Modelling Term and Inflation Risk Premia in the South African Bond Market

Luke van Schaik

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy

14 February, 2022

MPhil in Mathematical Finance,
University of Cape Town

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy at the University of Cape Town. It has not been submitted before for any degree or examination in any other University.

14 February, 2022

Abstract

A variety of approaches has been used to estimate the term premium of bond yields. Early attempts include linear regression models, such as those of Fama and Bliss (1987) and Cochrane and Piazzesi (2005), but these have been shown to be inconsistent and lacking in robustness (Kim and Orphanides (2007)). Affine term structure models developed by Duffie and Kan (1996) and extended by Duffee (2002) provide a more sophisticated framework for modelling bond yields and term premia, with improved results over the aforementioned regressions. However, parameters of these models have historically been estimated using maximum likelihood methods which are computationally inefficient and have been shown to have problems in finding the global maximum of the likelihood function (Hamilton and Wu (2012), Adrian *et al.* (2013)). The framework and estimation procedure of Adrian, Crump and Moench, or ACM, addresses the above problems by using ordinary least-square regressions exclusively to estimate the parameters of an affine term structure model (Adrian *et al.* (2013)). This dissertation applies the ACM procedure to South African zero-coupon nominal bond yields to estimate the term premium embedded in these yields. Performance of the ACM procedure is tested under Monte Carlo simulation and under applications to the smoothed United States nominal bond yield curves of Gurkaynak *et al.* (2006) and bootstrapped nominal bond yield curves from the South African market. Dynamics of the level, slope and curvature components of the US yield curves are compared to each other and to the estimated term premia. Results show that the ACM procedure generates very accurate fits to observed yield curves, but has some trouble capturing idiosyncratic features of the South African yield curve. Further, magnitudes of the term premium estimate are shown to be affected by the choice of time series of yield curves. Despite these limitations, the ACM procedure is shown to be a fast estimation procedure which generates term premium dynamics consistent with other approaches.

Acknowledgements

I would firstly like to thank my supervisors, Obeid Mahomed and Reza Ismail, for their roles in providing the motivation for this dissertation and for their insights and guidance throughout. I would also like to thank my family for their unconditional support, which made this dissertation possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The term structure of interest rates, depicted through the yield curve, represents the relationship between the yield and time-to-maturity of a given class of fixed income instrument. While the short end of the nominal bond yield curve is mostly determined by expectations of central bank policy, yields on longer-term bonds also incorporate compensation for the risk inherent in owning the bond to maturity. Thus, the yield on long-term bonds may be thought of as being comprised of two components: expectations of the short rate and the so-called term premium. While the term premium may be defined in subtly different ways, in general it represents the additional compensation that investors seek for investing their money for the long term as opposed to rolling over multiple short-term investments. The term premium is determined by risks including interest rate risk, inflation risk and liquidity, as well as the price of risk, which may vary with business cycles and other economic conditions (Kim and Orphanides (2007)). The magnitude of the term premium can therefore have important economic implications, such as investors' expectations of long-term inflation or sentiment regarding government or central bank policies. In the context of zero-coupon bonds (ZCBs), the risks associated with long-term investments tend to increase with time to maturity. The term premium correspondingly tends to increase with maturity of the ZCB, giving the typical upward-sloping shape of the yield curve, referred to as a "normal" yield curve. The slope of the yield curve is sometimes equated to the term premium for simplicity, but in reality expectations of the short rate and other factors also contribute to the slope.

The term premium takes on additional significance in the context of fixed income portfolio management. When there is meaningful movement in the slope of the yield curve, the term premium can be a valuable indicator in choosing the construction of a bond portfolio. Consider two cases of changing slope in a normal yield curve: i) the yield curve flattens (either the short rate increases, long rates

decrease, or both); ii) the yield curve steepens (the short rate decreases, long rates increase, or both).
In case i) a portfolio manager will benefit from holding a "barbell" portfolio, that is roughly equal holdings in long-dated and short-dated bonds, with few or no holdings in intermediate-dated bonds. This maximises the convexity of the portfolio for a given portfolio duration[1]. If long rates fall, the portfolio will benefit from a concentration in long-dated bonds. If short rates rise, losses are minimised by a concentrated holding in short-dated bonds, owing to the low sensitivity of the value of short-dated bonds to interest rate moves[2].
In case ii) a portfolio manager should hold a "bullet" portfolio, or holdings concentrated in intermediate-dated bonds. This configuration conversely minimises the convexity of the portfolio for a given duration. If long rates rise, the portfolio minimises losses by avoiding holdings in long-dated bonds. If short rates fall, little profit is lost owing to the low duration of short-dated bonds.

The term premium plays a significant role in determining whether the yield curve will steepen or flatten. For example, a decomposition of the yields of a given bond may show that the term premium is at an all-time high. If a fixed income portfolio manager determines that economic conditions do not justify the term premium sustaining this level, the yield curve will likely flatten through a reduction in long-term rates, and the portfolio would profit from a barbell construction.

Despite its relevance, the term premium is not directly observable and hence for this information to be interpreted, bond yields must be decomposed. Chapter 2 details some approaches to estimating the term premium, including simple, regression-based models such as those developed by Fama and Bliss (1987) and Piazzesi and Swanson (2008). These approaches generate somewhat intuitive results, but are shown to be inconsistent and lacking in robustness. Cochrane and Piazzesi (2005) develop a similar model, and achieve a high degree of forecasting accuracy.

More sophisticated affine term structure (ATS) models, developed by Duffie and Kan (1996) and extended by Duffee (2002), can also be used to estimate the term premium. This requires the parameters of the ATS model to be estimated, for example through maximum likelihood estimation (MLE). However, this approach has drawbacks, including numerical difficulties in optimisation. Hamilton and Wu

[1] The modified duration of a bond or portfolio of bonds represents the sensitivity of its value to changes in interest rates.

[2] The duration of zero-coupon bonds can be shown to increase linearly with the maturity of the bond.

(2012) and Joslin *et al.* (2011) improve on this method by increasing computational efficiency and reducing the aforementioned numerical difficulties.

Chapter 2 ends with a description of the ATS framework of Adrian, Crump and Moench (henceforth ACM). ACM use ordinary least-squares (OLS) regressions exclusively to estimate the model parameters, which avoids the aforementioned numerical problems completely, and additionally has greatly improved computational efficiency while being able to capture a large number of pricing factors (Adrian *et al.* (2013)). Chapter 2 also includes a brief overview of the ATS framework of Abrahams *et al.* (2016), which estimates the inflation risk premium of nominal bond yields, or the compensation that investors receive for bearing inflation risk. Appendix C discusses a naïve estimate of the inflation risk premium and compares it to that of Abrahams *et al.* (2016). Chapter 3 details the implementation of the ACM estimation procedure and Chapter 4 discusses the results of implementing the procedure in the US and South African bond markets. Lastly, Chapter 5 examines the performance of the procedure in detail, in particular accounting for unexpected results from Chapter 4.

Chapter 2

Modelling the Term Premium

This chapter outlines a variety of approaches to modelling the term premium. Early approaches include linear regressions on macroeconomic and market factors, but more sophisticated affine term structure (ATS) models have been shown to provide more consistent estimates of the term premium. After providing an outline of the theoretical framework of ATS models and specifications of the price of risk in such models, this chapter then describes the framework of Adrian *et al.* (2013), which enables model parameters to be estimated via OLS regressions exclusively. An overview of the framework of Abrahams *et al.* (2016) is also included, which estimates the inflation risk premium in nominal bond yields.

Since the mid-2000s, the long-term interest rates in many major economies' bond markets have exhibited non-intuitive behaviour, such as rising short-term rates being accompanied by low long-term rates following policy interventions by the US Federal Reserve in 2004 (Kim and Orphanides (2007)). These effects seem to provide evidence against the Expectations Hypothesis of the term structure. This hypothesis states that the term structure is simply the current expectation of future short rates, implying that a positively (negatively) sloped yield curve represents the market's expectation that short rates will rise (fall) with time (Fabozzi (2006)). Since the aforementioned Federal Reserve interventions did not warrant an expectation of falling short rates, the subsequent behaviour of the yield curve points to the fact that there must be additional factors, such as the term premium, influencing the shape of the yield curve. Additionally, the Expectations Hypothesis implies that current forward rates should predict future short rates, but there is little evidence for this (Fabozzi (2006)). Events such as those in 2004 have led to the term premium receiving much attention from practitioners and academics alike, with substantial progress being made in methods for its estimation.

Consider a discrete time series $t = \{t_0, t_1, \ldots, T\}$ and times to maturity $n = \{1, \ldots, N\}$.

Denoting the time-t price of a bond maturing in n periods as P_t^n, the continuously-compounded short rate is defined as $r_t := -\ln(P_t^1)$ and the time-t continuously-compounded forward rate accruing over the period $[t+n-1, t+n]$ as $f_t^{n-1,n} := \ln(P_t^{n-1}) - \ln(P_t^n)$. Kim and Orphanides (2007) provide three illustrative definitions of the term premium. The time-t term premium associated with maturity n, ϕ_t^n, may be defined in terms of forward rates as the difference between the one-period forward rate with maturity n and the expected short rate at the start of the accrual period:

$$\phi_{t,f}^n := f_t^{n-1,n} - \mathbb{E}_{\mathbb{P}}[r_{t+n-1}|\mathcal{F}_t]. \tag{2.1}$$

Here $\mathcal{F}_t$ represents the information observable at time t. Henceforth it will be assumed that expectations are under the objective measure $\mathbb{P}$, unless specifically denoted with the subscript $\mathbb{Q}$ for the risk-neutral measure.

In terms of excess returns, the term premium may be defined as the difference between the expected one-period excess return on a maturity-n bond and the one-period yield (i.e. the short rate). Letting $e_{t+1}^{n-1} := \ln(P_{t+1}^{n-1}) - \ln(P_t^n) - r_t$ denote the time-$(t+1)$ excess return on the bond maturing in n periods[1],

$$\phi_{t,e}^n := \mathbb{E}[e_{t+1}^{n-1}|\mathcal{F}_t] - r_t.$$

Lastly, the term premium may be defined in terms of bond yields as the difference between the yield on a maturity-n ZCB and the average of expected short rates up to the maturity of the bond. Denoting the time-t continuously-compounded yield on the ZCB maturing in n periods as y_t^n,

$$\phi_{t,y}^n := y_t^n - \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[r_{t+i}|\mathcal{F}_t].$$

The above definitions allow for the term premium to be estimated from observed market or economic data, and the following sections discuss some examples of such estimation procedures. The term premium may also be estimated in a model-independent approach as, for example, the difference between forward rates and surveyed expectations of short rates. While such estimates may be subject to measurement errors and are updated infrequently, they provide a benchmark against which model-implied estimates can be compared.

[1] The bond with time-t price P_t^n will have the price P_{t+1}^{n-1} in one period's time. Thus the one-period return on the bond is $\frac{P_{t+1}^{n-1}}{P_t^n}$ such that the log return over the short rate r_t (i.e. log excess return) is $\ln(P_{t+1}^{n-1}) - \ln(P_t^n) - r_t$.

2.1 Linear Regression Models

Under the Expectations Hypothesis, there should be no predictable difference between the time-t forward rate maturing in n periods and the ex-post realisation of the short rate at time $t+n$. A predictable difference between these two quantities implies a failure of the Expectations Hypothesis and consequently the existence of a term premium (Kim and Orphanides (2007)). As such, the predictable component of this difference can be used as a measure of the term premium. This component can be estimated via linear regressions, such as those of Fama and Bliss (1987) and Piazzesi and Swanson (2008). Fama and Bliss regress one-year bond returns over the period $[t, t+n-1]$ on a constant plus the spread between the forward rate accruing over $[t+n-1, t+n]$ and the current spot rate, with bond maturities n given in years:

$$r_{t+n-1} - r_t = \alpha + \beta(f_t^{n-1,n} - r_t) + \varepsilon.$$

For constants α, β and residual ε. Values of 0 and 1 for α and β respectively imply that the short rate realised at time $t+n-1$ is equal to the forward rate $f_t^{n-1,n}$, which is a representation of the Expectations Hypothesis.

Piazzesi and Swanson (2008) make use of a linear regression model with K factors:

$$f_t^{n-1,n} - r_{t+n-1} = \alpha^* + \beta_1^* X_{t,1} + \ldots + \beta_K^* X_{t,K} + \varepsilon,$$

for constants α^* and β_k^*, $k \in [1, K]$. $X_{t,k}$ represents the time-t value of the k^{th} pricing factor and r_t, $f_t^{n-1,n}$ are defined as above. The left-hand side of the equation is interpreted as the term premium according to equation (2.1). With a single factor $X_{t,1} = f_t^{n-1,n} - r_t$, letting $\alpha^* = -\alpha$ and $\beta_1^* = 1 - \beta$ gives a representation of the Fama-Bliss model. Piazzesi and Swanson (2008) make use of four-month Federal Funds Futures rates and year-on-year employment growth as factors in their model.

These approaches generated interesting results. For example, the Piazzesi-Swanson method shows countercyclical behaviour with high term premia during economic recessions. However, neither approach saw much use by central bank practitioners since the two methods give substantially different results to one another, with the estimated term premium being sensitive to the choice of regressors[2] and sample data (Kim and Orphanides (2007)). Kim and Orphanides (2007) compare the results of Piazzesi and Swanson (2008) to survey-implied estimates and find that there is

[2] These regressions make use of only a handful of explanatory variables, and it is questionable whether the term premium can be accurately estimated using such limited information.

a substantial difference between the two, which raises further questions about the accuracy of the Piazzesi-Swanson results.

Cochrane and Piazzesi (2005) (CP) extended the Fama-Bliss model in a popular 2005 paper. The CP model takes the form

$$e_{t+1}^{n-1} = \beta_0^n + \beta_1^n f_t^{0,1} + \ldots + \beta_5^n f_t^{4,5} + \varepsilon_{t+1}^n,$$

where e_{t+1}^{n-1} is the log excess return on a bond maturing in n periods as defined above and $f_t^{T-1,T}$ is the one-year continuously-compounded forward rate accruing over the period from year $T-1$ to year T. CP show that this model can be formulated with a single factor b_n for each maturity n:

$$e_{t+1}^{n-1} = b_n(\alpha_0 + \alpha_1 f_t^{0,1} + \ldots + \alpha_5 f_t^{4,5}) + \varepsilon_{t+1}^n.$$

This formulation allows for the factor b_n and coefficients α_i to be estimated easily via linear regressions. While the CP approach is computationally efficient and achieved impressive out-of-sample forecasting, the estimated five-year term premium exhibited larger variation than the four- to five-year forward rate itself, which suggests that the estimate may be unrealistically volatile (Kim and Orphanides (2007)). CP extended their term premium estimation approach to an ATS model in a 2008 paper (Cochrane and Piazzesi (2008)) which would go on to influence Adrian *et al.* (2013), whose framework and estimation procedure are discussed in Section 2.4 and Chapter 3 respectively.

2.2 Affine Term Structure Models

A popular alternative to the simple regression-based models above are term structure models, such as those developed by Vasiček (1977) and Cox *et al.* (1985) (CIR). These models are able to capture bond returns under variations in both state variables and times to maturity and have seen widespread application in pricing interest rate derivatives (Brigo and Mercurio (2007)). A further characterization of term structure models is the class of affine term structure models, with bond prices given as exponentially affine functions of the pricing factors. ATS models have likewise seen widespread application and success in instrument pricing and parameter estimation, with the latter often accomplished in the literature via MLE methods.
While Kim and Orphanides (2007) find that regression-based estimates of the term premium are inconsistent with model-independent approaches in both trends and magnitude, Cohen *et al.* (2018) show that ATS-based estimates of the term premium, while differing in magnitude, agree with trends in model-independent estimates.

ATS models have since become the preferred approach to estimating the term premium.

