text{nav}(t_j, t_i) e^{r^{PE}\Delta t} + \kappa(t_{i+1} - t_j)] \\ &= \sum_{j \leq (i+1)} a(t_j) \sum_{k=0}^{(i+1)-j} e^{r^{PE}k\Delta t} \kappa(t_{(i+1)-k} - t_j) \\ &= \sum_{j \leq (i+1)} a(t_j) \left[\sum_{k=1}^{(i+1)-j} e^{r^{PE}k\Delta t} \kappa(t_{(i+1)-k} - t_j) + \kappa(t_{i+1} - t_j) \right] \\ &= \sum_{j \leq i} a(t_j) e^{r^{PE}\Delta t} \left[\sum_{k=0}^{i-j} e^{r^{PE}k\Delta t} \kappa(t_{i-k} - t_j) + \kappa(t_{i+1} - t_j) \right] + a(t_{i+1}) [\kappa(t_{i+1} - t_j)] \\ &= e^{r^{PE}\Delta t} v(t_i) + \sum_{j \leq (i+1)} a(t_j) \kappa(t_{i+1} - t_j). \end{aligned}$$

We can therefore express the total portfolio value at time t_{i+1} as

$$p(t_{i+1}) = p(t_i) e^{r^*\Delta t} = (1 - \pi)p(t_i) e^{g\Delta t} + \pi p(t_i) e^{r^{PE}\Delta t} + \sum_{j \leq (i+1)} a(t_j) \kappa(t_{i+1} - t_j),$$

from which it follows that

$$\begin{aligned} r^* &= \frac{1}{\Delta t} \log \left[(1 - \pi) e^{g\Delta t} + \pi e^{r^{PE}\Delta t} + \sum_{j \leq (i+1)} \frac{a(t_j)}{p(t_i)} \kappa(t_{i+1} - t_j) \right] \\ &= \frac{1}{\Delta t} \log \left[(1 - \pi) e^{g\Delta t} + \pi e^{r^{PE}\Delta t} + \sum_{j \leq (i+2)} \frac{a(t_j)}{p(t_{i+1})} \kappa(t_{i+2} - t_j) \right]. \end{aligned}$$

We can conclude the relationship

$$\sum_{j \leq (i+1)} \frac{a(t_j)}{p(t_i)} \kappa(t_{i+1} - t_j) = \sum_{j \leq (i+2)} \frac{a(t_j)}{p(t_{i+1})} \kappa(t_{i+2} - t_j). \quad (6)$$

To simplify the above expression, let us impose the condition $\kappa(t_{i+1} - t_j) = 0 \forall j < (i + 1) - L$ for some common $L \in \mathbb{N}$. As funds have finite lifespan and $\kappa(t_{i+1} - t_j)$ is shift-invariant, this is not restrictive for large t_{i+1} . Then (6) can be restated as

$$\mathcal{K} := \sum_{(i+1)-L \leq j \leq (i+1)} \frac{a(t_j)}{p(t_i)} \kappa(t_{i+1} - t_j) = \sum_{(i+2)-L \leq j \leq (i+2)} \frac{a(t_j)}{p(t_{i+1})} \kappa(t_{i+2} - t_j).$$

The long-term growth rate of the entire portfolio r^* is clearly dependent on the long-term behaviour of the above quantity \mathcal{K} . We now distinguish three cases regarding the behavior of \mathcal{K} for large t_i : Should \mathcal{K} tend zero, we have $e^{r^* \Delta t} = (1 - \pi)e^{g \Delta t} + \pi e^{r^{PE} \Delta t}$, which clearly has a unique solution r^* and expresses the growth of the portfolio simply as the exponentially convex combination of the growth in the proportion holding of public equity and the growth of the proportion holding of private equity. This is what we expect given the long-term behaviour of \mathcal{K} .

Secondly, it is unrealistic that \mathcal{K} grows at a rate higher than the growth of the portfolio $p(t_i)$ as this quickly leads to the entire portfolio being committed to private equity contradicting our assumptions. This leaves us with the case

$$\mathcal{K} = c(r^*) \neq 0.$$

Then by shift-invariance

$$\mathcal{K} = \sum_{0 \leq j \leq L} \frac{a(t_j)}{p(t_{L-1})} \kappa(t_L - t_j) = c(r^*)$$

and

$$e^{r^* \Delta t} = (1 - \pi)e^{g \Delta t} + \pi e^{r^{PE} \Delta t} + c(r^*),$$

which again has a unique solution given

$$c(0) \geq 0 \quad \text{and} \quad c(\infty) < 1 - (1 - \pi)e^{g \Delta t} + \pi e^{r^{PE} \Delta t}. \quad (7)$$

As the terms $\kappa(t_L - t_j)$ are independent of r^* our assumptions further imply that $a(t_j) = a(t_0)e^{r^* t_j}$ for all $0 \leq j \leq L$. Then (7) is satisfied as soon as

$$\sum_{0 \leq j \leq L} \kappa(t_L - t_j) > 0 \quad \text{and} \quad \kappa(t_1) < 1 - (1 - \pi)e^{g \Delta t} + \pi e^{r^{PE} \Delta t}. \quad (8)$$

Conditions (8) are usually fulfilled, as they simply mean that a fund has higher distributions than contributions and there are fewer distributions than contributions at time t_1 .

As a proof of concept we fit an exponential function of $a_0 e^{0.25r^* \Delta t}$ to the target exposure and optimal commitment schedule for the Yale model computed in Section 2.4. We obtain the results 5.808% for the target exposure and 5.616% for the optimal commitment. Using a first order approximation for the formula $(1 - \pi)p(t_i)e^{g\Delta t} + \pi p(t_i)e^{r^{PE}\Delta t}$ yields a theoretical value of 5.8%.

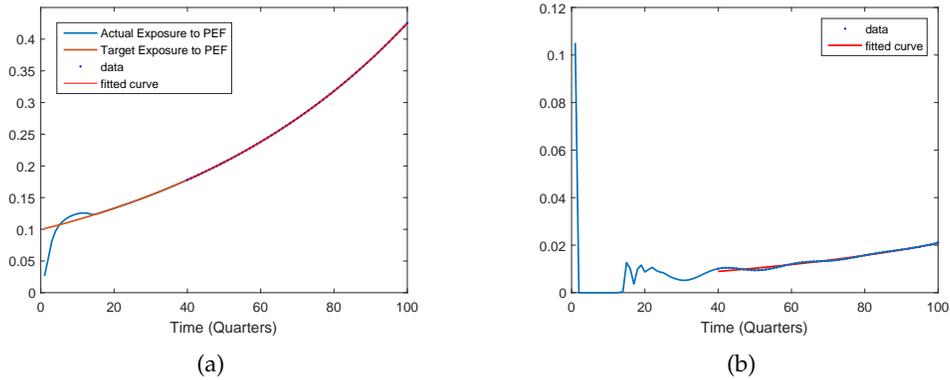


Figure 2.6: (a): Exponential fitting to the growth rate of private equity. (b): Exponential fitting to the growth rate of commitments

2.6 Sensitivity analysis

In order to gain an understanding of the interrelationship between model inputs and outputs we undertake a sensitivity analysis for the Yale model. Our basic approach is as follows: We fix all parameters in the Yale model apart from one and then compute the different actual exposures to private equity given by the optimal solution of (2). In this section, we vary the bow rate, number of funds, life time of funds, growth rate of private equity and target exposure to private equity. Furthermore we examine how the random removal of funds influences the performance of the optimal commitment schedule.

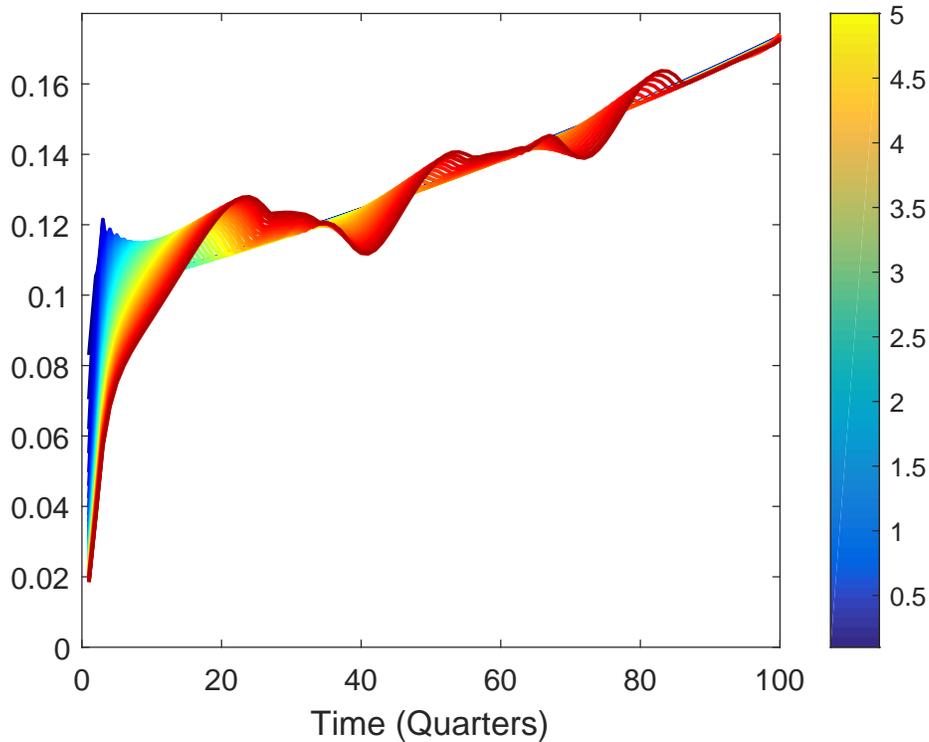


Figure 2.7: Impact of varying bow factor on actual exposure to private equity. The following model parameters were used: Private Equity Growth = 13%, Public Market Growth = 1%, Target Exposure = 1%, Number of Funds = 100, Fund Lifetime = 12 Years

Figure 2.7 shows the impact of varying the bow rate on the actual exposure to private equity. For a visual reference on how the bow factor changes the Yale model see Figure 2.1. A higher bow factor changes the distribution rate so it is more heavily concentrated in the later years of the fund.

In Figure 2.7 one can see that a higher bow factor increases the time amplitude of the oscillations before the steady state is reached. This can be attributed to two things; The greater time shift between contributions and distributions as bow rate is increased, and the distributions being paid out in larger sums meaning large amounts of private equity are converted to cash without immediate contributions to balance this out.

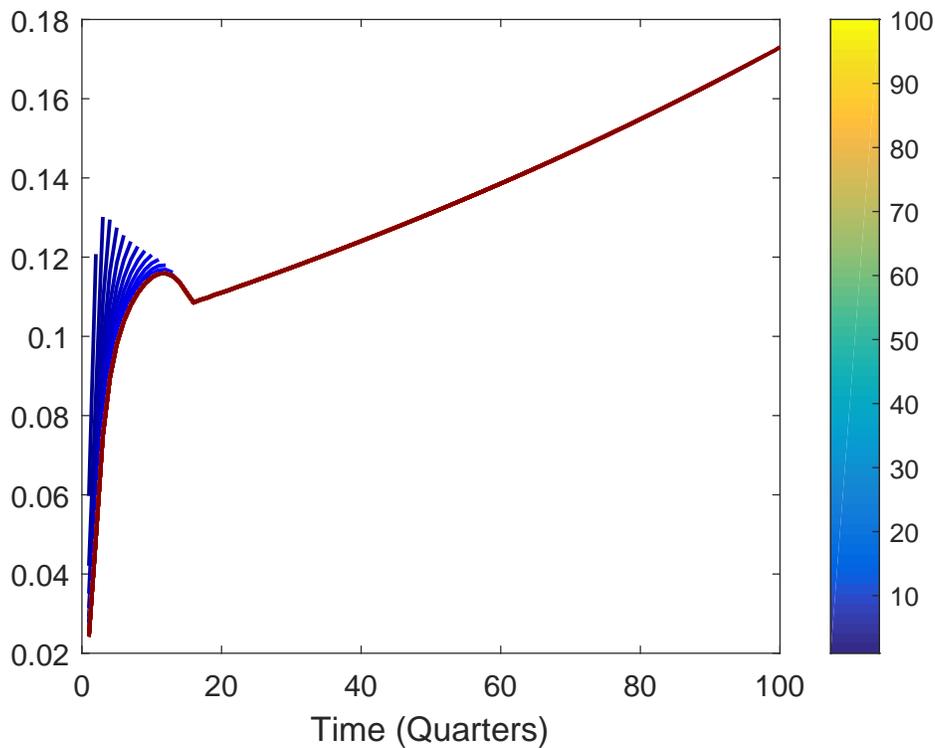


Figure 2.8: Impact of varying the number of funds on actual exposure to private equity. The following model parameters were used: Private Equity Growth = 13%, Public Market Growth = 1%, Target Exposure = 1%, Bow Factor = 2.5, Fund Lifetime = 12 Years

Figure 2.8 shows the impact of varying the number of funds on the actual exposure to PEF. In the simulated data a new fund is opened every quarter, so the number of funds is directly linked to the investment horizon of the investor. For a relatively small number of funds (less than 15) a steady state, in which the contributions and distributions remain constant over time, is not reached. This is why the initial tests do not reach the optimal line. For a higher number funds the steady state is reached, so all actual exposures to PEF follow the same path.

An important lesson from this graph is that the investment horizon of the investor plays a critical role for the choice of asset class. For private equity, an investment horizon of more than 5 years is required in order to reach and maintain a target exposure.

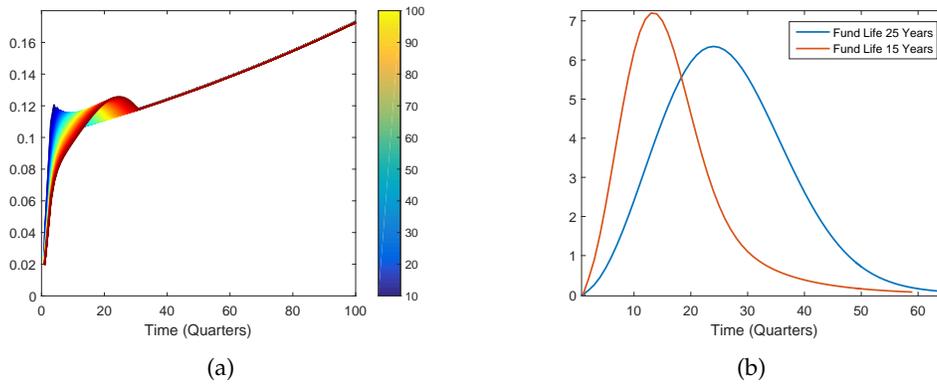


Figure 2.9: (a): Impact of varying fund life on actual exposure to private equity. The following model parameters were used: Private Equity Growth = 13%, Public Market Growth = 1%, Target Exposure = 1%, Number of Funds = 100, Bow Factor = 2.5. The initial portfolio size is 1. (b): Comparison of distribution rate for different fund life lengths.

Figure 2.9(a) shows the impact of varying fund life on actual exposure to private equity. Varying the fund life mainly impacts the distribution rate in the Yale model, as described in equations (2) and (3). A longer fund life increases the time until most contributions are paid out and decreases the peak distribution, as shown in Figure 2.9(b). This combination of changes in distributions accounts for the increased ramp speed but higher overshoot in the funds with shorter life span. Once the target exposure to PEF has been reached (post overshoot), all of the tests follow the same path.

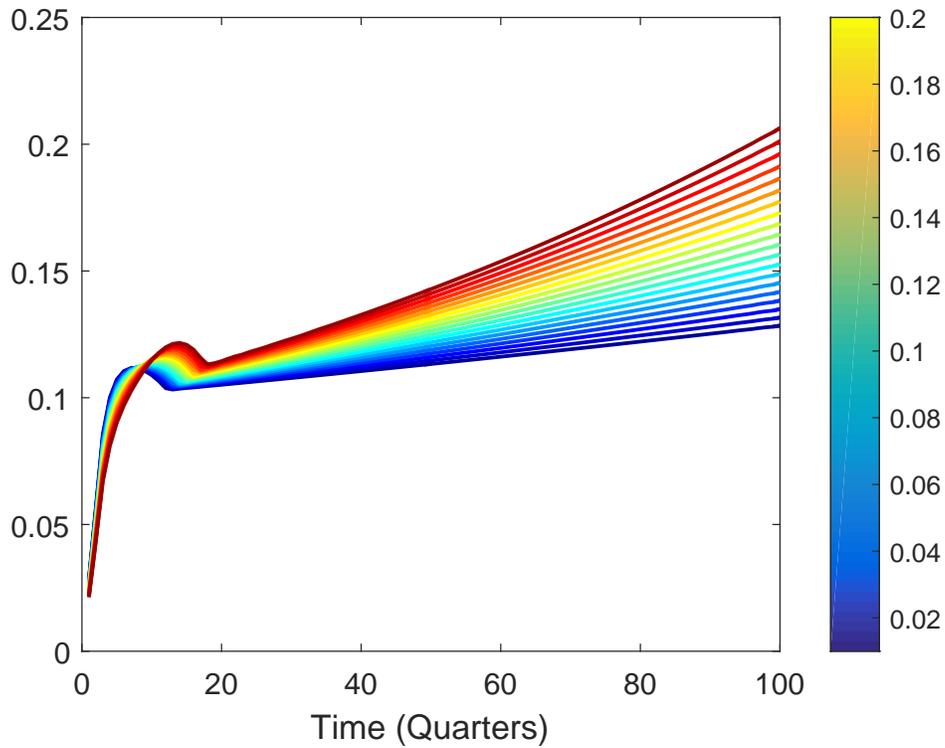


Figure 2.10: Impact of varying private equity growth on actual exposure to private equity. The following model parameters were used: Public Market Growth = 1%, Target Exposure = 1%, Number of Funds = 100, Fund Lifetime = 12 Years, Bow Factor = 2.5

Figure 2.10 shows the actual exposure to PEF as the private equity growth rate is varied between 1% and 20%. A higher growth rate obviously leads to a higher amount of raw exposure needed to private equity, which is most visible at the end of the graph. The growth rate also affects the ramp and overshoot dynamics, with a higher growth rate leading to a longer ramp period until the steady state is period is reached.

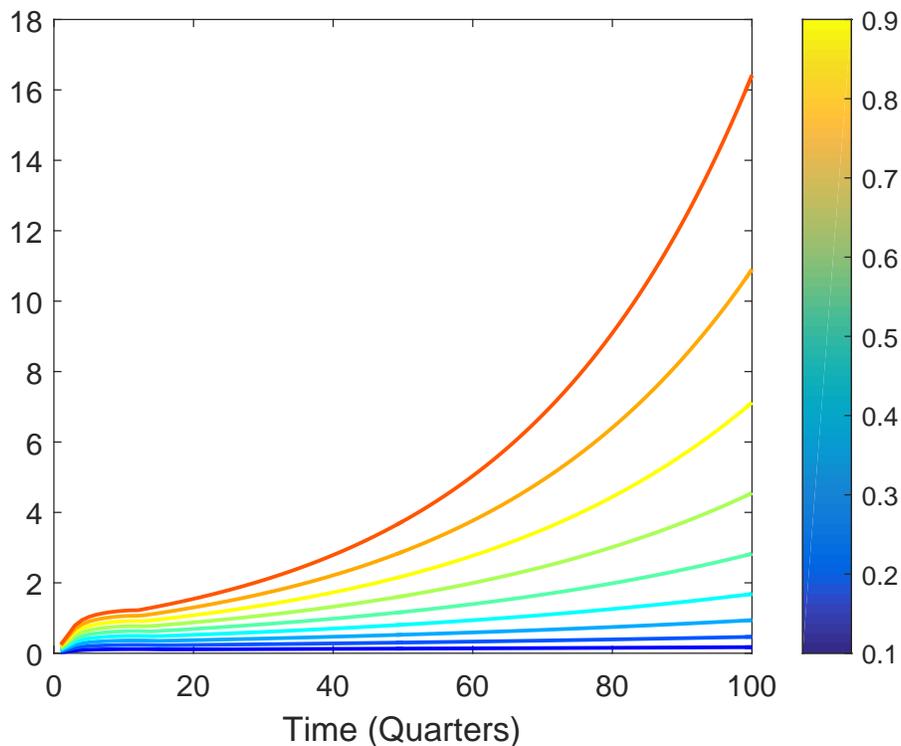


Figure 2.11: Impact of varying target exposure to private equity on actual exposure to private equity. The following model parameters were used: Private Equity Growth = 13%, Public Market Growth = 1%, Number of Funds = 100, Fund Lifetime = 12 Years, Bow Factor = 2.5

Figure 2.11 shows the dynamics of actual exposure to PEF as the target exposure to PEF is varied from 1% to 90%. This test behaved as expected, with higher target exposure leading to higher actual exposure. The ramp period for higher proportions is shorter than lower proportions. This difference in time can be explained as follows: to reach a higher proportion exposure, a larger initial commitment is required. Once the peak of the overshoot is reached, the target exposure is growing quicker from below for higher target exposure compared with lower target exposure. This means that the convergence from peak of overshoot to steady state is faster in these cases.

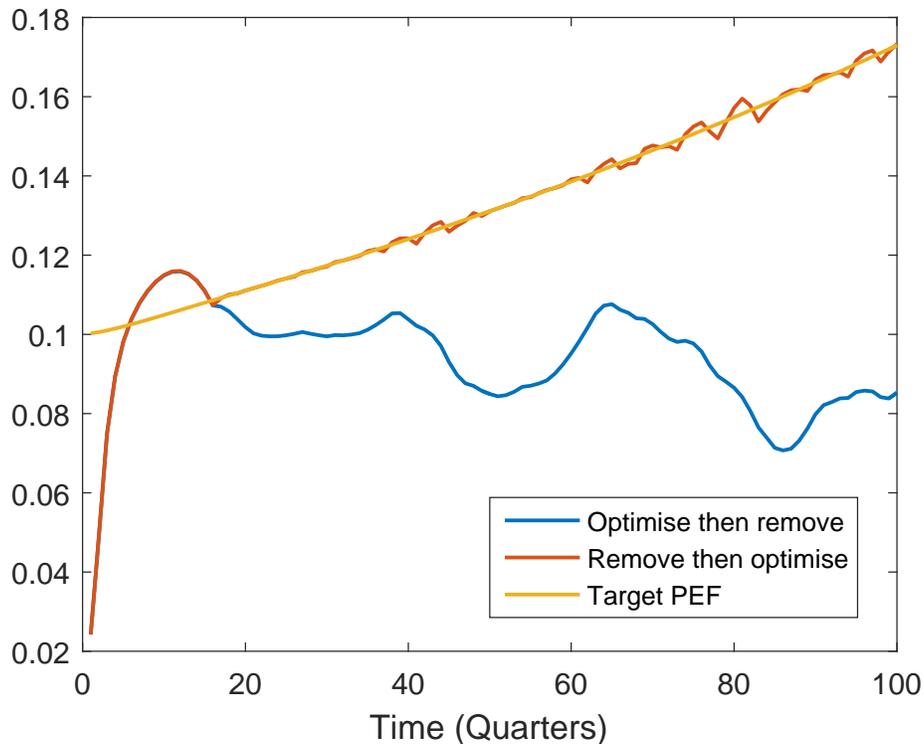


Figure 2.12: Impact of removing funds from certain quarters on actual exposure to private equity, using pre and post removal optimisation. The following model parameters were used: Private Equity Growth = 13%, Public Market Growth = 1%, Number of Funds = 100, Fund Lifetime = 12 Years, Bow Factor = 2.5

Figure 2.12 shows the impact of removing certain funds from the universe of funds. Normally in the model, a new fund is opened every quarter, allowing for a continuous stream of investments to be made if required. In this test, random funds are removed, so in certain quarters no investment can be made.

Two optimisations are run on Figure 2.12. In the first "remove then optimise", the funds are removed from the universe, then the optimisation is run for this new universe to produce the corresponding commitment schedule. There are minor deviations from the target PEF but overall the scheme performs quite well.

In the second optimisation shown in Figure 2.12, "optimise then remove", the optimisation is run first on the full universe, where a fund opens every quarter to obtain a commitment schedule. After this, the same funds are removed as in the

first optimisation. Finally the commitment schedule is applied to the modified universe. This scheme performs extremely poorly, with huge deviations from target. The ramp matches target quite well as the initial commitment is always large, followed by a period of low commitments while the target is reached. This means there is little effect from removing funds in this period. Once the ramp is complete however, larger deviations occur every time a fund is removed.

Overall these plots show that the actual exposure to PEF, is somewhat obviously, very sensitive to the dataset inputted. There could be extensions into the "optimise then remove" algorithm, potentially shifting missed commitments to the next available quarter.

2.7 Calibration

In this section we calibrate the deterministic model to the parameters derived from the data provided. We then backtest it in order to see how well the simulated contributions, distributions and net asset value curves fit the real data. Finally we compare the optimised commitment schedule to see how the actual exposure varies for the simulated and real data set. To be methodologically consistent, we use two thirds of the provided data as a training set while we backtest the performance of our strategies on the remaining third.

The data provided consists of contributions, distributions and net asset value for 1500 buyout funds. These funds are evenly spread over 100 vintage quarters, meaning that there are 15 funds per vintage quarter available. For each fund 100 quarters of contributions, distributions and net asset value data are provided.

To calibrate the (deterministic) Yale model, we first use least squares optimisation and the MATLAB function *fminbd* to find the rate of contributions rc , growth rate of private equity g^{yale} , and bow factor b . We train the model on the average of contributions, distributions and net asset value for each vintage year, as each vintage year was specified to have different dynamics.

For the rate of contribution, it is assumed that similar to Takahashi and Alexander (2002), the returns on capital in the first two years are independent of the subsequent years, which we set to be identical. Using equation (1) the first two equations used for return on capital are

$$rc(t_1) = \frac{c(t_1)}{\text{Commitment}},$$

and

$$rc(t_2) = \frac{c(t_2)}{\text{Commitment} (1 - rc(t_1))},$$

where as before $c(t_i)$, and $rc(t_i)$ are the contributions and rate of contributions at time t_i .

The minimisation problem for the rest of the (identical) return on commitments is,

$$\sum_{i=3}^n \left[rc(t_{3:n}) \text{ Commitment } (1 - rc(t_1)) (1 - rc(t_2)) (1 - rc(t_{3:n}))^{i-3} - c(t_i) \right]^2,$$

where $rc(t_{3:n})$ is a scalar value representing the return on capital on all years after year 2, and n is the number of timesteps (in quarters).

To find the growth rate we use the same method. According to equation (4) the minimisation problem is

$$\sum_{i=2}^n \left[\text{nav}(t_{i-1}) e^{g\Delta} + c(t_i) - d(t_i) \right]^2,$$

where n and $c(t_i)$ are defined as before. g is the growth rate, Δ is the time step, and $d(t_i)$ is the distribution at time t_i .

Finally, to find the bow factor we employ equations (2) and (3) with $y = 0$. Thus the minimisation problem is

$$\sum_{i=2}^n \left[\left(\frac{t_i}{L} \right)^b \text{nav}(t_{i-1}) e^{g\Delta} - d(t_i) \right]^2.$$

Recall that L is the lifetime of the fund, and b is the bow factor.

Performing a hyperparameter search on the data provided shows the goodness of fit cannot be substantially increased by introducing a higher number of modeled rates of contributions.

Once the parameters $rc(t_i)$, g and b are found, the Yale model is used to simulate dynamics of the contributions, distributions and net asset value. These are compared to the actual data below.

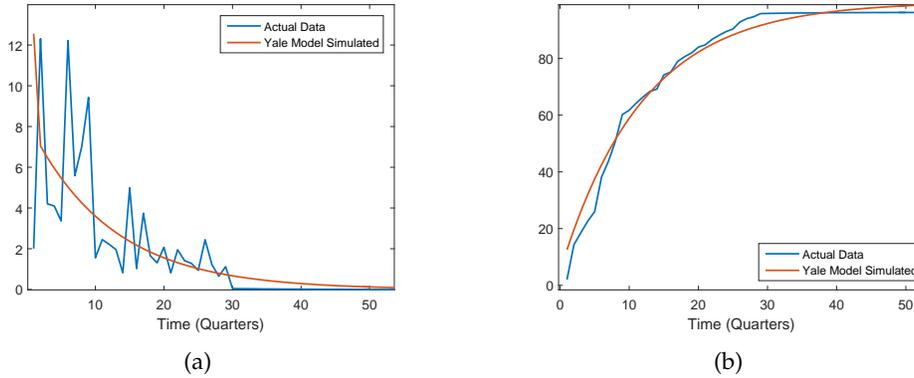


Figure 2.13: (a): Raw contributions for the average of a vintage year from the data, and the correspondingly calibrated Yale model. (b): Cumulative contributions for the average of a vintage year from the data, and the correspondingly calibrated Yale model.

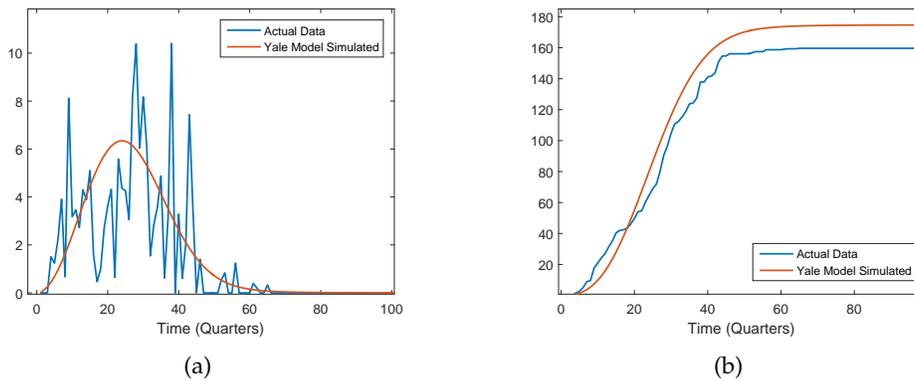


Figure 2.14: (a): Raw distribution for the average of a vintage year from the data, and the correspondingly calibrated Yale model. (b): Cumulative distribution for the average of a vintage year from the data, and the correspondingly calibrated Yale model.