2.2.1 Bond Price Dynamics

ATS models assume that spot interest rates are affine functions of pricing factors. In a one-factor model, this pricing factor is typically the short rate r_t. For a K-factor model, the pricing factors are denoted as a $K \times 1$ vector X_t such that the time-t value of the continuously-compounded spot rate R for maturity τ is given by:

$$R_t^\tau = p(t,T) + q(t,T)'X_t,$$

where the $K \times 1$ vectors $p(t,T)$ and $q(t,T)$ are deterministic and T is the time at maturity such that $\tau := T - t$. The time-t price of a ZCB maturing in τ periods is therefore

$$\begin{aligned} P_t^\tau &= \exp(-R_t^\tau(T-t)) \\ &= \exp(-(p(t,T) + q(t,T)'X_t)(T-t)). \end{aligned}$$

Letting $A(t,T) := -p(t,T)(T-t)$ and $B(t,T) := q(t,T)(T-t)$, ZCB prices can be represented as

$$P_t^\tau = \exp(A(t,T) - B(t,T)'X_t), \tag{2.2}$$

for K-dimensional vectors $A(t,T)$ and $B(t,T)$. To ensure an affine term structure, the pricing factors must have the following dynamics under the risk-neutral measure $\mathbb{Q}$ (Brigo and Mercurio (2007)):

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t^{\mathbb{Q}},$$

where $W_t^{\mathbb{Q}}$ represents a K-dimensional $\mathbb{Q}$-Brownian motion. The term structure is affine in X_t if $b(t,X_t)$ and $\sigma^2(t,X_t)$ are affine functions of X_t:

$$b(t,X_t) := \phi(t)X_t + \eta(t)$$

$$\sigma^2(t,X_t) := \gamma(t)X_t + \delta(t),$$

for deterministic functions $\phi(t)$, $\eta(t)$, $\gamma(t)$ and $\delta(t)$. Accordingly, Duffie and Kan (1996) model the $\mathbb{Q}$-dynamics of pricing factors as follows:

$$dX_t = \kappa(\theta - X_t)dt + \Sigma S_t dW_t^{\mathbb{Q}}, \tag{2.3}$$

where κ and Σ are $K \times K$ matrices, θ is a $K \times 1$ vector and S_t is defined as:

$$S_t := \begin{bmatrix} \sqrt{\alpha_1 + \beta_1' X_t} & 0 & \cdots & 0 \\ 0 & \sqrt{\alpha_2 + \beta_2' X_t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\alpha_K + \beta_K' X_t} \end{bmatrix}, \tag{2.4}$$

for constants α_k and β_k, $k \in [1, K]$. Duffie and Kan (1996) show that this model satisfies the requirements for an affine term structure. The short rate r_t is modelled as

$$r_t = \delta_0 + \delta_1' X_t, \tag{2.5}$$

where δ_0 and δ_1 are $K \times 1$ vectors.

Under the specification of Duffie and Kan (1996), the functions $A(t,T)$ and $B(t,T)$ from equation (2.2) satisfy the following ordinary differential equations (ODEs):

$$\frac{\partial A(t,T)}{\partial t} = -\theta'\kappa' B(t,T) + \frac{1}{2}\sum_i (\Sigma' B(t,T))_i^2 \alpha_i - \delta_0 \tag{2.6}$$

$$\frac{\partial B(t,T)}{\partial t} = -\kappa' B(t,T) - \frac{1}{2}\sum_i (\Sigma' B(t,T))_i^2 \beta_i + \delta_1. \tag{2.7}$$

$A(t,T)$ and $B(t,T)$ can be computed numerically from these ODEs given that $A(T,T) = B(T,T) = 0$, since $P_t^0 = 1$.

In order to model X_t under the objective measure $\mathbb{P}$, the dynamics of the pricing kernel M_t are defined as follows:

$$\frac{dM_t}{M_t} = -r_t dt + \Lambda_t' dW_t^{\mathbb{P}},$$

where $W_t^{\mathbb{P}}$ is a $K \times 1$ $\mathbb{P}$-Brownian motion and Λ_t is a $K \times 1$ vector specifying the price of risk associated with the Brownian motion. Applying Itô's Lemma to $\ln(M_t)$ gives the time-t value of the pricing kernel:

$$M_t = \exp\left(\int_0^t (-r_t - \frac{1}{2}\Lambda_t'\Lambda_t)dt - \int_0^t \Lambda_t' dW_t^{\mathbb{P}}\right). \tag{2.8}$$

This allows bond prices to be obtained as an expectation under the objective measure. For some future time $s \geq t$,

$$\begin{aligned} P_t^\tau &= \mathbb{E}_{\mathbb{Q}}[P_s^\tau | \mathcal{F}_t] \\ &= \mathbb{E}[M_s P_s^\tau | \mathcal{F}_t]. \end{aligned}$$

According to Girsanov's Theorem, $W_t^{\mathbb{Q}} = W_t^{\mathbb{P}} - \int_0^t \Lambda_s ds$, such that

$$dX_t = [\kappa(\theta - X_t) + \Sigma S_t \Lambda_t]dt + \Sigma S_t dW_t^{\mathbb{P}}. \tag{2.9}$$

Applying Itô's Lemma to equation (2.2) and inserting equations (2.6), (2.7) and (2.9) yields

$$\frac{dP(t,T)}{P(t,T)} = (r_t - B(t,T)'\Sigma S_t \Lambda_t)dt - B(t,T)'\Sigma S_t dW_t^{\mathbb{P}}. \tag{2.10}$$

The term $-B(t,T)'\Sigma S_t\Lambda_t$ can be interpreted as the expected instantaneous excess return above the short rate r_t of holding the bond, and $-B(t,T)'\Sigma S_t$ the volatility of the bond price. Thus it can be seen that excess bond holding returns are driven by both S_t and Λ_t. The specification of Λ_t therefore has important implications for the dynamics of bond returns.

2.2.2 Specifying the Price of Risk

Completely Affine

Dai and Singleton (2000) and Fisher and Gilles (1996) specify the price of risk Λ_t as

$$\Lambda_t = S_t\lambda_1, \tag{2.11}$$

where λ_1 is a $K \times 1$ vector. The prices of risk in the Vasiček (1977) and CIR models follow this specification in a one-factor setting. It follows that the k^{th} pricing factor has $\Lambda_{t,k} = \sqrt{\alpha_k + \beta_k X_t}\lambda_{1,k}$. Consequently, it can be seen from equations (2.3) and (2.9) that the dynamics of X_t are affine under both $\mathbb{Q}$ and $\mathbb{P}$, which simplifies empirical estimation (Duffee (2002)). The instantaneous volatility of the pricing kernel $\Lambda_t'\Lambda_t$ is also affine in X_t, which gives this specification the label "completely affine".

Duffee (2002) points out two major limitations of this formulation: first, the volatility of Λ_t is fully determined by the volatility of S_t; second, the elements of Λ_t and λ_1 must have the same sign, since S_t is by definition non-negative (see equations (2.11) and (2.4)). The former implies that the volatility of excess bond returns are entirely determined by the volatility of bond yields. This is a major drawback, as such behaviour is contradicted by empirical evidence (Duarte (2004)). Empirical evidence also shows that excess bond returns exhibit low means with high volatility and frequently take on negative values, with the distribution of excess returns having a low skew (Duffee (2002)). The latter point above implies that the completely affine specification cannot capture this behaviour. Consider a one-factor model. Let e_t^τ be the excess returns of the bond $P(t,T)$ with time-to-maturity $\tau := T - t$ and v_t^τ be the volatility of $P(t,T)$ such that $e_t^\tau := -B(t,T)\Sigma S_t\Lambda_t$ and $v_t^\tau := -B(t,T)\Sigma S_t$, according to equation (2.10). Thus,

$$e_t^\tau = v_t^\tau S_t\lambda_1.$$

Since v_t^τ is by definition non-negative, e_t^τ is therefore bounded by zero. Since e_t^τ exhibits a low mean and high variance, this means that the distribution of excess returns must be highly skewed. Hence the completely affine specification cannot produce values of e_t^τ that range from positive to negative, and consequently implies

an unrealistic skew in the distribution of excess returns. This specification therefore cannot capture the empirical behaviour of excess returns described above. This is less obvious in a multifactor setting, but the above intuition can be shown to hold, given that long-term yields tend to be driven predominantly by a single factor in a well-fitted multifactor ATS model (Duffee (2002)). Therefore, the completely affine price of risk cannot produce realistic behaviour of bond returns for all maturities.

Essentially Affine

In response to these limitations, Duffee (2002) introduces the "essentially affine" price of risk as

$$\Lambda_t = S_t\lambda_1 + S_t^-\lambda_2 X_t, \tag{2.12}$$

where λ_2 is a $K \times K$ matrix. For $k \in [1, K]$, S_t^- is defined as

$$S_t^- = \begin{cases} (\alpha_k + \beta_k' X_t)^{-\frac{1}{2}} & \text{if } \inf(\alpha_k + \beta_k' X_t) > 0 \\ 0 & \text{else.} \end{cases}$$

The second case ensures that the elements of S_t^- do not explode if the corresponding element of S_t approaches zero. This specification maintains the affine relationship between $S_t\Lambda_t$ and X_t under the objective measure, and hence the benefit to model estimation. However, $\Lambda_t'\Lambda_t$ is no longer affine in X_t, hence the label "essentially affine".
Equation (2.12) removes the strict relation between Λ_t and S_t as well as the restriction on the sign of the elements of Λ_t. The essentially affine specification therefore maintains the advantages of the completely affine price of risk, but eliminates the two drawbacks just discussed. Duffee (2002) shows that this specification improves forecasting results and obtains a better model fit to empirical bond yields.

ATS models allow for the term premium to be estimated as the difference between model-implied bond yields under the objective and risk-neutral measures, capturing the component of model-implied yields that correspond to the market's perception of risk. Given objective nominal bond yields y_t^n and risk-neutral yields $y_t^{n,\mathbb{Q}}$, the term premium ϕ_t^n can be estimated as

$$\phi_t^n := y_t^n - y_t^{n,\mathbb{Q}}.$$

An accurate estimate of the term premium therefore relies on a good choice of model and accurate parameter estimation.

2.3 Maximum Likelihood Estimation

Maximum likelihood estimation techniques are a popular approach to estimating model parameters in ATS frameworks. However, MLE typically suffers from two major drawbacks. Firstly, computational inefficiencies stemming from the numerical optimisation routines used tend to make estimation slow, especially when increasing the number of pricing factors in the model. Secondly, such approaches have been shown to suffer from problems relating to finding the global maximum of the log-likelihood function (Adrian *et al.* (2013), Hamilton and Wu (2012)).

In the traditional MLE approach, the unknown parameters are grouped into a vector θ which maximises the log-likelihood function. For some log-likelihood function L, parameter space Θ and data x:

$$\hat{\theta} = \arg\max_{\theta \in \Theta} L(\theta; x).$$

Kim (2008) points out that estimating the parameters of an ATS model using MLE can be challenging, since flexibly specified ATS models typically involve a large number of parameters with non-linear relationships between the parameters and bond yields. Hamilton and Wu (2012) (HW) demonstrate this with an example of a three-factor discrete-time ATS model as formulated by Ang and Piazzesi (2003). Under this formulation, pricing factors follow a first-order vector autoregression (VAR(1)) process:

$$X_{t+1} = \mu + \Phi X_t + \Sigma u_{t+1}, \tag{2.13}$$

where X_t is a 3×1 vector of pricing factors, μ is a 3×1 vector, Φ is a 3×3 matrix and u_{t+1} is a 3×1 vector of standard normal residuals with Σ representing the variance-covariance matrix of X_t. This model makes use of the essentially affine price of risk specification of the form $\Lambda_t := \lambda_0 + \lambda_1 X_t$, for K-vector λ_0 and $K \times K$ matrix λ_1. Risk-neutral dynamics are defined as:

$$X_{t+1} = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_t + \Sigma u^{\mathbb{Q}}_{t+1},$$

where HW show that the risk-neutral parameters are given by

$$\mu^{\mathbb{Q}} = \mu - \Sigma\lambda_0$$

and

$$\Phi^{\mathbb{Q}} = \Phi - \Sigma\lambda_1.$$

The short rate is modelled according to equation (2.5). Under the formulation of Ang and Piazzesi (2003), the parameters that must be estimated include those of

the above VAR(1) for both the objective and risk-neutral dynamics, the parameters specifying the price of risk and the parameters determining the short rate in equation (2.5), for a total of 23 parameters in the three-factor case. HW use an unconstrained numerical search algorithm to find the 23×1 vector $\hat{\theta}$ via MLE. Estimating θ 100 times using this algorithm, HW found that only one estimate corresponded to the global maximum of the log-likelihood function, and that 19 of the estimates failed to achieve convergence of the search algorithm. Such algorithms are also computationally inefficient, especially as the number of pricing factors (and hence model parameters) increases.

Joslin *et al.* (2011) (JSZ) propose a representation of the ATS framework which allows for more efficient estimation of model parameters. JSZ make use of pricing factor dynamics such as those in equation (2.13), with pricing factors given as linear combinations of observed zero-coupon bond yields (such as principal components). JSZ show that, in their representation, estimates of the parameters μ and Φ under the $\mathbb{P}$-measure can be obtained via OLS regression instead of MLE, increasing computational efficiency of the estimation procedure. However, JSZ still make use of MLE to estimate the parameters of the model's $\mathbb{Q}$-dynamics, and HW show that the JSZ procedure found the global maximum of the log-likelihood function in only 54 of 100 estimates of θ. HW improve on this approach. After estimating the pricing factor parameters in equation (2.13) using OLS (as per JSZ), HW make use of a minimum chi-square estimator (MCSE) function to estimate the parameters δ_1 and $\rho^{\mathbb{Q}}$. This estimator function has a much simpler form than a log-likelihood function, which greatly improves efficiency and addresses the numerical difficulties discussed[3]. HW then obtain $\hat{\rho}$ analytically from the aforementioned estimates, and finally obtain estimates of δ_0 and $c^{\mathbb{Q}}$ using a numerical search algorithm. HW therefore greatly reduce the reliance on numerical search algorithms as compared to standard MLE, and improve the efficiency of the numerical search they do use through the use of an MCSE function.

The estimation procedure of Adrian *et al.* (2013) builds on the approaches of HW and JSZ. The following section and Chapter 3 show that the ACM model is formulated such that all model parameters can be estimated via OLS regressions. The result is an estimation procedure that avoids the aforementioned numerical issues while being much more computationally efficient and able to handle a large number of pricing factors.

[3] HW find that their approach obtains the global maximum of their estimator function in 100 of 100 experiments.

2.4 The ACM Framework

The framework underlying the estimation procedure of Adrian *et al.* (2013) is an affine term structure with K pricing factors. The time series of yields t and times to maturity n are given in months. Similarly to Ang and Piazzesi (2003), Joslin *et al.* (2011) and Hamilton and Wu (2012), ACM formulate their pricing factor dynamics under the objective measure as:

$$X_{t+1} = \mu + \Phi X_t + v_{t+1},$$

where μ is a $K \times 1$ vector, Φ is a $K \times K$ matrix and v_{t+1} is a $K \times 1$ vector representing normally distributed pricing factor innovations with mean 0 and variance-covariance matrix Σ. ACM assume an exponentially affine pricing kernel M_t

$$M_{t+1} = \exp(-r_t - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\Sigma^{-\frac{1}{2}}v_{t+1}), \tag{2.14}$$

which is a discrete-time equivalent of equation (2.8). ACM define the short rate r_t as the continuously-compounded one-month treasury yield, i.e. $r_t := -\ln(P_t^1)$. The market price of risk Λ_t is defined according to the essentially affine specification of Duffee (2002):

$$\Lambda_t := \Sigma^{-\frac{1}{2}}(\lambda_0 + \lambda_1 X_t), \tag{2.15}$$

where λ_0 is a $K \times 1$ vector and λ_1 is a $K \times K$ matrix. Equations (2.14) and (2.15) allow for bond prices to be computed as a time-t expectation under the objective measure:

$$P_t^n = \mathbb{E}[M_{t+1}P_{t+1}^{n-1}|\mathcal{F}_t], \tag{2.16}$$

where P_{t+1}^{n-1} is the price of the bond P_t^n in one period's time. As per Cochrane and Piazzesi (2005), the time-t log excess holding return on a bond maturing in n periods is defined as

$$e_{t+1}^{n-1} := \ln(P_{t+1}^{n-1}) - \ln(P_t^n) - r_t. \tag{2.17}$$

Substituting equations (2.14) and (2.17) into equation (2.16) yields

$$1 = \mathbb{E}[\exp(e_{t+1}^{n-1} - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\Sigma^{-\frac{1}{2}}v_{t+1})|\mathcal{F}_t]. \tag{2.18}$$

Since innovations in the pricing factors drive excess returns, ACM assume e_{t+1}^{n-1} and v_{t+1} are jointly normally distributed. Cochrane and Piazzesi (2008) show that equations (2.18) and (2.15) then yield the following expression for expected excess returns:

$$\mathbb{E}[e_{t+1}^{n-1}|\mathcal{F}_t] = \mathbb{C}\text{ov}[e_{t+1}^{n-1}, v_{t+1}'|\mathcal{F}_t]\Sigma^{-1}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}\mathbb{V}[e_{t+1}^{n-1}|\mathcal{F}_t].$$

ACM then define $\beta_t^{n-1\prime} := \mathbb{C}\text{ov}[e_{t+1}^{n-1}, v_{t+1}'|\mathcal{F}_t]\Sigma^{-1}$ such that

$$\mathbb{E}[e_{t+1}^{n-1}|\mathcal{F}_t] = \beta_t^{n-1\prime}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}\mathbb{V}[e_{t+1}^{n-1}|\mathcal{F}_t]. \tag{2.19}$$

The unexpected excess return over the period t to $t+1$ is decomposed into a component correlated to the pricing factor innovations and an uncorrelated component ε_{t+1}^{n-1}:

$$e_{t+1}^{n-1} - \mathbb{E}[e_{t+1}^{n-1}|\mathcal{F}_t] = \beta_t^{n-1\prime} v_{t+1} + \varepsilon_{t+1}^{n-1}. \tag{2.20}$$

ε_{t+1}^{n-1} therefore represents the error observed in the return pricing process. These errors are assumed to be independently and identically distributed (i.i.d.) with $\mathbb{V}[\varepsilon_{t+1}^{n-1}|\mathcal{F}_t] =: \sigma^2$. Since ACM estimate β_t^n via a linear regression over the time series of yields, $\beta_t^n = \beta^n$ is constant over time. Combining equation (2.20) with equation (2.19),

$$\begin{aligned} e_{t+1}^{n-1} &= \mathbb{E}[e_{t+1}^{n-1}|\mathcal{F}_t] + \beta^{n-1\prime} v_{t+1} + \varepsilon_{t+1}^{n-1} \\ &= \beta^{n-1\prime}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}(\beta^{n-1\prime}\Sigma\beta^{n-1} + \sigma^2) + \beta^{n-1\prime} v_{t+1} + \varepsilon_{t+1}^{n-1}. \end{aligned} \tag{2.21}$$

This can be represented in matrix form across all time periods and times to maturity. For a time series $t = \{t_0, \ldots, T\}$ and times to maturity $n = \{1, \ldots, N\}$,

$$e = \beta'(\lambda_0 + 1_T' + \lambda_1 X_l) - \frac{1}{2}(B^*\text{vec}(\Sigma) + \sigma^2 1_N)1_T' + \beta' V + E, \tag{2.22}$$

where 1_T and 1_N are $T \times 1$ and $N \times 1$ vectors of ones, respectively, e is defined as the $N \times T$ matrix of log excess returns, β is the $K \times N$ matrix $[\beta^1 \ldots \beta^N]$, X_l is the $K \times T$ matrix of lagged pricing factors $[X_{t_0} \ldots X_{T-1}]$, E is the $N \times T$ matrix of return pricing errors ε_{t+1}^{n-1}, and B^* is defined as the $N \times K^2$ matrix

$$B^* := \begin{bmatrix} \text{vec}(\beta^1\beta^{1\prime}) & \cdots & \text{vec}(\beta^N\beta^{N\prime}) \end{bmatrix}, \tag{2.23}$$

such that $B^*\text{vec}(\Sigma)$ is an $N \times 1$ vector with entries

$$\begin{aligned} [B^*\text{vec}(\Sigma)]_n &= (\beta_1^n)^2\Sigma_{11} + \beta_1^n\beta_2^n\Sigma_{21} + \ldots + (\beta_n^n)^2\Sigma_{nn} \\ &= \beta^{n\prime}\Sigma\beta^n. \end{aligned}$$

Equation (2.22) forms the basis of the linear regressions that are used to estimate the parameters of the ACM framework, described in detail in Chapter 3.

2.4.1 Computing Bond Pricing Parameters

Under the assumption of an affine term structure, bond yields are affine in the pricing factors:

$$\ln(P_t^n) = A_n - B_n' X_t + u_t^n, \tag{2.24}$$

where A_n is a constant, B_n is a $K \times 1$ vector and u_t^n represents the error in the yield pricing process. Substituting equation (2.24) into equation (2.17) gives

$$\begin{aligned} e_{t+1}^{n-1} &= A_{n-1} - B'_{n-1}X_{t+1} + u_{t+1}^{n-1} - A_n + B'_n X_t - u_t^n + A_1 - B'_1 X_t + u_t^1 \\ &= A_{n-1} - B'_{n-1}(\mu + \Phi X_t + v_{t+1}) + u_{t+1}^{n-1} - A_n + B'_n X_t - u_t^n + A_1 - B'_1 X_t + u_t^1, \end{aligned}$$

and equating this to (2.21),

$$\begin{aligned} A_{n-1} - B'_{n-1}(\mu + \Phi X_t + v_{t+1}) + u_{t+1}^{n-1} - A_n + B'_n X_t - u_t^n + A_1 - B'_1 X_t + u_t^1 \\ = \beta^{n-1\prime}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}(\beta^{n-1\prime}\Sigma\beta^{n-1} + \sigma^2) + \beta^{n-1\prime}v_{t+1} + \varepsilon_{t+1}^{n-1}, \quad (2.25) \end{aligned}$$

which holds for all t. The parameters A_0 and B_0 correspond to $P_t^0 = 1$, hence $A_0 = B_0 = 0$. Since $n = 1$ corresponds to the one-month yield, i.e. r_t, $A_1 = -\delta_0$ and $B_1 = \delta_1$, according to equation (2.5). Matching terms in equation (2.25) yields the following relations:

$$A_{n-1} - B'_{n-1}\mu - A_n + A_1 = \beta^{n-1\prime}\lambda_0 - \frac{1}{2}(\beta^{n-1\prime}\Sigma\beta^{n-1} + \sigma^2) \quad (2.26)$$

$$- B'_{n-1}\Phi + B'_n - B'_1 = \beta^{n-1\prime}\lambda_1 \quad (2.27)$$

$$B_{n-1} = -\beta^{n-1} \quad (2.28)$$

$$u_{t+1}^{n-1} - u_t^n + u_t^1 = \varepsilon_{t+1}^{n-1}. \quad (2.29)$$

Equation (2.28) implies that β^n is equivalent to the loading B_n on the pricing factor X_t which prices the bond P_t^n. Inserting this result from (2.28) into equations (2.26) and (2.27) gives recursive formulae for computing the pricing parameters A_n and B_n:

$$A_n = A_{n-1} - B'_{n-1}(\mu - \lambda_0) + \frac{1}{2}(B'_{n-1}\Sigma B'_{n-1} + \sigma^2) - \delta_0 \quad (2.30)$$

$$B'_n = B'_{n-1}(\Phi - \lambda_1) - \delta'_1. \quad (2.31)$$

Equation (2.29) implies that if ε_t^n is i.i.d. then u_t^n is correlated across t and n, and vice versa. Since ACM assume return pricing errors to be i.i.d., yield pricing errors are implicitly assumed to be correlated in this framework. This reflects the empirical behaviour of zero-coupon bond returns. In comparison, maximum likelihood approaches (including those of Joslin *et al.* (2011) and Hamilton and Wu (2012)) tend to assume i.i.d. yield errors and correlated return errors (Adrian *et al.* (2013)). The next chapter describes how equation (2.22) is used to estimate the model parameters via OLS regressions.

2.5 Modelling the Inflation Risk Premium

Abrahams *et al.* (2016) extend the work of Adrian *et al.* (2013) to a decomposition of the real yield curve in order to jointly price nominal bonds and Treasury Inflation-Protected Securities (TIPS) and extract an inflation risk premium. Abrahams *et al.* (2016) make use of an ATS framework closely resembling that of ACM, while incorporating the Consumer Price Index (CPI) to allow for the pricing of inflation-linked zero-coupon bonds. As in the ACM framework, bond yields are affine in the pricing factors X_t, in this case three principal components of treasury yields, two principal components of TIPS yields and a liquidity factor (discussed below). Break-even inflation for horizon n, π_t^n, can therefore be estimated in this framework as

$$\pi_t^n := y_t^n - y_{t,R}^n = -\frac{1}{n}\left(A_n - B_n' X_t - (A_{n,R} - B_{n,R}' X_t)\right),$$

where y_t^n represents the time-t continuously-compounded yield on the ZCB maturing in n periods, and the subscript R denotes the yields and pricing factors associated with real bonds. Similarly to the ACM framework, these pricing parameters are computed recursively using estimated price of risk parameters according to equations (2.30) and (2.31), with adjustments made to the recursions for $A_{n,R}$ and $B_{n,R}$ in order to incorporate CPI into the real bond prices. Setting the price of risk parameters to zero in the recursive equations yields risk-neutral break-even inflation $\pi_t^{n,\mathbb{Q}}$:

$$\pi_t^{n,\mathbb{Q}} := y_t^{n,\mathbb{Q}} - y_{t,R}^{n,\mathbb{Q}} = -\frac{1}{n}\left(A_n^{\mathbb{Q}} - B_n^{\mathbb{Q}\prime} X_t - (A_{n,R}^{\mathbb{Q}} - B_{n,R}^{\mathbb{Q}}{}' X_t)\right).$$

The inflation risk premium $\phi_{t,i}^n$ is then given as the difference between break-even inflation under the two measures, i.e.

$$\phi_{t,i}^n := \pi_t^n - \pi_t^{n,\mathbb{Q}}. \tag{2.32}$$

This formulation implies that the nominal term premium ϕ_t^n can be decomposed into the inflation risk premium and a "real term premium" $\phi_{t,R}^n$, according to the definition $\phi_t^n := y_t^n - y_t^{n,\mathbb{Q}}$:

$$\begin{aligned} \phi_{t,i}^n &= (y_t^n - y_{t,R}^n) - (y_t^{n,\mathbb{Q}} - y_{t,R}^{n,\mathbb{Q}}) \\ &= \phi_t^n - \phi_{t,R}^n \\ \iff \phi_t^n &= \phi_{t,i}^n + \phi_{t,R}^n. \end{aligned}$$

While in theory equation (2.32) should represent the inflation risk premium, in reality the relative illiquidity of inflation-linked securities means that there is a large liquidity premium embedded in real yields. To adjust for this, Abrahams *et al.*

(2016) incorporate a TIPS liquidity factor constructed from TIPS yield curve fitting errors from a Nelson-Siegel-Svensson (NSS) model[4] and a moving average of the ratio of Treasury Bill transaction volumes to TIPS transaction volumes. This factor is included in the vector X_t in the model, and bond pricing parameters A_n and B_n are computed as usual according to recursive formulae (2.30) and (2.31), and the equivalent recursions for $A_{n,R}$ and $B_{n,R}$. Liquidity-adjusted yields are then computed by subtracting the yield component associated with this liquidity factor from the unadjusted yields. Unlike the estimation procedure of ACM, the model parameters in this framework must be estimated using MLE, so this decomposition does not benefit from the efficiency of the ACM procedure. Abrahams *et al.* (2016) find that their estimate of the inflation risk premium is closely aligned with relevant macroeconomic variables such as consumer confidence, and that model-implied inflation expectations closely match surveys of professional forecasters. While implementation of this model is out of the scope of this dissertation, Appendix C discusses a naïve approach to estimating the inflation risk premium using the ACM procedure.

[4] See Appendix A for an overview of the Nelson-Siegel-Svensson framework.

Chapter 3

Estimation in the ACM Framework

The ACM framework is constructed such that model parameters can be estimated via a three-step OLS regression process. This chapter describes the estimation procedure in detail.
Let the time series over which yields are observed be given as $t = \{t_0, t_1, \ldots, T\}$ and the cross-section of maturities in the yield curve as $n = \{1, 2, \ldots, N\}$. ACM make use of monthly observations of yield curves with maturities n in months.

3.1 Principal Component Analysis

ACM's approach allows for the use of observable or unobservable pricing factors. In their original paper, ACM make use of K principal components (PCs) of continuously-compounded zero-coupon nominal bond yields as pricing factors. These PCs are obtained through a principal component analysis (PCA) of the $(T + 1) \times N$ time series of yield curves. Let Y be the matrix of observed ZCB yields such that

$$Y := \begin{bmatrix} y_{t_0}^1 & \cdots & y_{t_0}^N \\ \vdots & \ddots & \vdots \\ y_T^1 & \cdots & y_T^N \end{bmatrix},$$

where y_t^n represents the time-t yield to maturity of the ZCB maturing in n periods. The PCA proceeds as follows. First, the normalised matrix $\bar{Y}$ is obtained. For column Y_n of the matrix Y with mean μ_n and variance σ_n^2, the column $\bar{Y}_n$ of $\bar{Y}$ is given as $\bar{Y}_n := \frac{Y_n - \mu_n}{\sigma_n}$. The $N \times N$ variance-covariance matrix $\Sigma_{\bar{Y}}$ is computed from $\bar{Y}$.

An eigendecomposition is then performed on $\Sigma_{\bar{Y}}$ to obtain N $N \times 1$ eigenvectors v_n and N eigenvalues λ_n. The $N \times N$ matrices Λ and V are constructed as $\Lambda := \text{diag}(\lambda_1, \ldots, \lambda_N)$ and $V := [v_1 \cdots v_N]$, respectively. By the Spectral Theorem,

since $\Sigma_{\bar{Y}}$ is symmetric and V is orthogonal,

$$\begin{aligned} \Sigma_{\bar{Y}} &= V\Lambda V' \\ \iff \Lambda &= V'\Sigma_{\bar{Y}}V. \end{aligned}$$

The normalised yield matrix is then transformed by the eigenvector matrix V to create the transformed yield matrix Z:

$$Z := \bar{Y}V,$$

such that the columns of Z are orthogonal and the variance-covariance matrix Σ_Z of Z is given as

$$\begin{aligned} \Sigma_Z &= Z'Z \\ &= V'\bar{Y}'\bar{Y}V \\ &= V'\Sigma_{\bar{Y}}V = \Lambda. \end{aligned}$$

Thus the variance of the n^{th} column of Z is given as $\mathbb{V}[Z_n] = \Lambda_{n,n} = \lambda_n$. Since $\bar{Y} = ZV'$, $\bar{Y}$ is a linear combination of the columns of Z. Hence the columns of Z with the highest eigenvalues contribute the most to the variation observed in the data Y[1]. The K pricing factors are chosen to be the columns of Z with the K highest eigenvalues.

3.2 Step 1

The columns of the PC matrix Z are placed in descending order of their respective eigenvalues to give the matrix Z^*. For Z^*_k as the k^{th} column of Z^*, the time series of pricing factor X_k is then taken to be the $1 \times (T+1)$ vector $Z^{*\prime}_k$ for $k = [1, 2, \ldots, K]$. The vectors X_k are stacked into the $K \times (T+1)$ matrix X. Letting $X_c := \begin{bmatrix} X_{t_1} & \cdots & X_T \end{bmatrix}$ and $X_l := \begin{bmatrix} X_{t_0} & \cdots & X_{T-1} \end{bmatrix}$ denote the $K \times T$ matrices of current and lagged pricing factors respectively, the VAR(1) parameters are estimated via OLS regression according to equation (2.4). Parameters $\hat{\mu}_k$ and $\hat{\Phi}_k$ are estimated for each of the k factors using the k^{th} rows of the matrices X_c and X_l. Defining the $T \times 2$ regressor matrix for lagged pricing factors as $R_{1,k} := \begin{bmatrix} 1_T & X'_{l,k} \end{bmatrix}$ and letting $X_{c,k}$ and $X_{l,k}$ denote the k^{th} rows of the respective matrices,

$$\begin{bmatrix} \hat{\mu}_k \\ \hat{\Phi}_k \end{bmatrix} = (R'_{1,k}R_{1,k})^{-1}R'_{1,k}X_{c,k}, \tag{3.1}$$

[1] PCA of the US zero-coupon yield curve from 1961 to 2021 shows that the first three PCs account for over 99% of the observed variation in yields.

such that $\hat{\mu} := \begin{bmatrix}\hat{\mu}_1 & \cdots & \hat{\mu}_K\end{bmatrix}'$ is a $K \times 1$ vector and $\hat{\Phi} := \text{diag}(\hat{\Phi}_1, \ldots, \hat{\Phi}_K)$ is a $K \times K$ diagonal matrix. The pricing factor innovations from this regression are computed as the $K \times T$ matrix $\hat{V} := X_c - (\hat{\mu} + \hat{\Phi} X_l)$, and the variance-covariance matrix of the pricing innovations is estimated as $\hat{\Sigma} = \hat{V}\hat{V}'/T$.