Figure 2.13 shows the plots for the contributions from the average of a single vintage year in the data against the correspondingly calibrated Yale model. In Figure 2.13(a) the raw actual data varies greatly around the simulated Yale model. This is due to the fragmented nature of the actual data, which is a simulation of real private equity contributions.

Figure 2.13(b) is a plot of the cumulative contributions, which shows how well

the Yale model fits the supplied data. In this plot it is visible that the total contributions from the actual data are lower than those from the Yale model. This is due to the nature of the Yale model, which assumes that 100% of the commitment is drawn down by the GP, whereas in the real data an average of 96% of the initial commitment is drawn.

Figure 2.14 shows similar plots to Figure 2.13 for the distributions. Again, the choppiness of the raw actual data in comparison to the Yale model is apparent, but the cumulative distributions show a much better comparison as before.

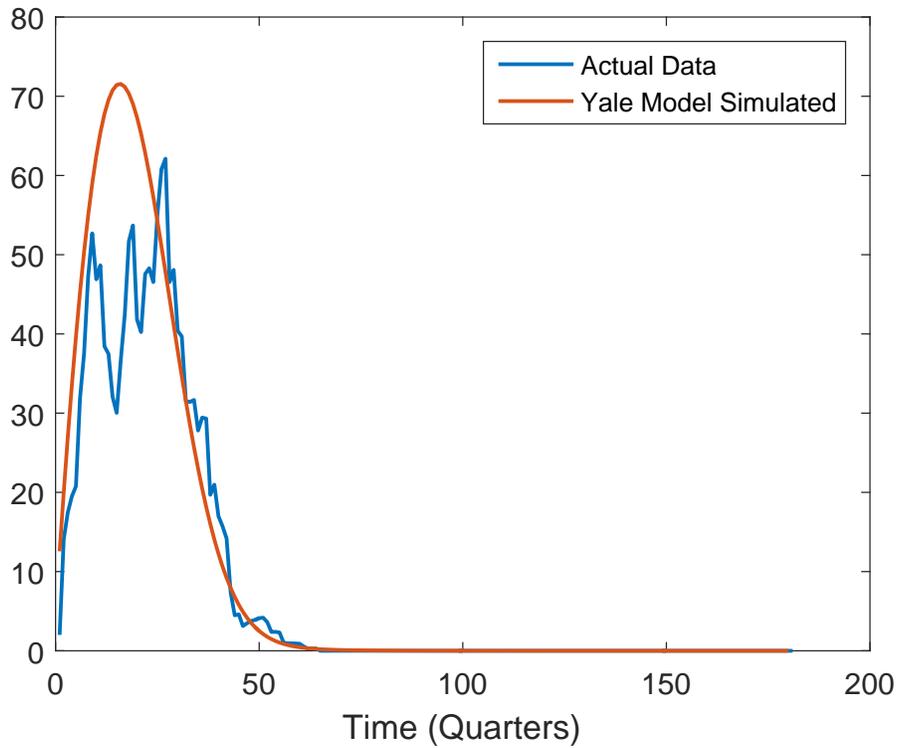


Figure 2.15: Raw net asset value for the average of a vintage year from the data, and the correspondingly calibrated Yale model.

Figure 2.15 shows the raw net asset value for the actual data in comparison to the calibrated Yale model. Similarly to the contribution and distribution plots, the main difference between the actual and simulated data is the fragmented nature of the actual data. In this graph the difference in peak values is also visible. This difference can mostly be accounted for by the way the data values were calibrated, i.e. by minimising the squared value between the two curves and the fact that the

net asset value in the Yale model is unimodal.

We now test the commitment scheduling optimiser for the raw data and calibrated Yale model.

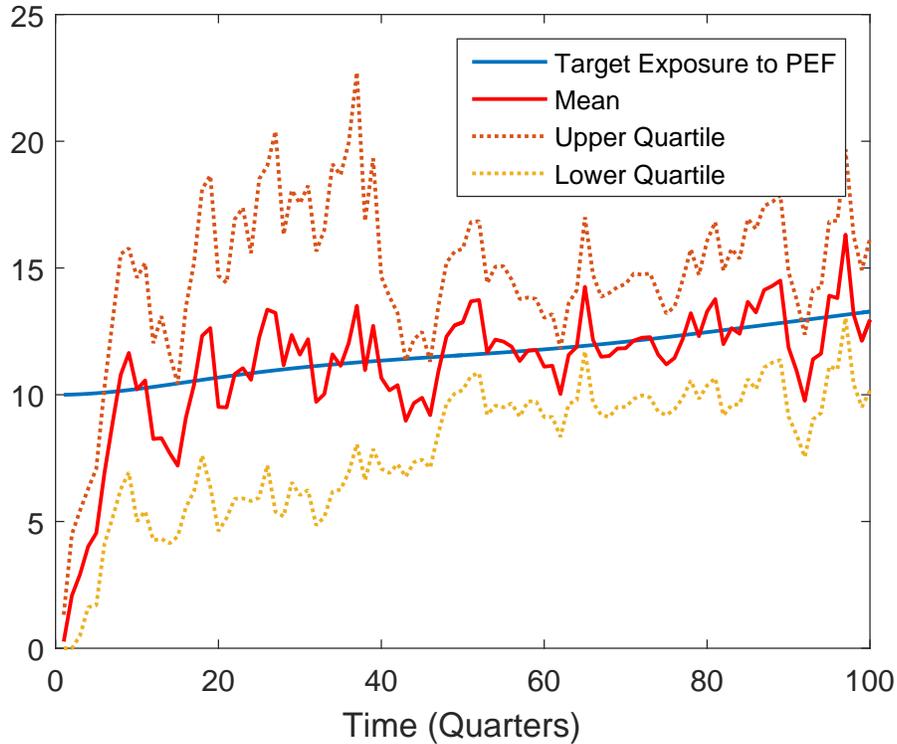


Figure 2.16: Actual exposure to PEF for the mean of the data, with commitment schedule optimised on the data. The upper (90%) and lower (10%) quartiles are also plotted for the data

Figure 2.16 shows the actual exposure to PEF for the mean of the data. The mean of the data is the average contributions, distributions and returns for each vintage year. The optimal commitment schedule is obtained from this and then used to plot the mean exposure to PEF. It is apparent that even the mean of the data produces a much higher variance of exposure than the Yale model examples shown earlier.

Figure 2.16 also shows the 10% and 90% quartiles of the data. These are computed by taking random combinations of funds from each quarter to create a universe of one fund per quarter, then applying the commitment schedule obtained from the

mean of all funds to each of these sample universes. The quantiles are then computed from these sample universes.

The range of the quantiles varies greatly, with the greatest amount in roughly the first 40 quarters of the data. This can be attributed to the difference between contributions and distributions at each time step, as more funds are opening each quarter. Once 40 quarters have passed, the number of funds opening and closing remains constant, producing a "steady state" of contributions and distributions, which allows for the confidence intervals to reduce in size.

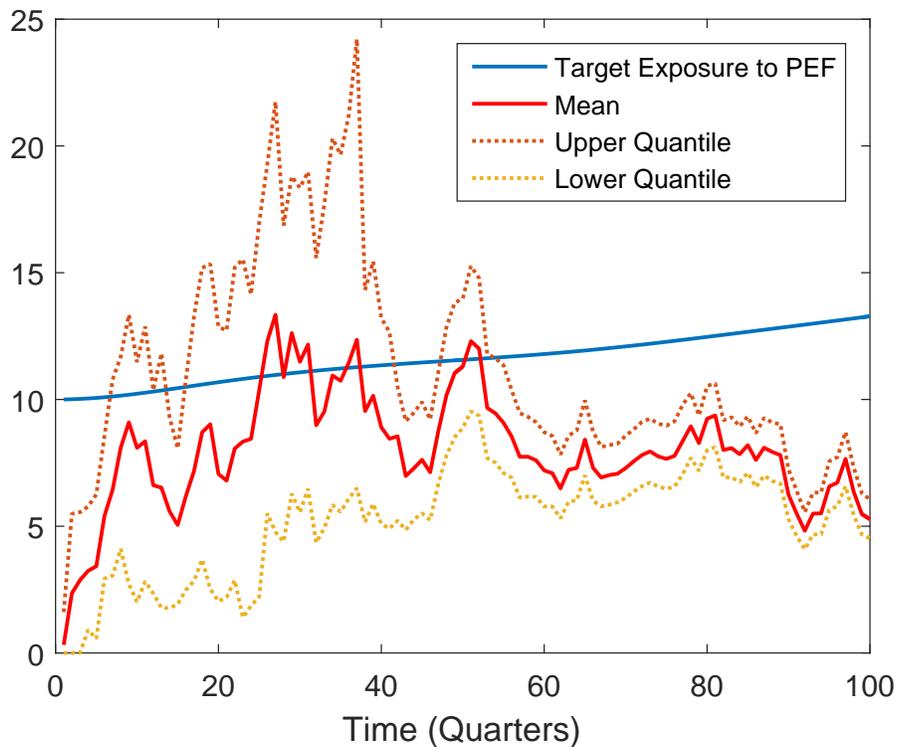


Figure 2.17: Actual exposure to PEF for the mean of the data with commitment schedule optimised on the calibrated Yale model. The upper (90%) and lower (10%) quantiles are also plotted for the data

Figure 2.17 shows the mean exposure to PEF using the commitment schedule optimised on the calibrated Yale model. This commitment schedule is then applied to the fund net asset values from the real data. It is apparent that while initially the commitment schedule reaches and loosely matches the target exposure, it has very high variance and wide confidence intervals. As time progresses the calibrated

model matches the real data less closely leading to a large deviation between target and actual exposure to PEF. We conjecture that the reason for this is the finite life time of funds assumed for the Yale model, which does not seem to be reflected in the data.

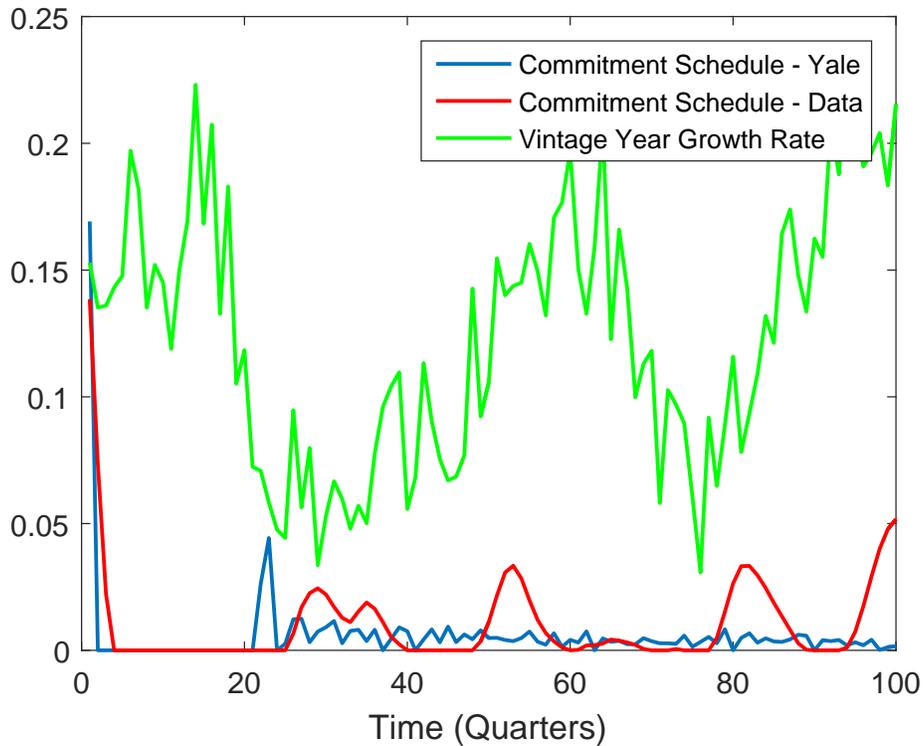


Figure 2.18: Optimal commitment schedule for the mean of the data and correspondingly calibrated Yale model. The growth rate for each vintage quarter is also plotted.

Figure 2.18 shows the optimal commitment schedule for the mean of the data and the correspondingly calibrated Yale model. The optimiser on the data has a large smoothing factor as described in Section 2.3, as without it the optimal commitment schedule is extremely rough. It is clear that even with the smoothing factor, the optimiser for the data produces much larger and less frequent commitments than the Yale model. This is again due to the rough nature of the real data, compared with the Yale model which produces smooth outputs for contributions and distributions.

In Figure 2.18 it is also evident that the optimiser tends to increase the size and frequency of commitments when the growth rate of the vintage quarters declines.

While this produces more optimal actual exposure to PEF, it may not be applicable as the optimiser is essentially anticipating future returns of each vintage quarter. This is explored in more detail in Figure 2.19, which plots the anticipative and non-anticipate exposure to PEF for the mean of the data. The anticipative line shows the exposure to PEF if the optimiser is able to know the return of each vintage year at the beginning of that vintage. The non-anticipative line assumes each vintage year follows the mean dynamics of all the funds, across vintages. As expected, the optimal schedule for anticipative data produces better actual exposure to PEF, with less variance from the target line than the corresponding non-anticipative schedule.

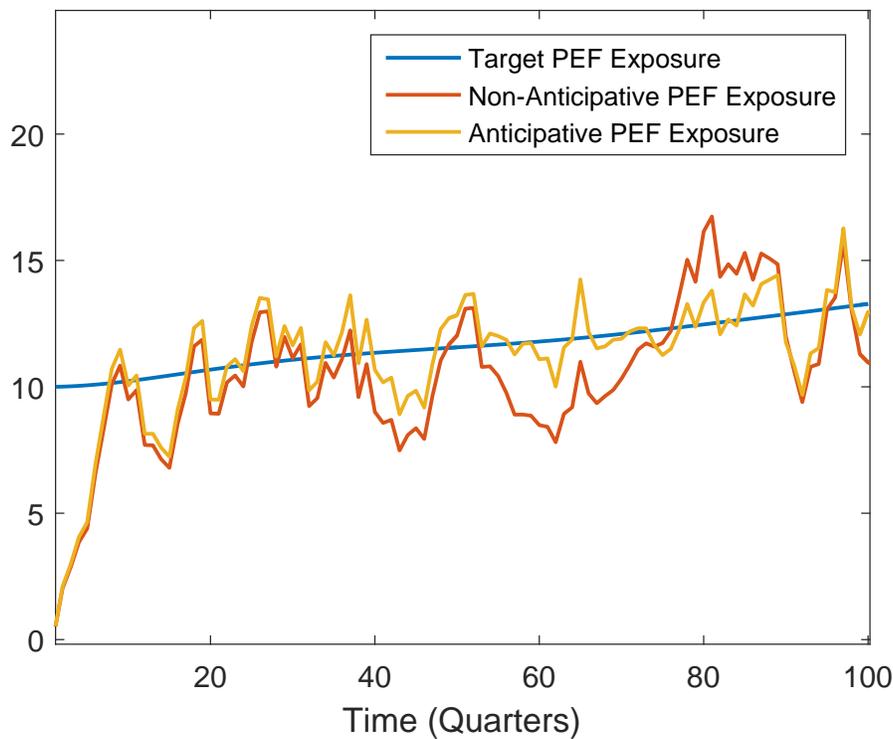


Figure 2.19: The anticipative and non-anticipate exposure to PEF for the mean of the data.

3 Continuous-Time Stochastic Modelling

In this section we detail the stochastic modelling of private equity funds (PEFs), and follow the approach taken by de Malherbe (2004), in which a three stage model is discussed: the rate of contributions and distributions are respectively modeled

by a squared Bessel process, while the net asset value (NAV) is modeled as a log-normal process. We retain some of the notation used in (de Malherbe, 2004) to avoid possible confusion with the deterministic model (see Section 2).

We begin with the simple case of modelling a single PEF, and later generalise this model to a collateralised fund obligation (CFO) of PEFs of the same type (e.g. PEFs in the same sector etc.). Derivations of the results below may be found in (de Malherbe, 2004).

3.1 The de Malherbe Model

We assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ large enough to support all stochastic processes defined in this section, and define the natural filtration satisfying the usual conditions. The fund in question has a contribution period T_* , maturity T and current time t . It has contributions and distributions payable continuously at rates δ_t and ρ_t , respectively, with immediate investment of available funds. Due to the peculiar nature of PEFs, in addition to the distribution rate ρ_t , the contribution rate δ_t will be decided by the fund manager.

Without loss of generality we consider initial commitments equal to one with zero initial contributions. Cumulative contributions up to t are denoted by C_t . Throughout the contribution period, contributions are payable at rate δ_t on the undrawn commitment amount $1 - C_t$ so that

$$dC_t = \delta_t(1 - C_t)1_{\{t \leq T_*\}}dt.$$

This is an ordinary differential equation which can be solved to yield

$$C_t = 1 - e^{-\int_0^{t \wedge T_*} \delta_s ds}.$$

For a finite T_* , the functional form of C implies that $C_{T_*} < 1$ (i.e. a portion of the committed amount is not drawn by the manager), which is not a problem since a large contribution period or a large contribution rate yields $C_{T_*} \approx 1$.

Similar to the above, we denote cumulative distributions up to t by D_t . Throughout the life of the fund, distributions are payable at rate ρ_t on the net asset value V_t so that

$$R_t = \int_0^t \rho_s V_s ds + V_T 1_{\{t=T\}}.$$

The indicator term allows for a final distribution of the assets in the fund at maturity, which leads to a possible jump in the process.

Changes in the portfolio net asset value are given by the return earned on the private equity investment, inflow of contributions into the fund, and outflow of distributions from the fund. Hence for $t < T$, we model the portfolio NAV as

$$dV_t = \mu_t V_t dt + \sigma_t V_t dW_t^V + dC_t - dD_t.$$

Upon using the dynamics of contributions and distributions above, a closed-form expression for V_t can be obtained by solving a linear SDE. We refer again to (de Malherbe, 2004) for technical details.

To model the uncertainty in contributions and distributions, we utilize two independent standard Bessel processes with the following dynamics:

$$\begin{aligned} d\delta_t &= (c_1 + c_2\delta_t)dt + c_3\sqrt{\delta_t}dW_t^\delta, \\ d\rho_t &= (q_1 + q_2\rho_t)dt + q_3\sqrt{\rho_t}dW_t^\rho. \end{aligned}$$

where c_1, c_3, q_1, q_3 are positive constants. To ensure that the rates δ_t and ρ_t remain positive, we impose the Feller condition on these parameter constants so that $c_1 > c_3^2/2$ and $q_1 > q_3^2/2$. This will prove important for the discretisation schemes we use later.

Below we plot the first moments over time of the random variables defined in Sections 3.1.1-3.1.3 above.

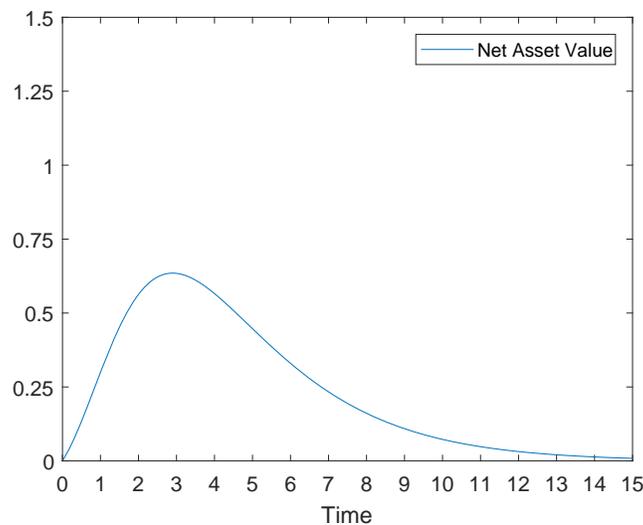


Figure 3.1: Expected net asset value over time under the stochastic model with parameters $\delta_0 = 0.2$, $c = (0.4, -0.1, 0.4)$, $\rho_0 = 0.2$, $q = (0.1, -0.1, 0.15)$, $\mu = 0.15$, $\sigma = 0.25$

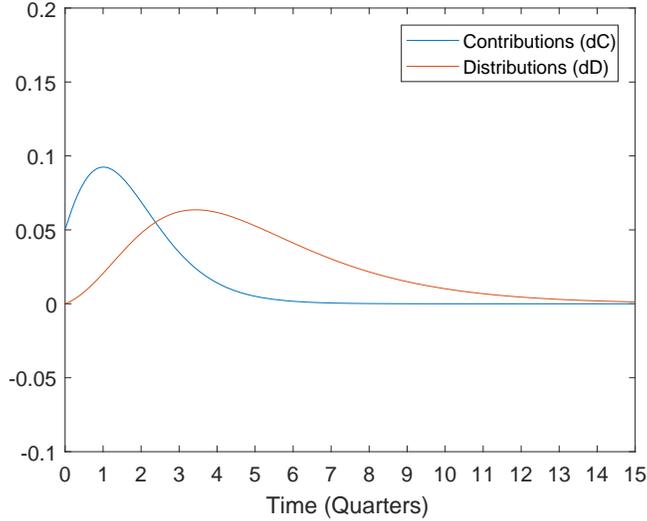


Figure 3.2: Expected contributions and distributions over time under the stochastic model with parameters $\delta_0 = 0.2$, $c = (0.4, -0.1, 0.4)$, $\rho_0 = 0.2$, $q = (0.1, -0.1, 0.15)$, $\mu = 0.15$, $\sigma = 0.25$

The shape of the NAV plot in Figure 3.1 is typical for PEFs because the earlier years are associated to the contribution period, and later years with the distribution period. Figure 3.2 illustrates this situation more clearly where contributions exceed distributions in the first three years, and the reverse holds in the later years. The net effect is a negative cashflow within the first three years, and a positive cashflow out of the fund in later years.

We now generalise the above setting to the case of a portfolio of PEFs of the same type, namely buyout funds and make the following additional assumptions:

- The contribution and distribution rates of all funds are subject to a systemic random component together with a specific random component.
- The contribution and distribution rates of all funds have the same parameters, namely the vectors (c_1, c_2, c_3) and (q_1, q_2, q_3) respectively. We further assume these parameters satisfy the Feller condition.
- The parameters of fund performance are identical across all funds, and these are μ and σ .

Now assuming that the above assumptions hold, for $T_{\bullet}^{(j)} \leq t < T^{(j)}$, the dynamics

of fund $j \in \{1, \dots, m\}$ are given by

$$\begin{aligned}
dV_t^{(j)} &= [\delta_t^{(j)}(1 - C_t^{(j)})1_{\{t \leq T_*^{(j)}\}} + (\mu - \rho_t^{(j)})V_t^{(j)}]dt + \sigma V_t^{(j)}dW_t^{\mu(j)}, \\
d\delta_t^{(j)} &= (c_1^{(j)} + c_2^{(j)}\delta_t^{(j)})dt + c_3^{(j)}\sqrt{\delta_t^{(j)}}\left(\sqrt{1 - \beta^2}dW_t^{\delta(j)} + \beta dW_t^\delta\right), \\
d\rho_t^{(j)} &= (q_1^{(j)} + q_2^{(j)}\rho_t^{(j)})dt + q_3^{(j)}\sqrt{\rho_t^{(j)}}\left(\sqrt{1 - \gamma^2}dW_t^{\rho(j)} + \gamma dW_t^\rho\right),
\end{aligned} \tag{9}$$

where m is the total number of funds in the portfolio; $W^\delta, W^\rho, \{W^{\mu(j)}\}_{j=1}^m, \{W^{\delta(j)}\}_{j=1}^m$ and $\{W^{\rho(j)}\}_{j=1}^m$ are independent Brownian motions; $T_\bullet^{(j)}$ is the investment period start date, $T_*^{(j)}$ is the investment period end date, and $T^{(j)}$ is the fund's maturity date. The correlation between contribution and distribution policies of different funds is reflected by the β and γ parameters respectively.

To simulate the CFO model, for each fund $j \in \{1, \dots, m\}$, we utilize a second-order Taylor approximation of the SDEs above using the reflected Milstein scheme. Given initial conditions $\{V_{T_\bullet^{(j)}}^{(j)}\}_{j=1}^m, \{\delta_{T_\bullet^{(j)}}^{(j)}\}_{j=1}^m$ and $\{\rho_{T_\bullet^{(j)}}^{(j)}\}_{j=1}^m$ for times $\{t_i\}_{i=1}^n$ and uniform time step $\Delta = t_i - t_{i-1}$ the Milstein scheme yields an approximate solution to (9) given by

$$\begin{aligned}
V_{t_{i+1}}^{(j)} &= V_{t_i}^{(j)} + [\delta_{t_i}^{(j)}(1 - C_{t_i}^{(j)})1_{\{t \leq T_*^{(j)}\}} + (\mu - \rho_{t_i}^{(j)})V_{t_i}^{(j)}]\Delta + \sigma V_{t_i}^{(j)}\sqrt{\Delta}Z_{i+1}^{\mu(j)} \\
&\quad + \frac{1}{2}\sigma^2 V_{t_i}^{(j)}\Delta \left[\left(Z_{i+1}^{\mu(j)} \right)^2 - 1 \right] \\
\delta_{t_{i+1}}^{(j)} &= \left| \delta_{t_i}^{(j)} + (c_1 + c_2\delta_{t_i}^{(j)})\Delta + c_3\sqrt{\delta_{t_i}^{(j)}}\Delta \left(\sqrt{1 - \beta^2}Z_{i+1}^{\delta(j)} + \beta Z_{i+1}^\delta \right) \right. \\
&\quad \left. + \frac{1}{4}c_3^2\Delta \left[\left(\sqrt{1 - \beta^2}Z_{i+1}^{\delta(j)} + \beta Z_{i+1}^\delta \right)^2 - 1 \right] \right| \\
\rho_{t_{i+1}}^{(j)} &= \left| \rho_{t_i}^{(j)} + (q_1 + q_2\rho_{t_i}^{(j)})\Delta + q_3\sqrt{\rho_{t_i}^{(j)}}\Delta \left(\sqrt{1 - \gamma^2}Z_{i+1}^{\rho(j)} + \gamma Z_{i+1}^\rho \right) \right. \\
&\quad \left. + \frac{1}{4}q_3^2\Delta \left[\left(\sqrt{1 - \gamma^2}Z_{i+1}^{\rho(j)} + \gamma Z_{i+1}^\rho \right)^2 - 1 \right] \right|,
\end{aligned} \tag{10}$$

provided the times points satisfy $T_\bullet^{(j)} \leq t_i < T^{(j)}$. Outside this interval the three processes vanish to zero. The sets $\{Z_i^\delta\}, \{Z_i^\rho\}, \{Z_i^{\mu(j)}\}, \{Z_i^{\delta(j)}\}$ and $\{Z_i^{\rho(j)}\}$ are independent sets of simulated standard normal variables. Now even though the c 's and q 's may satisfy the Feller condition, it is still theoretically possible that the resultant rates δ and ρ are negative for some time points due to discretisation error, at least in the case when Δ is quite large. Of course when Δ is "small enough",

this will lead to more accurate results and decrease the probability of ever hitting the boundary zero; however, this comes at a computational cost. An alternative solution to this problem, termed the reflection scheme, which we have undertaken, amounts to simply reflecting the negative rates as illustrated by the absolute value. See Diop (2004) for a detailed analysis of the reflection scheme, together with its convergence properties.

3.2 The stochastic optimisation problem

To find the optimal commitment schedule in the stochastic model, we simulate N realisations of the portfolio of m funds. This yields the two sets $\{H_l\}_{l=1}^N$ and $\{f_l\}_{l=1}^N$ where

$$\begin{aligned} H_l &:= 2(\text{nav}_l \cdot \text{nav}_l^T + 2\pi \text{nav}_l \cdot C_l^T - \pi^2 C_l \cdot C_l^T), \\ f_l &:= 2\pi(p(t_0)_l - v(t_0)_l)k^T + \pi^2(p(t_0)_l - v(t_0)_l)C_l \cdot k^T, \end{aligned}$$

where we use the same notation as in Section 2.2 and denote the dependence on the realisations by the subscript l . The optimal commitment schedule a is then found by minimising

$$\frac{1}{2} a \cdot \hat{H} \cdot a^T + a \cdot \hat{f} + \text{const.}$$

subject to $a_j \geq 0$ for all $j \in \{1, \dots, m\}$, where

$$\begin{aligned} \hat{H} &:= \frac{1}{N} \sum_{l=1}^N H_l, \\ \hat{f} &:= \frac{1}{N} \sum_{l=1}^N f_l. \end{aligned}$$

3.3 Implementation

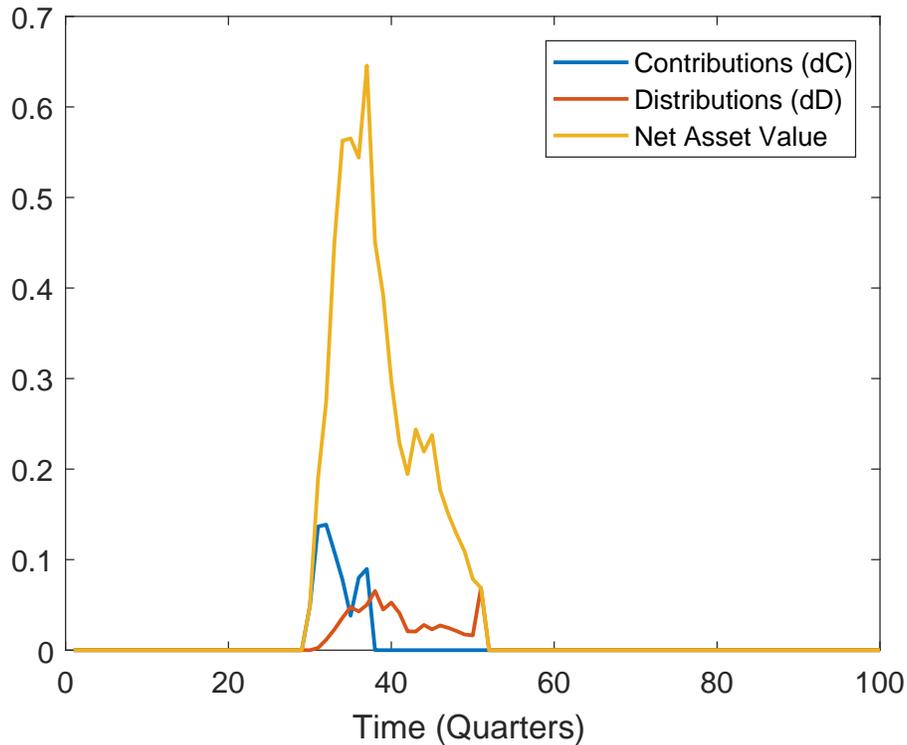


Figure 3.3: Sample dynamics of contributions, distributions and net asset value of a simulated fund with $T_{\bullet} = 30$, $T_{\star} = 37$, $T = 51$, $\delta_0 = 0.2$, $c = (0.35, -0.1, 0.6)$, $\rho_0 = 0.2$, $q = (0.1, -0.1, 0.15)$, $\mu = 0.08$, $\sigma = 0.3$, $\beta = 0.1$, $\gamma = 0.15$, $m = n = 100$, $L = 100$

Figure 3.3 above displays a sample path for contributions, distributions and NAV for a fund simulated using the parameters given in the figure. The shapes of the plots above are similar to those shown in Figure 3.1. A visible difference is the roughness of the plots in Figure 3.3, which is not surprising since only one sample path was taken. A smoother plot can be obtained by averaging different sample paths. The sharp rise in NAV is due to the high initial commitments that are required in order to achieve a 10% portfolio exposure.

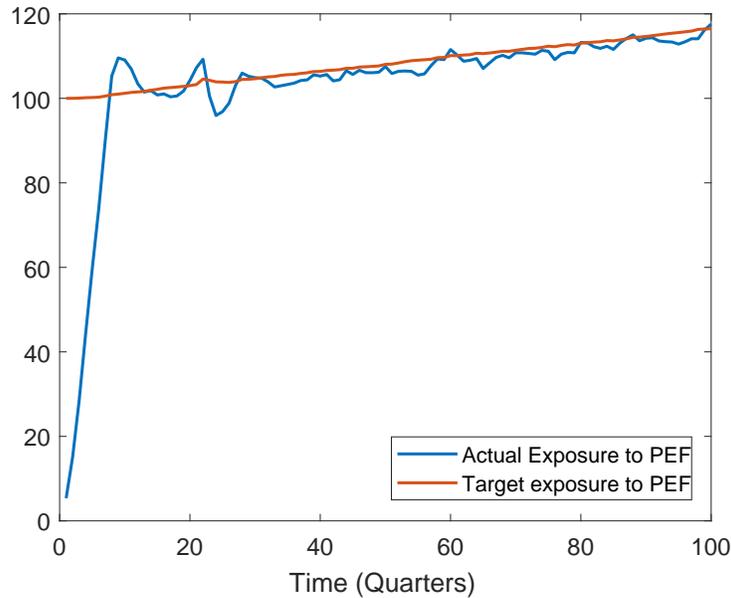


Figure 3.4: Average actual exposure to PEF vs average target exposure to PEF under the stochastic model with parameters $\delta_0 = 0.2$, $c = (0.4, -0.1, 0.4)$, $\rho_0 = 0.2$, $q = (0.1, -0.1, 0.15)$, $\mu = 0.15$, $\sigma = 0.25$

Similarly to Figure 2.4 from the Yale Model, the actual exposure to PE ramps up quickly and overshoots the target exposure due to the high initial commitments required to achieve the 10% portfolio exposure. A striking difference is the convergence speed of the two strategies: In the stochastic case, convergence is rather slow and tends to oscillate about the target exposure. The oscillatory nature is due to the presence of a stochastic component in the dynamics of contributions, distributions and NAV. Note also that Figure 3.4 shows expected values while we are minimising the squared distance between the random variables. Thus the Bias-Variance tradeoff introduces an additional error term, which explains the disparity of convergence compared with the deterministic case shown in Figure 2.4.

3.4 Sensitivity analysis

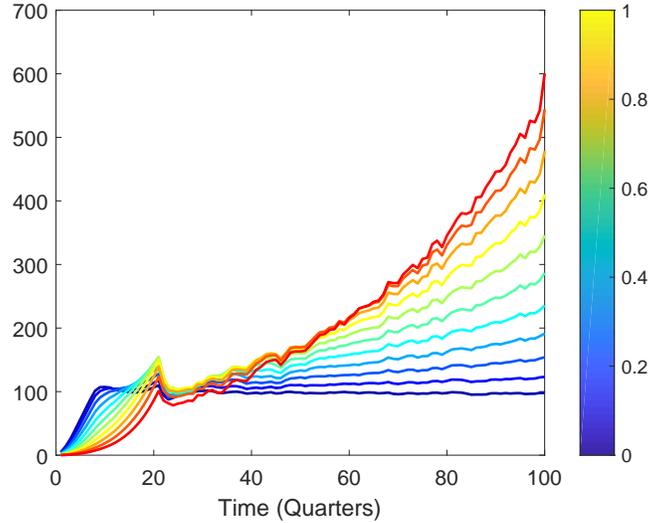


Figure 3.5: Effect of varying the drift parameter μ on the actual exposure to PEF under the stochastic model with parameters $\delta_0 = 0.2$, $c = (0.4, -0.1, 0.4)$, $\rho_0 = 0.2$, $q = (0.1, -0.1, 0.15)$, $\sigma = 0.25$, $\beta = 0.1$, $\gamma = 0.15$, $m = n = 100$, $L = 100$

With a higher drift parameter μ , the portfolio of PEFs yields more value per unit invested. This results in a higher total portfolio value, and hence a higher investment in private equity would be required to maintain the 10% target exposure.

3.5 Calibration

In order to calibrate the stochastic model (9) we estimate the parameters

$$\{c_1, c_2, c_3, q_1, q_2, q_3, \mu, \sigma, \beta, \gamma\}$$

by the method of moments. For this we deduce the following formulae for $j \neq k \in \{1, \dots, m\}$ assuming integrability of the processes:

$$\begin{aligned}
\mathbb{E}(dV_t^{(j)}) &= \mathbb{E}(dD_t^{(j)}) + \mathbb{E}((\mu - \rho_t^{(j)})V_t^{(j)} dt) \\
\mathbb{E}(d\delta_t^{(j)}) &= \mathbb{E}((c_1 + c_2\delta_t^{(j)})dt) \\
\mathbb{E}(d\rho_t^{(j)}) &= \mathbb{E}((q_1 + q_2\delta_t^{(j)})dt) \\
\mathbb{E}\left(dV_t^{(j)}dV_t^{(k)}\right) &= \sigma^2\mathbb{E}(V_t^{(j)}V_t^{(k)} dt) \\
\mathbb{E}\left(d\delta_t^{(j)}d\delta_t^{(k)}\right) &= c_3^2\beta^2\mathbb{E}\left(\sqrt{\delta_t^{(j)}\delta_t^{(k)}} dt\right) \\
\mathbb{E}\left(d\rho_t^{(j)}d\rho_t^{(k)}\right) &= q_3^2\rho^2\mathbb{E}\left(\sqrt{\delta_t^{(j)}\delta_t^{(k)}} dt\right) \\
\mathbb{E}\left(\left(d\delta_t^{(j)}\right)^2\right) &= (c_3)^2\mathbb{E}\left(\delta_t^{(j)} dt\right) \\
\mathbb{E}\left(\left(d\rho_t^{(j)}\right)^2\right) &= (q_3)^2\mathbb{E}\left(\rho_t^{(j)} dt\right)
\end{aligned}$$

Discretising the equations above yields the following estimators:

$$\begin{aligned}
\hat{c}_3 &:= \sqrt{\frac{1}{nm} \sum_{j=1}^m \sum_{i=1}^{n-1} \frac{\mathbb{E}((\delta_{t_{i+1}}^{(j)})^2) - \mathbb{E}((\delta_{t_i}^{(j)})^2)}{\mathbb{E}(\delta_{t_i}^{(j)})\Delta}} \\
\hat{q}_3 &:= \sqrt{\frac{1}{nm} \sum_{j=1}^m \sum_{i=1}^{n-1} \frac{\mathbb{E}((\rho_{t_{i+1}}^{(j)})^2) - \mathbb{E}((\rho_{t_i}^{(j)})^2)}{\mathbb{E}(\rho_{t_i}^{(j)})\Delta}} \\
\hat{\mu} &:= \frac{1}{nm} \sum_{j=1}^m \sum_{i=1}^{n-1} \frac{\mathbb{E}(V_{t_{i+1}}^{(j)}) - \mathbb{E}(V_{t_i}^{(j)}) - (\mathbb{E}(V_{t_{i+1}}^{(j)}) - \mathbb{E}(V_{t_i}^{(j)})) + \mathbb{E}(\rho_{t_i}^{(j)}V_{t_i}^{(j)})\Delta}{\Delta\mathbb{E}(V_{t_i}^{(j)})} \\
\hat{\sigma}^2 &:= \frac{1}{nm} \sum_{j=1}^m \sum_{i=1}^{m-1} \frac{\mathbb{E}(V_{t_{i+1}}^{(j)}V_{t_{i+1}}^{(k)}) - \mathbb{E}(V_{t_i}^{(j)}V_{t_i}^{(k)})}{\Delta\mathbb{E}(V_{t_i}^{(j)}V_{t_i}^{(k)})} \\
\hat{\beta} &:= \sqrt{\frac{1}{nm^2} \sum_{j,k=1}^m \sum_{i=1}^{n-1} \frac{\mathbb{E}\left(\delta_{t_{i+1}}^{(j)}\delta_{t_{i+1}}^{(k)}\right) - \mathbb{E}\left(\delta_{t_i}^{(j)}\delta_{t_i}^{(k)}\right)}{c_3^j c_3^k \mathbb{E}\sqrt{\delta_{t_i}^{(j)}\delta_{t_i}^{(k)}}\Delta}} \\
\hat{\gamma} &:= \sqrt{\frac{1}{nm^2} \sum_{j,k=1}^m \sum_{i=1}^{n-1} \frac{\mathbb{E}\left(\rho_{t_{i+1}}^{(j)}\rho_{t_{i+1}}^{(k)}\right) - \mathbb{E}\left(\rho_{t_i}^{(j)}\rho_{t_i}^{(k)}\right)}{c_3^j c_3^k \mathbb{E}\sqrt{\rho_{t_i}^{(j)}\rho_{t_i}^{(k)}}\Delta}}
\end{aligned}$$

We are left with determining estimators for the parameters c_1, c_2, q_1, q_2 . We note

that

$$d\mathbb{E}(\delta_t^{(j)}) = \mathbb{E}((c_1^{(j)} + c_2^{(j)} \delta_t^{(j)}))dt$$

has the solution

$$\mathbb{E}(\delta_t^{(j)} - \delta_u^{(j)}) = \int_u^t c_1 e^{c_2(t-s)} ds = \frac{c_1}{c_2} (e^{c_2(t-u)} - 1).$$

Thus determining c_1 and c_2 can be carried out jointly by solving the equations

$$\begin{aligned} \mathbb{E}(\delta_{t_{i+1}}^{(j)}) - \mathbb{E}(\delta_{t_i}^{(j)}) &= \frac{c_1}{c_2} (e^{\Delta c_2} - 1) \\ \mathbb{E}(\delta_{t_{i+2}}^{(j)}) - \mathbb{E}(\delta_{t_i}^{(j)}) &= \frac{c_1}{c_2} (e^{2\Delta c_2} - 1), \end{aligned}$$

which yield the equations

$$c_1 = \frac{c_2 \left(\mathbb{E}(\delta_{t_{i+1}}^{(j)}) - \mathbb{E}(\delta_{t_i}^{(j)}) \right)}{e^{2\Delta c_2} - 1}, \quad \frac{\mathbb{E}(\delta_{t_{i+2}}^{(j)}) - \mathbb{E}(\delta_{t_i}^{(j)})}{\mathbb{E}(\delta_{t_{i+1}}^{(j)}) - \mathbb{E}(\delta_{t_i}^{(j)})} = \frac{e^{2\Delta c_2} - 1}{e^{\Delta c_2} - 1}. \quad (11)$$

Equations (11) can be solved numerically using the “fmin” function in MATLAB. As the provided data is relatively sparse, an appropriate winsorisation has to be applied to account for non-positive entries in the computation of these estimators. We obtain the values given in Figure 3.5, which we compare to de Malherbe (2004). We also bootstrap 10% confidence intervals. From the table it is evident that our calibrated values have the same order of magnitude as the one obtained by de Malherbe (2004). Strikingly the values obtained for σ are quite high, which makes an estimation of μ using method of moments nearly impossible.

Figure 3.6: Comparison of calibration results of the stochastic model for the provided data

Parameter	Model calibration	Confidence intervals	de Malherbe (2004)
c_1	3.156	(0.942, 5.100)	2.821
c_2	-13.479	(-18.494, -7.951)	-8.740
c_3	1.785	(1.293, 2.025)	1.463
q_1	2.582	(1.234, 4.147)	3.508
q_2	-20.945	(-27.337, -13.997)	-17.468
q_3	1.693	(1.390, 1.918)	1.929
μ	0.002	(-0.861, 0.929)	0.043
σ	0.421	(0, 1.704)	0.293
β	0.624	(0.0148, 0.718)	NA
γ	0.265	(0.0244, 1.265)	NA

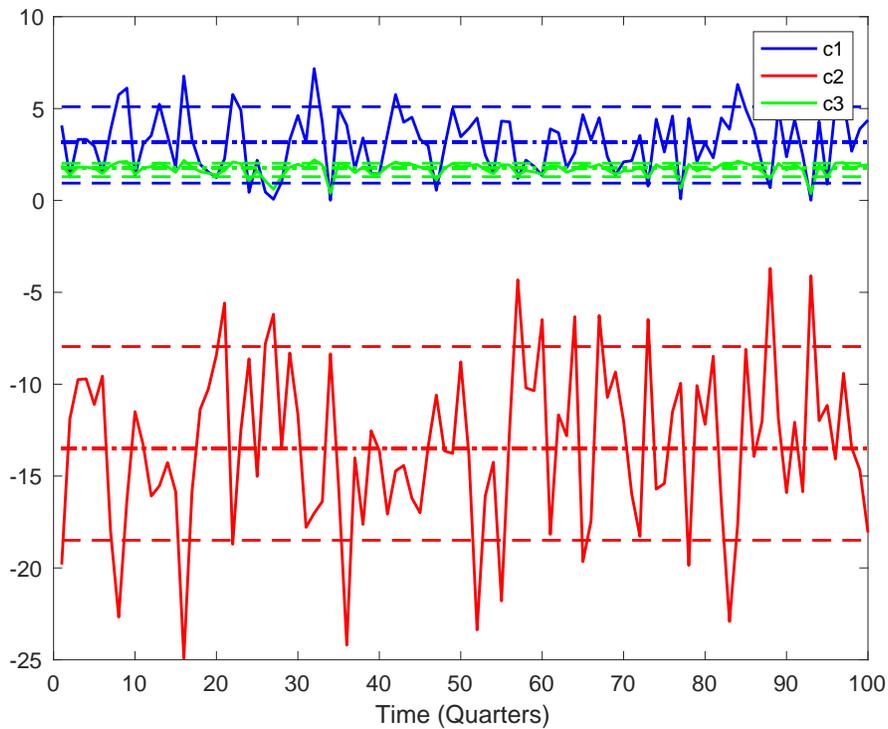


Figure 3.7: Calibration of c_1, c_2, c_3 for the stochastic model

Figure 3.7 shows different values of (c_1, c_2, c_3) calibrated for the different vintages individually. Contrary to the deterministic case, where a relationship between the

optimal commitment schedule $a(t_i)$ and the growth rate for the different funds was evident, we cannot observe such a relationship between the optimal commitment schedule and the different values for $c = (c_1, c_2, c_3)$ in this case. It thus seems justified to assume that the values of c do not depend on the vintage of the fund. Figure 3.8 shows the calibrated distributions, contributions and net asset values for a specific fund. Evidently the performance of the fitting is not convincing. This might be the reason why de Malherbe (2004) opts for more sophisticated estimation procedure: In fact they use maximum likelihood estimation adapted to the Milstein discretisation of the CIR processes.

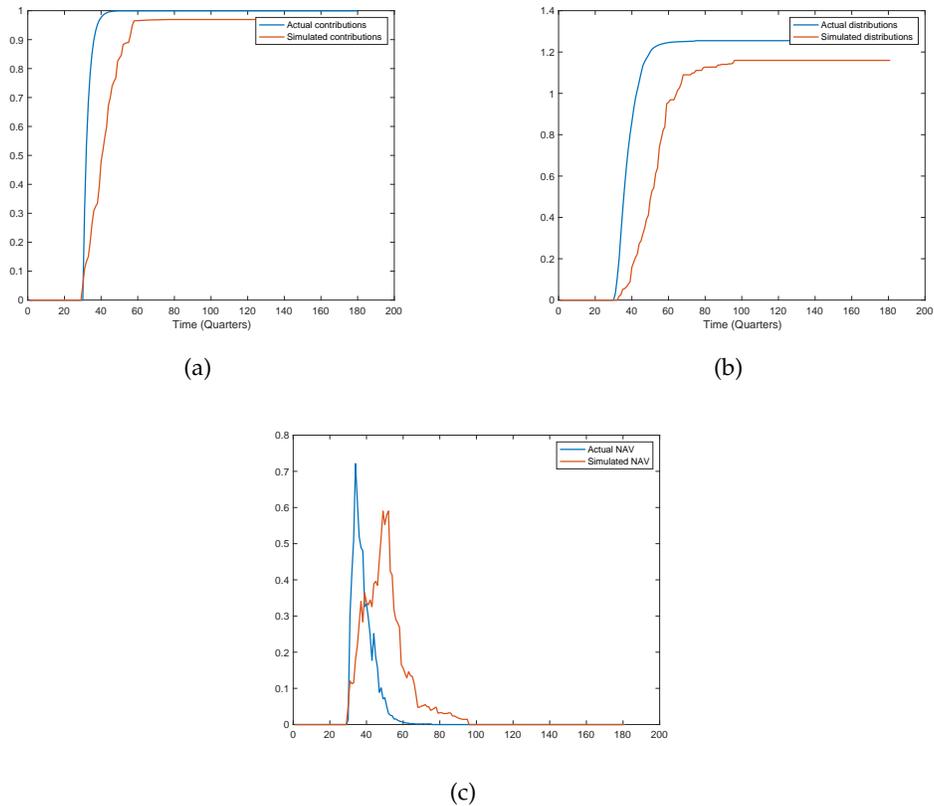


Figure 3.8: Comparison of fitted contributions (a), distributions (b) and net asset values (c) according to the stochastic model and contributions, distributions and net asset values provided in the data for one vintage

Lastly a comparison of the optimal commitment strategies according to the deterministic Model and stochastic model is presented in Figure 3.9. The characteristic spike of the commitment rate according to the deterministic model yields a lower overall level reached in the long run compared to the stochastic model. Both

optimal commitment schedules show the same seasonality, albeit it is more pronounced in the stochastic model.

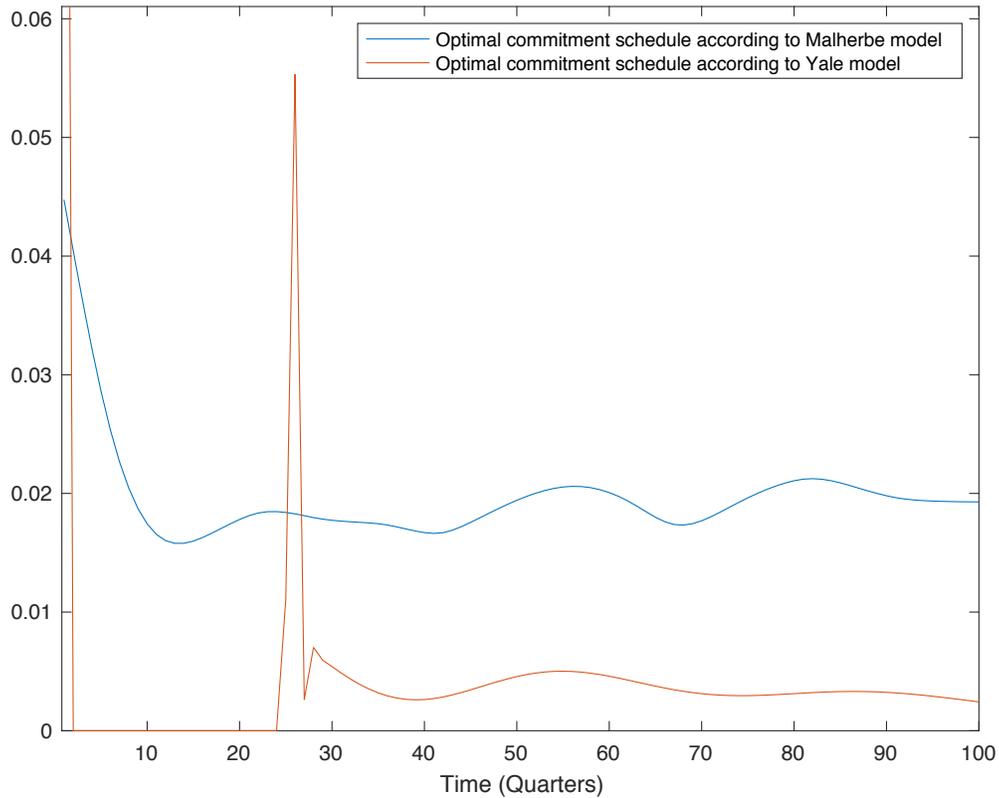


Figure 3.9: Comparison of the optimal commitment schedule of the deterministic and the stochastic model calibrated to the data provided.

4 Conclusion

In this report we introduce the methodology and state of the art valuation tools of the private equity industry. Furthermore we elaborate on the most common practices employed. An understanding of the operational features of a particular fund is key to modelling its cash flow dynamics. The number of capital commitments is contractually agreed upon at the closing, but is only partially drawn down during the investment period. Furthermore, the timing and size of distributions is uncertain, while the value of the private equity stake is dynamic and not reflected in current market prices. This makes the construction of a model that accurately

captures the dynamics of the contributions, distributions and net asset value of a private equity portfolio quite intricate.

In the first part of this report we discuss the deterministic Yale Model introduced in Takahashi and Alexander (2002). We then set up and solve a quadratic optimisation problem in order to find a commitment schedule, such that the proportion of private equity held in our portfolio is constant over time. The optimized commitment schedule shows high variability. This is why we apply a smoothing procedure via variance penalization. Furthermore we undertake a sensitivity analysis of the underlying parameters.

We calibrate our model by estimating the required parameters using the data provided and then assess the goodness of fit by comparing it to the observed values. The long-term behaviour of the commitment is studied in the hope of finding a steady state.

Finally, we turn our attention to the continuous-time stochastic model introduced in de Malherbe (2004). We discuss the dynamics of the model before turning to the calibration of the parameters required to fit the model to the data. Our parameter estimates are close to the parameter values given in de Malherbe (2004).

Our analysis shows there is hardly any convincing performance improvement of the stochastic model in terms of calibration and robustness compared with the deterministic model. Furthermore the deterministic model has better fitting properties and is computationally less expensive than the stochastic one. The Yale model is also more tractable as it has an easier underlying model structure, which leads to a better understanding of its robustness properties. We leave an adaptive optimisation of the commitment schedule obtainable by the Dynamic Programming Principle for further research.

References

- Cumming, D., 2009. Private equity: Fund types, risks and returns, and regulation. Vol. 10. John Wiley and Sons.
- de Malherbe, E., 2004. Modeling private equity funds and private equity collateralised fund obligations. *International Journal of Theoretical and Applied Finance* 7 (03), 193–230.
- Diop, M. B.-A., 2004. An efficient discretisation scheme for one dimensional sdes with a diffusion coefficient function of the form— $x^{-\alpha}$, $\alpha \in [1/2, 1]$.
- Kocis, J. M., Bachman IV, J. C., Long III, A. M., Nickels, C. J., 2009. Inside Private Equity: The Professional Investor's Handbook. Vol. 495. John Wiley & Sons.
- Takahashi, D., Alexander, S., 2002. Illiquid alternative asset fund modeling. *Journal of Portfolio Management* 28 (2), 90–100.

Portfolio Optimisation Under Uncertainty

TEAM 3

HOLLY BRANNELLY, University College London
JAMES NEVIN, University of Cape Town
DUMISANI NGWENZA, University of Cape Town
YIRAN WANG, University College London

Supervisor:
SAM COHEN, University of Oxford

Contents

1	Introduction	3
2	Modern Portfolio Theory	5
2.1	The MV optimal portfolio	5
2.2	Singularity of the covariance matrix	7
2.3	The Market Model	8
3	Parameter Estimation	10
3.1	Principal component analysis	10
3.2	Bayesian shrinkage	11
3.2.1	MV portfolio using shrinkage techniques	11
4	Optimisation Under Uncertainty	14
4.1	DR-expectation	14
4.2	Nested optimisation problem	16
4.3	Minimax approximation	18
4.4	Data	18
4.5	Numerical implementation	19
4.6	Comparison	24
5	Extensions	29
5.1	Moving-window approach	29
5.2	Expected shortfall portfolio optimisation	30
5.2.1	Non-Gaussian extension	34
6	Conclusion	35
7	Appendix	38

1 Introduction

One of the most frequently practiced areas of research in financial mathematics is that of portfolio optimisation; in particular the mean-variance (MV) framework of modern portfolio theory, as introduced by Markowitz (1952). By imposing certain assumptions on the asset expected return vector and the variance-covariance matrix, the model allows us to solve a quadratic programming problem to calculate optimal portfolios based on maximising the expected return subject to a given, minimal, level of risk and certain budget constraints. However, the parameters required to solve this problem must be estimated using historical data, and commonly used estimation approaches such as frequentist and Bayesian expected loss approaches fail to incorporate any parameter uncertainty into decision making. Due to the statistical uncertainty in parameter estimates and the sensitivity of solutions to resulting perturbations in the parameters, the optimal portfolios obtained are often unreliable. Additionally, the MV portfolio problem assumes returns are independent and identically distributed (iid) Gaussian, which in practice may not always seem realistic, and can fail if there is not enough data or too large a number of assets under consideration.

This motivates the exploration into alternative ways to compute optimal portfolios, in particular those that manage to incorporate statistical error in the estimated parameters into the portfolio selection problem. In this report, we will make use of the ‘divergence robust’ or ‘data-driven robust’ expectation (DR-expectation), as introduced by Cohen (2017), to extend the existing MV portfolio optimisation framework in order to overcome the difficulty of incorporating uncertainty into the valuation of decisions regarding optimal portfolios. The DR-expectation is a non-linear expectation that is closely related to a risk measure and, given a data set, can be used to obtain a prediction interval for a random variable that will incorporate the certainty of an estimate of that variable.

The data we will work with throughout will be that of the 30 constituents of the Dow Jones Industrial Average (DJIA). In Section 2 we introduce the theory underlying modern portfolio theory and compute the optimal portfolio, given the data, under the maximum likelihood estimator (MLE) for the parameters. All optimisation frameworks require estimates of the parameters involved; Section 3 introduces two ways in which to estimate these parameters - principal component analysis (PCA) and Bayesian shrinkage techniques - the first of which we use to obtain the MLE when there are large amounts of data involved. Section 4 introduces the DR-expectation, which will then be used to compute the optimal portfolio for in-sample data from the original data set, however this time incorporating the uncertainty in the parameter estimation into the decision through a ‘penalty’ function. The optimal portfolio that is obtained in this setting is then compared to the naive

MV portfolio (where statistical error is not taken into consideration) in the two cases where the parameters are estimated by the MLE and Bayesian shrinkage, an equally weighted portfolio, and a market weighted portfolio for the same data.

Finally, in Section 5 we extend the results of Section 4 to first consider a moving-window approach in which we varied both the length of the historical data used for calibration, and the frequency with which we rebalanced the portfolio holdings, in order to find the optimal combination of both factors. Secondly, we introduce the idea of optimisation under an alternative risk measure - expected shortfall. Expected shortfall portfolio optimisation allows us to overcome the problems that may be associated to assuming iid normality of the data, as the framework accommodates for flexibility in both the model fitted to each individual asset's returns, as well as the dependence structure between assets (we describe a copula approach here). Section 6 concludes this project.

2 Modern Portfolio Theory

In this chapter we introduce the mean-variance (MV) portfolio optimisation theory of Markowitz (1952), which will underpin the main optimisation problem of this project. The model assumes that an investor faces a risk-return trade-off, and must find the strategy that maximises their portfolio return without increasing the level of risk they take on past a certain threshold (dependent on the risk preferences of the investor, which are captured by a ‘risk aversion’ parameter).

2.1 The MV optimal portfolio

The general set-up assumes that an investment is made into N assets whose returns are modelled by the random matrix $\mathbf{R} \in \mathbb{R}^{N \times T}$, given by

$$\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_N)^\top = \begin{pmatrix} R_1^{(1)} & R_2^{(1)} & \dots & R_T^{(1)} \\ R_1^{(2)} & R_2^{(2)} & \dots & R_T^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ R_1^{(N)} & R_2^{(N)} & \dots & R_T^{(N)} \end{pmatrix} \quad (1)$$

where $\mathbf{r}_n = (R_1^{(n)}, \dots, R_T^{(n)})$, $1 \leq n \leq N$ denotes the return of the n^{th} asset in the time period $\{1, \dots, T\}$ for some $T > 0$. The assets are assumed to be iid Gaussian and their distribution is characterised by their expected return vector, given by $\boldsymbol{\mu} = \mathbb{E}[\mathbf{R}] \in \mathbb{R}^N$, and covariance matrix, given by $\mathbf{V} = \mathbb{E}[(\mathbf{R} - \boldsymbol{\mu})(\mathbf{R} - \boldsymbol{\mu})^\top] \in \mathbb{R}^{N \times N}$, which must both be estimated using historical data.