3.3 Step 2

It can be seen from equation (2.22) that excess bond returns e_t^n are modelled as a linear function of V, X_l and E. In matrix form:

$$\begin{aligned} e &= \beta'(\lambda_0 + 1_T' + \lambda_1 X_l) - \frac{1}{2}(B^* \text{vec}(\Sigma) + \sigma^2 1_N)1_T' + \beta' V + E \\ &= (\beta'\lambda_0 - \frac{1}{2}(B^* \text{vec}(\Sigma) + \sigma^2 1_N))1_T' + \beta' V + \beta'\lambda_1 X_l + E \\ &=: a 1_T' + \beta' V + c X_l + E. \end{aligned}$$

where $a := \beta'\lambda_0 - \frac{1}{2}(B^* \text{vec}(\Sigma) + \sigma^2 1_N)$ is an $N \times 1$ vector, β is the $K \times N$ matrix of pricing factor loadings $\begin{bmatrix}\beta^1 & \cdots & \beta^N\end{bmatrix}$ and $c := \beta'\lambda_1$ is an $N \times K$ matrix. ACM compute the $N \times T$ matrix of excess returns e directly from observed bond yields, taking the short rate as the one-month treasury yield $r_t := y_t^1 = -\ln(P_t^1)$:

$$\begin{aligned} e_{t+1}^{n-1} &= \ln(P_{t+1}^{n-1}) - \ln(P_t^n) + \ln(P_t^1) \\ &= -(n-1)y_{t+1}^{n-1} + n y_t^n - r_t. \end{aligned}$$

The regressors 1_T, $\hat{V}$ and X_l, computed in Step 1, are grouped into a $(2K+1) \times T$ matrix $R_2 := \begin{bmatrix}1_T & \hat{V}' & X_l'\end{bmatrix}$. The estimators are then obtained via OLS regression as:

$$\begin{bmatrix}\hat{a} & \hat{\beta}' & \hat{c}\end{bmatrix} = e R_2'(R_2 R_2')^{-1}. \tag{3.2}$$

Residuals from this regression, representing the return pricing errors ε_t^n, are stored in the $N \times T$ matrix $\hat{E}$. The variance of return pricing errors, defined in Chapter 2 as σ^2, is estimated as the variance of these residuals. Since the errors are assumed to be i.i.d., the total variance may be estimated using the trace of the variance-covariance matrix $\hat{E}\hat{E}'/NT$:

$$\hat{\sigma}^2 = \frac{\text{trace}(\hat{E}\hat{E}')}{NT}.$$

Lastly, the matrix B^* is estimated from $\hat{\beta}$ as per equation (2.23):

$$\hat{B}^* = \begin{bmatrix}\text{vec}(\hat{\beta}^1 \hat{\beta}^{1\prime}) & \cdots & \text{vec}(\hat{\beta}^N \hat{\beta}^{N\prime})\end{bmatrix}.$$

3.4 Step 3

Since the coefficients a and c defined above are linear functions of the price of risk parameters λ_0 and λ_1, these parameters are estimated using a final OLS regression on the estimates computed in Steps 1 and 2:

$$\hat{\lambda}_0 = (\hat{\beta}\hat{\beta}')^{-1}\hat{\beta}(\hat{a} + \frac{1}{2}(\hat{B}^* \text{vec}(\hat{\Sigma}) + \hat{\sigma}^2 1_N))$$

$$\hat{\lambda}_1 = (\hat{\beta}\hat{\beta}')^{-1}\hat{\beta}\hat{c}.$$

3.5 Estimating the Term Premium

Once the price of risk parameters have been estimated, fitted bond yields can be obtained. Equations (2.30) and (2.31) allow for the bond pricing parameters A_n and B_n to be computed recursively, given $A_0 = B_0 = 0$ and $A_1 = -\delta_0$, $B_1 = \delta_1$. The parameters δ_0 and δ_1 are estimated via an OLS regression on the short rate according to equation (2.5). Defining the $T \times (K+1)$ matrix $R_3 := \begin{bmatrix} 1_T & X_c' \end{bmatrix}$ and $T \times 1$ vector $r_c := \begin{bmatrix} r_{t_1} & \cdots & r_T \end{bmatrix}'$,

$$\begin{bmatrix} \hat{\delta}_0 \\ \hat{\delta}_1 \end{bmatrix} = (R_3' R_3)^{-1} R_3' r_c,$$

where $\hat{\delta}_0$ is a constant and $\hat{\delta}_1$ is a $K \times 1$ vector. The pricing parameters are then estimated recursively:

$$A_n = A_{n-1} - B_{n-1}'(\hat{\mu} - \hat{\lambda_0}) + \frac{1}{2}(B_{n-1}'\hat{\Sigma}B_{n-1}' + \hat{\sigma}^2) - \hat{\delta}_0 \tag{3.3}$$

$$B_n' = B_{n-1}'(\hat{\Phi} - \hat{\lambda}_1) - \hat{\delta}_1', \tag{3.4}$$

giving model-implied bond prices as

$$P_t^n = \exp(A_n - B_n' X_t).$$

Setting the price of risk parameters $\hat{\lambda}_0$ and $\hat{\lambda}_1$ to zero in equations (3.3) and (3.4) generates risk-neutral pricing parameters $A_n^{\mathbb{Q}}$ and $B_n^{\mathbb{Q}}$. These parameters can be used to compute risk-neutral continuously-compounded yields as $y_t^{n,\mathbb{Q}} := -\frac{1}{n}(A_n^{\mathbb{Q}} - B_n^{\mathbb{Q}\prime} X_t)$, which represent the time-t expectation of future short rates averaged over the next n periods. Subtracting this risk-neutral yield from the model-implied yield under the objective measure gives an estimate of the term premium:

$$\phi_t^n = -\frac{1}{n}\left((A_n - B_n' X_t) - (A_n^{\mathbb{Q}} - B_n^{\mathbb{Q}\prime} X_t)\right).$$

3.6 Data

To validate implementation, the ACM procedure was first applied to the data of Gurkaynak *et al.* (2006) (GSW), as per ACM's original implementation. GSW provide the parameters of an NSS model of the US zero-coupon treasury yield curve for June 1961 to present, giving a smoothed term structure of continuously-compounded yields. These parameters are used to compute yield curves over this time series for maturities of $n = 1, 2, \ldots, 120$ months, with yield curve observations on the last day of each month.

The ACM procedure was applied to zero-coupon continuously-compounded nominal bond yields obtained from the South African market using the same maturities for a time series from 30 September 2012 to 30 September 2021. Yields were obtained from Bloomberg and bootstrapped via the Johannesburg Stock Exchange's (JSE) methodology (Modise (2019)), with yields observed on the last day of each month. This data makes use of the South African Futures Exchange (SAFEX) overnight rate, treasury bills for $n = 3, 6, 9, 12$ months, and nominal government bonds thereafter. Bootstrapped South African yield curves provided continuously-compounded zero-coupon yields for maturities out to 30 years, with maturities at a range of intervals. Linear interpolation was performed between the bootstrapped yields to obtain yields for the required maturities of $n = 1, 2, \ldots, 120$ months.

Chapter 4

Results

Results of implementing the estimation procedure described in Chapter 3 are discussed in this chapter. The procedure is first applied in the context of Monte Carlo simulation to assess parameter estimation, then to data from US and South African zero-coupon nominal yield curves to assess model fits and behaviour of the term premium and other features of the yield curve.

4.1 Monte Carlo Simulation

Monte Carlo simulations were used to assess the performance of ACM's estimation procedure. In order to assess the accuracy of the procedure's estimates, 1 000 realisations of an ACM model were simulated and parameter estimates were evaluated. Monte Carlo simulation of a single model was also performed to examine the fit of the model-implied yield curves and the ability of the framework to capture pricing factor dynamics.

4.1.1 Parameter Estimation

Using model parameters obtained from an application of the ACM procedure to the US NSS yield curves, 1 000 simulated yield curves were obtained via equations (2.30) and (2.31) for one- and three-factor models. Results of the parameter estimation are given in Appendix B in Tables B.1 and B.2. The results of the simulation generally show slight bias with standard deviation in the order of 10% of the estimated value. The exception is the price of risk coefficient λ_1, which is significantly biased in the one-factor model and in the diagonals of the three-factor model, with large variation.

Comparing the one- and three-factor models, it can be seen that the three-factor model exhibits greater variation in every parameter. The three-factor model produces less biased estimates for δ_1 and μ, but a larger bias for the other parame-

ters. Both models show poor precision in the estimation of the prices of risk, with standard deviations larger than the true value of λ_1 in both cases. However, this difficulty in estimating the price of risk is not unique to the ACM framework or this estimation process. Poor estimation of price of risk parameters has also been documented by Duffee and Stanton (2012) and Bolder (2001) when estimating the parameters of ATS models using MLE.

In general, the precision of the ACM estimation procedure is in line with that of MLE-based methods assessed in the literature. Bolder (2001) and Duffee and Stanton (2012) both find that in the context of a single-factor ATS model, MLE obtains parameter estimates with standard deviation in the order of 10% of the true value, except for the price of risk parameters as discussed.

4.1.2 Yield Curve and Pricing Factor Fits

The ability of the ACM procedure to capture pricing factors and fit the observed yield curves was evaluated through the simulation of a single three-factor ACM model. Figures 4.1 and 4.2 show model estimates of the pricing factors (estimates shown as red dotted lines) and absolute error between the model-implied and observed yield curves, respectively. It is evident that there is some bias in the estimates of the principal components. Additional tests showed that the model has some trouble distinguishing between pricing factors, especially for models with a larger number of factors. Such behaviour has also been documented by Pillay (2016) in the context of parameter estimation using the Kalman filter, suggesting that this may not be unique to ACM's estimation procedure. However, the weighted sum of the PCs tended to have a low bias and variation. The result is an accurate fit to the overall yield curve as shown in Figure 4.2. This figure shows that model error is concentrated at the short end of the yield curve with a maximum error of 10.68 basis points and an almost perfect fit at longer maturities. This concentrated error may be attributable to the fact that the short end of the yield curve tends to be more volatile than the long end, since the model may have difficulty capturing this variation.

4.2 US Yield Curve

The ACM procedure was applied to the data of Gurkaynak *et al.* (2006), as per the original specification of Adrian *et al.* (2013). Zero-coupon treasury yield curves were generated from the published NSS parameters according to equation (A.2) for maturities $n = 1, 2, \ldots, 120$ months. Figure 4.3 shows the absolute error for the

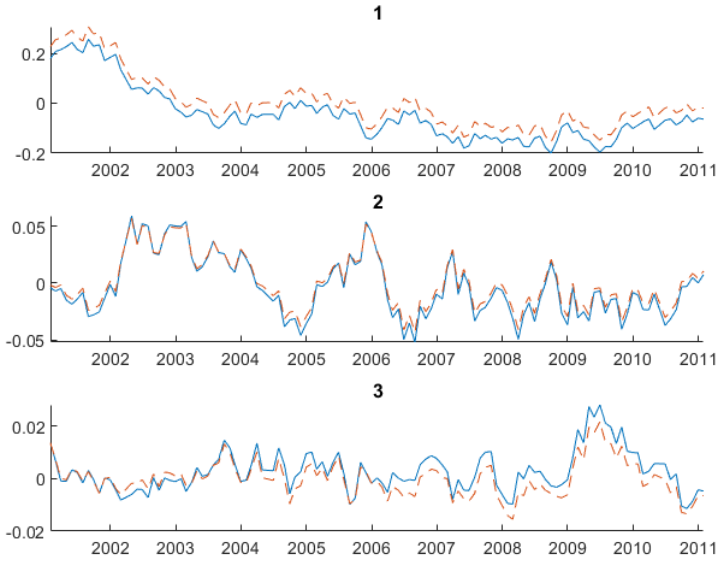

Fig. 4.1: Pricing factor estimates - three-factor Monte Carlo simulation

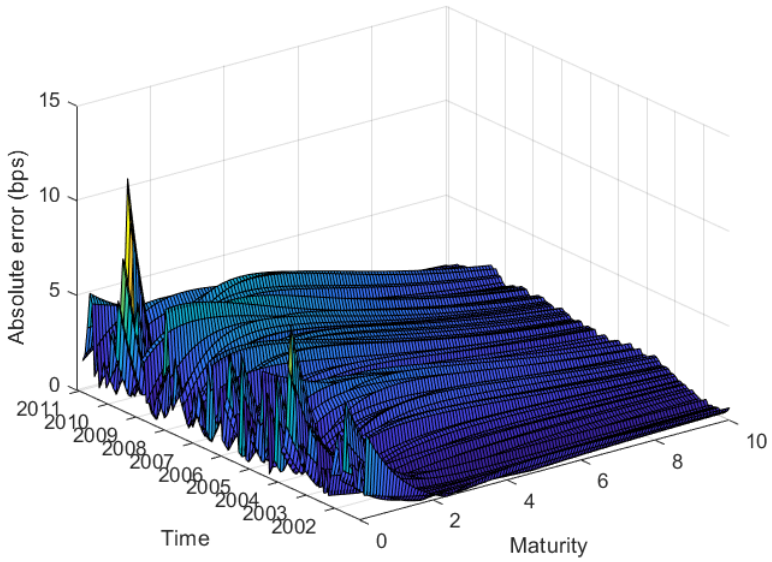

Fig. 4.2: Yield curve model error - three-factor Monte Carlo simulation

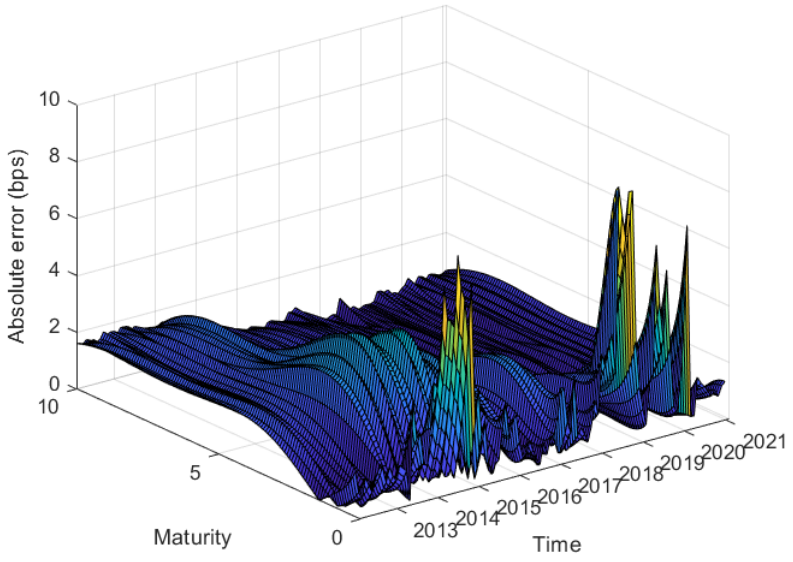

Fig. 4.3: Absolute error in model-implied US yield curves

model-implied time series of yield curves from January 2012 to January 2021, as compared to the observed yield curves. The application to NSS curves generates very accurate results, with a maximum absolute error of less than 10 basis points. Figure 4.4 shows estimates of the US nominal term premium for maturities of 24 and 120 months over the period from January 1987 to December 2011. These results are shown to closely match the term premia obtained by Adrian *et al.* (2013), published by the New York Federal Reserve Board (2022). Chapter 5 briefly discusses the discrepancies observed between the two sets of results.