A portfolio consisting of a combination of the N assets can be constructed, and is characterised by the weight vector $\boldsymbol{\pi} = (\pi^{(1)}, \pi^{(2)}, \dots, \pi^{(N)})^\top$, where $\pi^{(i)}$ is the fraction of the total amount of capital that is invested into the i^{th} asset. If we initialise the initial capital of the investor to 1, we have the constraint that all portfolio weights must sum to 1, i.e. $\boldsymbol{\pi}^\top \mathbf{1} = 1$ - we will refer to this as the investor’s budget constraint. It is important to note that in this set-up, we allow for negative weights, as investors are permitted to short-sell.

The total return of the portfolio is therefore given by $\boldsymbol{\pi}^\top \mathbf{R}$, and hence the expected portfolio return by $\boldsymbol{\pi}^\top \boldsymbol{\mu}$ and the variance of the return by $\boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi}$. An investor assesses the value of a portfolio by combining its mean and variance in the following way

$$\boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi} \quad (2)$$

where $\lambda \in \mathbb{R}$ is the risk aversion parameter. Assuming that we are only considering

fully-invested portfolios, we can use the method of Lagrange multipliers to solve the optimisation problem

$$\max_{\boldsymbol{\pi}} (\boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi}) \quad (3)$$

under the budget constraint $\boldsymbol{\pi}^\top \mathbf{1} = 1$, and obtain a closed-form expression for the optimal portfolio weights, given by

$$\boldsymbol{\pi}^* = \frac{1}{2\lambda} \mathbf{V}^{-1} \left(\boldsymbol{\mu} + \frac{2\lambda - \boldsymbol{\mu}^\top \mathbf{V}^{-1} \mathbf{1}}{\mathbf{1}^\top \mathbf{V}^{-1} \mathbf{1}} \mathbf{1} \right). \quad (4)$$

As mentioned above, in practice the expected returns vector and the covariance matrix of the returns are not known and are therefore must be estimated from historical data. For a portfolio of N assets, this leaves us with N expected returns, N variances and $N(N - 1)/2$ covariances to estimate. Due to the large number of parameters, this can prove problematic. For example, if we wanted to construct a portfolio consisting of a combination of the constituents of the S&P500, there are 125,000 covariances, giving 126,000 total parameters to estimate, and hence solving these estimation problems numerically will be both tricky and time consuming.

Assume we use the method of maximum likelihood estimation to estimate the parameters, $\boldsymbol{\theta} = (\boldsymbol{\mu}, \mathbf{V})$. Given some data observations \boldsymbol{x} , the model uses the likelihood function $\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{x})$ and finds the values of the parameters that maximise it, given the data, i.e.

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} (\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{x})). \quad (5)$$

In practice, it is often convenient to work with the log-likelihood function,

$$\ell(\boldsymbol{\theta}; \boldsymbol{x}) = \ln \mathcal{L}(\boldsymbol{\theta}; \boldsymbol{x})$$

and maximise this instead.

If the sample data are assumed to be iid normally distributed, as in the MV set-up, MLEs for the mean and covariance matrix are given by

$$\hat{\boldsymbol{\mu}} = \frac{1}{n} \sum_{k=1}^n \boldsymbol{x}_k$$

and

$$\hat{\mathbf{V}} = \frac{1}{n} \sum_{k=1}^n (\boldsymbol{x}_k - \hat{\boldsymbol{\mu}})(\boldsymbol{x}_k - \hat{\boldsymbol{\mu}})^\top$$

respectively.

As an example, and to put this into practice, we start by considering the first 3 stocks from the DJIA (see Table 7 in the appendix) and the following three cases:

- The standard MV portfolio problem, the closed-form solution of which is given by Eq.(4);
- The MV portfolio problem, however now with the added constraint of long-only portfolios, i.e. $\pi^{(i)} \geq 0$ for all $1 \leq i \leq N$;
- Equally weighted portfolios, i.e. $\pi^{(i)} = 1/N$.

Figure (1) shows the value of the optimal portfolios in each of the three settings. While we note that Case 1 has the largest value over all time periods, as expected it is significantly more volatile due to the lack of restriction on the hedging positions that an investor may take (i.e. both long and short positions).

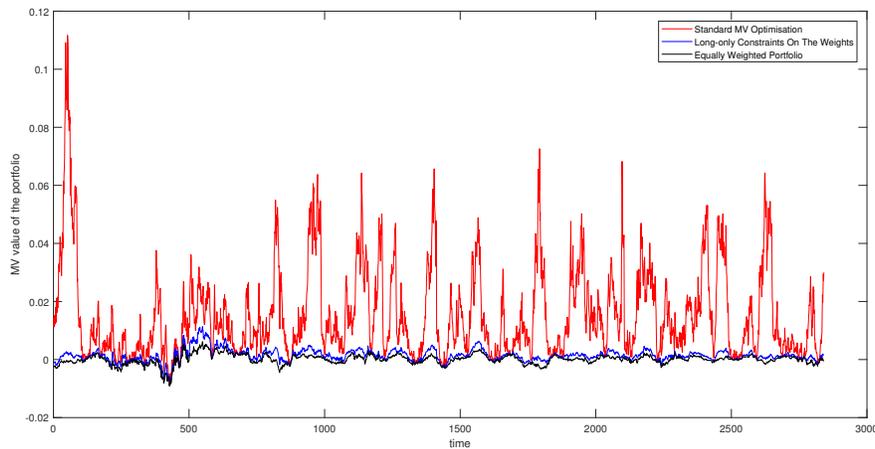


Figure 1: The value of the optimal portfolios over time under the assumptions of (1) standard MV optimisation, (2) long-only constraints on the weights, and (3) an equally weighted portfolio of the three assets.

2.2 Singularity of the covariance matrix

One of the main problems that one may encounter in solving the above portfolio optimisation problem is singularity of the covariance matrix, as the optimal closed-form solution given by Eq. (4) requires a matrix inverse. The return matrix, given by Eq. (1), can be written as $\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_T)$, where $\mathbf{r}_t = (R_t^{(1)}, \dots, R_t^{(N)})^\top$ is the vector of returns of all assets at time $t > 0$.

It follows that the expected return can be written as

$$\hat{\boldsymbol{\mu}} = \frac{1}{T} \sum_{t=1}^T \mathbf{r}_t = \frac{1}{T} \mathbf{R} \cdot \mathbf{1}, \quad (6)$$

and the covariance matrix as

$$\hat{\mathbf{V}} = \frac{1}{T} \sum_{t=1}^T (\mathbf{r}_t - \hat{\boldsymbol{\mu}})(\mathbf{r}_t - \hat{\boldsymbol{\mu}})^\top = \frac{1}{T} \mathbf{R}(\mathbf{I} - \frac{1}{T} \mathbf{1}\mathbf{1}^\top) \mathbf{R}^\top \quad (7)$$

where T represents the number of data observations available (i.e. the number of days considered if we are working with daily returns). Regardless of whether the true covariance matrix is invertible, we observe from Eq. (7) that the sample covariance can never be invertible in the case where $N \geq T$. This follows from the fact that the maximum rank of \mathbf{V} will be the rank of $(\mathbf{I} - \frac{1}{T} \mathbf{1}\mathbf{1}^\top)$ which is $T - 1$:

$$\text{rank}(\mathbf{V}) \leq \text{rank}(\mathbf{I} - \frac{1}{T} \mathbf{1}\mathbf{1}^\top) = T - 1 < T \leq N. \quad (8)$$

To summarise, if the number of assets exceeds the number of available observations, the sample covariance matrix will be non-invertible.

In cases where we are considering a large number of assets, the covariance matrix \mathbf{V} will be close to singular. Natural approaches to overcome this problem are to use some additional regularization of \mathbf{V} and/or \mathbf{V}^{-1} , or to impose further constraints on $\boldsymbol{\pi}$. These approaches will be discussed in Section 4.

2.3 The Market Model

One of the downfalls of modern portfolio theory is the requirement for extensive data to estimate the parameters reliably, which is not always available. An extension of modern portfolio theory is the multi-factor model, which assumes that returns are driven by a number of underlying, observable factors, each with some economic interpretation. In this context, of particular interest is the model which asserts market returns should be the sole factor determining asset returns. This is the renowned capital asset pricing model (CAPM), pioneered by Sharpe (1964), and is a single-index model with the market portfolio as the sole factor.

Due to the fact that assets in the market are held by a number of investors in various quantities, it is assumed that financial assets are held in proportion to the market capitalisation. Let $R_t^{(i)}$ denote the return at time $t \in [0, T]$ for asset i , then the single-index model for this asset is given by

$$R_t^{(i)} = \alpha_t^{(i)} + \beta R_t^M + \epsilon_t^{(i)} \quad (9)$$

or in matrix form for N assets by

$$\mathbf{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{R}_t^M + \boldsymbol{\epsilon}_t \quad (10)$$

where $\epsilon_t^{(i)}$ are the zero mean uncorrelated elements of the $N \times 1$ residual matrix, $\boldsymbol{\epsilon}_t$. The asset return covariance matrix, implied by the market model is

$$\mathbf{V} = \sigma_M^2 \boldsymbol{\beta}\boldsymbol{\beta}^T + \boldsymbol{\Sigma}_\epsilon \quad (11)$$

where σ_M^2 is the market portfolio variance.

Ledoit and Wolf (2003) state that the covariance matrix can be estimated by performing a multivariate regression and can be denoted as,

$$\mathbf{F} = s_M^2 \mathbf{b}\mathbf{b}^T + \hat{\boldsymbol{\Sigma}}_\epsilon \quad (12)$$

where s_M^2 is the sample variance of the market portfolio and $\hat{\boldsymbol{\Sigma}}_\epsilon$ the sample residual covariance matrix.

It should be noted that in choosing to work with the market model framework, we only need to estimate $2N + 1$ parameters, which is significantly less than the $N(N + 1)/2$ parameters required to be estimated in the standard MV framework.

3 Parameter Estimation

As mentioned in Section 2, portfolio optimisation problems require estimation of the necessary parameters; in the MV framework, this is the expected return vector and covariance matrix of the iid Gaussian asset returns. The expected return vector rarely causes any problem when being estimated from historical data, however this is not the case with the covariance matrix, often due to its large dimensionality. In this section, we discuss two methods of estimating the covariance matrix that can deal with a high number of assets.

3.1 Principal component analysis

Principal component analysis (PCA) is a statistical procedure that can be used to describe the covariance structure using only a few linear combinations of the original stochastic variables. As a result, PCA reduces the amount of data required and allows interpretability of the model. Moreover, PCA can also bring to light some relationships in the data that might not always be noticed when using alternative estimation methods. Consider an $N \times N$ covariance matrix. In order to reproduce all the variability in the system in a standard estimation framework, N principal components are required. However, in reality, most of the variability can be explained with $K < N$ principal components without significant loss of information. This reduction of the data characterises PCA.

Assume that we have N assets with returns $\mathbf{R} = (r_1, \dots, r_N)$, as before, where \mathbf{S} is the sample covariance matrix for the assets. The matrix \mathbf{S} will have eigenvalue-eigenvector pairs $(e_1, \lambda_1), (e_2, \lambda_2), \dots, (e_N, \lambda_N)$, where $e_i = [e_{1,i}, e_{2,i}, \dots, e_{N,i}]^\top$ and $\lambda_1 > \lambda_2 > \dots > \lambda_N$. The principal components for the i^{th} sample will be given by

$$g_i = e_i^\top \mathbf{R} = \sum_{n=1}^N e_{n,i} r_n, \quad i = 1, 2, \dots, N. \quad (13)$$

As a result, and as shown by Johnson and Wichern (1992), \mathbf{S} can be decomposed as

$$\begin{aligned} \mathbf{S} &= \lambda_1 e_1 e_1^\top + \lambda_2 e_2 e_2^\top + \dots + \lambda_N e_N e_N^\top \\ &= [\sqrt{\lambda_1} e_1 \quad \sqrt{\lambda_2} e_2 \quad \dots \quad \sqrt{\lambda_N} e_N] \begin{bmatrix} \sqrt{\lambda_1} e_1 \\ \sqrt{\lambda_2} e_2 \\ \dots \\ \sqrt{\lambda_N} e_N \end{bmatrix} = \mathbf{L} \mathbf{L}^\top. \end{aligned} \quad (14)$$

As mentioned above, one of the desirable characteristics of PCA is the ability to reduce the amount of data whilst maintaining important information. It can be shown that the covariance structure can be described by using only a few of the

principal components and neglecting the contributions of the smallest eigenvalues, i.e. the first K eigenvalues. It then follows that Eq. (14) can be rewritten as

$$\begin{aligned} \mathbf{S} &\approx \lambda_1 \mathbf{e}_1 \mathbf{e}_1^\top + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^\top + \dots + \lambda_K \mathbf{e}_K \mathbf{e}_K^\top \\ &= \begin{bmatrix} \sqrt{\lambda_1} \mathbf{e}_1 & \sqrt{\lambda_2} \mathbf{e}_2 & \dots & \sqrt{\lambda_K} \mathbf{e}_K \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} \mathbf{e}_1 \\ \sqrt{\lambda_2} \mathbf{e}_2 \\ \dots \\ \sqrt{\lambda_K} \mathbf{e}_K \end{bmatrix} = \mathbf{L} \mathbf{L}^\top \end{aligned} \quad (15)$$

where $\mathbf{L} \in \mathbb{R}^{N \times K}$.

3.2 Bayesian shrinkage

A second method for estimating the covariance matrix when there is a large number of assets under consideration (and hence too many parameters to be estimated in order to use the sample covariance matrix) is to use a Bayesian statistical procedure called shrinkage. The resulting estimator will be a typical Stein estimator. This technique also overcomes the problem of singularity of the covariance matrix when the number of assets, N , exceeds the number of historical observations per asset, T , as discussed in Section 2.2.

As with any method in Bayesian statistics, a prior must be assumed for the covariance matrix, which in turn imposes structure to the estimation problem to reduce the dependency on purely estimated parameters (as the uncertainty in the sample estimate increases, the prior takes on greater importance). Stein (1956) proposed a method of obtaining an optimal balance between estimation error and bias by taking an appropriate weighting of the unbiased (sample covariance matrix) and the biased (the prior covariance matrix) estimator. This weighting factor assigned to the prior, $\alpha \in [0, 1]$, is called the ‘shrinkage intensity’ as it shrinks the unbiased estimator to the biased estimator (also referred to as the ‘shrinkage target’), which will also generally be non-singular.

3.2.1 MV portfolio using shrinkage techniques

In this context, we take the prior to be the single-index CAPM covariance matrix, \mathbf{F} , and weight between the two extremes: this matrix, and the N -factor sample covariance matrix \mathbf{S} . It should be noted in making this choice of prior that the CAPM covariance matrix is not an unbiased estimator, due to stringent structural assumptions, but it is not exposed to extreme estimation error. Additionally, on the other hand, the sample covariance matrix is an asymptotically unbiased estimator, yet has a large amount of estimation error. Fundamentally, a trade-off between bias

and estimation error is needed. The shrinkage intensity is denoted by $\alpha \in [0, 1]$, and the resultant shrinkage estimator is given by

$$\hat{\Sigma}_{Shrink} = \alpha \mathbf{F} + (1 - \alpha) \mathbf{S}. \quad (16)$$

In this set-up, we assume that the asset returns are iid, have finite fourth moments, and that N is fixed whilst $T \rightarrow \infty$. This allows \mathbf{S} to be consistent whilst \mathbf{F} is not, and the shrinkage intensity to asymptotically tend to 0 over time.

The objective will now be to find a shrinkage estimator that gives an optimal solution to Eq. (16) and that does not break down when $N \geq T$, i.e. the objective function we must solve to obtain such an estimator does not require the covariance matrix to be non-singular. Ledoit and Wolf (2003) define the optimal α as the solution that minimises the expected value of the quadratic loss function between the shrinkage estimator and the true covariance matrix (based on the Frobenius norm), given by

$$L(\alpha) = \|\alpha \mathbf{F} + (1 - \alpha) \mathbf{S} - \Sigma\|^2. \quad (17)$$

Setting the expected value of the first derivative of Eq. (17) equal to zero, we can obtain the resultant optimal shrinkage intensity as

$$\alpha^* = \frac{\sum_{i=1}^N \sum_{j=1}^N (Var[s_{i,j}] - Cov[f_{i,j}, s_{i,j}])}{\sum_{i=1}^N \sum_{j=1}^N (Var[f_{i,j} - s_{i,j}] + (\phi_{i,j} - \sigma_{i,j})^2)}, \quad (18)$$

which can be rewritten as

$$\alpha^* = \frac{1}{T} \frac{\pi - \rho}{\gamma} + \mathcal{O}\left(\frac{1}{T^2}\right), \quad (19)$$

where

$$\pi = \sum_{i=1}^N \sum_{j=1}^N AsyVar(\sqrt{T} s_{i,j}) \quad (20)$$

$$\rho = \sum_{i=1}^N \sum_{j=1}^N AsyVar(\sqrt{T} f_{i,j}) \quad (21)$$

$$\gamma = \sum_{i=1}^N \sum_{j=1}^N (\phi_{i,j} - \sigma_{i,j})^2. \quad (22)$$

At this point, it should be noted that the true parameters ϕ , ρ and γ are not directly observable. Thus, they need to be estimated. Ledoit and Wolf (2003) propose the consistent estimator of $\kappa = (\phi - \rho)/\gamma$ as $k = (p - r)/c$, and it follows that the asymptotic optimal shrinkage estimator is given by,

$$\hat{\Sigma}_{Shrink} = \frac{k}{T} \mathbf{F} + \left(1 - \frac{k}{T}\right) \mathbf{S}. \quad (23)$$

The consistent estimator of ϕ is $p = \sum_{i=1}^N \sum_{j=1}^N p_{i,j}$, where

$$p_{i,j} = \frac{1}{T} \sum_{t=1}^T ((r_{i,t} - \bar{r}_i)(r_{j,t} - \bar{r}_j) - s_{i,j})^2. \quad (24)$$

Similarly, the consistent estimator of ρ is given by

$$r = \sum_{i=1}^N \sum_{j=1}^N r_{i,j} \quad (25)$$

for $i \neq j$ and $r_{i,j} = \frac{1}{T} \sum_{t=1}^T r_{i,j,t}$ where

$$r_{i,j,t} = \frac{s_{j,M} s_{M,M} (r_{i,t} - \bar{r}_i) + s_{i,M} s_{M,M} (r_{j,t} - \bar{r}_j) - s_{i,M} s_{j,M} (r_{M,t} - \bar{r}_M)}{s_{M,M}^2} (r_{M,t} - \bar{r}_M) \\ \times (r_{i,t} - \bar{r}_i)(r_{j,t} - \bar{r}_j) - f_{i,j} s_{i,j}. \quad (26)$$

For the diagonal elements, we have that $r_{i,i} = p_{i,i}$. Finally, we have the sample equivalent

$$c = \sum_{i=1}^N \sum_{j=1}^N c_{i,j}, \quad (27)$$

which can be shown to be a consistent estimator of γ where

$$c_{i,j} = (f_{i,j} - s_{i,j})^2. \quad (28)$$

In Section 4, we use these shrinkage techniques to obtain an estimator for the covariance matrix of our data, and then compare the performance of MV portfolios using this estimator, to those using the MLE.

4 Optimisation Under Uncertainty

In this section we present the nested optimisation problem that, given a data set (here, the log-returns of the constituents of the DJIA), will allow us to obtain optimal portfolios that take statistical uncertainty into consideration. We begin by introducing the DR-expectation, as this will be central to the nested convex optimisation problem that we aim to solve.

4.1 DR-expectation

The motivation behind the DR-expectation, as introduced by Cohen (2017), is to construct a means of incorporating statistical uncertainty into valuation problems that involve unknown parameters. We begin this section by introducing the general framework of non-linear expectations, which is often used to model uncertainty in a random setting (i.e. Knightian uncertainty as in Föllmer and Schied (2002)), and then use these concepts, however, now connecting them to statistical estimation to define the DR-expectation. We refer to Cohen (2017) for the definitions and theorems given below.

Definition 4.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $L^\infty(\mathcal{F})$ denote the space of \mathbb{P} -essentially bounded \mathcal{F} -measurable random variables. A non-linear expectation on $L^\infty(\mathcal{F})$ is a mapping

$$\mathcal{E} : L^\infty(\mathcal{F}) \rightarrow \mathbb{R}$$

satisfying the assumptions

- **Strict monotonicity:** for any $\xi_1, \xi_2 \in L^\infty(\mathcal{F})$, if $\xi_1 \geq \xi_2$ a.s. then $\mathcal{E}(\xi_1) \geq \mathcal{E}(\xi_2)$ and if in addition $\mathcal{E}(\xi_1) = \mathcal{E}(\xi_2)$ then $\xi_1 = \xi_2$ a.s.
- **Constant triviality:** for any constant $k \in \mathbb{R}$, $\mathcal{E}(k) = k$
- **Translation equivariance:** for any $k \in \mathbb{R}$, $\xi \in L^\infty(\mathcal{F})$, $\mathcal{E}(\xi + k) = \mathcal{E}(\xi) + k$.

A convex expectation will satisfy the above conditions, as well as the following condition:

- **Convexity:** for any $\lambda \in [0, 1]$, $\xi_1, \xi_2 \in L^\infty(\mathcal{F})$,

$$\mathcal{E}(\lambda\xi_1 + (1 - \lambda)\xi_2) \leq \lambda\mathcal{E}(\xi_1) + (1 - \lambda)\mathcal{E}(\xi_2).$$

We consider a class of convex expectations that satisfy lower semicontinuity, i.e. for a sequence $\{\xi_n\}_{n \in \mathbb{N}} \subset L^\infty(\mathcal{F})$ with $\xi_n \uparrow \xi \in L^\infty(\mathcal{F})$ pointwise, $\mathcal{E}(\xi_n) \uparrow \mathcal{E}(\xi)$, and denote the space of all probability measures on (Ω, \mathcal{F}) that are absolutely continuous with respect to \mathbb{P} by \mathcal{M}_1 to give the following theorem by Föllmer and Schied (2002) and Frittelli and Gianin (2002).

Theorem 4.2. Suppose \mathcal{E} is a lower semi-continuous convex expectation. Then there exists a ‘penalty’ function $\alpha : \mathcal{M}_1 \rightarrow [0, \infty]$ such that

$$\mathcal{E}(\xi) = \sup_{\mathbb{Q} \in \mathcal{M}_1} \{ \mathbb{E}_{\mathbb{Q}}[\xi] - \alpha(\mathbb{Q}) \}. \quad (29)$$

If $\alpha(\mathbb{Q}) < \infty$, for some $\mathbb{Q} \sim \mathbb{P}$, we can restrict our attention to measures in \mathcal{M}_1 that are equivalent to \mathbb{P} without loss of generality.

The definition of the DR-expectation, which itself is a non-linear (convex) expectation, follows by specifying a form of the penalty function $\alpha(\cdot)$ that incorporates the likelihood function of the data set. This penalty term indicates how reasonable the estimated value is based on the data observations, and is defined as follows.

Definition 4.3. For a model $\mathbb{Q} \in \mathcal{M}_1$, let the likelihood of the data \mathbf{x} under \mathbb{Q} be denoted by $L(\mathbb{Q}|\mathbf{x})$. Let $\mathcal{Q} \subset \mathcal{M}_1$ be a set of models under consideration (e.g. a parametric set of distributions). We define the $\mathcal{Q}|\mathbf{x}$ -divergence as the negative log-likelihood ratio, i.e.

$$\alpha_{\mathcal{Q}|\mathbf{x}}(\mathbb{Q}) := -\log(L(\mathbb{Q}|\mathbf{x})) + \sup_{\tilde{\mathbb{Q}} \in \mathcal{Q}} \left\{ \log(L(\tilde{\mathbb{Q}}|\mathbf{x})) \right\}. \quad (30)$$

Now, we can define the DR-expectation.

Definition 4.4. For fixed observations \mathbf{x} , an uncertainty aversion parameter $k > 0$ and exponent $k' \in [1, \infty]$, we define the convex expectation, which we will refer to as the “ $\mathcal{Q}|\mathbf{x}$ -divergence robust expectation” or “data-driven robust expectation” (DR-expectation), as

$$\mathcal{E}_{\mathcal{Q}|\mathbf{x}}^{k,k'}(\xi) := \sup_{\mathbb{Q} \in \mathcal{Q}} \left\{ \mathbb{E}_{\mathbb{Q}}[\xi(\omega, \mathbf{x})] - \left(\frac{1}{k} \alpha_{\mathcal{Q}|\mathbf{x}}(\mathbb{Q}) \right)^{k'} \right\} \quad (31)$$

where $\xi : \Omega \times \mathbb{R}^N \rightarrow \mathbb{R}$ is a Borel-measurable function with respect to the data observations and we take $x^\infty = 0$ for $x \in [0, 1]$ and $+\infty$ otherwise.

In the above definition, ξ explicitly depends on the data observations \mathbf{x} . From now on, assume that under each $\mathbb{Q} \in \mathcal{Q}$, we know X , $\mathbf{x} = \{X_n\}_{n=1}^N$ are iid random variables and $\xi = \phi(X)$ for some ϕ a Borel-measurable function. This allows us to write $\mathbb{E}_{\mathbb{Q}}[\xi(\omega, \mathbf{x})] = \mathbb{E}_{\mathbb{Q}}[\xi]$.

The paper presents the two extremal cases where $k' = 1$ and $k' = \infty$, and states that the intervening cases are natural interpolations between these:

$$\mathcal{E}_{\mathcal{Q}|\mathbf{x}}^{k,1}(\xi) := \sup_{\mathbb{Q} \in \mathcal{Q}} \left\{ \mathbb{E}_{\mathbb{Q}}[\xi] - \frac{1}{k} \alpha_{\mathcal{Q}|\mathbf{x}}(\mathbb{Q}) \right\} \quad (32)$$

$$\mathcal{E}_{\mathcal{Q}|\mathbf{x}}^{k,\infty}(\xi) := \sup_{\{\mathbb{Q} : \alpha_{\mathcal{Q}|\mathbf{x}}(\mathbb{Q}) < k\}} \{ \mathbb{E}_{\mathbb{Q}}[\xi] \}. \quad (33)$$

Essentially, given a random variable ξ , we can use the DR-expectation to obtain a prediction interval for its estimated value. The expectation $\mathcal{E}(\xi)$ (which is convex) can be thought of as an ‘upper’ expectation, depending on the certainty of the estimate of ξ given the sample, i.e. we consider all possible values of ξ and use the data to determine how reasonable we think they are. Additionally, we can then define the corresponding ‘lower’ expectation (which is concave) by $-\mathcal{E}(-\xi)$, thus giving the prediction interval for ξ by

$$[-\mathcal{E}(-\xi), \mathcal{E}(\xi)].$$

Note that since \mathcal{E} is a convex expectation, then $\rho(\xi) = \mathcal{E}(-\xi)$ is a convex risk measure. The term that takes into consideration how reasonable the estimate is based on data observations is the penalty term, and thus we can use the DR-expectation to retain knowledge of levels uncertainty in our estimates and feed this knowledge into decision making. In the following subsection, we apply this set-up to a portfolio optimisation problem.

In what follows, we denote the penalty function $\alpha_{\mathcal{Q}|\mathbf{x}}(\mathbb{Q})$ by $\mathcal{R}(\boldsymbol{\theta}; \mathbf{x})$.

4.2 Nested optimisation problem

In this subsection, we address the main problem of this report: **how can we incorporate statistical uncertainty into decision making involving optimal portfolio selection?** The problem extends the classical results, as discussed in Section 2, by incorporating the DR-expectation, and whilst there exists a lot of literature on convex optimisation, there has been little research into applying this nested optimisation problem to the optimal portfolio theory setting.

Recall the classical MV portfolio optimisation problem where we want to maximise the expected returns, $\boldsymbol{\pi}^\top \boldsymbol{\mu}$, subject to minimising the variance, $\boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi}$, i.e. we use a Lagrangian multiplier with risk aversion parameter λ to solve

$$\sup_{\boldsymbol{\pi}} \left\{ \boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi} \right\} \quad (34)$$

with the constraint $\boldsymbol{\pi}^\top \mathbf{1} = 1$.

To extend this problem to incorporate statistical uncertainty, first consider the following DR-expectation where the data used is denoted by \mathbf{r} . We can use this to define the following risk measure:

$$\mathcal{E}(\boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi}) = \inf_{\boldsymbol{\theta}} \left\{ \boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi} + \mathcal{R}(\boldsymbol{\theta}; \mathbf{r}) \right\}. \quad (35)$$

The form of the penalty function $\mathcal{R}(\boldsymbol{\theta}; \mathbf{r})$ follows from Definition (4.3), and requires the likelihood function of the data:

$$\mathcal{R}(\boldsymbol{\theta}; \mathbf{r}) = \left(\frac{1}{k} \left(-\ell(\boldsymbol{\theta}; \mathbf{r}) + \sup_{\boldsymbol{\theta}'} \ell(\boldsymbol{\theta}'; \mathbf{r}) \right) \right)^{k'}. \quad (36)$$

This convex expectation gives us an upper expectation for the return on the portfolio given the minimum variance constraint. To find the portfolio $\boldsymbol{\pi}^*$ that will maximise this return, we therefore have to solve the following nested optimisation problem:

$$\sup_{\boldsymbol{\pi}} \mathcal{E}(\boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi}) = \sup_{\boldsymbol{\pi}} \left(\inf_{\boldsymbol{\theta}} \left\{ \boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi} + \mathcal{R}(\boldsymbol{\theta}; \mathbf{r}) \right\} \right) \quad (37)$$

with the constraint $\boldsymbol{\pi}^\top \mathbf{1} = 1$.

To begin, we will assume the 30 DJIA log-returns in our data set are independent and identically distributed (iid) Gaussian with parameters $\boldsymbol{\theta} = (\boldsymbol{\mu}, \mathbf{V})$ which, using the corresponding likelihood function, gives us the following optimisation problem to solve:

$$\begin{aligned} & \sup_{\boldsymbol{\pi}} \inf_{\boldsymbol{\theta}} \left\{ \boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi} + \left(\frac{1}{k} \left(\frac{1}{2} \log |\mathbf{V}| - \frac{1}{2} \sum_{i=1}^N (\mathbf{r} - \boldsymbol{\mu})^\top \mathbf{V}^{-1} (\mathbf{r} - \boldsymbol{\mu}) \right. \right. \right. \\ & \left. \left. \left. - \frac{1}{2} \log |\hat{\mathbf{V}}| - \frac{1}{2} \sum_{i=1}^N (\mathbf{r} - \hat{\boldsymbol{\mu}})^\top \hat{\mathbf{V}}^{-1} (\mathbf{r} - \hat{\boldsymbol{\mu}}) \right) \right)^{k'} \right\} \\ & = \sup_{\boldsymbol{\pi}} \inf_{\boldsymbol{\theta}} \left\{ \boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi} + \left(\frac{1}{k} \left(\frac{1}{2} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^\top (\mathbf{I}_{\hat{\boldsymbol{\theta}}}(\mathbf{r}) + \mathcal{O}(N^{-1/2}) (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})) \right) \right)^{k'} \right\} \end{aligned} \quad (38)$$

where $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\mu}}, \hat{\mathbf{V}})$ is the MLE and $\mathbf{I}_{\hat{\boldsymbol{\theta}}}(\mathbf{r})$ is the observed information evaluated at the MLE or, in other words, the Hessian matrix.

The main challenge lies in solving this nested pair of optimisation problems. Practically, it is not simple to consider them simultaneously, and to calculate each step in turn would be extremely time consuming. In the following subsection we discuss an approximation to simplify the problem, however this can only be applied under certain conditions.

4.3 Minimax approximation

If the function $\mathcal{R}(\boldsymbol{\theta}; \mathbf{r})$ in Eq. (37) is convex, we have a concave-convex problem and can interchange the sup and inf, leaving us with the following convex optimisation problem to solve:

$$\begin{aligned} & \inf_{\boldsymbol{\theta}} \sup_{\boldsymbol{\pi}} \left\{ \boldsymbol{\pi}^\top \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^\top \mathbf{V} \boldsymbol{\pi} + \mathcal{R}(\boldsymbol{\theta}; \mathbf{r}) \right\} \\ & = \inf_{\boldsymbol{\theta}} \left\{ \boldsymbol{\pi}^{*\top} \boldsymbol{\mu} - \lambda \boldsymbol{\pi}^{*\top} \mathbf{V} \boldsymbol{\pi}^* + \mathcal{R}(\boldsymbol{\theta}; \mathbf{r}) \right\} \end{aligned} \quad (39)$$

where $\boldsymbol{\pi}^*$ is the optimal portfolio obtained by solving the MV problem. This reduces the difficulty of the problem, as many ways to solve convex optimisation problems exist. Another case in which we can interchange the order of the sup and inf in the nested optimisation problem is when we have a large sample, and $\mathcal{R}(\boldsymbol{\theta}; \mathbf{r})$ will be large whenever $\boldsymbol{\theta} \neq \hat{\boldsymbol{\theta}}$ and $\mathcal{R}(\boldsymbol{\theta}; \mathbf{r})$ is twice differentiable, giving a local minimum.

However, in practice these conditions often fall through. In the Gaussian case for example the penalty function is only convex in the parameters $(\mathbf{V}^{-1}\boldsymbol{\mu}, \mathbf{V}^{-1})$ and not in $(\boldsymbol{\mu}, \mathbf{V})$. Additionally, the approximation fails when the estimated value of the covariance matrix is singular.

When solving the optimisation problem, it is important to evaluate the trade-off between the error induced by using a minimax approximation to simplify the problem, and the computational time and difficulty of solving the nested optimisation without such an approximation. We find that the error associated to using the approximation in our problem is of order $1/N^2$, and hence for large sample sizes it seems justifiable to use a minimax approximation.

4.4 Data

In what follows, the data we use¹ consists of the daily price returns as well as log-returns of the Dow-Jones Industrial Average constituents, which are given in Table 4 in the appendix. The data relates to the period 2015/01/02 to 2018/06/27. Moreover, the data was filtered for the adjusted closing prices. Despite the reliability of Bloomberg, there were some missing values in the data set and due to the fact that the time period under consideration was relatively short and the sample size of stocks was small, we had to resort to linear interpolation to deal with these missing values.

¹The data was obtained through the Bloomberg Terminal.

4.5 Numerical implementation

In order to solve the nested optimisation problem, we used the ‘optim’ routine in R. This routine minimises an objective function, using a given algorithm, which for the purposes of this project we chose to be the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, unless stated otherwise. The ‘optim’ routine only allows for optimisation over a vector, and permits only lower and upper bounds for the vector as restrictions. This meant that the matrix V needed to be transformed into a vector, where no symmetry restrictions are placed on the vector. To do this, PCA was used, with the two characterisations given by Cholesky and eigendecomposition were considered.

We began by using the characterisation given by the Cholesky decomposition, starting with the case where we only considered three stocks in the portfolio. Whilst in this set-up the characterisation solved the problem efficiently, the parameters we needed to estimate scaled proportionately to n^2 , where n is the number of stocks under consideration. As expected, when extending the problem to consider a larger number of stocks, i.e. all 30 of the DJIA, this presented issues in terms of both a rapid increase in computing time, as well as machine error becoming significant. As a result, we concluded that the Cholesky decomposition was unfeasible, and decided to resort to the eigendecomposition with the restriction that the number of eigenvalues used in the decomposition would be given by K for

$$\begin{aligned} K &= \min(20, |\hat{\lambda}|) \\ \hat{\lambda} &= \{\lambda : \lambda > 10^{-8}\} \end{aligned} \tag{40}$$

where λ is an eigenvalue of V . This ensured the number of parameters to optimise over for the variance matrix was limited to at most 20. The assumption was made that eigenvalues below 10^{-8} suggested little to no movement along the eigenvectors associated with these eigenvalues, and so these eigenvalues were treated as zero. Additionally, the eigenvectors were treated as constant, and were not optimised over. As such, we then approximated the covariance matrix V by

$$\begin{aligned}
\mathbf{V} &= [e_1 \ e_2 \ \dots e_N] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_N \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix} \\
&\approx [e_1 \ e_2 \ \dots e_N] \begin{bmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & & \\ & & \dots & & & \\ & & & \lambda_K & & \\ & & & & 0 & \\ & & & & & \dots \\ & & & & & & 0 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix} \\
&= [e_1 \ e_2 \ \dots e_K] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_K \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_K \end{bmatrix}
\end{aligned}$$

where $K \leq 20$ is the number of eigenvalues used. Note that this method of characterising \mathbf{V} requires the same number of parameters to be estimated, regardless of how many stocks the portfolio needs to be allocated between. This also extends to the estimated values used in the expected return vector, μ . Normally, this vector would require N estimates, however based on the assumption that the eigenvalues after the K^{th} biggest are small and hence set to zero, the mean parameters can be expressed as a linear combination of the first K eigenvectors as follows

$$\begin{aligned}
\boldsymbol{\mu} &= [e_1 \ e_2 \ \dots e_N] \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix} \\
&\approx [e_1 \ e_2 \ \dots e_K \ 0 \ \dots 0] \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix} \\
&= [e_1 \ e_2 \ \dots e_K] \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_K \end{bmatrix}.
\end{aligned}$$

The total number of parameters to estimate will always be $2K$: K eigenvalues and K mean scalars, μ_i for $1 \leq i \leq K$. This parametrisation was used for the vector of parameters in our model, and is what is being referred to by ' θ ' for the remainder of Section 4.

We began by using the above framework to solve the optimisation problem required to calculate the MLE under the assumption that the data follows a multivariate Gaussian distribution. The 'optim' routine discussed above also allows the user to extract the Hessian matrix from the objective function, which was used in calculating the DR-expectation penalty function $\mathcal{R}(\theta; r)$, as we used the approximation given by the second line of Eq. (38).

At this point, a long-only restriction was placed on the values of π , i.e. requiring that both $\mathbf{1}^\top \pi = 1$ where $\mathbf{1}$ is a vector of 1's, and $\pi_i \geq 0, \forall \pi_i \in \pi$. This recognises the fact that it is often difficult in practice to short assets. An added benefit of this approach is that the portfolio weights are all necessarily less than or equal to 1, which creates a scenario where the total capital exposure the investor is subject to is limited to the initial amount invested, i.e. the most they can lose is their total investment. After implementing this added constraint in our optimisation problem, we saw that it decreased the discrepancy between the naive and DR-expectation portfolios versus the evenly weighted and market weighted portfolios, relative to the set-up where the investor is also permitted to short assets.

To summarise, after incorporating this additional restriction, the problem statement now becomes

$$\begin{aligned} \max_{\pi} & \left(\pi^\top \mu - \lambda \pi^\top V \pi \right) \\ \text{s.t.} & \quad \mathbf{1}^\top \pi = 1 \\ & \quad \pi_i \geq 0; \quad \forall \pi_i \in \pi \end{aligned} \tag{41}$$

for the naive portfolio, and

$$\begin{aligned} \sup_{\pi} \inf_{\theta} & \{ \pi^\top \mu - \lambda \pi^\top V \pi + \mathcal{R}(\theta; r) \} \\ \text{s.t.} & \quad \mathbf{1}^\top \pi = 1 \\ & \quad \pi_i \geq 0; \quad \forall \pi_i \in \pi \end{aligned} \tag{42}$$

for the portfolio problem that incorporates uncertainty via the DR-expectation. Additionally, note that in this case, the optimal naive portfolio can no longer be computed as

$$\pi^* = \frac{1}{2\lambda} V^{-1} \left(\mu + \frac{2\lambda - \mu^\top V^{-1} \mathbf{1}}{\mathbf{1}^\top V^{-1} \mathbf{1}} \mathbf{1} \right).$$

As a result, an optimisation routine needed to be implemented to calculate the portfolio weights under the quadratic programming problem given by Eq. (41). We worked with the optimisation routine ‘solve.QP’², which allows one to solve problems of the form

$$\min \left(-\mathbf{d}^\top \mathbf{b} + \frac{1}{2} \mathbf{b}^\top \mathbf{D} \mathbf{b} \right)$$

subject to $\mathbf{A}^\top \mathbf{b} \geq \mathbf{b}_0$

where \mathbf{D} and \mathbf{d} are a matrix and a vector, respectively, appearing in the quadratic function to be minimised. \mathbf{A} is a matrix defining the constraints under which we want to minimise the quadratic function, and \mathbf{b} is a vector holding the values of \mathbf{b}_0 . This routine also gives the additional option of enforcing strict equality on the first constraint, allowing us to easily implement our long-only portfolio constraint, and hence was used to calculate the naive long-only portfolio π^* . The default for ‘solve.QP’ requires the matrix \mathbf{D} to be invertible, however since \mathbf{D} is not of full rank, this is not the case. The function ‘solve.QP’ also gives the option of supplying the inverse of \mathbf{D} instead of \mathbf{D} itself, and therefore to take advantage of this and to overcome the requirement of invertibility of \mathbf{D} if the inverse is not provided in the function, we calculated the pseudoinverse of \mathbf{D} and provided this instead.

The notion of a pseudoinverse was first introduced by Moors (1920) and later pioneered by Penrose (1955). Suppose \mathbf{A} is an $N \times N$ matrix. The pseudoinverse of \mathbf{A} , denoted as \mathbf{A}^+ will have the following properties:

1. $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$,
2. $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$.

However it is not always the case that $(\mathbf{A}\mathbf{A}^\top)^{-1}$ exists. Barata and Hussein (2012) propose that $\mathbf{A}\mathbf{A}^\top + \mu \mathbf{1}$ will be invertible for non-vanishing μ and a small $|\mu|$. Thus, $\mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top + \mu \mathbf{1})^{-1}$ and $(\mathbf{A}^\top \mathbf{A} + \mu \mathbf{1})^{-1} \mathbf{A}^\top$ will be well-defined for $\mu \neq 0$ and will converge to \mathbf{A}^+ as $\mu \rightarrow 0$. This process is defined as Tikhonov’s regularisation, as introduced by Tikhonov (1963).

Now, relating this back to our problem, if $\mathbf{V} = \mathbf{E}\hat{\mathbf{V}}\mathbf{E}^\top$ is the eigenvalue decomposition of \mathbf{V} , then the matrix \mathbf{D} in Eq. (41) can be characterised by

$$\mathbf{D}^\top \mathbf{D} = 2\lambda \mathbf{V}.$$

Note that the risk aversion parameter here, λ , is not to be confused with any eigenvalues. Simple matrix algebra can be used to show that $\mathbf{D}^{-1} = (1/\sqrt{2\lambda})\mathbf{E}\mathbf{V}^*\mathbf{E}^\top$,

²solve.QP is a function under *quadprog*. More detail can be found at <https://cran.r-project.org/web/packages/quadprog/index.html>

where V^* is the diagonal matrix with elements $1/\sqrt{\text{eigenvalues}}$.

We now return to the nested optimisation problem given by Eq. (37), which we initially attempted to solve using two stacked instances of ‘optim’. This was effective for a smaller number of stocks, such as the three stock case, but had difficulty with larger portfolios, in the form of significant machine error, and hence did not provide the flexible solution that we required. The issue here stemmed from the fact that when a large number of stocks were considered, ‘optim’ was unable to differentiate between small changes in the value of either π or θ , since the value of

$$\pi^\top \mu - \lambda \pi^\top V \pi + \mathcal{R}(\theta; r)$$

would often take on values greater than the machine was able to track. This would then result in the routine being unable to vary the parameters in any intuitive way, essentially resulting in no change in the portfolio weights π from whichever initial guess was provided.

The first way this was dealt with was by scaling the variables in θ by a factor of 10^4 . This allowed ‘optim’ to measure small changes in θ , and the difference this made on functions of it, hence allowing the routine to deviate from the initial guess of π . Secondly, tolerance levels were lowered from the default of 10^{-8} , and values of 10^{-6} , 10^{-4} , and 10^{-3} were tested. Despite these efforts however, the nested optimisation was still not workable, and as a result the minimax approximation had to be used. With this approach, tolerances could be left at the machine default, but the upscaling of the θ variable was still necessary, and the number of iterations used had to be increased from the default 100 to 20000. The minimax approach gave reasonable estimates when using a large sample size to estimate θ , as suggested by the argument in Section 4.3. The effectiveness of the DR-expectation portfolio over the naive portfolio was severely reduced when using smaller sample sizes, as expected, as when we have a low number of parameters to estimate, we expect to be less uncertain of our estimated values - this was quantified by tracking the returns of each optimal portfolio through time in the two cases where the number of stocks was 3 and 30.

Using this minimax approximation, we begin solving Eq. (42) by interchanging the infimum and supremum and then using the ‘optim’ routine in conjunction with the ‘solve.QP’ routine to minimise

$$\sup_{\pi} \{ \pi^\top \mu - \lambda \pi^\top V \pi \} + \mathcal{R}(\theta; r) \quad (43)$$

over θ , while ‘solve.QP’ solved the inner supremum (subject to the given budget and long-only constraints). In practice, we found that ‘optim’ attempted to utilise negative eigenvalues in searching through possible values of θ , which resulted in

the matrix D not being positive definite — a problem for the ‘solve.QP’ function. To correct this, the “L-BFGS-B” algorithm was used in place of the “BFGS” algorithm, as this allows one to implement lower bounds for θ ; a lower bound of 10^{-12} was used for the eigenvalues.

When implementing this, portfolio balancing was done on a once off basis (i.e. we computed the optimal portfolio weights and held this position over time), and results were collected for a period of 250 days. The following set of parameters was tested:

$$\begin{aligned} & \{\text{Days used to estimate } \theta \in D_\theta, k \in k_{\text{used}}, k' \in k'_{\text{used}}, \lambda \in \lambda_{\text{used}}\} \\ & D_\theta = \{7, 21, 63, 126, 252\} \\ & k_{\text{used}} = \{5, 10, 15, 30\} \\ & k'_{\text{used}} = \{5, 6, 7, 8\} \\ & \lambda_{\text{used}} = \{10, 20, 30\}. \end{aligned}$$

4.6 Comparison

To make comparisons between our long-only DR-expectation portfolio, the naive MV portfolio (where both the sample covariance matrix and the shrinkage covariance matrix estimator are used), the equally weighted portfolio and the market weighted portfolio, we ran the code over a variety of different choices of the parameters λ, k, k' , as well as using various lengths of historical data, as shown above. In this section, we review the findings from a variety of these cases.

1. In Figure 2, we plotted the returns on the naive portfolio over time (measured in days) for three values of the risk aversion parameter, $\lambda = 10, 20, 30$. These changes in λ did not have a significant impact on the portfolio however, as demonstrated by the similarity in the three lines in the figure, suggesting that MV portfolios do not have high sensitivity to the risk preferences of investors. Since we used the minimax approximation in computing the DR-expectation portfolios, it follows that λ also will have little effect on these portfolios, and therefore in what follows we will not take the choice of λ to be of particular importance in our analysis.
2. Sharpe ratios are often perceived to be an accurate method of comparing the risk-adjusted return between portfolios when assets are assumed to be normally distributed (as is the case in our set-up). The Sharpe ratios for the evenly weighted and market weighted portfolios were calculated as 0.05207 and 0.5037, respectively. Considering all of the different combinations of the parameters described above, we computed the Sharpe ratio in each case. Table 1 shows the largest of these Sharpe ratios for the different portfolio types

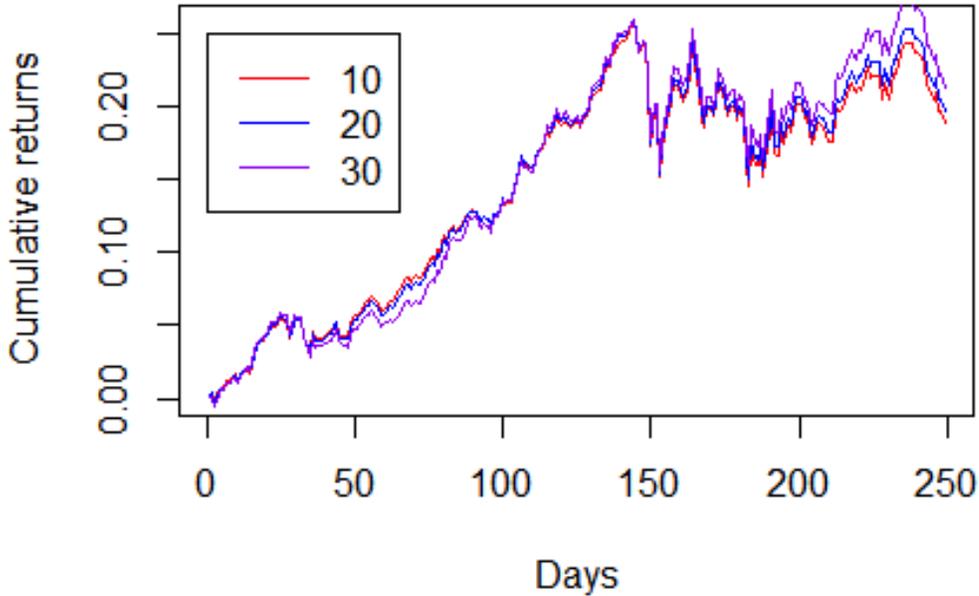


Figure 2: Returns on the naive portfolio for a variety of values of the risk aversion parameter $\lambda = 10, 20, 30$.

over a variety of calibration periods. In all cases, the three portfolio construction methods under consideration outperformed the evenly weighted and market weighted portfolios. This holds even under the shortest calibration period of 7 days, where only 7 historical data points would for each stock have been used to determine the optimal portfolio. We concluded that one possible explanation for this was that the data was taken from a highly stable and developed market, leading to strong data and the ability to draw reliable inferences using only small amounts of it.

Table 1: The maximum Sharpe ratios over all parameter combinations for each portfolio type under a variety of calibration windows.

Calibration Period	Naive	DR	Shrinkage
7	0.06914	0.05884	0.06576
21	0.05294	0.06347	0.05598
63	0.09788	0.09765	0.09775
126	0.11362	0.11002	0.11364
252	0.09206	0.08794	0.08968

3. Figure 3 allows us to visualise these Sharpe ratio results for the DR-expectation portfolio vs the naive portfolio, as we have plotted the difference in their Sharpe ratios for the five calibration periods. When the value is above zero the DR portfolio is outperforming the naive. We see that when the calibration period being used is 7 or 126 days, the DR portfolio under performs significantly. This is expected in the case of a 7 day learning period as the interchange between the sup and inf in the nested optimisation is less justified than when we are using longer learning periods, and hence we do not expect strong performance from the resulting DR portfolio. Excluding the 7 day calibration period case, we also note that as this period increases, there is less variability between the Sharpe ratios of the two portfolios i.e. altering the learning period once we exceed a certain number of days will have less impact on the relative performance of portfolios.

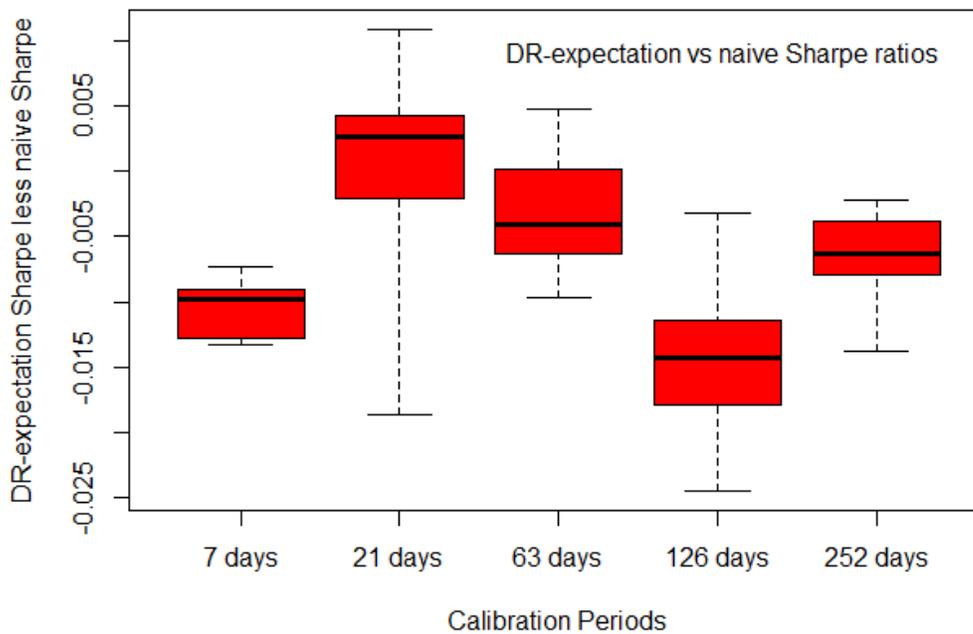


Figure 3: The difference between the Sharpe ratios of the DR-expectation portfolio and the naive MV portfolio over a variety of different calibration periods.

Calibration Period	Parameters			max(DR Sharpe - naive Sharpe)	Percentage change
	k	k'	λ		
7	5	5	10	-0.00725	-10.974%
21	30	6	20	0.01089	23.959%
63	10	7	10	0.00474	5.148%
126	5	7	10	-0.00325	-2.873%
252	5	5	30	-0.00218	-2.508%

Table 2: The maximum difference of the Sharpe ratios for various combinations of the parameters under various calibration periods for the DR–expectation vs naive portfolio.

- Using the Sharpe ratios that we computed for the DR and naive portfolios over all different combinations of the parameters, in Table 2 we find the combination of λ, k, k' under each learning window that gives the best improvement to the Sharpe ratio of the DR portfolio over the naive portfolio, i.e. the furthest right column of the table shows $\max(\text{DR Sharpe ratio} - \text{naive Sharpe ratio})$ over all parameters for a given calibration period. In some cases, there is no combination of parameters that improve the DR Sharpe ratio over that of the naive portfolio, and hence we just minimise how much worse the difference is (this corresponds to negative values in the last column). We observe that when the calibration period is larger (126 and 252 days) or on the smaller end (7 days), the performance of the DR portfolio is worse than that of the naive in all cases of parameter combinations. Our hypothesis as to why this is the case for the larger two calibration periods lies in the fact that the large amount of data provides enough information that the naive portfolio can come up with accurate estimates, and hence the penalty that captures our lack of estimation confidence isn't particularly necessary and hence doesn't make much of a difference to the portfolio performance. Intuitively this makes sense, as we are only estimating 40 parameter values (eigenvalues and mean scalars), but using either 156 or 252 data samples to do so. We believe that the out-performance of the DR expectation breaks down in the 7 day calibration period case as the interchange between the sup and inf functions in the nested optimisation problem is less justified when using such a small number of data samples - this also suggests that any results corresponding to a 7 day calibration window may be outliers. In the intermediate cases (21 days and 63 days), we can find values of the parameters for which the DR portfolio will outperform, and note that the best performance of the DR portfolio over the naive occurs in the case of a 21 day calibration period. Table 3 demonstrates similar results for the Sharpe ratios of the MV naive portfolio constructed using the shrinkage estimate and the usual naive

portfolio that makes use of the sample covariance matrix. Figure 4 shows the cumulative returns over the three portfolios for the 21 day calibration period, with the parameters given in Tables 2 and 3 for the DR and Shrinkage cases. Whilst the DR portfolio outperforms over time, there is an unusual period starting at around 30 days into holding the portfolio where the returns on the DR portfolio drop significantly, but those on the other two portfolios do not.

Table 3: The maximum of the differences of the Sharpe Ratios for various values of λ over various calibration periods for the Shrinkage portfolio vs the naive portfolio.

	Parameter		
Calibration Periods	λ	max(Shrinkage Sharpe - naive Sharpe)	Percentage change
7	10	-0.0003	-5.17%
21	20	0.01054	23.185%
63	20	0.0002	2.07%
126	10	0.000039	0.034%
252	20	0.00105	1.191%

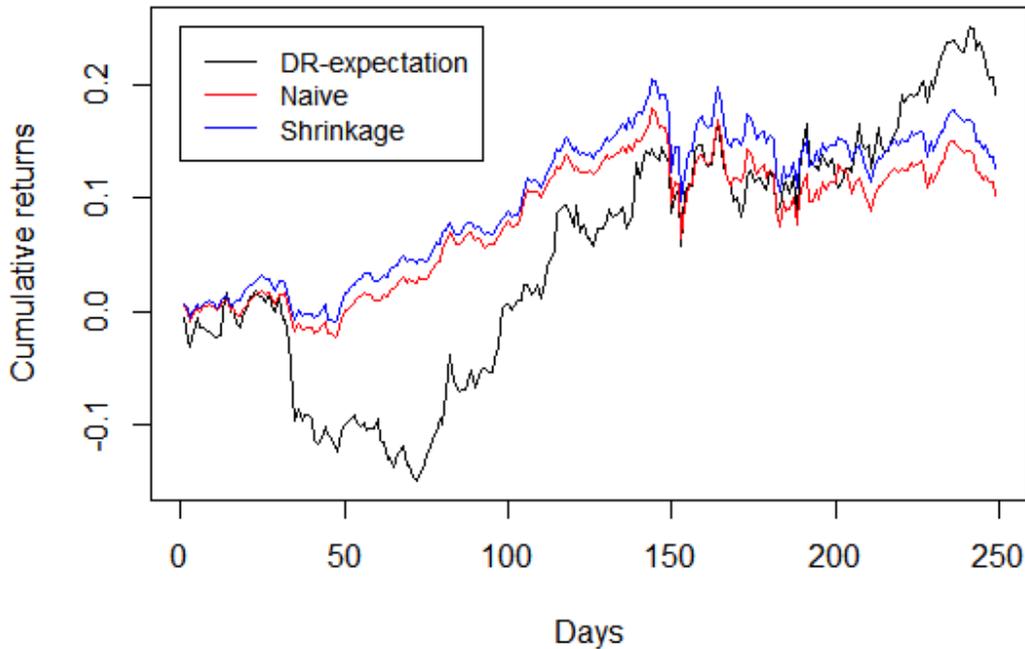


Figure 4: Cumulative returns through time of the three different portfolios, based on the optimal parameters shown in Tables 2 and 3 for the DR and Shrinkage cases, and a 21 day calibration period.

5 Extensions

In this section we consider a few ways in which we can extend the results discussed previously.

5.1 Moving-window approach

Here, we started by removing the long-only restriction and constructed portfolios that also allowed for a rebalancing over time. Recall that, since the long-only restriction has been removed, the closed-form solution for the optimal portfolio under a given θ is given by

$$\pi^* = \frac{1}{2\lambda} \mathbf{V}^{-1} \left(\boldsymbol{\mu} + \frac{2\lambda - \boldsymbol{\mu}^\top \mathbf{V}^{-1} \mathbf{1}}{\mathbf{1}^\top \mathbf{V}^{-1} \mathbf{1}} \mathbf{1} \right)$$

where the pseudoinverse (also referred to as the generalised inverse) of \mathbf{V} was used where necessary (usually whenever the number of stocks considered exceeded the number of eigenvalues use in the eigendecomposition, K).

The MLE for θ was used to calculate the naive portfolio holdings. In construction of the DR-expectation portfolio, the minimax approximation was used to interchange the sup and inf, and we optimised

$$\hat{\pi}^\top \boldsymbol{\mu} - \lambda \hat{\pi}^\top \mathbf{V} \hat{\pi} + \mathcal{R}(\boldsymbol{\theta}; \mathbf{r})$$

over $\boldsymbol{\theta}$ where $\hat{\pi}$ is the closed form optimal solution to the maximisation.