Figure 4.5 shows the relationship between the 10-year US term premium and the yield curve's slope. Diebold and Li (2006) define the level, slope and curvature of the yield curve as:

$$L_t := y_t^{120}$$

$$S_t := y_t^{120} - y_t^3$$

$$C_t := 2y_t^{24} - (y_t^3 + y_t^{120})$$

where L_t, S_t and C_t represent time-t level, slope and curvature respectively and superscripts denote the maturity of yields in months. These quantities are shown to be closely related to the parameters of the NSS framework (see Appendix A for further details on level, slope and curvature in the NSS framework and their relation to the above definitions). The results shown in Figure 4.5 confirm the assertion that the term premium is closely related to the slope of the yield curve. Figure 4.6

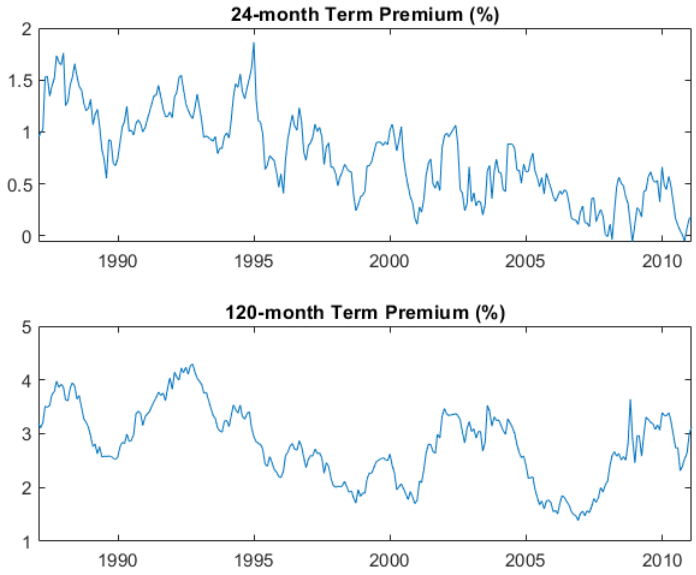

Fig. 4.4: US nominal term premia

shows the relationship between level, slope and curvature in the US yield curve.

There is an evident inverse relationship between the slope and curvature since 1990, with curvature typically falling as slope rises. Correlation between the two factors was -60.45% over this period. It is also noted that slope has a positive mean and general skew towards positive values, while curvature has a negative mean and skew towards negative values. All else equal, a negative curvature, as defined above, implies a more convex yield curve. However, an increasingly positive slope can also create a concavity in the yield curve, as visible from Figure A.1[1]. Additionally, the slope and curvature seem to be stationary and mean-reverting, as can be seen in Figure 4.7. This was confirmed with augmented Dickey-Fuller tests, which confirm stationarity for the two features at 95% significance. The result of these two effects is that the yield curve tends to move between a flat (i.e. low slope and low curvature) shape to a highly sloped, highly concave shape. The highly-sloped yield curve often corresponds to a high term premium, as per Figure 4.5, although the slope and term premium can also differ substantially. For example, a negative slope can exist simultaneously with a positive term premium, and vice versa, as shown in this Figure. While the term premium is itself not a mean-reverting

[1] Note that the "slope" coefficient β_1 under the NSS framework corresponds to the negative of the real-world slope as defined above (see Appendix A). Hence the convex shape of the slope loading factor under the NSS framework shown in Figure A.1 corresponds to a concave shape in the observed yield curve.

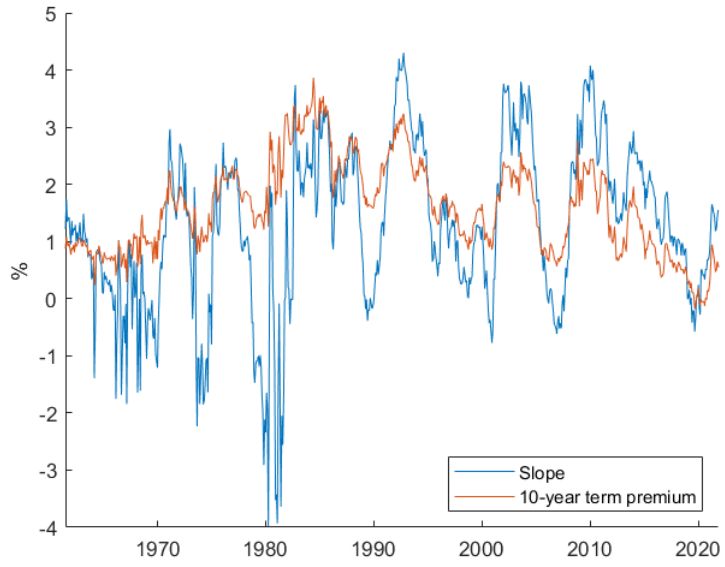

Fig. 4.5: 10-year US term premium vs yield curve slope

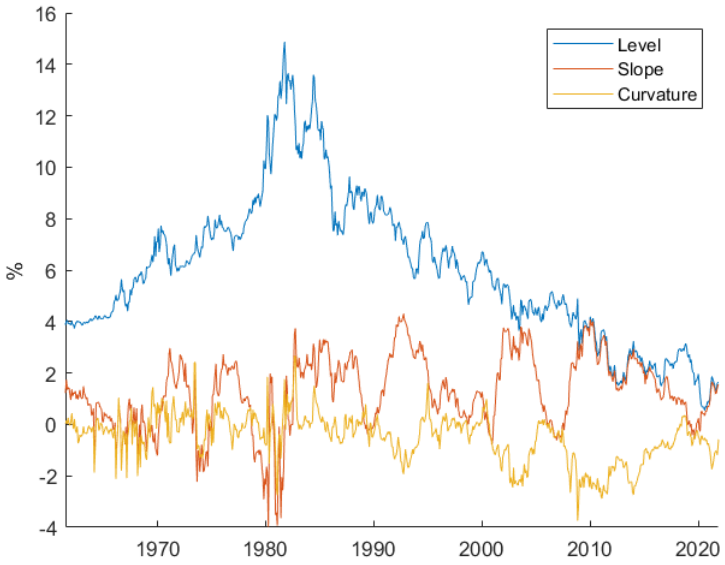

Fig. 4.6: Level, slope and curvature of the US yield curve, as defined by Diebold and Li (2006)

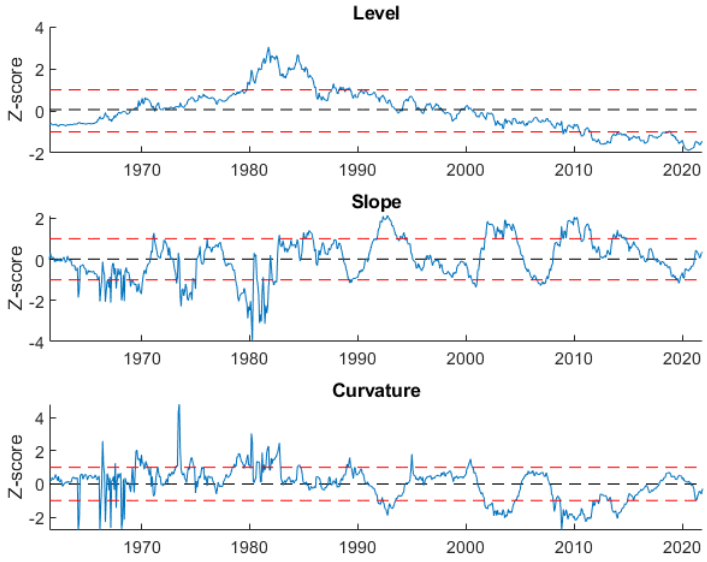

Fig. 4.7: Z-scores of level, slope and curvature in the US yield curve, means and standard deviations shown as dashed lines

process, Figure 4.5 shows that a high term premium tends to be associated with a subsequent flattening of the yield curve and may therefore be a signal to hold a barbell portfolio of bonds, as described in Chapter 1.

4.3 South African Yield Curve

Figure 4.8 shows the absolute error between the observed and model-fitted time series of yield curves for a five-factor ACM model. A maximum error of 145 basis points was observed, with errors concentrated in the short end of the yield curve. This is discussed further in Chapter 5.

4.3.1 Nominal Term Premium

Figure 4.9 shows the nominal term premium in South African bonds for maturities of 24 and 120 months over the time period from September 2012 to September 2021, and risk-free term rates are shown in Figure 4.10. Figure 4.11 superimposes the JSE All-Share Index (ALSI), obtained from Bloomberg, onto the 10-year nominal term premium.

Theoretically, the term premium should increase during times of economic uncertainty with longer maturities having a more volatile term premium (Soobyah

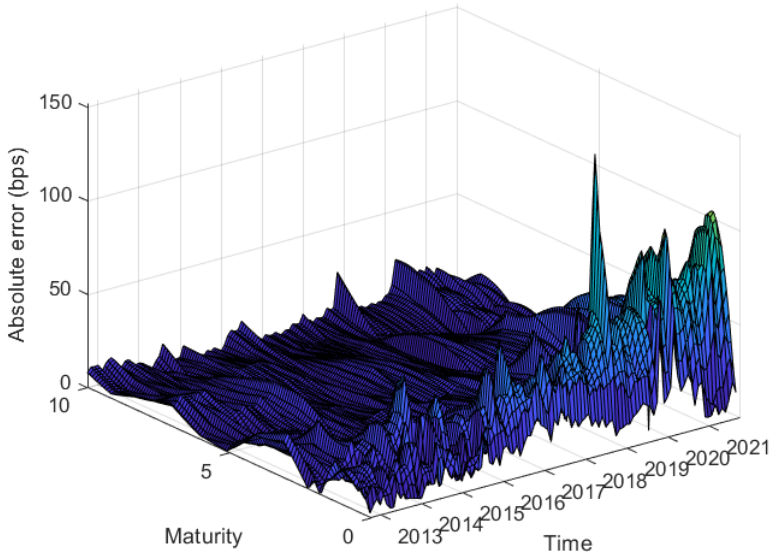

Fig. 4.8: Yield curve model error - SA data

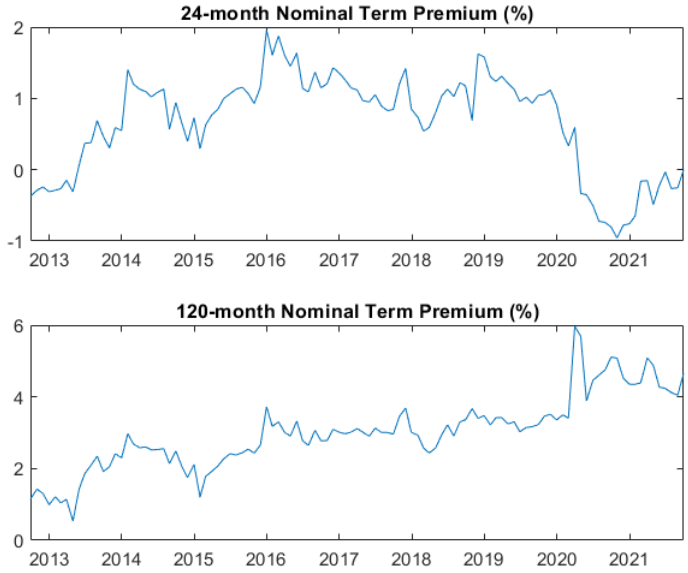

Fig. 4.9: SA nominal term premium

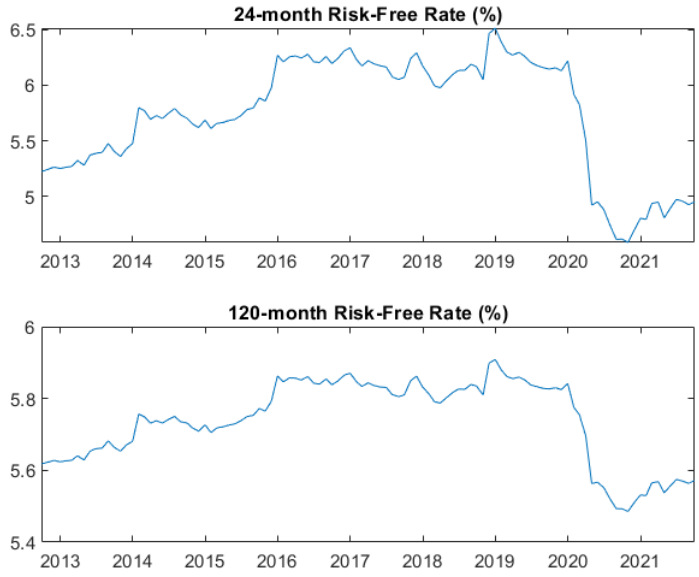

Fig. 4.10: SA risk-free term rates

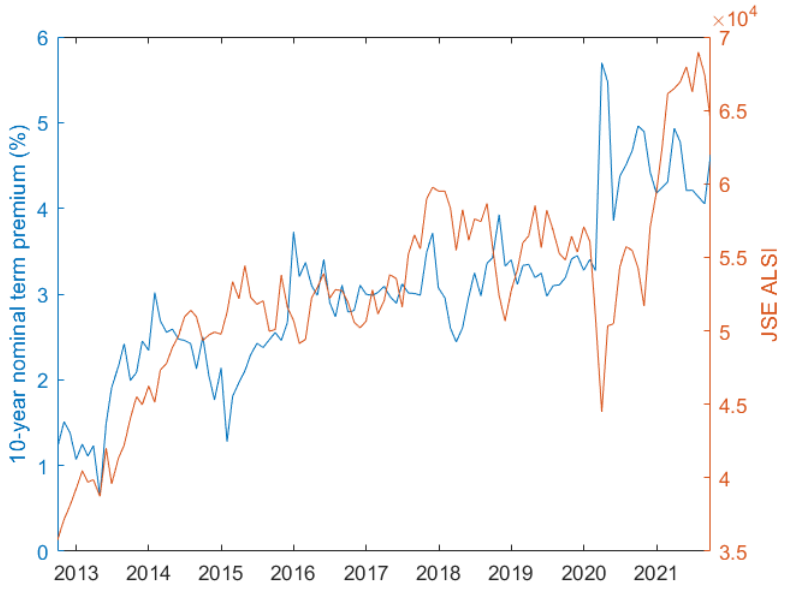

Fig. 4.11: SA 10-year nominal term premium vs JSE ALSI

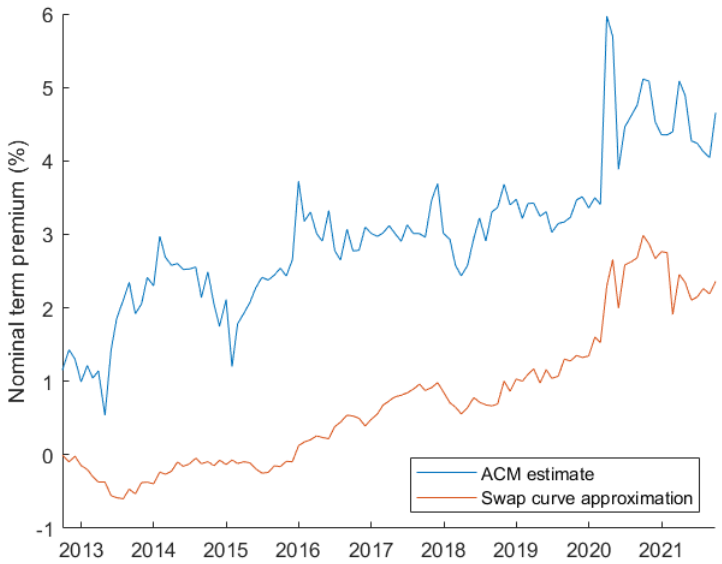

Fig. 4.12: SA 10-year nominal term premium and difference between 10-year nominal yield and swap rate

and Steenkamp (2020)), while the short rate tends to decrease during economic downturns and falling inflation. The results obtained confirm this. For example, sharp increases in the 10-year term premium are accompanied with dips in the ALSI in December 2015 and March 2020, corresponding to the surprise firing of South Africa's finance minister and the start of COVID-19 lockdowns, respectively. Falling risk-free rates are also observed in 2020, caused by the South African Reserve Bank (SARB) cutting repo rates in response to the economic slowdown caused by COVID-19. Using bootstrapped swap zero curves obtained from Bloomberg as a proxy for risk-free rates, the 10-year estimate was also compared to a model-independent approximation of the term premium as the difference between the 10-year nominal yield and the 10-year swap rate. This comparison is shown in Figure 4.12. The ACM estimate exhibits a larger magnitude and volatility than the model-independent approach, but trends in the two estimates appear consistent and agree with the economic interpretations discussed above.