In the analysis, time constraints prevented us from considering all combinations of parameters, and hence just the following parameters were used:

$$\begin{aligned} k &= 22 \\ k' &= 8 \\ \lambda &= 30. \end{aligned}$$

These parameters describe a moderately risk-averse individual, who places a high value on the confidence of parameter estimates based on the data. Moving windows were constructed under the following combinations:

Days used to estimate θ	Number of days between rebalances
7	7, 21
21	7, 21, 63
63	7, 21, 63, 126
126	7, 21, 63, 126, 252
252	7, 21, 63, 126, 252

Figure 5 shows the cumulative return of the DR-expectation, naive, and evenly weighted portfolios, where there is no long-only holding restriction, a calibration period of 126 days, and rebalancing every 126 days. In this scenario, the evenly weighted portfolio out-performed both the DR-expectation and naive portfolio. The most likely explanation for this is that an aggressive short investment might be undertaken as a result of the optimisation problem. In the worst performing period, the DR-expectation drifted further away from the naive portfolio — likely due to uncertainty limiting highly aggressive allocation, and suggesting that this poor performance of the DR portfolio would be reduced if we were to reintroduce the long-only constraint.

5.2 Expected shortfall portfolio optimisation

Here, we consider an extension of the MV optimisation problem, and then analogously to before, we incorporate statistical uncertainty into our optimisation problem through the use of the DR-expectation. Any optimisation problem requires a risk measure; in MV optimisation, the risk measure chosen by Markowitz was the volatility of the returns. The drawbacks to measuring risk in terms of the variance

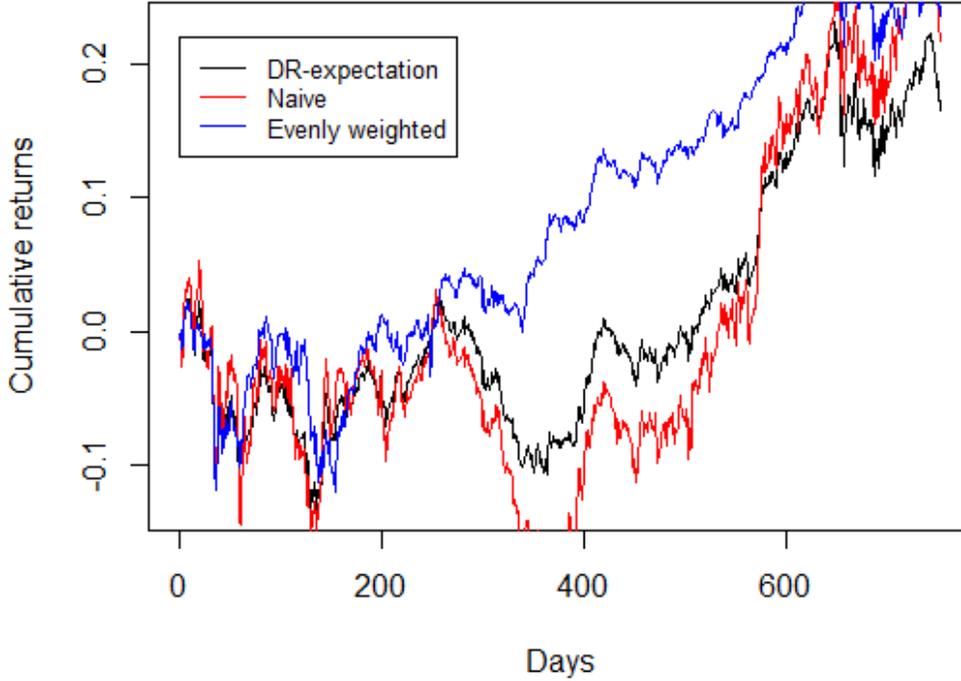


Figure 5: Cumulative returns through time of three different portfolios, with 126 day calibration period and 126 days between rebalances.

however, is that large negative and large positive fluctuations in the returns are equally penalised, which, from the point of view of a financial investor may not seem justified and hence motivated the introduction of downside risk measures (focussing on the lower tail risk, i.e. downside losses) in financial modelling.

Definition 5.1. For a random variable X with distribution function $F_X(\cdot)$, and a given confidence level $\alpha \in [0, 1]$, the value at risk (VaR) of X is defined by

$$\text{VaR}_\alpha(X) = \inf \{x \in \mathbb{R} | F_X(x) \geq \alpha\}. \quad (44)$$

In the case where X is a continuous random variable and its distribution function has a well-defined inverse (the quantile function, $F_X^{-1}(\cdot)$), we have $\text{VaR}_\alpha(X) = F_X^{-1}(\alpha)$.

Value at Risk, as defined above, essentially gives you the threshold that your portfolios losses will not exceed with probability α . One of the main drawbacks to VaR, however, was its lack of ability to describe tail behaviour beyond the α -quantile level. This lead to the construction expected shortfall (ES), which averages over the portfolio losses that exceed VaR.

Definition 5.2. For a random variable X with distribution function $F_X(\cdot)$, and a given confidence level $\alpha \in [0, 1]$, the expected shortfall (ES) of X is defined by

$$\text{ES}_\alpha(X) = -\frac{1}{\alpha} \int_0^\alpha \text{VaR}_\gamma(X) d\gamma. \quad (45)$$

The alternative robust portfolio optimisation problem we consider in this section is to find the optimal portfolio weights subject to minimising the ES of portfolio returns for a given level α . Excluding parameter uncertainty, the naive equivalent to this optimisation problem is given by

$$\begin{aligned} \inf_{\boldsymbol{\pi}} \quad & \left\{ \text{ES}_\alpha(\boldsymbol{\pi}^\top \mathbf{R}; \hat{\boldsymbol{\theta}}) \right\} \\ \text{s.t.} \quad & \boldsymbol{\pi}^\top \mathbf{1} = 1 \end{aligned} \quad (46)$$

where the ES will be a function of the parameter estimates $\hat{\boldsymbol{\theta}}$ as it is determined by the distribution of the future portfolio returns $\boldsymbol{\pi}^\top \mathbf{R}$. Since $\boldsymbol{\pi}^\top \mathbf{R}$ exhibits a multivariate distribution, calculate the ES is non-trivial due to the complexity of constructing multivariate quantile functions, and unlike the MV optimisation problem, a closed form solution for $\boldsymbol{\pi}^*$ does not exist, regardless of the assumption in underlying distribution of the returns. We compute the ES as follows. Assume we have measures corresponding to $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$ (the MLE) and n samples of $\boldsymbol{\pi}^\top \mathbf{R}$ under $\hat{\boldsymbol{\theta}}$ given by $\mathbf{r}_i \sim f(\mathbf{r}_i; \boldsymbol{\theta})$ for $1 \leq i \leq n$, to which we assign an equal weighting, $w_i = 1/n$. We can compute the sample weights under $\boldsymbol{\theta}$ by

$$\bar{w}_i = w_i \frac{f(\mathbf{r}_i; \boldsymbol{\theta})}{f(\mathbf{r}_i; \hat{\boldsymbol{\theta}})}. \quad (47)$$

Now, fixing a choice of $\boldsymbol{\pi}$ throughout, we use the following algorithm to compute the ES of $\boldsymbol{\pi}^\top \mathbf{R}$ numerically:

1. Order $\boldsymbol{\pi}^\top \mathbf{r}_i \rightarrow \boldsymbol{\pi}^\top \mathbf{r}_{(i)}$
2. Calculate $\text{cumsum}(\bar{w}_{(i)})$
3. Find i^* such that $\text{cumsum}(\bar{w}_{(i)}) = \alpha$
4. Set $\hat{\text{VaR}}_\alpha(\boldsymbol{\theta}) = \boldsymbol{\pi}^\top \mathbf{r}_{(i^*)}$

5. Set $\text{ES}_\alpha(\boldsymbol{\theta}) = \text{mean}(\boldsymbol{\pi}^\top \mathbf{r}_{(j)})$ for all j such that $\boldsymbol{\pi}^\top \mathbf{r}_{(j)} > \boldsymbol{\pi}^\top \mathbf{r}_{(i^*)}$.

Now, taking statistical error into consideration, we use the DR-expectation to introduce a penalty function to quantify this uncertainty, and aim to solve the following nested optimisation problem

$$\begin{aligned} \inf_{\boldsymbol{\pi}} \sup_{\boldsymbol{\theta}} \left\{ \text{ES}_\alpha(\boldsymbol{\pi}^\top \mathbf{R}; \boldsymbol{\theta}) - \mathcal{R}(\boldsymbol{\theta}; \mathbf{r}) \right\} \\ \text{s.t. } \boldsymbol{\pi}^\top \mathbf{1} = 1 \end{aligned} \quad (48)$$

where, as before,

$$\mathcal{R}(\boldsymbol{\theta}; \mathbf{r}) = \left(\frac{1}{k'} \left(-\ell(\boldsymbol{\theta}; \mathbf{r}) + \sup_{\boldsymbol{\theta}'} \ell(\boldsymbol{\theta}'; \mathbf{r}) \right) \right)^{k'}. \quad (49)$$

Note that as a higher ES is considered worse for a portfolio, there is a sign change before the penalty function and the order of the sup and inf are interchanged in Eq. (48).

The main advantage of the ES optimisation problem over the MV approach is that it allows us to consider non-Gaussian log-returns (assuming normality for asset returns is often scrutinised, one of the reasons being that returns data usually exhibits fatter tails than a Gaussian model allows for) as well as non-linear portfolios. This introduces flexibility in the distribution fitted to the data, for instance allowing the tails of log-returns to be modelled asymmetrically, and hence expanding the general scope of investment scenarios to which the optimisation problem can be applied. On the other hand, however, a drawback to this approach is that the computation of the ES of multivariate data is nontrivial, and hence ES optimisation is a simulation based optimisation problem which in turn introduces more statistical uncertainty. Additionally, if a copula is used to model the multivariate dependence between assets, the computational time increases as drawing samples from a copula proves more time consuming than drawing from an elliptical distribution.

Despite the fact that the full advantages of using ES as a risk measure are not utilised unless log-returns are assumed non-Gaussian, for simplicity you could begin by remaining in the Gaussian setting and computing the ES optimal portfolio to compare to the MV case. Isaksson (2016) demonstrates that under the assumption of elliptically distributed log-returns, there is in fact a connection between the ES and MV optimisation problems that is characterised through the risk aversion parameter, λ . This set-up allows you to use the same penalty function as Section 4.5.

Due to time constraints we did not implement this idea numerically, however if we were to do so, we would first need to compute the naive ES optimal portfolio, and

then again by implementing a minimax approximation, we could largely make use of our existing code to solve the ES DR-expectation nested optimisation problem.

5.2.1 Non-Gaussian extension

Assume we now want to use a distribution beside the Gaussian to model our data, and a copula to describe the multivariate dependence structure between each stock.

Definition 5.3. A function $C : [0, 1]^d \rightarrow [0, 1]$ is a d -dimensional copula if C is the joint distribution function of a d -dimensional random vector on the unit cube with standard uniform marginals and it holds that

- $C(u_1, \dots, u_d) = 0$ whenever $u_i = 0$ for at least one $i \in \{1, \dots, d\}$;
- $C(u_1, \dots, u_d) = u_i$ if $u_j = 1$ for all $j = 1, \dots, d$ and $j \neq i$;
- C is non-decreasing on its support.

For a random vector $\mathbf{X} = (X_1, \dots, X_d)$ with continuous marginals where the i^{th} marginal has distribution function $F_i(x)$, the joint distribution of the X_i can be constructed using a d -copula as

$$\begin{aligned} F(x_1, \dots, x_d) &= C(F_1(x_1), \dots, F_d(x_d)) \\ &= C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)). \end{aligned} \quad (50)$$

The corresponding copula density is given by

$$c(u_1, \dots, u_d) = \frac{\partial^d}{\partial u_1 \dots \partial u_d} C(u_1, \dots, u_d) = \frac{f(x_1, \dots, x_d)}{f_1(x_1) \dots f_d(x_d)}. \quad (51)$$

Assuming that the distribution function we fit to the i^{th} stock's log-return data is given by $F_i(r_i)$ (parameters are estimated via the maximum likelihood method) and the chosen copula by $C_{\theta}(u_1, \dots, u_N)$, the optimisation problem becomes

$$\begin{aligned} &\inf_{\boldsymbol{\pi}} \sup_{\boldsymbol{\theta}} \left\{ \text{ES}_{\alpha} \left(\boldsymbol{\pi}^{\top} \mathbf{R}; \boldsymbol{\theta} \right) \right. \\ &\quad \left. - \left(\frac{1}{k} \left(-c_{\boldsymbol{\theta}}(F_1(x_1), \dots, F_N(x_N)) + \sup_{\boldsymbol{\theta}'} c_{\boldsymbol{\theta}'}(F_1(x_1), \dots, F_N(x_N)) \right) \right)^{k'} \right\} \end{aligned} \quad (52)$$

subject to $\boldsymbol{\pi}^{\top} \mathbf{1} = 1$.

To solve the nested optimisation problem under this setting, we only have to change the penalty function (as this is determined by the likelihood function of the data), and can use the same algorithm to calculate the ES as a function of the parameters and portfolio weights ($\text{ES}_{\alpha}(\boldsymbol{\pi}^{\top} \mathbf{R}; \boldsymbol{\theta})$).

6 Conclusion

In this project, we considered a new way of constructing optimal portfolios of a set of assets that incorporate statistical error resulting from parameter estimation (this being one of the leading reasons as to why naive portfolios that do not account for this error fail in practice) into the optimisation problem. This set-up resulted in a complex nested optimisation problem which, in order to construct these portfolios numerically, required numerous estimation and approximation techniques to be employed. A number of problems were encountered along the way, stemming from both the fact that whilst convex optimisation is a well-researched area, there is little existing literature or algorithms on solving convex optimisation problems of our type, as well as both time and computational power constraints. Given that we had to leave our code running overnight in order to obtain the spread of results required to accurately compare the performance of the different portfolios, we were restricted in the number of times we could tweak and re-run our code.

Despite these concessions, however, in the case where we constructed portfolios based on a long-only investment restriction that account for uncertainty (what we referred to throughout as the DR-expectation portfolios), results indicated an improvement in portfolio performance over the evenly weighted and market weighted portfolios; this analysis was based off both the Sharpe ratio and the cumulative returns through time of each of the portfolios. When dropping the long-only restriction, the performance of the DR-expectation did not beat that of the evenly weighted or market weighted naive portfolios, results that were both disappointing and slightly perplexing.

Throughout, our analysis was limited to using data from the DJIA, an index consisting of only 30 stocks that trade in a well-developed and highly liquid market, suggesting that the quality of the data is likely to be high. It would be interesting to extend these results to a portfolio consisting of a significantly larger number of assets (for instance, the constituents of the S&P 500), as the method employed in our code carries scalability. Additionally, a further extension could be to consider different markets where, for example, the quality of the data may be lower than that of the DJIA, in order to observe whether the performance of the DR-expectation portfolio would differ significantly. Intuitively, we would assume that when the quality of data is lower, the confidence one may have in statistical estimates of parameters drawn from the data would be lower, and hence it would be more beneficial to construct a DR portfolio over a naive portfolio in order to account for this lack of confidence. The same idea holds when the number of assets increases, as as the number of parameters to estimate in the naive case increases, the confidence in each estimate should be expected to decrease. Exploring such options, however, was restricted due to time constraints and a lack of computational power.

Bibliography

- João Carlos Alves Barata and Mahir Saleh Hussein. The Moore–Penrose pseudoinverse: A tutorial review of the theory. *Brazilian Journal of Physics*, 42(1-2):146–165, 2012.
- Samuel N Cohen. Data-driven nonlinear expectations for statistical uncertainty in decisions. *Electronic Journal of Statistics*, 11(1):1858–1889, 2017.
- Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and stochastics*, 6(4):429–447, 2002.
- Marco Frittelli and Emanuela Rosazza Gianin. Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486, 2002.
- Daniel Isaksson. Robust portfolio optimization with expected shortfall, 2016.
- Richard Johnson and Dean Wichern. Applied multivariate statistical methods. *Prentice Hall, Englewood Cliffs, NJ*, 1992.
- Olivier Ledoit and Michael Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5):603–621, 2003.
- Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.
- E Moors. On the reciprocal of the general algebraic matrix. *Bull. Amer. Math. Soc.*, 26:394–395, 1920.
- Roger Penrose. A generalized inverse for matrices. In *Mathematical proceedings of the Cambridge philosophical society*, volume 51, pages 406–413. Cambridge University Press, 1955.
- William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964.
- Charles Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Technical report, Stanford University, Stanford, United States, 1956.

Andrei N Tikhonov. Solution of incorrectly formulated problems and the regularization method. In *Dokl. Akad. Nauk.*, volume 151, pages 1035–1038, 1963.

7 Appendix

Table 4: Constituents of the Dow-Jones Industrial Average

Company Name	Company Name	Company Name
3M	American Express Co	Apple Inc
Boeing Co	Caterpillar Inc.	Chevron Corp.
Cisco Systems, Inc	Coca-Cola Co.	DuPont de Nemours & Co.
Exxon Mobil Corp.	General Electric Co.	Goldman Sachs Group Inc
Home Depot Inc.	Intel Corp.	International Business Machines
Johnson & Johnson	JP Morgan Chase	Co McDonald's Corp.
Merck & Co. Inc.	Microsoft Corp.	NIKE Inc
Pfizer Inc.	Procter & Gamble Co	Travelers Companies Co.
United Technologies Corp.	UnitedHealth Group	Verizon Communications Inc.
Visa Inc.	Wal-Mart Stores Inc.	Walt Disney Co.

An Assessment on the Appropriateness of the use of the LFMM in South Africa

TEAM 4

WADE CRESSWELL University of Cape Town

XI KLEISINGER-YU, Eidgenössische Technische Hochschule Zürich

AUORE DE SAVIGNY, University College London

CHRIS STERLEY, Univeristy of Cape Town

Supervisors:

OBEID MAHOMED, University of Cape Town

THOMAS MCWALTER, University of Cape Town

Contents

1	Introduction	3
2	Lognormal Forward-LIBOR Market Model	4
2.1	Structure	4
2.2	Pricing caps	5
2.3	Pricing swaptions	7
3	Calibration	9
3.1	Data	9
3.1.1	Data description	9
3.1.2	Data thinning	10
3.1.3	Data analysis	10
3.1.4	Yield rate adjustments	14
3.2	Determining South African forward rates	15
3.3	Estimating caplet volatilities	15
3.4	Calibrating instantaneous volatilities	17
3.5	Estimating historical correlation	20
3.6	Swaption price approximation	20
3.6.1	Note on implementation	22
3.7	Correlation parametrisation	22
4	Methodology	24
4.1	Fixing volatilities and calibrating correlation	24
4.2	Fixing correlation and calibrating volatility	25
4.3	A note on implementation of Rebonato's Angle Parameterisation	25
5	Results	26
6	Conclusion	31

1 Introduction

Since the 2008 financial crisis, South Africa's derivatives market has continued to grow and develop. It is one of the largest economies on the continent and has benefited from a strong and well developed supervisory financial framework. In South Africa, primitive financial markets exist for bonds and equity while derivative financial markets exist for swaps, futures and options written on all underlying asset classes. The inter-bank interest rate market is a financial market in which participants are able to trade vanilla funding products as well as interest rate derivatives.

In this report, we will discuss the applicability of the Lognormal Forward-LIBOR Market Model (LFMM) to the South African inter-bank market. We will discuss pricing Caps, Floors and Swaptions within this framework. In particular, we aim to answer the question: Is there a parsimonious instantaneous volatility and correlation parameterization for long-term modelling with the LFMM after joint calibration?

The LFMM was introduced simultaneously by Brace, Gatarek & Musiela (1997), and Miltersen, Sandmann & Sondermann (1997). It is widely used for pricing interest rates derivatives by modelling the forward inter-bank rates that are directly observable in the market rather than the short rates or instantaneous forward rates. The LFMM assumes that forward rates are lognormally distributed under its forward measure. This has the advantage of being consistent with the Black 1976 model when pricing caplets and floorlets and, hence, caps and floors. However, it is incompatible with pricing swaptions using the same approach. Indeed, we will show in section 2 that the assumption of lognormal forward rates under its forward measure used under the LFMM leads to non-lognormal swap rates under the same measure. We therefore require approximations to price swaptions under the LFMM, such as the Rebonato formula and the Hull-White formula.

We first present the LFMM model and the valuation of caps and swaptions under this framework. Next, we will describe the data obtained from the market. We will discuss two different calibration procedures in great detail. Then, we will detail the different approximations we used to recover the swaption prices. Finally, we will conclude whether or not there exists a parsimonious instantaneous volatility and correlation parameterization that provide adequate fit and can thus be used for long-term modelling, furthermore, we aim to assess the appropriateness of the LFMM in the South African inter-bank market.

2 Lognormal Forward-LIBOR Market Model

2.1 Structure

In this section we introduce the Lognormal Forward-LIBOR Market Model (LFMM), its assumptions and the notation we will use throughout the paper.

Let $P(t, T)$ be the value at time t of a zero-coupon bond delivering $P(T, T) = 1$ at time $T > t$. The simply-compounded forward LIBOR rate is defined as:

$$F(t, T, S) := \frac{1}{\tau(T, S)} \left(\frac{P(t, T)}{P(t, S)} - 1 \right), \quad 0 \leq t < T < S. \quad (1)$$

where $\tau(T, S)$ is the year fraction of the period $[T, S]$. Let t be the current time. Consider a set $T_0 < \dots, T_M < T_{M+1}$ of bond tenor dates. Times T_i are expressed in years from the current time. The corresponding simple forward rate $F_i(t)$ of the LIBOR rate $L(T_i, T_{i+1})$ from T_i to T_{i+1} at some time $t \leq T_i$ is given by $F(t, T_i, T_{i+1}) =: F_i(t)$. We have $F_i(T_i) = L(T_i, T_{i+1})$.

Consider the probability measure \mathbb{Q}^{i+1} associated with the numeraire $P(t, T_{i+1})$, i.e. with the price of the ZCB whose maturity coincides with the maturity of the forward rate. \mathbb{Q}^{i+1} is often called the forward adjusted measure for maturity T_{i+1} . Let τ_i be the year fraction of the period $[T_i, T_{i+1}]$. By definition, we have $F_i(t)P(t, T_{i+1}) := \frac{1}{\tau_i}(P(t, T_i) - P(t, T_{i+1}))$. It follows that $F_i(t)P(t, T_{i+1})$ is the price of a tradable asset (difference between two zero-coupon bonds with a nominal amount of $\frac{1}{\tau_i}$) and it has to be a martingale under the probability measure \mathbb{Q}^{i+1} associated with the numeraire $P(t, T_{i+1})$ when divided by this numeraire. Hence, $F_i(t)$ is martingale under \mathbb{Q}^{i+1} . The LFMM assumes the following driftless geometric Brownian for the forward LIBOR rate $F_i(t)$ under \mathbb{Q}^{i+1} :

$$dF_i(t) = \sigma_i(t)F_i(t) dW_i^{i+1}(t), \quad t \leq T_k \quad (2)$$

where $\sigma_i(t)$ is the instantaneous volatility at time t of the forward LIBOR rate $F_i(t)$ and $W_i^{i+1}(t)$ is the i^{th} component of the M -dimensional Brownian motion $W^{i+1}(t)$ under \mathbb{Q}^{i+1} with instantaneous covariance given by:

$$d \left\langle W_i^{i+1}, W_j^{i+1} \right\rangle_t = \rho_{i,j} dt. \quad (3)$$

The forward LIBOR rate $F_i(t)$ is lognormally distributed under \mathbb{Q}^{i+1} . However, the dynamics of $F_i(t)$ under a measure \mathbb{Q}^{m+1} different from \mathbb{Q}^{i+1} are not martingales and are given by:

$$dF_i(t) = \mu_i(t) dt + \sigma_i(t)F_i(t) dW_i^{m+1}(t), \quad t \leq T_i \quad (4)$$

$$\text{where } \mu_i(t) := \begin{cases} -\sigma_i(t)F_i(t) \sum_{j=i+1}^m \frac{\rho_{i,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)} & \text{for } i < m; \\ 0 & \text{for } i = m; \\ \sigma_i(t)F_i(t) \sum_{j=m+1}^i \frac{\rho_{i,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)} & \text{for } i > m, \end{cases} \quad (5)$$

where $W_i^{m+1}(t)$ is the i^{th} component of the M -dimensional vector Brownian motion $W^{m+1}(t)$ under \mathbb{Q}^{m+1} with instantaneous correlations given by

$$d\langle W_i^{m+1}, W_j^{m+1} \rangle_t = \rho_{i,j} dt. \quad (6)$$

The matrix formed by elements $\rho_{i,j}$ is denoted ρ .

2.2 Pricing caps

A caplet (floorlet) is essentially a call (put) option on a LIBOR rate $L(T_i, T_{i+1})$. Recall τ_i is the year fraction of the period $[T_i; T_{i+1}]$. The payoffs of such options are given by:

$$\text{Caplet payoff} := \tau_i (F_i(T_i) - K)^+ \quad (7)$$

$$\text{Floorlet payoff} := \tau_i (K - F_i(T_i))^+ \quad (8)$$

Hence the value of caplet prices at time 0 is:

$$\begin{aligned} \text{Caplet}(0, T_i, K) &:= \mathbb{E} \left[\tau_i D(0, T_{i+1}) (F_i(T_i) - K)^+ \right] \\ &= \mathbb{E}^{i+1} \left[\tau_i P(0, T_{i+1}) (F_i(T_i) - K)^+ \right] \\ &= \tau_i P(0, T_{i+1}) \mathbb{E}^{i+1} \left[(F_i(T_i) - K)^+ \right], \end{aligned} \quad (9)$$

where $\mathbb{E}^{i+1}[\cdot]$ is the expectation (at time 0) under the T_{i+1} forward measure \mathbb{Q}^{i+1} , $D(0, T_{i+1})$ is the stochastic discount factor and $P(0, T_{i+1})$ is the zero-coupon bond price.

Define the average percentage variance of the lognormally distributed forward rate $F_i(t)$ for $t \in [0, T_i]$ (the integrated instantaneous variance standardized with respect to the time difference) as:

$$\begin{aligned} v_{F_i}^2 &:= \frac{1}{T_i} \mathbb{V}\text{ar}^{i+1} \left[\int_0^{T_i} \frac{dF_i(t)}{F_i(t)} \right] = \frac{1}{T_i} \mathbb{V}\text{ar}^{i+1} \left[\int_0^{T_i} d \log F_i(t) \right] \\ &= \frac{1}{T_i} \mathbb{E}^{i+1} \left[\int_0^{T_i} d \langle \log F_i \rangle_t \right] \\ &= \frac{1}{T_i} \int_0^{T_i} \sigma_i^2(t) dt, \end{aligned} \quad (10)$$

where $\mathbb{V}\text{ar}^{i+1}$ is the variance under the T_{i+1} forward measure \mathbb{Q}^{i+1} . We can then compute the expectation using Black's formula:

$$\mathbb{E}^{i+1} \left[\left(F_i(T_i) - K \right)^+ \right] =: \mathbf{Bl} \left(F_i(0), v_{F_i} \sqrt{T_i}, K, 1 \right) \quad (11)$$

where v_{F_i} is the caplet volatility for a given strike K , and \mathbf{Bl} is the Black's function defined by:

$$\mathbf{Bl} \left(F, v, K, \xi \right) = \xi \left[F \Phi \left(\xi d_1(F, v, K) \right) - K \Phi \left(\xi d_2(F, v, K) \right) \right], \quad (12)$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function, with

$$d_1(F, v \sqrt{T}, K) = \frac{\log \left(\frac{F}{K} \right) + \frac{v^2 T}{2}}{v \sqrt{T}} \quad (13)$$

and

$$d_2(F, v \sqrt{T}, K) = d_1(F, v \sqrt{T}, K) - v \sqrt{T}. \quad (14)$$

Similarly, floorlet prices are given by:

$$\begin{aligned} \mathbf{Floorlet} \left(0, T_i, K \right) &:= \mathbb{E} \left[\tau_i D \left(0, T_{i+1} \right) \left(K - F_i(T_i) \right)^+ \right] \\ &= \tau_i P \left(0, T_{i+1} \right) \mathbf{Bl} \left(F_i(0), v_{F_i} \sqrt{T_i}, K, -1 \right) \end{aligned} \quad (15)$$

A cap (floor) is a collection of caplets (floorlets) with a common strike. Hence the value of the cap (floor) covering the period T_α to T_β with payment dates $\mathcal{T} = \{T_{\alpha+1}, \dots, T_\beta\}$ is simply the sum of the values of the caplets (floorlets):

$$\begin{aligned} \mathbf{Cap} \left(0, \mathcal{T}, K \right) &= \sum_{i=\alpha}^{\beta-1} \mathbf{Caplet} \left(0, T_i, K \right) \\ &= \sum_{i=\alpha}^{\beta-1} \tau_i P \left(0, T_{i+1} \right) \mathbb{E}^{i+1} \left[\left(F_i(T_i) - K \right)^+ \right] \end{aligned} \quad (16)$$

$$= \sum_{i=\alpha}^{\beta-1} \tau_i P \left(0, T_{i+1} \right) \mathbf{Bl} \left(F_i(0), v_{F_i} \sqrt{T_i}, K, 1 \right) \quad (17)$$

Similarly, the price of a floor with the same payment dates \mathcal{T} is given by:

$$\mathbf{Floor} \left(0, \mathcal{T}, K \right) = \sum_{i=\alpha}^{\beta-1} \tau_i P \left(0, T_{i+1} \right) \mathbf{Bl} \left(F_i(0), v_{F_i} \sqrt{T_i}, K, -1 \right) \quad (18)$$

2.3 Pricing swaptions

An Interest-Rate Swap (IRS) is a contract according to which two parties exchange payment streams, typically fixed interest rates against floating interest rates. When the fixed leg is paid and the floating leg is received the IRS is termed Payer IRS (PFS), whereas, in the other case, we have a Receiver IRS (RFS). Payments occur at predetermined dates $\mathcal{T} = \{T_{\alpha+1}, \dots, T_{\beta}\}$. At each time T_{i+1} an PFS investor will pay the amount $\tau_i K$ and receive $\tau_i L(T_i, T_{i+1}) = \tau_i F_i(T_i)$. Hence, the value at time $t < T_{\alpha}$ of his cash flow is equal to:

$$\mathbf{PFS}(t, \mathcal{T}, K) := \sum_{i=\alpha}^{\beta-1} \mathbb{E} \left[\tau_i D(t, T_{i+1}) (F_i(T_i) - K) \middle| \mathcal{F}_t \right] \quad (19)$$

$$= \sum_{i=\alpha}^{\beta-1} \tau_i P(t, T_{i+1}) (F_i(t) - K) \quad (20)$$

$$= \sum_{i=\alpha}^{\beta-1} \tau_i P(t, T_{i+1}) \frac{1}{\tau_i} \left(\frac{P(t, T_i)}{P(t, T_{i+1})} - 1 \right) - \sum_{i=\alpha}^{\beta-1} \tau_i P(t, T_{i+1}) K \quad (21)$$

$$= P(t, T_{\alpha}) - P(t, T_{\beta}) - K \sum_{i=\alpha}^{\beta-1} \tau_i P(t, T_{i+1}) \quad (22)$$

The forward swap rate $S_{\alpha, \beta}(t)$ at time t for the sets of times \mathcal{T} is the rate used in the fixed leg of the an IRS that makes it a fair contract at inception, t . In other words, it is the fixed rate K for which $\mathbf{PFS}(t, \mathcal{T}, K)$ is equal to 0. With this convention we obtain:

$$S_{\alpha, \beta}(t) = \frac{P(t, T_{\alpha}) - P(t, T_{\beta})}{\sum_{i=\alpha}^{\beta-1} \tau_i P(t, T_{i+1})} \quad (23)$$

Dividing both the numerator and denominator by $P(t, T_{\alpha})$ and noticing that $\frac{P(t, T_k)}{P(t, T_{\alpha})} = \prod_{i=\alpha}^{k-1} \frac{P(t, T_{i+1})}{P(t, T_i)} = \prod_{i=\alpha}^{k-1} \frac{1}{1 + \tau_i F_i(t)}$ for all $k > \alpha$, we can express the forward swap rate $S_{\alpha, \beta}(t)$ in terms of the forward-LIBOR rates as:

$$S_{\alpha, \beta}(t) = \frac{1 - \prod_{j=\alpha}^{\beta-1} \frac{1}{1 + \tau_j F_j(t)}}{\sum_{i=\alpha}^{\beta-1} \tau_i \prod_{j=\alpha}^i \frac{1}{1 + \tau_j F_j(t)}} \quad (24)$$

Hence, under the LFMM, the swap rate is not lognormally distributed.