Chapter 5

Model Diagnostics

This chapter investigates the speed and performance of ACM's estimation procedure and accounts for some of the results observed in the previous chapter. The application to US data is compared to the term premia published by Adrian *et al.* (2013), showing the effects of varying the length of the time series of yield curves and jumps in the time series of parameters published by Gurkaynak *et al.* (2006). Errors from the application to South African data are also investigated, showing the effect of the increased variation in the short end of the curve.

5.1 Estimation Speed

The speed of the estimation procedure was tested for models with $K = 1, 2, \ldots, 50$ pricing factors, using NSS curves generated from the parameters of Gurkaynak *et al.* (2006) as input data. Computation time was measured as the average run time of 1000 estimation procedures for a set number of pricing factors. Figure 5.1 plots these average computation times in milliseconds, demonstrating that increasing the number of pricing factors tends to increase computation times exponentially, albeit with very fast estimation for any number of factors. Considering that a typical application of the ACM procedure may use as little as five factors and that increasing K too high can cause problems in the estimation procedure (discussed in Section 5.3), the exponential-time efficiency of the procedure is not a practical concern.

5.2 US Data

Implementation of the ACM algorithm was validated against the term premium estimates published by ACM (New York Federal Reserve Board (2022)). The length of the time series of data was also varied to assess model performance. Under an application of a five-factor ACM procedure with $n = 1, 2, \ldots, 120$ months and the time series from June 1961 to November 2021, the resulting model-implied yield curves

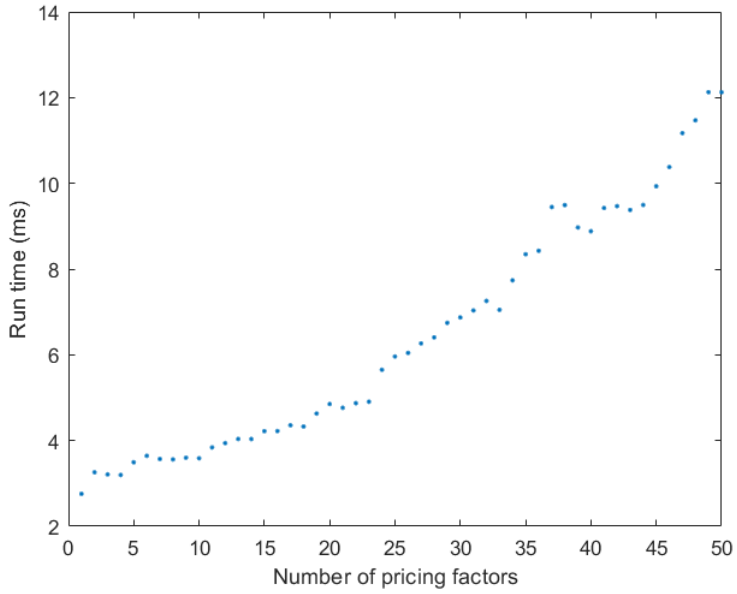

Fig. 5.1: Average of 1000 run times of ACM estimation procedure under varying number of pricing factors

match the inputted Nelson-Siegel-Svensson curves with a maximum absolute error of less than 10 basis points, as shown in Figure 4.3. However, it was observed that there is a discrepancy between these curves and the published curves of Adrian *et al.* (2013). Figure 5.2 shows this error. After some investigation it was found that the errors correspond to discontinuities in the NSS data. Figure 5.3 shows a plot of the model-implied short rate as obtained from the GSW data versus the short rate from ACM's published yield curves, as well as the first two NSS coefficients from Gurkaynak *et al.* (2006). It is evident that there are unrealistic discontinuities in latter data, and that these correspond to similar discontinuities in the model-implied short rate, which impacts regression (3.2) through equation (2.17). These discontinuities correspond to errors in the resulting model-implied yield curves, suggesting that ACM smoothed out these anomalies in their implementation.

By varying the length of the time series of data it was found that shorter time series created a positive bias in the term premium when compared to the initial time series of June 1961 to November 2021. Figure 5.4 shows how reducing the length of time series of the observed data drastically changes estimates of the term premium cross-section for January 2011. The plot shows a cross-section of the January 2011 term premium for models trained on three time series: June 1961 to January 2011 ($\sim$50 years of data), January 1981 to January 2011 (30 years of data) and January

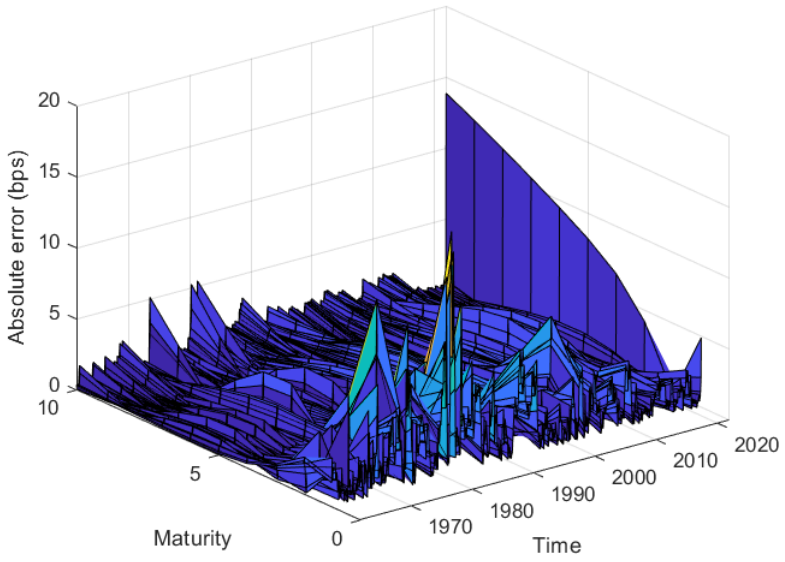

Fig. 5.2: Discrepancy between ACM's published fitted yields and NSS yield curves obtained via GSW's parameters

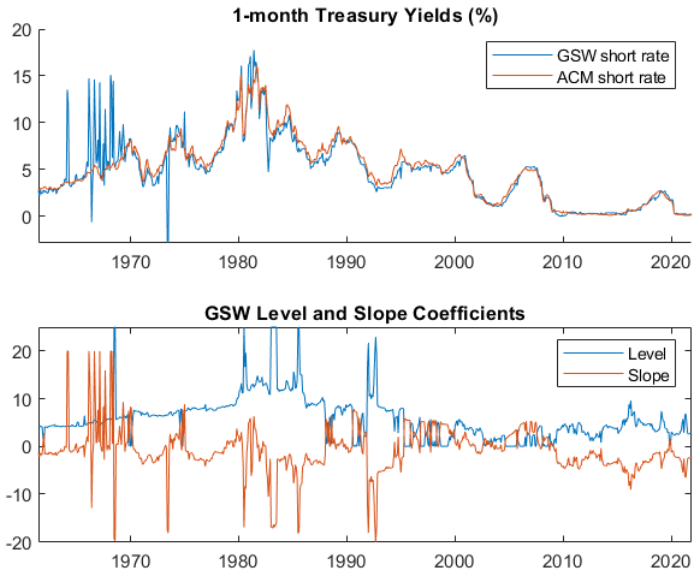

Fig. 5.3: Comparison of ACM published short rate and short rate implied by GSW parameters

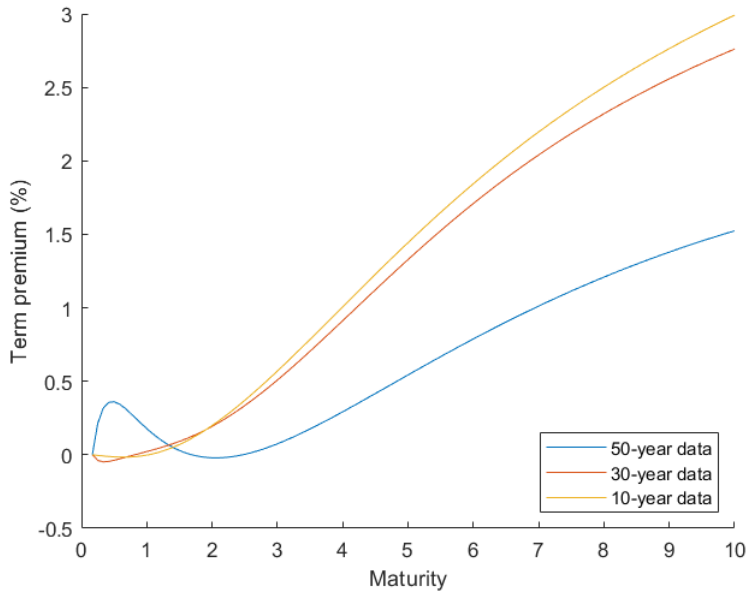

Fig. 5.4: Effect of varying length of time series on the January 2011 term premium cross-section

2001 to January 2011 (10 years of data). Figure 5.5 shows a longitudinal view of the same phenomenon. It is clear that there is a significant variation with respect to the choice of time series, which has implications for the application to South African data. This is discussed further in the next section.

Figure 5.6 shows the relationship between the regression coefficient β (shown as blue dots), obtained from equation (3.2), and the pricing factor B (shown as a red line) obtained from equation (2.31), from an application of the ACM procedure to the smoothed curves published by Gurkaynak *et al.* (2006). As shown in Chapter 2, both B and β correspond to the loadings on pricing factors which determine bond yields at different maturities. Since this relationship is not imposed directly, these plots provide an indication of how accurately the model generates bond returns (Adrian *et al.* (2013)). Figure 5.6 shows that the application to US data generates very accurate bond returns.

5.3 South African Data

A five-factor ACM procedure was applied to the South African data described in Section 3.6. Figure 5.7 shows that the application to South African bootstrapped yields provides significantly less accurate results than the application to the data of Gurkaynak *et al.* (2006), albeit with a fairly accurate fit for PC1 and PC2.

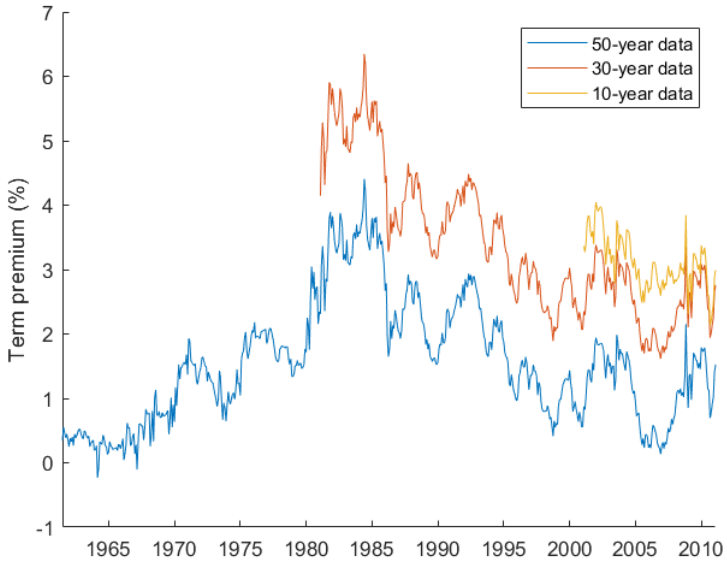

Fig. 5.5: Effect of varying length of time series on 10-year term premium time series

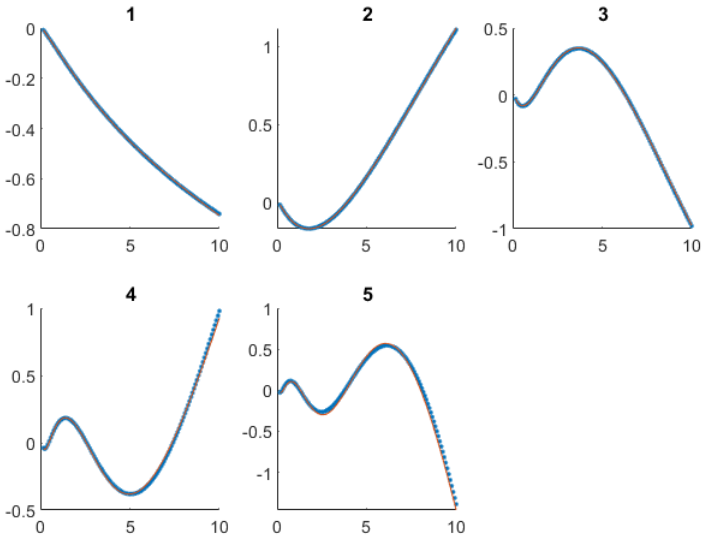

Fig. 5.6: Pricing factor B vs regression coefficient β - GSW data

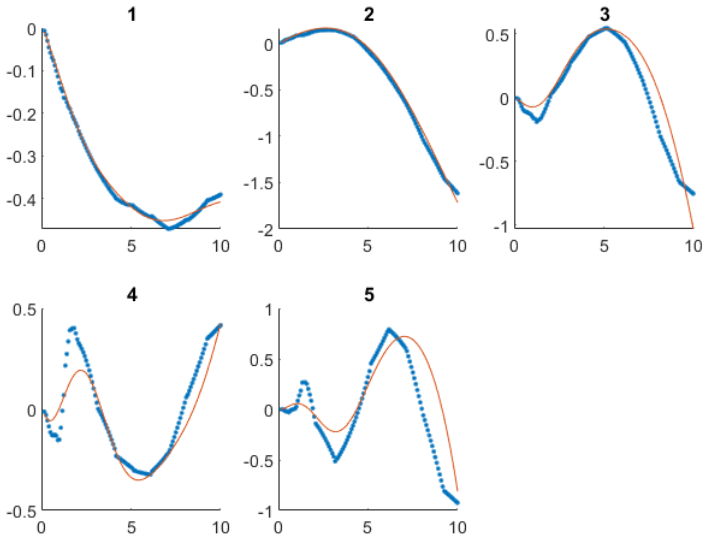

Fig. 5.7: Pricing factor B vs regression coefficient β - SA data

These errors point to problems with the linear regressions used to obtain β. The estimation problems here are most evident at a maturity of one month, as can be seen in panels 4 and 5 of Figure 5.7. This also corresponds to the location of maximum error in Figure 4.8. Observing the bootstrapped yield curves from the South African market shows noticeable idiosyncratic variation in the short end of the curve, especially around the one-month yield, which the model has trouble capturing (shown in Figure 5.8).

The variation in the short end of the curve appears to be the source of the error observed in Figure 5.7, and consequently the main source of error in the model-implied yield curves. In theory, such variation can be captured by additional pricing factors. Figure 5.9 shows that adding pricing factors does indeed capture additional variation in the yield curve and hence reduces the maximum model error. Model error in Figure 5.9 is measured as the logarithm of the maximum absolute error in basis points between the observed and model-implied yield curves for $K = 1, 2, \ldots, 30$.

Figure 5.9 also shows that there is a limit to the accuracy that can be achieved. Increasing K too high (24 in the application to SA data) results in the matrix $R'_{1,k}R_{1,k}$ in regression (3.1) tending towards singularity as the principal components X_k become progressively smaller in magnitude. This results in poor estimation of the VAR(1) coefficients and increasing model error. Optimal performance is achieved with South African data for $K = 18$.