A swaption with strike K is the option to enter an IRS with fixed rate K at a fixed future date T_{α} , with payments occurring at dates $T_{\alpha+1}, \dots, T_{\beta}$. Hence, if we assume

unit notional amount, the payer swaption payoff can be written as:

$$\mathbf{PS}(t, \mathcal{T}, K) := \mathbb{E} \left[\sum_{i=\alpha}^{\beta-1} \tau_i D(t, T_{i+1}) (S_{\alpha, \beta}(T_\alpha) - K)^+ | \mathcal{F}_t \right] \quad (25)$$

Introducing the numeraire $C_{\alpha, \beta}(t) = \sum_{i=\alpha}^{\beta-1} \tau_i P(t, T_{i+1})$, we can consider the probability measure $\mathbb{Q}^{\alpha, \beta}$ associated with it, called the forward-swap measure, and write the swaption price as follows:

$$\mathbf{PS}(t, \mathcal{T}, K) = C_{\alpha, \beta}(t) \mathbb{E}^{\alpha, \beta} \left[(S_{\alpha, \beta}(T_\alpha) - K)^+ | \mathcal{F}_t \right] \quad (26)$$

where $\mathbb{E}^{\alpha, \beta}[\cdot]$ is the expectation (at time t) under the measure $\mathbb{Q}^{\alpha, \beta}$.

As we noticed earlier, the forward swap rate $S_{\alpha, \beta}(t)$ is a function of the forward LIBOR rates and it is not lognormally distributed under the swap measure $\mathbb{Q}^{\alpha, \beta}$.

However, if we assume it is, we can write:

$$dS_{\alpha, \beta}(t) = \sigma_{S_{\alpha, \beta}}(t) S_{\alpha, \beta}(t) dW^{\alpha, \beta}(t), \quad (27)$$

where $\sigma_{S_{\alpha, \beta}}(t)$ is the instantaneous volatility at time t of the forward swap rate $S_{\alpha, \beta}(t)$ and $W^{\alpha, \beta}(t)$ a Brownian motion under the forward swap measure $\mathbb{Q}^{\alpha, \beta}$. Define the average percentage variance of the lognormally distributed forward swap rate $S_{\alpha, \beta}(t)$ for $t \in [0, T_\alpha]$ (the integrated instantaneous variance standardized with respect to the time amount) as:

$$\begin{aligned} v_{S_{\alpha, \beta}}^2 &:= \frac{1}{T_\alpha} \mathbb{V}\text{ar}^{\alpha, \beta} \left[\int_0^{T_\alpha} \frac{dS_{\alpha, \beta}(t)}{S_{\alpha, \beta}(t)} \right] = \frac{1}{T_\alpha} \mathbb{V}\text{ar}^{\alpha, \beta} \left[\int_0^{T_\alpha} d \log S_{\alpha, \beta}(t) \right] \\ &= \frac{1}{T_\alpha} \mathbb{E}^{\alpha, \beta} \left[\int_0^{T_\alpha} d \langle \log S_{\alpha, \beta} \rangle_t \right] \\ &= \frac{1}{T_\alpha} \int_0^{T_\alpha} \sigma_{S_{\alpha, \beta}}^2(t) dt, \end{aligned} \quad (28)$$

where $\mathbb{V}\text{ar}^{\alpha, \beta}$ is the variance under the forward swap measure $\mathbb{Q}^{\alpha, \beta}$. This model for the evolution of the forward swap rate is also known as lognormal forward swap market model (LSMM).

With this assumption of lognormal distribution for the forward swap rate, we can then compute the swaption price using Black's formula :

$$\mathbf{PS}(t, \mathcal{T}, K) = C_{\alpha, \beta}(t) \mathbf{BI} \left(S_{\alpha, \beta}(t), v_{S_{\alpha, \beta}} \sqrt{T_\alpha}, K, 1 \right) \quad (29)$$

Although this assumption of lognormally distributed forward swap rates is not compatible with the LFMM assumptions that lead to expression (24), the swaption volatility $v_{S_{\alpha,\beta}}^2$ can be approximated from the LFMM parameters so that the swaptions prices can be computed from (29).

The following sections detail the different assumptions and approximations we used to calibrate our LFMM to compute the swaption volatilities from the LFMM parameters.

For further details, see Chapter 6 of Brigo & Mercurio (2006).

3 Calibration

3.1 Data

3.1.1 Data description

South African swap curve data was sourced from the JSE, and market cap and swaption volatilities were sourced from Bloomberg. The swap curve data included yield rates up to 30 years, with a total of 165 points per yield curve. In total, there were 3559 business days worth of yield curve data, starting at 2 January 2004, up until 29 March 2018.

The cap market data included data from 26 January 2012 up until 30 March 2018, which included a total of 1612 observations. These observations were of caps with terms to maturity that ranged between 1 year to 10 years, each with a common tenor of a quarter of a year. The data included at the money caps, as well as caps with a range of absolute strikes.

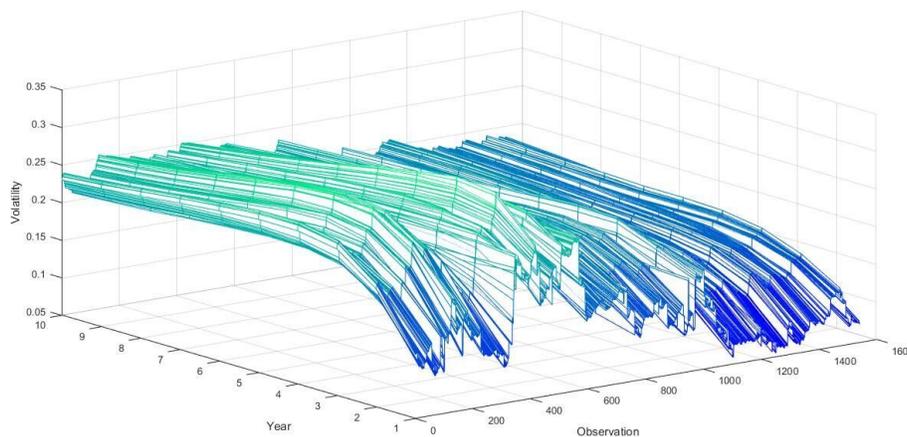


Figure 1: South African Swap Curve

The swaption data included volatilities from 24 November 2010 until 30 March 2018. This data excluded a large number of values throughout, between the start of the data, and until 24 December 2010. Thus, it included a total of 1892 viable observations. Both the terms and the tenor of the swaption observations varied, with the tenor ranging between 1 and 10 (in steps of 1), and the term ranging between 0.08333 and 10.

3.1.2 Data thinning

Not all data that was collected was relevant to the final outcome. As a result, a data thinning algorithm was generated in order to remove data points from the final outcome which were not necessary.

The cap and swap data included data from days which were public holidays in South Africa. This meant that there wasn't any corresponding yield curve data from the JSE. In order to manage this problem, there were two possible avenues. Either estimate the yield rates on the public holidays through some linear interpolation scheme, or exclude the additional Bloomberg data from further analysis. The latter option was deemed more appropriate given that data from a public holiday could not be actual market data, and the aim of using the time series data was simply to determine trends in the market over time. There was enough data without these additional points to get a general idea of these trends.

In addition to this, there was one data point in the yield curve data which appeared to deviate drastically from all others. For example, the 30 year yield rate was more than a third that of the 30 year yield rate for the previous day and the following day.

3.1.3 Data analysis

Figure 2 shows a plot of the yield curves across time. This shows that the majority of South African yield curves have a similar structure: yield rates increase for shorter maturities and then taper off for longer maturities. Yield rates themselves have not necessarily been consistent which is to be expected given the period of time that the data covers. The most recent data points, however, appear to have relatively consistent yield rates in terms of the structure of the yield curve. These trends are consistent amongst the interpolated rates. This implies that the interpolation scheme is sufficient to capture the market trends, whilst being more appropriate for the assumptions made by the LFMM model.

On the whole, the South African forward rates have maintained a consistent trend over time. Figure 7 suggests that there are pockets of deviation from this trend, but it appears that these deviations occur over short enough periods of time that they do not constitute a shift in the underlying term structure of the forward rates themselves. Promisingly, the most recent forward rates appear to be inline with

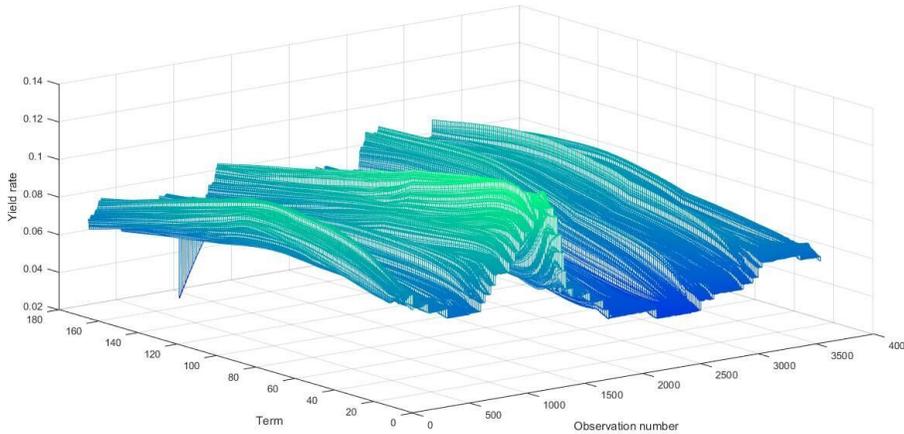


Figure 2: South African Swap Curve

this general trend, implying that they would be relatively good for estimation. Figure 8, which shows the caplet volatility estimate results, suggested a shift in the structure of caplet volatilities. Initially, the trend in caplet volatilities with respect to term to maturity was to increase then to gradually decrease beyond some local maximum. However, more recent caplet volatility data suggests that caplet volatilities generally increase with term to maturity, although the rate of increase is generally steeper initially, then decreases with term to maturity. The recent consistency of this trend would suggest that more recent caplet volatilities would be better suited to estimating the future volatility rates than a longer historical average.

As the swaption data was essentially non-existent between the start of the data, and until 24 December 2010, these points were excluded from analysis. In addition, there were many observations which were incomplete; in order to manage these observations, the empty instances within each observation was estimated by the previous day's observation. In order to analyse the resulting data, we decided to get an idea for market trends through analysing the 1, 5 and 10 year term swaptions along all their tenors. This was deemed acceptable as these terms give an idea of the short, medium and long term nature of the swaption market. The plots of these can be found in Figure 3, Figure 4 and Figure 5. From these plots it is clear that there is a great deal of oscillation in the market structures. This would likely suggest that it would be particularly difficult to model these volatilities, as well as to appropriately use past data for the sake of calibrating any model that would be used to model these volatilities (and prices). Across all tenors, there tends to be a great deal of change in the quoted volatilities, whereas they all tend to converge to some common volatility as the tenor increases. This would suggest that

the modelling of longer tenor swaps may be reasonable.

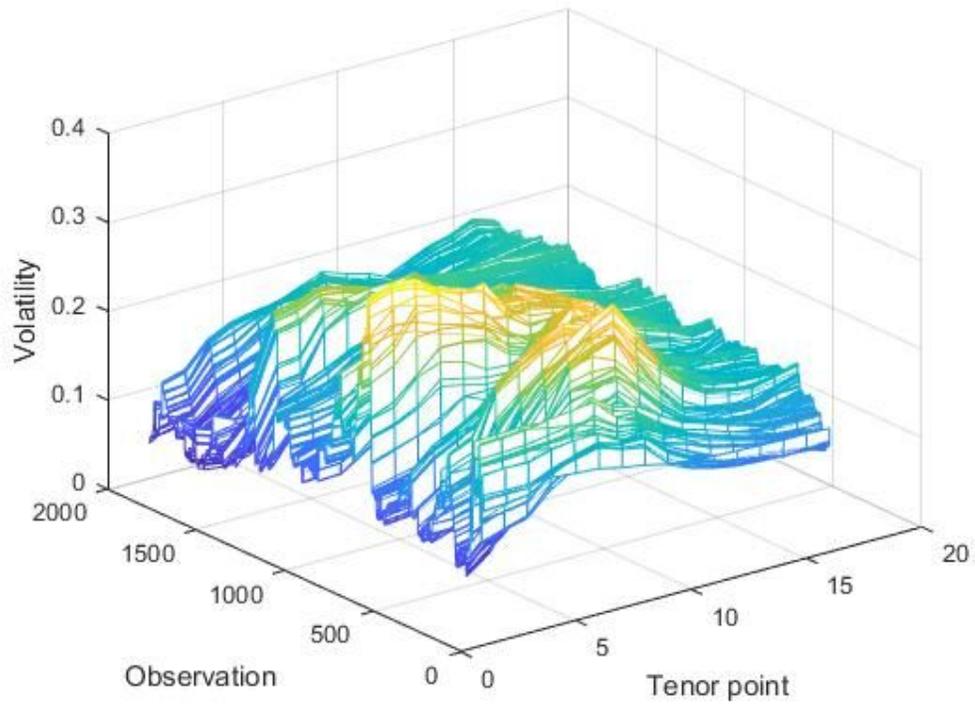


Figure 3: South African Swap Volatilities (1 Year Term)

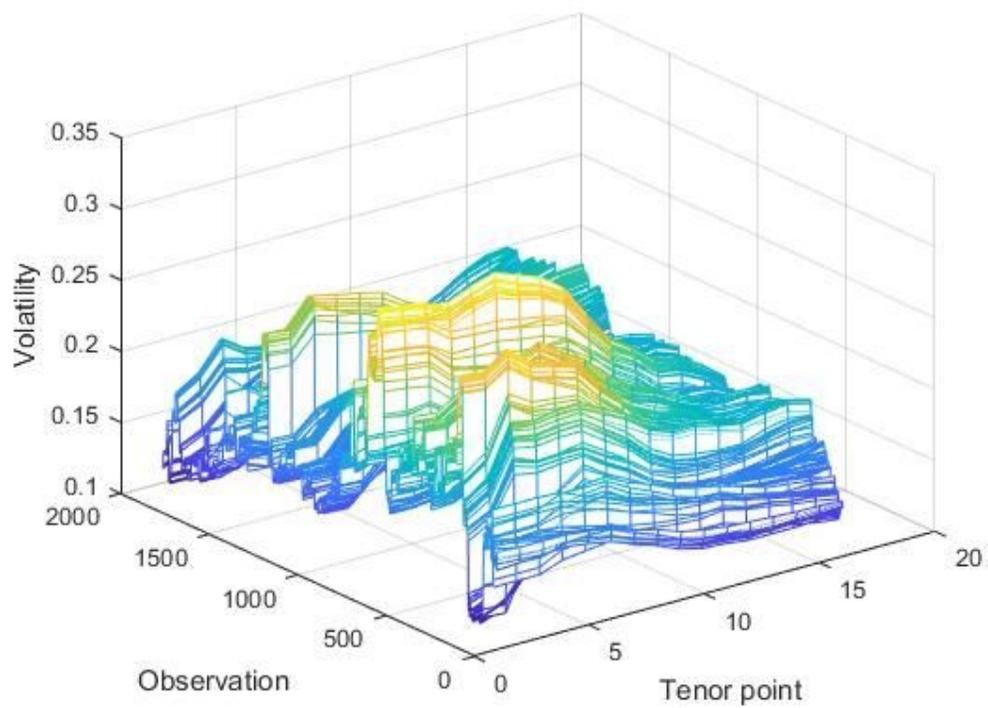


Figure 4: South African Swap Volatilities (5 Year Term)

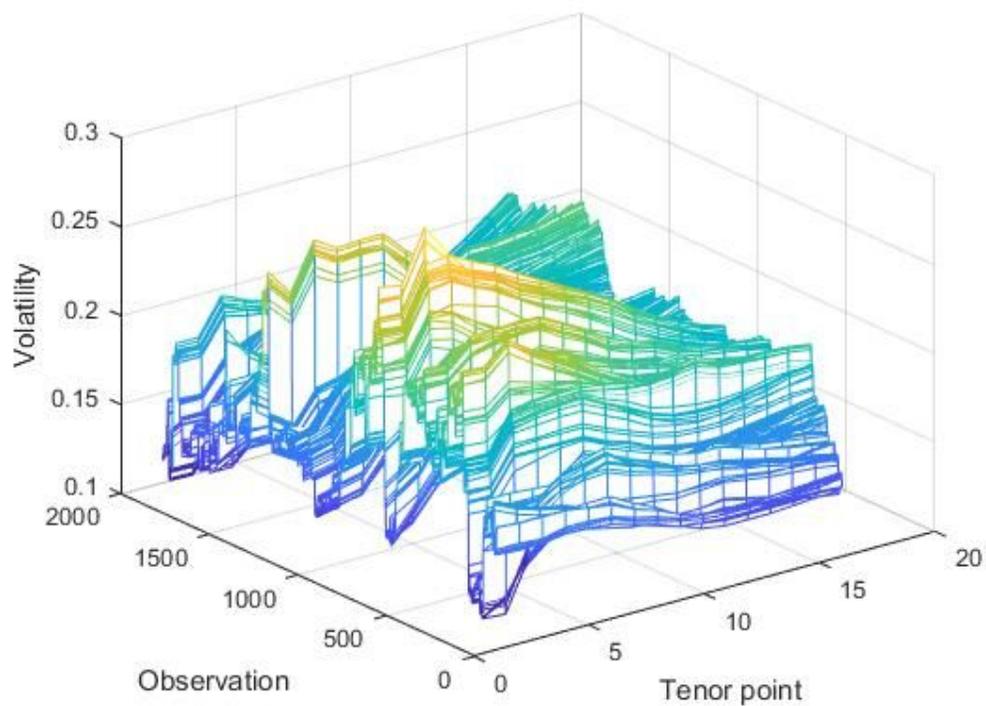


Figure 5: South African Swap Volatilities (10 Year Term)

3.1.4 Yield rate adjustments

One of the assumptions of the LFMM is that the swaption term is exactly a quarter of a year. This implies that, for example, that the first payment occurs after 91.25 days. Now, the yield rates that occur at this sub-day periods, are not directly observable in the market. In order to estimate these rates from the data available, a linear interpolation scheme was used.

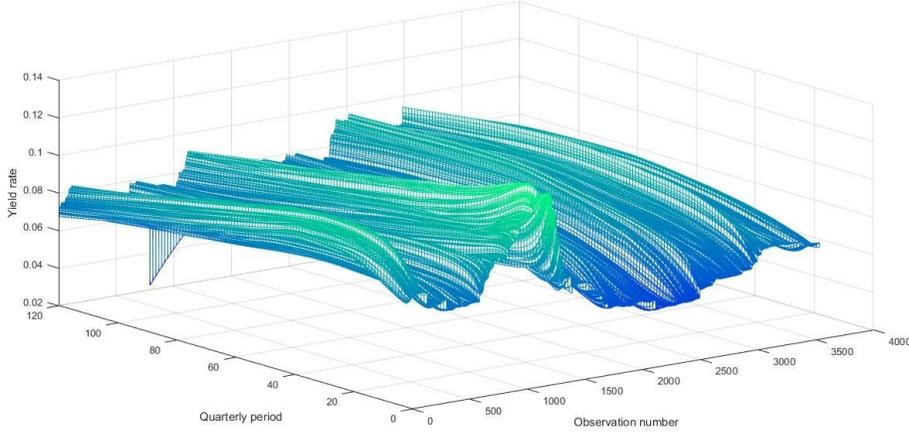


Figure 6: Interpolated South African Swap Curve

3.2 Determining South African forward rates

There are several methods to determine the forward rates implied by the South African swap curve. The method used in this paper is as follows:

- Determine the zero-coupon bond values implied by the South African swap curve. This can be done by using the following equation:

$$P(0, T_i) = e^{-r(0, T_i)T_i} \quad (30)$$

- Determine the forward rates implied by the zero-coupon bond values through the following equation:

$$F_i(0) := F(0; T_i, T_{i+1}) = \frac{1}{\tau_i} \left(\frac{P(0, T_i)}{P(0, T_{i+1})} - 1 \right). \quad (31)$$

The resulting forward rates can be seen in Figure 7.

3.3 Estimating caplet volatilities

The cap data includes the cap volatilities for caps with annual term to maturities, up until a term to maturity of 10 years. However, there is no direct observation of caplet volatility data. As a result of the structure of this data, there are 4 caplet volatilities to estimate between calendar dates. There are several methods of estimating these caplet volatilities, which range in terms of the degree of their complexity. The simplest method was elected for use in this paper; the reason for this

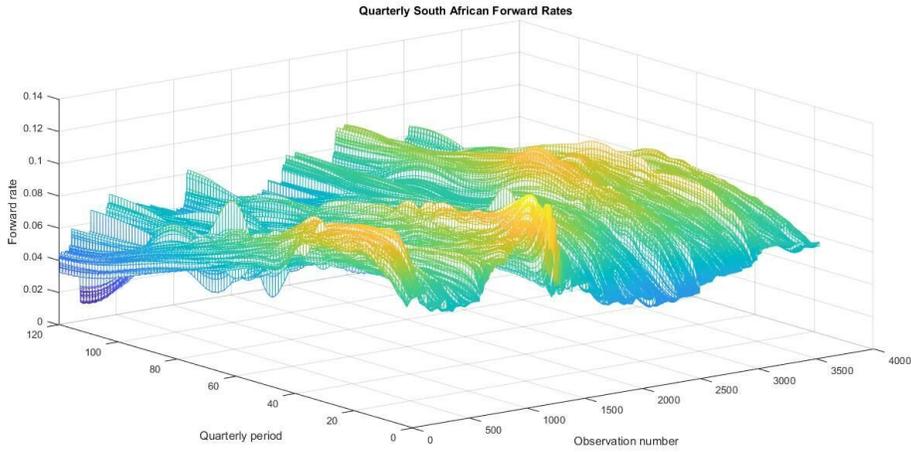


Figure 7: Implied South African Forward Rates

is that our focus is the implementation of the LFMM. Simpler methods of caplet volatility estimation give sufficient estimates, and are appropriate for the sake of determining model adequacy. The cap data available for the sake of caplet estimation was a set of at-the-money caps, as well as caps with the same absolute strikes and terms to maturity. Bloomberg appears to use its own yield curve in determining the at-the-money strike for these caps, and thus the corresponding volatility. As a result, it was deemed sufficient to consistently use the at-the-money caps for the sake of estimating the at-the-money volatility. The method elected for use in this paper is bootstrapping. This method's strength, despite its simplicity, is its efficiency and that it will always provide a solution, given that a solution exists for the system. It was further assumed that caplet volatilities were constant over calendar years. The algorithm for this method can be found in Iwashita & White (2014), and is as follows:

- Sort the caps into increasing order of maturity
- Determine the price of the caps implied by their market volatilities using the Black-Scholes formula for caps
- Generate a series of differences between these cap prices, setting the first element equal to the price of the cap with the shortest term to maturity
- Assign caplets to the period corresponding to the relevant price differences
- For each disjoint period, assign a common volatility to the caplets and solve for the volatility that can be substituted into the Black-Scholes formula, which gives the relevant price difference

The last line essentially breaks down into solving the following equation:

$$\mathbf{Cap}(0, \mathcal{T}_j, K) - \mathbf{Cap}(0, \mathcal{T}_l, K) = \sum_{i=\beta_l}^{\beta_j-1} \tau_i P(0, T_{i+1}) \mathbf{Bl}(F_i(0), v_{\text{Caplet}} \sqrt{T_i}, K, 1)$$

Where $\mathcal{T}_j = \{T_{\alpha+1}, \dots, T_{\beta_j}\}$ and $\mathcal{T}_l = \{T_{\alpha+1}, \dots, T_{\beta_l}\}$, with $\beta_j < \beta_l$.

Iwashita & White (2014) suggest the use of a one dimensional root finding algorithm in order to solve for the caplet volatilities from the price differences. The *fzero* Matlab function was used to solve these equations.

A simple check as to whether the bootstrapping algorithm was applied correctly is to reprice the caps as a sum of the caplet prices, using the bootstrapped caplet volatilities. If the bootstrapping algorithm was applied correctly, then these prices will be equal. This was the case when we estimated the caplet volatilities, and we can conclude that, given the parameter inputs, the bootstrapping algorithm was applied correctly.

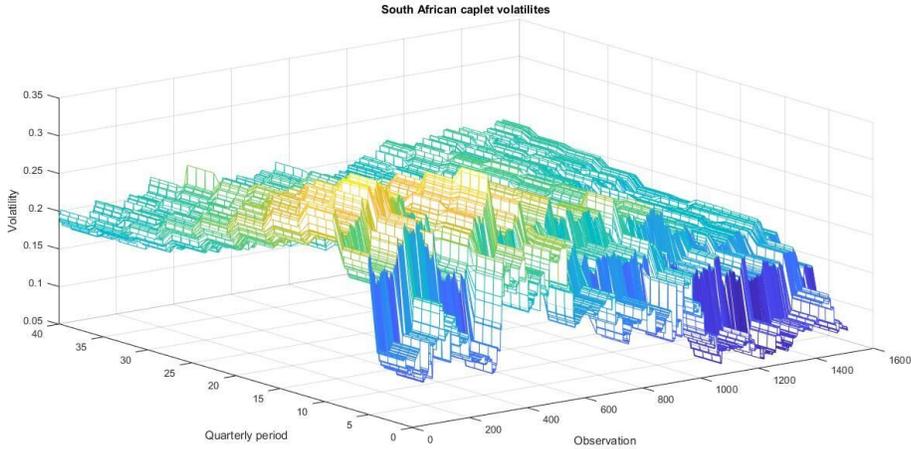


Figure 8: Time Series of Estimated Caplet Volatilities

3.4 Calibrating instantaneous volatilities

The instantaneous volatilities can be calibrated from the caplet volatilities. This is essentially finding the curve which best fits the caplet volatility data. One of the problems that is required to be solved in this paper is to determine the optimal structure of the instantaneous volatilities. Thus, using no *a priori* knowledge of the exact structure of the instantaneous volatilities, it was deemed appropriate to

use more general parameterisations rather than less, whilst maintaining a parsimonious approach to this volatility estimation process. One such parameterisation, suggested by Brigo & Mercurio (2006) is:

$$\sigma_i(t) = [a(T_{i-1} - t) + d]e^{-b(T_{i-1}-t)} + c$$

Where the values of a, b, c and d can be estimated through the following relationship:

$$v_{i-Caplet}^2 = \frac{1}{T_{i-1}} \int_0^{T_{i-1}} ([a(T_{i-1} - t) + d]e^{-b(T_{i-1}-t)} + c)^2 dt$$

Where $v_{i-Caplet}^2$ is the estimated volatility of the i^{th} caplet, and T_{i-1} is the exercise date of the i^{th} caplet. The least-squares method was applied to the above in order to derive estimates for the values of a, b, c and d in the context of the market data. An issue that arises when applying such a general formula is that the values for the parameters are not guaranteed to be unique, within some level of tolerance for the minimisation of the sum of squares. This would reduce the ability to compare these parameter values over time. Whilst this isn't necessarily the biggest drawback, it does diminish the general strength of this type of parameterisation over that of a simpler parameterisation.

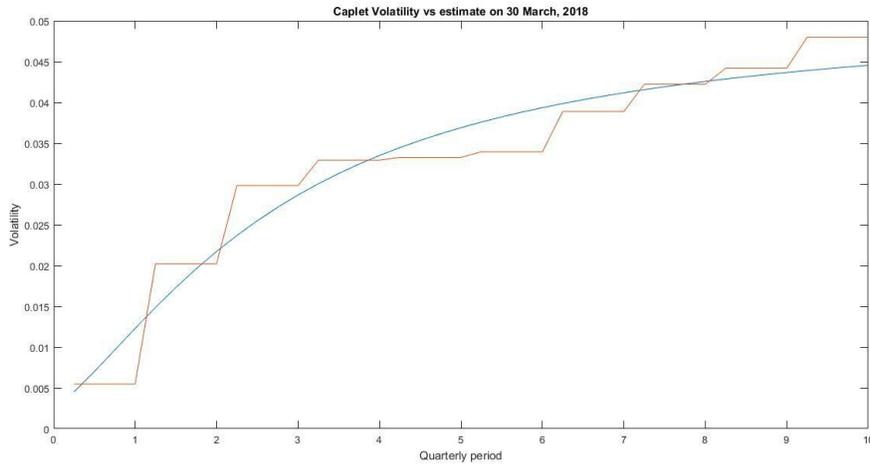


Figure 9: Instantaneous Volatility vs Continuous caplet Volatilities

The other instantaneous volatility structure used, also suggested by Brigo & Mercurio (2006) was based off the assumption that the volatilities depend only the term

to maturity. Through this assumption, one can construct the following table of instantaneous volatilities:

Table 1: Matrix of instantaneous volatilities $\sigma_{i,j}(t)$

Instantaneous Volatilities	Time $t \in (0; T_0]$	$(T_0; T_1]$	$(T_1; T_2]$...	$(T_{M-2}; T_{M-1}]$
Forward Rate $F_1(t)$	η_1	Dead	Dead	...	Dead
$F_2(t)$	η_2	η_1	Dead	...	Dead
...
$F_M(t)$	η_M	η_{M-1}	η_{M-2}	...	η_1

The above η values allow for an exact fit of the instantaneous volatilities to the estimated caplet volatilities. Given the evaluation method used to find the caplet volatilities, this instantaneous volatility structure constitutes a piece-wise constant instantaneous volatility assumption.

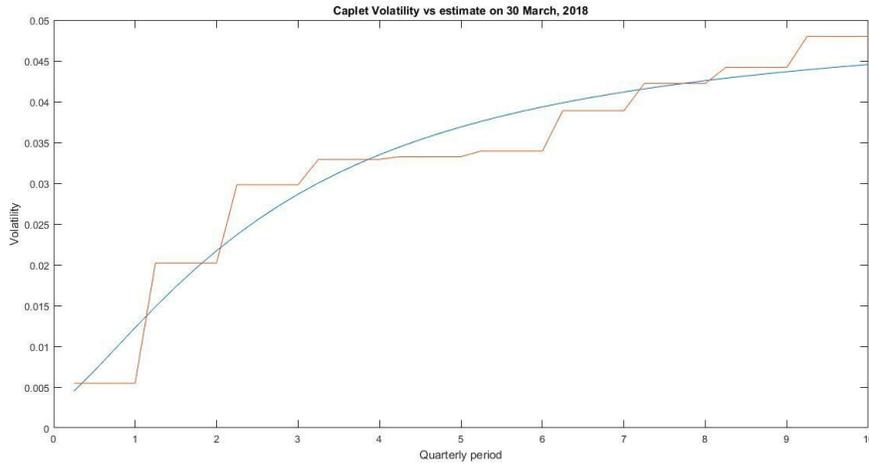


Figure 10: Instantaneous Volatility vs Constant caplet Volatilities

Both sets of instantaneous volatility assumptions allow for a reasonably good illustration of the effect of the structure of the instantaneous volatility on the final swaption and caplet prices predicted by the LFMM. Thus, they are sufficient to achieve the goal of this paper: to check whether the LFMM is a reasonably good fit for the South African inter-bank market.

3.5 Estimating historical correlation

Historical correlation of the forward rates can be estimated by the correlation of the log returns of the forward rates. This is a result of the lognormal result of the LFMM. The estimated historical correlation has the characteristics that one would generally expect from forward correlations, namely a so-called tent shape. An interesting feature of the calculated historical correlation matrix is that it has some negative correlations. These negative correlations tend to occur beyond the 10 year term points.

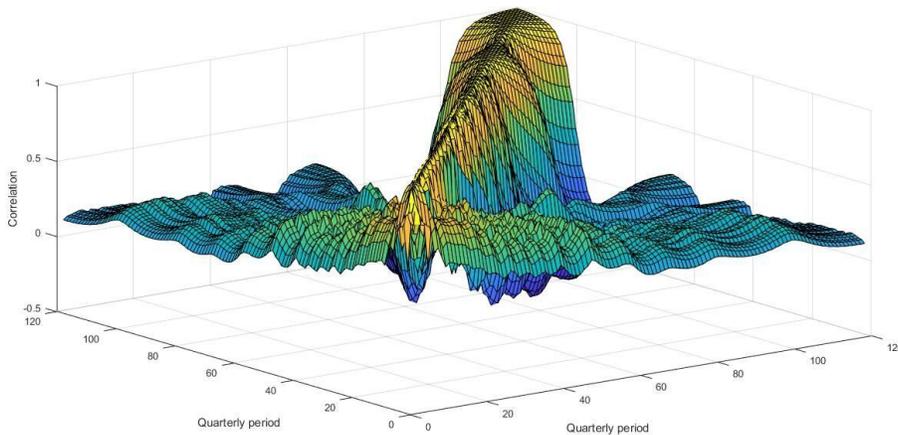


Figure 11: Historical Forward Correlations for South Africa

3.6 Swaption price approximation

In this section we briefly review two approximation formulae for swaption pricing, namely the Rebonato formula and the Hull-White formula. For further details see Brigo & Mercurio (2006) and McWalter & Van Appel (2018)

Rebonato noted that the swap rate can be seen as a linear combination of forward rates (Rebonato, 1999). In other words, a swap rate is a weighted sum of forward rates between maturity (α) and expiry (β), where the weight of each Libor forward rate w_i is dependent on the time t and all rates F_i from maturity to this rate, i.e.

$$S_{\alpha,\beta}(t) = \sum_{i=\alpha}^{\beta-1} w_i(t) F_i(t), \quad (32)$$

with weights defined by

$$w_i(t) = \frac{\tau_i \prod_{j=\alpha}^i \frac{1}{1+\tau_j F_j(t)}}{\sum_{k=\alpha}^{\beta-1} \tau_k \prod_{j=\alpha}^k \frac{1}{1+\tau_j F_j(t)}}. \quad (33)$$

Moreover, Rebonato made the following simplifying assumptions for calculating a swaption rate:

- each LIBOR forward rate $F_i(t)$ and its weight $w_i(t)$ are independent;
- each weight $w_i(t)$ stays stable over time and thus can be approximated by its initial value, i.e. $w_i(t) \approx w_i(0)$;
- each $F_i(t)$ evolution is stable over time and thus its volatility is negligible, i.e. $F_i(t) \approx F_i(0)$.

All assumptions together lead to the following approximation

$$S_{\alpha,\beta}(t) = \sum_{i=\alpha}^{\beta-1} w_i(t) F_i(t), \quad (34)$$

which can be further specified as the Rebonato formula below.

Approximation 3.1 (The Rebonato formula). *The squared Black swaption volatility is*

$$(v_{S_{\alpha,\beta}})^2 \approx \frac{1}{T_\alpha} \sum_{i,j=\alpha}^{\beta-1} \frac{w_i(0)w_j(0)F_i(0)F_j(0)}{S_{\alpha,\beta}^2(0)} \rho_{ij} \int_0^{T_\alpha} \sigma_i(t)\sigma_j(t) dt. \quad (35)$$

Hull & White (2000) extended the Rebonato formula by using a Taylor expansion of first order to approximate the weight w_i .

Approximation 3.2 (The Hull-White formula). *The squared Black swaption volatility is*

$$(v_{S_{\alpha,\beta}})^2 \approx \frac{1}{T_\alpha} \sum_{h,j=\alpha}^{\beta-1} G_{h,j}(0) \rho_{hj} \int_0^{T_\alpha} \sigma_h(t)\sigma_j(t) dt. \quad (36)$$

where

$$G_{h,j}(t) = \frac{\tilde{w}_h(t)\tilde{w}_j(t)F_h(t)F_j(t)}{S_{\alpha,\beta}^2(t)}. \quad (37)$$

and

$$\tilde{w}_h(t) = w_h(t) + \sum_{i=\alpha}^{\beta-1} F_i(t) \frac{\partial w_i(t)}{\partial F_h} \quad (38)$$

and

$$\frac{\partial w_i(t)}{\partial F_h} = \frac{w_i(t)\tau_h}{1 + \tau_h F_h(t)} \left[\frac{\sum_{k=h}^{\beta-1} \tau_k \prod_{j=\alpha}^k \frac{1}{1+\tau_j F_j(t)}}{\sum_{k=\alpha}^{\beta-1} \tau_k \prod_{j=\alpha}^k \frac{1}{1+\tau_j F_j(t)}} - \mathbb{I}_{\{i \geq h\}} \right] \quad (39)$$

In the remainder of this paper, for all calibration purposes, we consider only approximations generated using the Hull-White formula.

3.6.1 Note on implementation

There are several important considerations to make when implementing both the Rebonato and Hull-White Formula Hull & White (2000) for approximating Black swaption volatilities under the LFMM. For the purpose of calibrating the LFMM to the given market data one needs to generate, as efficiently as possible, a grid of Black swaption volatilities with corresponding term (α) and tenor ($\beta - \alpha$). Thus, one would have to structure the underlying code in such a way that is best suited to calculate these grids in aggregate.

Calculating these approximations from scratch for each specific α and β results in much of the same information being generated repetitively, ultimately rendering the calibration process far too time consuming. It is possible to separate much of the information common to all α and β combinations so as to only generate this once, which greatly improves the efficiency of generating the grid of Black swaption volatilities.

For example, when calculating the weights used in the Hull-White Formula given by 38, one would (only once, used for all combinations of α and β) create a ZxM matrix with entries given by

$$\text{where } d_{i,j} = \begin{cases} \tau \prod_{j=\alpha_i}^j \frac{1}{1+\tau_j F_j(t)} & \text{for } j > \alpha_i; \\ 0 & \text{otherwise;} \end{cases} \quad (40)$$

where $i = 1, \dots, Z$ and $j = \alpha_i, 2, \dots, M$ where Z is total number of α_i 's considered. From this, it is possible (through various selective summation schemes) to extract the information necessary to calculate the respective weights for each considered α and β combination.

3.7 Correlation parametrisation

For a matrix to be considered a valid correlation matrix, the following properties must hold:

- $|\rho_{i,j}| \leq 1$ for all i, j
- ρ positive semi-definite
- $\rho_{i,i} = 1$ for all i

Furthermore, correlation of Forward Rates should yield the following desired properties:

- $i \rightarrow \rho_{i,j}$ increasing for all $i \geq j$ "decreasing along row"
- $i \rightarrow \rho_{i+p,i}$ increasing for fixed p "increasing along sub-diagonals"

In the following we consider several parsimonious correlation parametrisations:

- A Single-Parameter parametrization

$$\rho_{i,j} = \exp(-\beta|T_i - T_j|), \quad \beta \geq 0. \quad (41)$$

- A stable Two-Parameter parametrization

$$\rho_{i,j} = \exp \left[-\frac{|i-j|}{M-1} \left(\log \rho_\infty + \eta \frac{M-1-i-j}{M-2} \right) \right] \quad (42)$$

where $\rho_\infty = \rho_{1,M}$ is the correlation between the farthest forward rates in the family considered, and $-\log(\rho_\infty) > \eta$.

- A Three-Parameter parametrization:

$$\rho_{i,j} = \exp \left[-|i-j| \left(\beta - \frac{\alpha_2}{6M-18} (i^2 + j^2 + ij - 6i - 6j - 3M^2 + 15M - 7) \right. \right. \\ \left. \left. + \frac{\alpha_1}{6M-18} (i^2 + j^2 + ij - 3Mi - 3Mj + 3i + 3j + 3M^2 - 6M + 2) \right) \right]$$

for $i = 2, 3, \dots, M-1$ and $i = M$ where the parameters should be constrained to be non-negative in order to ensure all desired characteristics for correlation.

- An improved Two-Parameter parametrization

$$\rho_{i,j} = \exp \left[-\frac{|i-j|}{M-1} \left(-\log \rho_\infty \right. \right. \\ \left. \left. + \eta \frac{i^2 + j^2 + ij - 3Mi - 3Mj + 3i + 3j + 3M^2 - M - 4}{(M-2)(M-3)} \right) \right]$$

with the same constraint as before. This version of Two-Parameter parametrisation has the additional advantage that it is increasing along the sub-diagonal, so further smoothing the correlation structure at the diagonal.

There exists a reduced-rank parameterisation, as suggested by Rebonato (1999) and discussed by Brigo & Mercurio (2006), which we shall denote as the *Rebonato's Angles* parameterisation. We know that

$$\rho = PHP' \quad (43)$$

where P is a real orthogonal matrix, and H is a diagonal matrix of positive eigenvalues of ρ as ρ is positive semi-definite. If we set $A := P\Lambda$, where Λ is the diagonal matrix whose entries are the square roots of those in H , we yield the decomposition $\rho = AA'$. We can mimic this by means of a suitable $M \times n$ matrix B such that $\rho^B = BB'$. However, for the sake of simplicity of calibration, we consider the specific case of $n = 2$. There are certain conditions that B has to meet for ρ to be a valid correlation matrix. Rebonato (1999) suggests a general form for the rows of such B , which in the two-factor case reduces to

$$\rho_{i,j} = \cos(\theta_i - \theta_j) \quad 0 \leq \theta \leq \frac{\pi}{2} \quad (44)$$

This parameterisation ensures that all the necessary characteristics of a correlation matrix are maintained. However, the desired characteristics may not be present. A least squares method was used to calibrate the above models to the historical market correlations. The *lsqnonlin* Matlab function was used to achieve this, which was built specifically to handle least squares minimisation for non-linear functions.

4 Methodology

There exist two central approaches for calibrating the LFMM. We consider both avenues, assessing appropriateness of fit and implications in each situation.

4.1 Fixing volatilities and calibrating correlation

Here, instantaneous volatilities of forward rates are fixed to that implied by the market under the continuous parameterisation. Then, the following correlation parameterisations are calibrated to the swaption prices:

- Single-Parameter
- Two-Parameter
- Two-Parameter Improved
- Rebonato's Angles (M-Parameter)

A test of the quality of fit to market swaption prices is then conducted, using the absolute relative error.

4.2 Fixing correlation and calibrating volatility

Instantaneous correlations are fixed using the markets historical correlation data. Then, the continuous parameterisation of instantaneous volatilities is calibrated to:

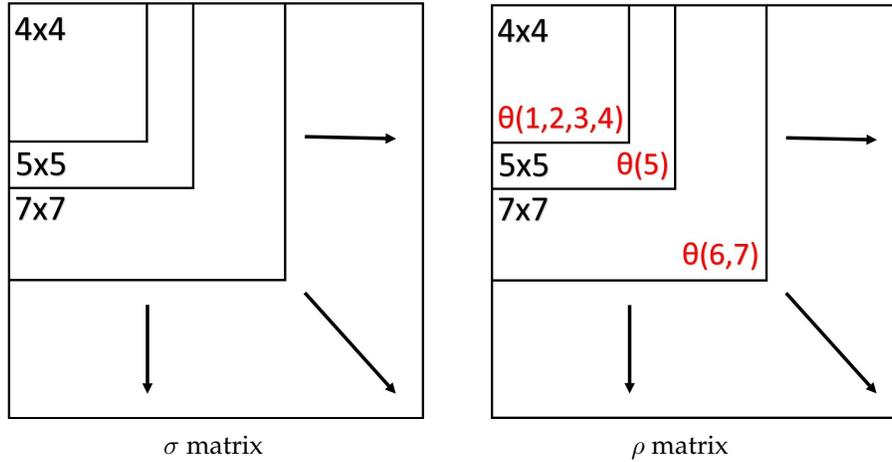
- Swaption prices exclusively
- A weighted combination of swaption and cap prices

A test of the quality of fit to market swaption prices is then conducted, using the absolute relative error. Furthermore, the resultant instantaneous volatilities of the forward rates are compared to those implied by the cap market data.

4.3 A note on implementation of Rebonato's Angle Parameterisation

A hybrid bootstrapping technique is used to calibrate Rebonato's Angle parameterisation. Initially, the first four θ_i 's (corresponding to the first four forward rates) are calibrated by optimising them to fit the swaption price which incorporates only the first four forward rates. This is done by selecting the upper-left 4x4 segment of the σ -matrix and ρ -matrix (as seen below), which is then used to generate the approximation of the swaption price.

One would then calibrate θ_5 using the swaption which incorporates the first five forward rates by selecting the upper-left 5x5 segment of both matrices, with $\theta_1, \theta_2, \theta_3, \theta_4$ fixed from the previous calibration. It is notable that there exist gaps in the swaption expiries (for example, when the previous swaption has ten relevant forward rates, and the next available swaption has twelve) such that several θ 's may need to be calibrated simultaneously. This process continues cascading down the matrices, until one encounters the problem when there are two swaptions incorporating the same forward rates (for example, the 2x2 and 1x3 swaptions). In this case, one could consider a weighting of these two swaption prices. However, for the sake of simplicity, we select the swaption with the longer term to maturity as it includes the least incremental information about the θ 's.



5 Results

In this section we present the results of the two different calibration approaches mentioned in the previous section. In all calibration approaches we aim to minimise the average relative errors between model-based swaption volatility and the swaption volatility implied by the market. Recall that the relative error of calibrating a swaption volatility $\sigma_{\alpha,\beta}^{\text{market}}$ is defined as $\epsilon_{\alpha,\beta} := \frac{|\sigma_{\alpha,\beta}^{\text{market}} - \sigma_{\alpha,\beta}^{\text{model}}|}{\sigma_{\alpha,\beta}^{\text{market}}}$ where $\sigma_{\alpha,\beta}^{\text{model}}$ denotes the corresponding model-based swaption volatility.

In the first approach, we fix the instantaneous volatilities of forward rates and calibrate the correlations to the swaption prices in the market. The fitted instantaneous correlations of each parametrisation are presented in Figure 12, Figure 13, Figure 14 and Figure 15, and the fitting quality (relative errors of swaption calibration) are summarized in Table 2. We see that the first three parametrization give a poor fit, but the calibrated correlation structures are rigid, smooth and contains all the desired characteristics. The Rebonato's angle parameterisation has the smallest calibration error for approximating the market swaptions due to the large number of parameters that allow for extra freedom in the calibration. The calibrated correlation structure suffices the necessary conditions, i.e. positive semi-definite matrix. However it oscillates with large amplitudes as it might have incorporated too much market noise of the swaption volatilities from the market.

In the second approach, we fix the instantaneous correlation and calibrate the instantaneous volatilities of the forward rates to either only the swaption prices or to a weighted combination of Caps and Swaptions prices. Figure 16 and Figure 17 shows the relative errors of fitting only swaptions and jointly fitting swaptions and caps respectively. We see in both cases significant calibration errors. We notice

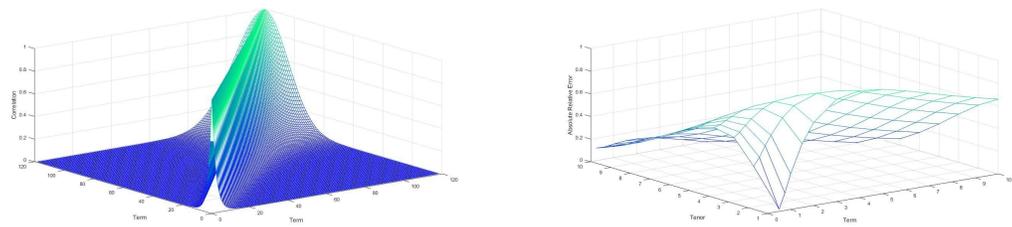


Figure 12: Swaption calibration using the One-factor parameterisation. Left: instantaneous correlation; Right: relative error of the swaption calibration. Each point on the grid represents the relative error of calibrating to that swaption with the corresponding term and tenor.

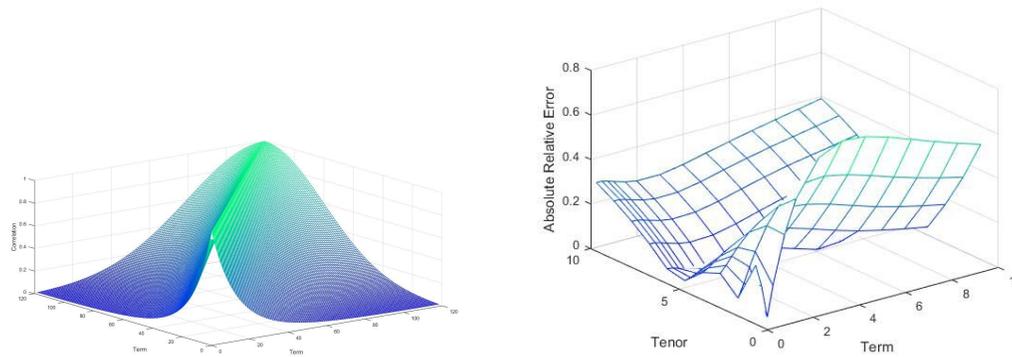


Figure 13: Swaption calibration using the Two-factor parameterisation. Left: instantaneous correlation; Right: relative error of the calibration.

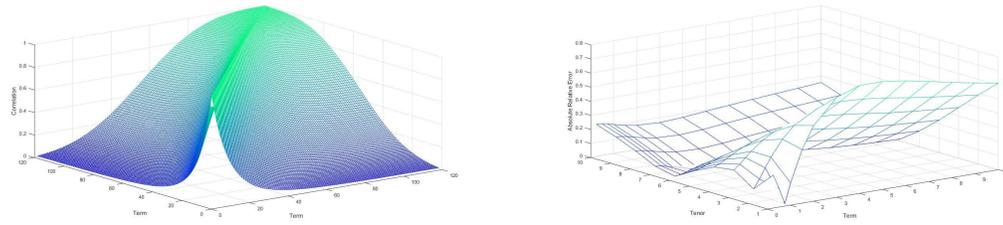


Figure 14: Swaption calibration using the improved Two-Parameter parameterisation. Left: instantaneous correlation; Right: relative error of the calibration.

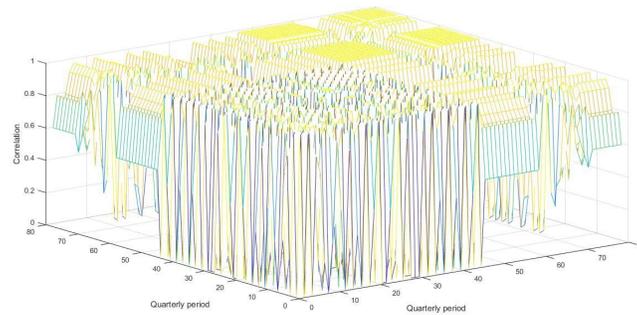


Figure 15: Swaption calibration using the Rebonato's angle parametrization.

that the first one is slightly better, as it completely ignored the volatility structure of the caps. Indeed, the calibrated volatility structure differs significantly from the volatility structure implied from the Caps in the market.

	Single-Parameter	Two-Parameter	improv. Two-Parameter	Rebonato's
	β	ρ_∞, η	$\tilde{\rho}_\infty, \tilde{\eta}$	angle θ
average				
relative error	36.14376%	35.28341%	34.37879%	5.46626%

Table 2: Relative error of calibrating correlation to Swaption prices from the market. The best calibration is obtained by using Rebonato's angles.

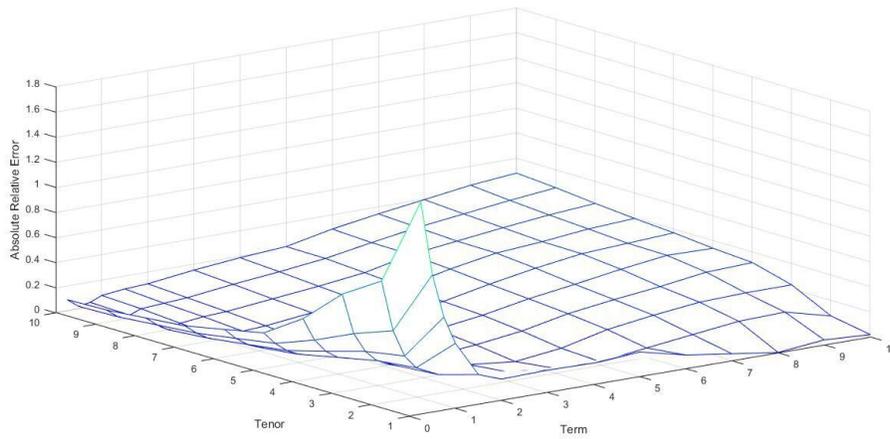


Figure 16: Relative error of calibrating swaption prices by assuming historical correlation between the underlying forwards. Here we only assume continuous parametrization for volatility of Caps. The average relative error is 20.85398%.

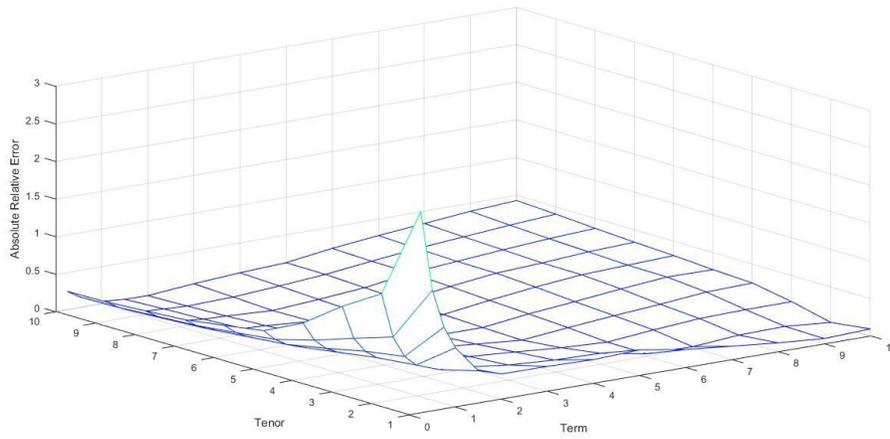


Figure 17: Relative error of calibrating swaption prices. Here we calibrate the instantaneous volatilities to the cap and swaption prices using the fixed instantaneous correlations (historical correlations). Further we assume equal weights for minimizing the relative errors of both the cap and swaption prices. The average relative error is 25.09454%.

6 Conclusion

Within the two approaches of calibration of the LFMM, the results above give a clear indication as to the appropriateness of the model within the South African inter-bank interest rate market.

Using fixed instantaneous volatilities implied by the cap data:

- The correlation parameterisations that maintained the desirable characteristics resulted in a low quality of fit to the swaption prices. (30%+ mean absolute relative error)
- Using a more flexible parameterisation (Rebonato's Angles) a much higher quality of fit is achieved, but the correlation matrix oscillates beyond the point of interpretation.

This corroborates the findings of Brigo & Mercurio (2006) (where their swaption volatility structure was akin to that of a developed market) in which they state that in order to have a good calibration to the swaption data we have to allow at least partial oscillations in our correlation matrix. *A priori* it was conjectured that the upward sloping volatility structure in South Africa may be conducive to a rigid parametric correlation structure. However, our results refute this idea.

In order to implement the more flexible Rebonato's Angles parameterisation we have proposed a novel hybrid calibration algorithm to deal with the much higher dimensionality of the problem in the South African market, given that all South African interest rate derivatives reference three month JIBAR (which results in 79 forward rates). Please refer to subsection 4.2 for further information. Brigo & Mercurio (2006) present a cascade algorithm which provides exact fitting of the LFMM to the swaption market via a piecewise constant and non-homogenous volatility matrix. This was not conducive to our problem because the non-homogeneity of the volatility structure does not permit long-term modelling. Therefore, we have sought a similar result with a parsimonious homogenous parameterisation of the volatility structure.

Using fixed instantaneous correlations from historical data:

- Calibrating the instantaneous volatilities to cap and swaption data results in vastly different instantaneous volatilities in comparison to those implied by the market.
- Further, the resultant swaption prices deviate significantly from those present in the market.
- This is a particularly large problem within the context of the South African market, and may be as a result of the unusual swaption volatility structure.

When using parsimonious instantaneous volatility and correlation parameterisations (which would be used for long-term modelling), the resultant fit is of poor quality. Only once more flexible structures are introduced can the model yield an adequate fit. We conclude that the LFMM is not appropriate for this purpose within the context of the South African inter-bank interest rate market.

Bibliography

- J. Jacques Van Appel & Thomas A. McWalter Efficient long-dated swaption volatility approximation in the forward-LIBOR model (2018).
- A. Brace & D. Gatarek & M. Musiela. "The market model of interest rate dynamics." *Mathematical finance* 7.2 (1997): 127-155.
- R. White & Yukinori Iwashita. "Caplets: An Introduction to Caplet stripping." (2014)
- K. Miltersen & K. Sandmann & D. Sondermann. "Closed form solutions for term structure derivatives with lognormal interest rates." *The Journal of Finance* 52.1 (1997): 409-430.
- D. Brigo & F. Mercurio (2006) *Interest Rate Models — Theory and Practice*, second edition. Springer.
- J. Hull & A. White (2000) Forward rate volatilities, swap rate volatilities and the implementation of the LIBOR market model, *Journal of Fixed Income* **10** (2), 46–62.
- R. Rebonato (1999) On the pricing implications of the joint lognormal assumption for the swaption and cap markets, *Journal of Computational Finance* **2** (3), 57–76.
- J. Schoenmakers & B. Coffey (2003) Systematic generation of parametric correlation structures for the LIBOR market model, *International Journal of Theoretical and Applied Finance* **6** (5), 507–519.