Using 18 pricing factors, errors from the application to South African data are still

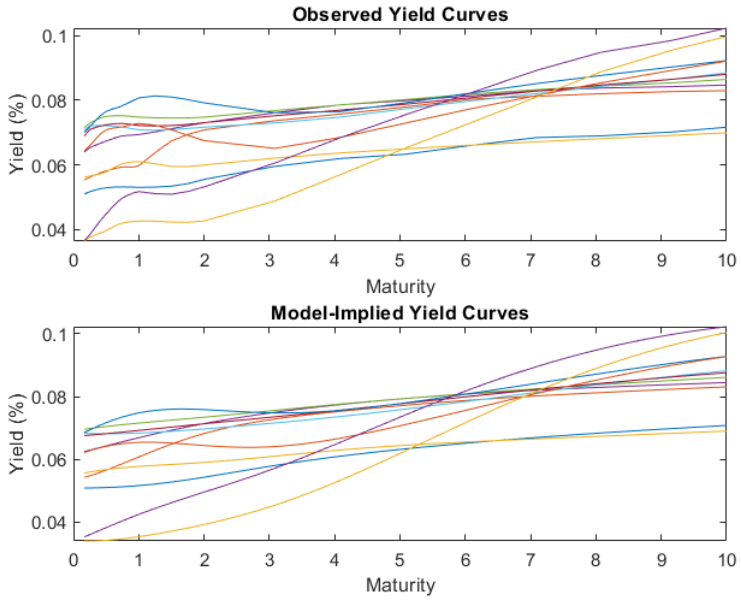

Fig. 5.8: Comparison of a selection of yield curves showing idiosyncratic variation in short end

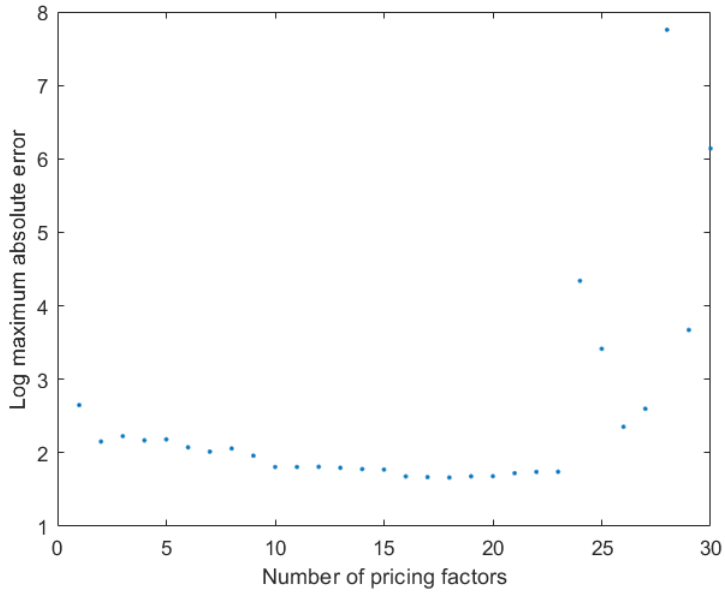

Fig. 5.9: Log maximum absolute model error for varying number of pricing factors

substantially higher than those from the five-factor application to NSS curves from the US market, owing to the fact that such curves smooth out some of this variation from the short end of the curve.

Comparing the term premia obtained from South African data to those obtained by Soobyah and Steenkamp (2020) (SS) of the SARB shows a significant bias. For example, the SS 10-year term premium peaks at about 2.25% in 2016, whereas the term premium estimated in this implementation peaks at 3.7%. This bias is evident at all maturities and seems to mirror the bias observed in Figure 5.5. Given that SS obtain their term premium estimates from a 20-year time series of yield curves beginning in 2000, whereas the data used for this dissertation contains nine years of yield curves from 2012 to 2021, training the model on a longer time series of data may reduce this bias.

Observations from Cohen *et al.* (2018) corroborate this conjecture. Cohen *et al.* (2018) compare term premium estimates from Adrian *et al.* (2013), Hördahl and Tristani (2014) and Kim and Wright (2005), and find that while the shape of the three estimates closely match, their magnitudes vary substantially, by up to 2%. Cohen *et al.* (2018) also estimate the term premium as the difference between the 10-year yield and surveyed 10-year expectations of the short rate obtained from the Survey of Professional Forecasters. This estimate is therefore model-independent and should not depend on the time series of data used in the estimation. The ACM estimate appeared more volatile, with generally higher levels than the other three estimates. Cohen *et al.* (2018) note that all three model-based approaches had very close fits to yield data, such that fitting error is not the cause of this discrepancy. The implication is that model-based approaches to estimating the term premium, as in any model-based estimation, necessarily introduce biases stemming from parameter estimation and modelling assumptions. By varying the time series of data used for estimation, the model assumes dynamics based on varying market regimes, which may differ substantially from one another. Despite these biases, the shape of the term premium estimates are consistent across choices of model and data, and trends in the term premium obtained by Soobyah and Steenkamp (2020) closely match those obtained in this dissertation. Similarly, trends in the longitudinal views of the three scenarios shown in Figure 5.4 are also closely matched, despite obvious biases.

Depending on the context, it may be beneficial to estimate the term premium using a joint term structure and macroeconomic model such as that developed by

Hördahl and Tristani (2014), which provides term premium values more closely in line with survey estimates. However, trends in the term premium, rather than its magnitude, are the most important variable for analyses such as that of portfolio management described in Chapter 1.

Chapter 6

Conclusion

The ACM procedure is shown to generate model-implied yield curves that closely fit observed yields, with parameter estimation being completed in milliseconds. Adding pricing factors to the model was shown to have negligible practical impact on estimation speed. Under Monte Carlo simulation, model parameters are estimated with a fair degree of accuracy, albeit with greater variation in the estimation of models with increasing numbers of factors, and poor accuracy for the price of risk estimates. Duffee and Stanton (2012) and Bolder (2001) find that MLE-based techniques also produce poor estimates of these parameters in ATS frameworks, suggesting that such difficulties are not unique to the ACM procedure.

Applying the ACM procedure to Nelson-Siegel-Svensson curves fitted to US zero-coupon nominal bond yields produces very accurate fits to the observed yields. Application to bootstrapped South African zero-coupon yield curves produces substantially less accurate fits, attributable to the idiosyncratic variation observed in the short end of the yield curve, some of which is smoothed in the US data through the use of NSS curves. Adding additional factors to the ACM model allows for much of this variation to be captured, with optimal model accuracy achieved using 18 factors when applied to South African yield curves; however, increasing the number of pricing factors above 23 was found to cause problems with the procedure's linear regressions. It was also found that the choice of yield curve time series substantially impacts the magnitude of the estimated term premium, a result which is corroborated by the findings of Cohen *et al.* (2018). Despite the varying magnitudes of the estimate, trends in the term premium are preserved across choice of sample data, which is arguably the most important variable in analysing the term premium.

Estimates of the nominal term premium in both the US and South African markets are shown to reflect the market's perception of risk, showing sharp increases during economic downturns and periods of uncertainty. The ACM term premium is

shown to closely match trends in a model-independent approximation of the term premium and to be closely related to the slope of the yield curve.

While this dissertation focuses on the performance of the ACM estimation procedure, there is scope for a detailed practitioner-oriented analysis of the estimated nominal term premium in the South African bond market, and of the inflation risk premium through an implementation of the framework of Abrahams *et al.* (2016). Such analyses may include further investigation of the relationship between the term premium and macroeconomic variables, similarly to Soobyah and Steenkamp (2020), and backtesting of bond portfolio strategies using the term premium as a trading signal.

Bibliography

Abrahams, M., Adrian, T., Crump, R., Moench, E. and Yu, R. (2016). Decomposing real and nominal yield curves, *Journal of Monetary Economics* **84**: 182–200.

Adrian, T., Crump, R. and Moench, E. (2013). Pricing the term structure with linear regressions, *Journal of Financial Economics* **110**(1): 110–138.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics* **50**(4): 745–787.

Bolder, D. (2001). Affine term structure models: Theory and implementation, *Bank of Canada Working Paper* (2001-15).

Brigo, D. and Mercurio, F. (2007). *Interest Rate Models - Theory and Practice*, 2 edn, Springer.

Bureau for Economic Research (2022). Survey of inflation expectations, 4th quarter 2022. Accessed January 2022.
URL: *https://www.ber.ac.za/knowledge/pkviewdocument.aspx?docid=15111*

Cochrane, J. and Piazzesi, M. (2005). Bond risk premia, *American Economic Review* **95**: 138–160.

Cochrane, J. and Piazzesi, M. (2008). Decomposing the yield curve, *Unpublished Working Paper, University of Chicago* .

Cohen, B., Hördahl, P. and Xia, D. (2018). Term premia: Models and some stylised facts, *BIS Quarterly Review September 2018* .

Cox, J., Ingersoll, J. and Ross, S. (1985). A theory of the term structure of interest rates, *Econometrica: Journal of the Econometric Society* **53**(2): 385–407.

Dai, Q. and Singleton, K. (2000). Specification analysis of affine term structure models, *Journal of Finance* **55**: 1943–1978.

Diebold, F. and Li, C. (2006). Forecasting the term structure of government bonds, *Journal of Econometrics* **130**(2): 337–364.

Duarte, J. (2004). Evaluating an alternative risk preference in affine term structure models, *Review of Fincancial Studies* **17**(2): 379–404.

Duffee, G. (2002). Term premia and interest rate forecasts in affine models, *The Journal of Finance* **57**(1): 405–443.

Duffee, G. and Stanton, R. (2012). Estimation of dynamic term structure models, *The Quarterly Journal of Finance* **2**(2).

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates, *Mathematical Finance* **6**(4): 379–406.

Fabozzi, F. (2006). *Fixed Income Analysis*, 2 edn, John Wiley and Sons, New Jersey.

Fama, E. and Bliss, R. (1987). The information in long-maturity forward rates, *The American Economic Review* **77**(4): 680–692.

Fisher, M. and Gilles, C. (1996). Estimating exponential-affine models of the term structure, *Federal Reserve Board Working Paper Series* .

Frankel, J. and Lown, C. (1994). An indicator of future inflation extracted from the steepness of the interest rate yield curve along its entire length, *Quarterly Journal of Economics* **109**: 517–530.

Gurkaynak, R., Sack, B. and Wright, J. (2006). The US treasury yield curve: 1961 to present, *Finance and Economics Discussion Series, Federal Reserve Board* (28).

Gurkaynak, R., Sack, B. and Wright, J. (2010). The TIPS yield curve and inflation compensation, *American Economic Journal* **2**(1): 70–92.

Hamilton, J. and Wu, C. (2012). Identification and estimation of affine term structure models, *Journal of Econometrics* (168): 315–331.

Hördahl, P. and Tristani, O. (2014). Inflation risk premia in the Euro area and the United States, *International Journal of Central Banking* **10**(3): 1–47.

Joslin, S., Singleton, K. and Zhu, H. (2011). A new perspective on Gaussian dynamic term structure models, *Review of Financial Studies* (24): 926–970.

Kim, D. and Orphanides, A. (2007). The bond market term premium: What is it, and how can we measure it?, *BIS Quarterly Review January 2007* pp. 27–40.

Kim, D. and Wright, J. (2005). An arbitrage-free three-factor term structure model and the recent behaviour of long-term yields and distant-horizon forward rates, *Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System* (33).

Kim, H. (2008). Challenges in macro-finance modelling, *BIS Working Paper* **240**.

Modise, T. (2019). *Interest Rates Products User Manual*, JSE.

Nelson, C. and Siegel, A. (1987). Parsimonious modeling of yield curves, *The Journal of Business* **60**(4): 473–489.

New York Federal Reserve Board (2022). Treasury term premia. Accessed December 2021.
URL: *https://www.newyorkfed.org/research/data_indicators/term_premia.html*

Piazzesi, M. and Swanson, E. (2008). Futures prices as risk-adjusted forecasts of monetary policy, *The Journal of Monetary Economics* (55): 677–691.

Pillay, D. (2016). Robustness of bond portfolio optimisation, *AIFMRM, University of Cape Town* .

Raubenheimer, H. and Kruger, M. (2010). Generating interest-rate scenarios for fixed-income portfolio optimisation, *South African Actuarial Journal* **10**: 1–42.

Soobyah, L. and Steenkamp, D. (2020). Term premium and rate expectation estimates from the south african yield curve, *South African Reserve Bank Working Paper Series* .

Statistics SA (2021). Consumer price index. Accessed January 2022.
URL: *https://www.statssa.gov.za/publications/P0141/P0141September2021.pdf*

Svensson, L. (1994). Estimating and interpreting forward interest rates: Sweden 1992-1994, *NBER Working Paper Series* (4871).

Vasiček, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**(2): 177–188.

Appendix A

The Nelson-Siegel-Svensson Framework

This appendix discusses the Nelson-Siegel-Svensson framework of the yield curve, as formulated by Diebold and Li (2006). The NSS framework provides an intuitive description of the main drivers of the shape of the yield curve, namely, level, slope and curvature.

A.1 Level, Slope and Curvature

Level is commonly calculated as the maximum maturity yield of a given curve, and under the NSS framework. Changing this quantity results in a parallel shift of the curve. Diebold and Li (2006) measure level as $L_t := y_t^{120}$, i.e. the time-t yield of the 10-year bond. Slope is calculated practically as the difference between a long-term and short-term yield, such as the 10-year yield minus the short rate, and corresponds to the increase in yield as maturity increases in a normal yield curve. Diebold and Li (2006) measure this as $S_t := y_t^{120} - y_t^3$. This is related to the term premium, which is described similarly. However, the slope of the yield curve necessarily includes the component of yields arising from expectations of future yields. As such, a high slope may be reflective of high short rate expectations rather than a high term premium. Curvature describes the relationship between the short, middle and long end of the curve. Visually this can be thought of as a "hump" in the curve, measured as $C_t := 2y_t^{24} - (y_t^3 + y_t^{120})$. Figure 4.6 in Chapter 4 shows the relationship between the three shape factors for the US yield curve, while Figure 4.7 shows the Z-scores and means of the factors.

A.2 The Framework

In certain contexts it may be beneficial to have access to yield curve data which smooths out idiosyncratic pricing issues such as liquidity, hedging demand, bid-ask spreads, etc. While such anomalies may be of interest to practitioners such as traders, in the context of a macroeconomic analysis it is often helpful to smooth the yield curve since such idiosyncratic factors are not explained by macroeconomic variables (Gurkaynak *et al.* (2006)). ACM use data published by Gurkaynak *et al.*

(2006), who address this problem by using the framework of Svensson (1994), an extension to the original framework developed by Nelson and Siegel (1987).

Diebold and Li (2006) formulate the framework of Nelson and Siegel (1987) in equation (A.1), with forward rates described by the four parameters β_{0t}, β_{1t}, β_{2t} and λ_t:

$$f_t(\tau) = \beta_{0t} + \beta_{1t}e^{-\lambda_t t} + \beta_{2t}\lambda_t e^{-\lambda_t \tau}, \tag{A.1}$$

which implies the following form for yields:

$$y_t(\tau) = \beta_{0t} + \beta_{1t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right). \tag{A.2}$$

Diebold and Li (2006) point out that the coefficients β_{0t}, β_{1t} and β_{2t} can be interpreted as the level, slope and curvature of the yield curve, respectively. It can easily be shown that the loadings on the factors β_{1t} and β_{2t} asymptote to zero such that $\lim_{\tau\to\infty} y_t^\tau = \beta_{0t}$, and hence that β_{0t} corresponds to the level of the yield curve. Frankel and Lown (1994) define the slope as $y_t^\infty - y_t^0$, and $-\beta_{1t}$ is exactly equal to this quantity. Lastly, the curvature of the yield curve is defined by Diebold and Li (2006) as $2y_t^{24} - y_t^3 - y_t^{120}$, with maturities in months. Diebold and Li (2006) show that this quantity equals $0.00053\beta_{2t} + 0.37\beta_{3t}$, such that yield curve curvature is equivalent to β_{3t}.

Figure A.1 gives a visual intuition for the above. The loading on the slope coefficient can be seen to decrease linearly with maturity, giving a slope to the curve by increasing short yields more than long yields. Loading on the curvature coefficient can be seen to add a hump to the yield curve.

Gurkaynak *et al.* (2006) point out that the original Nelson and Siegel (1987) specification has trouble capturing the effect of convexity at longer maturities. Accordingly, they make use of the extension of Svensson (1994), which adds an extra two parameters to equation (A.1). This extension adds greater flexibility by adding a second hump to the model, hence allowing for a more realistic representation of convexity in the yield curve:

$$f_t(\tau) = \beta_{0t} + \beta_{1t}e^{-\lambda_{1t}t} + \beta_{2t}\lambda_{1t}e^{-\lambda_{1t}\tau} + \beta_{3t}\lambda_{2t}e^{-\lambda_{2t}\tau}$$

$$y_t(\tau) = \beta_{0t} + \beta_{1t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau}\right) + \beta_{2t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau} - e^{-\lambda_{1t}\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau} - e^{-\lambda_{2t}\tau}\right).$$

This second hump typically appears at the far long-end of the curve, with the maturity of its peak determined by the parameter λ_{2t}. This extension has been shown to generate closer fits to observed yields while maintaining the smoothness of a Nelson-Siegel curve, and has seen widespread adoption in the literature (Gurkaynak *et al.* (2006), Raubenheimer and Kruger (2010)).

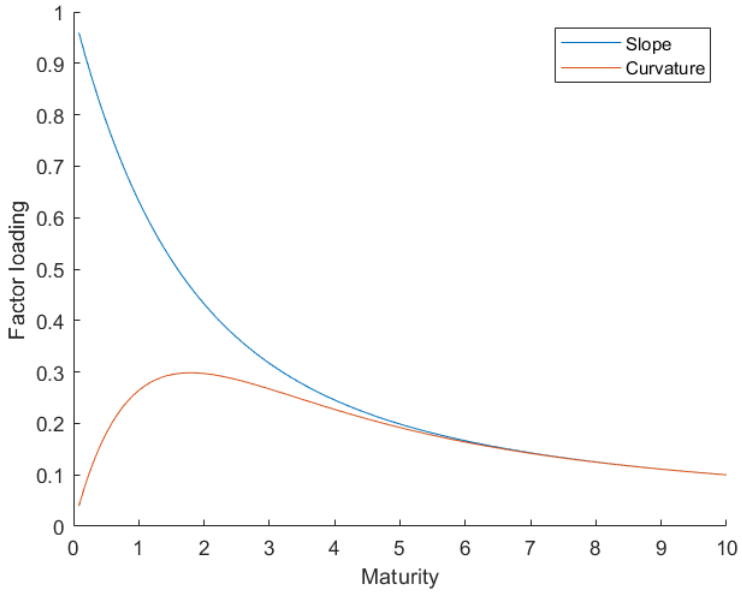

Fig. A.1: NSS slope and curvature loadings as a function of maturity

Appendix B

Simulation Results

Parameter	True Value	Estimate	Std Deviation
δ_0	3.355×10^{-3}	2.697×10^{-3}	1.100×10^{-4}
δ_1	7.843×10^{-3}	8.429×10^{-3}	2.291×10^{-17}
μ	-2.175×10^{-3}	-1.806×10^{-3}	9.732×10^{-5}
Φ	0.9972	0.9970	1.916×10^{-3}
λ_0	-5.528×10^{-3}	-5.318×10^{-3}	2.199×10^{-4}
λ_1	8.284×10^{-4}	6.698×10^{-4}	1.918×10^{-3}

Tab. B.1: Parameter estimates from one-factor model

Parameter	True Value	Estimate	Std Deviation
δ_0	3.355×10^{-3}	2.697×10^{-3}	1.100×10^{-4}
δ_1	$1 \times 10^{-3} \begin{bmatrix} 7.843 \\ 16.69 \\ 29.58 \end{bmatrix}$	$1 \times 10^{-3} \begin{bmatrix} 7.480 \\ 14.77 \\ 30.67 \end{bmatrix}$	$1 \times 10^{-4} \begin{bmatrix} 8.618 \\ 43.07 \\ 13.38 \end{bmatrix}$
μ	$1 \times 10^{-3} \begin{bmatrix} -2.175 \\ 0.1767 \\ -0.04269 \end{bmatrix}$	$1 \times 10^{-3} \begin{bmatrix} -1.960 \\ 0.003430 \\ -0.002788 \end{bmatrix}$	$1 \times 10^{-4} \begin{bmatrix} 3.861 \\ 0.9411 \\ 0.08082 \end{bmatrix}$
$\Phi \times 10$	$\begin{bmatrix} 9.972 & 0 & 0 \\ 0 & 9.690 & 0 \\ 0 & 0 & 8.210 \end{bmatrix}$	$\begin{bmatrix} 9.953 & 0 & 0 \\ 0 & 9.440 & 0 \\ 0 & 0 & 7.977 \end{bmatrix}$	$\begin{bmatrix} 0.06449 & 0 & 0 \\ 0 & 2.344 & 0 \\ 0 & 0 & 3.674 \end{bmatrix}$
λ_0	$1 \times 10^{-3} \begin{bmatrix} -6.435 \\ -0.2140 \\ -0.1537 \end{bmatrix}$	$1 \times 10^{-3} \begin{bmatrix} -5.992 \\ 0.1562 \\ 0.2304 \end{bmatrix}$	$1 \times 10^{-4} \begin{bmatrix} 7.805 \\ 5.496 \\ 2.6581 \end{bmatrix}$
$\lambda_1 \times 10^3$	$\begin{bmatrix} -3.089 & 74.86 & 157.4 \\ -6.881 & -12.21 & 188.8 \\ -4.318 & -3.240 & -79.31 \end{bmatrix}$	$\begin{bmatrix} -6.173 & 63.54 & 158.5 \\ -6.431 & -44.84 & 176.7 \\ -4.335 & -7.132 & -93.67 \end{bmatrix}$	$\begin{bmatrix} 7.394 & 18.92 & 11.44 \\ 2.462 & 29.32 & 55.61 \\ 0.7182 & 14.35 & 40.83 \end{bmatrix}$

Tab. B.2: Parameter estimates from three-factor model

Appendix C

Naïve Estimation of the Inflation Risk Premium

As discussed in Chapter 2, Abrahams *et al.* (2016) estimate the inflation risk premium using an ATS model as the difference between model-implied break-even inflation under the objective and risk-neutral measures. By extracting a "real term premium" through an application of the ACM procedure to real bond yields, an inflation risk premium can be estimated naïvely as the difference between the nominal term premium and this real term premium. This estimate has some obvious flaws. For example, the larger liquidity premium present in real yields means that this estimate of the inflation risk premium is biased by the difference between the liquidity premia in the nominal and real bond yields. Differences between supply and demand for the two classes of bonds and other idiosyncratic effects also affect the accuracy of this estimate.

An application of this naïve estimation using nominal and real yield curves obtained from the NSS parameters of Gurkaynak *et al.* (2006) and Gurkaynak *et al.* (2010) for the period from January 1999 to January 2014 is shown in Figure C.1, with break-even inflation estimated as the difference between model-implied nominal and real yields and risk-neutral break-even inflation given as the difference between the risk-neutral nominal and real yields. This last component is discussed in detail below.

A comparison to the results of Abrahams *et al.* (2016) (Figure C.2) shows that while the estimates of break-even inflation under the objective measure closely match, showing that the ACM procedure is able to accurately price real bonds under the this measure, the naïve inflation risk premium exhibits much more variation and a stronger upward trend than the original estimate. However, the naïve estimate does capture some features of the original estimate, including a downward trend between 2004 and 2007 and a sharp increase leading into 2008.

Abrahams *et al.* (2016) decompose break-even inflation into three components. As per equation (2.5), the time-t break-even inflation for horizon n is denoted π_t^n, with $\pi_t^{n,\mathbb{Q}}$ representing break-even inflation under $\mathbb{Q}$, interpreted as the risk-neutral expectation of average future inflation. Letting $\phi_{t,i}^n$ and $\phi_{t,L}^n$ denote the corresponding inflation risk premium and liquidity component respectively, this decomposi-

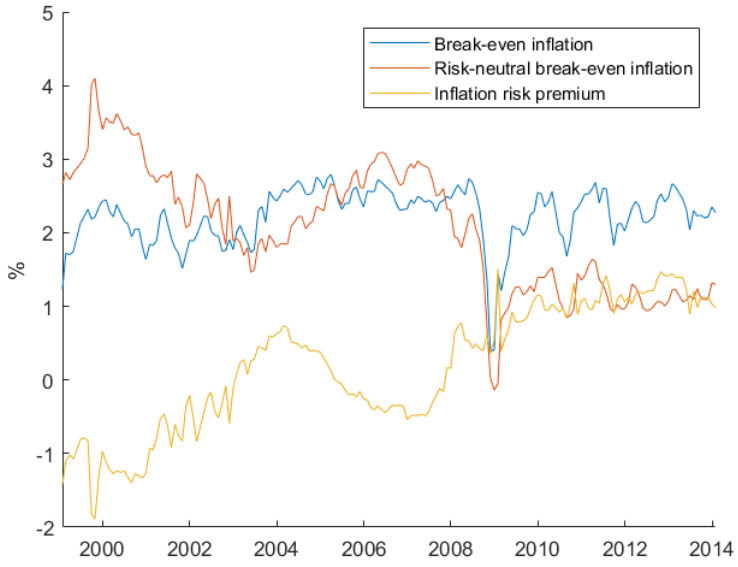

Fig. C.1: Naïve decomposition of break-even inflation

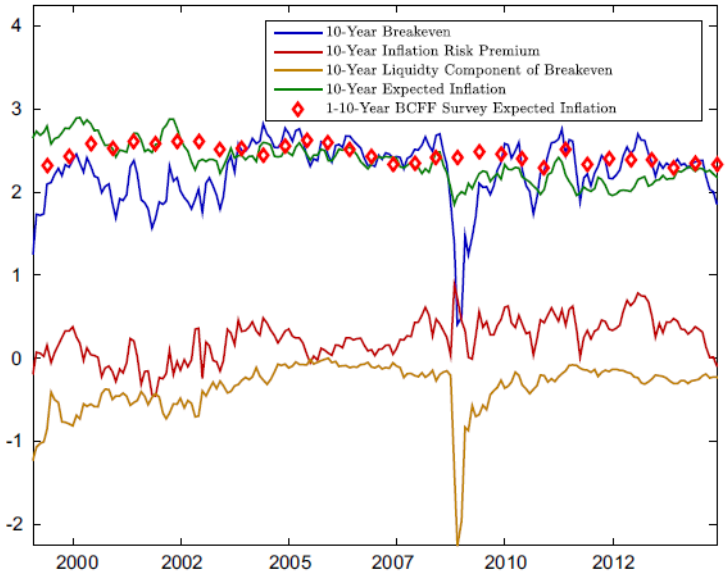

Fig. C.2: Decomposition of break-even inflation. Source: Abrahams *et al.* (2016)

tion can be represented as

$$\pi_t^n = \pi_t^{n,\mathbb{Q}} + \phi_{t,i}^n + \phi_{t,L}^n. \tag{C.1}$$

In contrast, the naïve approach decomposes break-even inflation as follows:

$$\begin{aligned}\pi_t^n &= y_t^n - y_{t,R}^n \\ &= y_t^{n,\mathbb{Q}} + \phi_t^n - y_{t,R}^{n,\mathbb{Q}} - \phi_{t,R}^n,\end{aligned} \tag{C.2}$$

as per an ACM decomposition, with y_t^n representing nominal bond yields, ϕ_t^n the nominal term premium and subscripts R denoting the same quantities associated with real yields. The naïve inflation risk premium is defined as

$$\bar{\phi}_{t,i}^n := \phi_t^n - \phi_{t,R}^n.$$

Thus, under the assumption that $\bar{\phi}_{t,i}^n$ includes the liquidity component of nominal and real yields, relating (C.1) and (C.2) yields

$$\pi_t^{n,\mathbb{Q}} \equiv y_t^{n,\mathbb{Q}} - y_{t,R}^{n,\mathbb{Q}} =: \bar{\pi}_t^{n,\mathbb{Q}}.$$

That is, the risk-neutral inflation expectations of Abrahams *et al.* (2016) correspond to the difference between the risk-neutral nominal and real yields obtained from the ACM estimation. Contrasting these two quantities explains much of the discrepancy between the naïve inflation risk premium and that of Abrahams *et al.* (2016). Figure C.1, plotting the decomposition $\pi_t^n = \bar{\pi}_t^{n,\mathbb{Q}} + \bar{\phi}_{t,i}^n$, shows that the discrepancies observed between the two approaches stem mainly from the fact that $\bar{\pi}_t^{n,\mathbb{Q}}$ does not accurately represent the inflation expectations component $\pi_t^{n,\mathbb{Q}}$. The naïve expectations component $\bar{\pi}_t^{n,\mathbb{Q}}$ exhibits much stronger volatility and an exaggerated downward trend, corresponding to the increased volatility and upward trend in the naïve inflation risk premium $\bar{\phi}_{t,i}^n$.
Capturing this component accurately requires the joint pricing of real and nominal bonds, as per the framework of Abrahams *et al.* (2016). In the naïve approach, price of risk parameters are necessarily estimated separately for the nominal and real bonds, and the two procedures result in substantially different estimates of said parameters. The joint pricing procedure of Abrahams *et al.* (2016) estimates one market price of risk which is used to price both bonds, while the difference between the nominal and real bond prices is derived from the inclusion of CPI into the real bond pricing parameters and the liquidity adjustment. Under the naïve approach, these different prices of risk result in inconsistent dynamics of the risk-neutral real and nominal bond prices, giving rise to the inaccurate dynamics of $\bar{\pi}_t^{n,\mathbb{Q}}$. However, the naïve estimate is still able to capture many features of the rigorously estimated inflation risk premium, and is shown below to be closely related to year-on-year changes in CPI as well as surveyed inflation expectations.

Figure C.3 shows the naïve inflation risk premium associated with the 24- and 120-month nominal bonds in the South African market. This application made use of a five-factor ACM model trained on real and nominal bond zero curves from

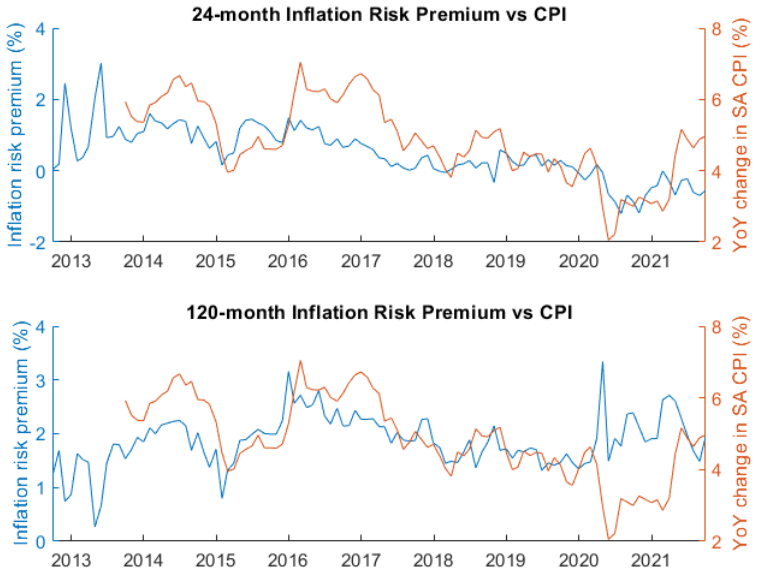

Fig. C.3: SA naïve inflation risk premium and year-on-year CPI

September 2012 to September 2021. Yields were obtained from Bloomberg and bootstrapped according to the JSE methodology. CPI data published by Statistics SA (Statistics SA (2021)) was used to obtain year-on-year inflation rates, which are superimposed on the inflation risk premia in Figure C.3.

This plot shows that the naïve inflation risk premium captures trends in the CPI, showing a sharp increase in late 2015 and a downward trend between 2016 and 2020. These trends are also in line with analysts' decreasing surveyed inflation expectations from the Bureau for Economic Research (2022) over that period, shown in Figure C.4.

While the naïve inflation risk premium has some clear drawbacks stemming from the lack of a rigorous theoretical framework, the ease and speed of implementation given the efficient ACM procedure may give the approach some utility in quickly obtaining a rough estimate of trends in the inflation risk premium without the need of MLE.

Fig. C.4: One year ahead surveyed inflation expectations