Implications of a generic no-arbitrage condition on restructuring missing data

Master thesis - August 2019

Author:
Paul Kemper BSc.
S1589717

Supervisors:
University of Twente
dr. Berend Roorda
dr. Wouter van Heeswijk

Triple A – Risk Finance
dr. Jakob Bosma

University of Twente
Drienerlolaan 5
7522 NB Enschede

Triple A – Risk Finance
Hoogoorddreef 54
1101 BE Amsterdam

I. Acknowledgement

This paper is written based on the idea, and initial research performed by dr. Jakob Bosma. Part of chapter 3 and appendix 11 in full are taken from his unpublished paper: "*Generic Proxy Methodology for Risk Factors".* In addition to elaborating on his theory, we focus on implementing and testing his theory on several models.

II. Management summary

In this paper we determine the impact of a no-arbitrage condition on parameter estimation using models which otherwise possibly yield arbitrage opportunities. By establishing a positive impact on the overall out-of-sample fit, we aim to assist banks in their dual goal to implement a proxy methodology which generates sufficiently realistic market risk scenarios, and to gain the approval of the regulator to apply the methodology (BCBS, 2016) (BCBS, 2018).

Our generic proxy is defined as follows:

Definition 1. Generic proxy methodology – *In the event we observe at a given historical date* t *insufficient market prices to infer all risk factor values* x_t *we can derive a proxy value* $\hat{x}_t$ *from risk factors and a historical overnight risk-free rate* r *observed at* $t-1$ *with the following expression:*

$$\boldsymbol{M}_{t-1} * \left(\boldsymbol{F}_t(\widehat{\boldsymbol{x}}_t, \boldsymbol{\tau}_t) - \boldsymbol{F}_{t-1} * (1 + r_{t-1}) + diag\left\{\frac{\delta \boldsymbol{F}_{t-1}}{\delta \boldsymbol{\tau}'}\right\}\right) = 0. \quad (3.3)$$

With historical data observed at date $t-1$ we can solve the linear system (3. 3) for the instruments' theoretical prices at time t and subsequently solve for the risk factor values $\widehat{\boldsymbol{x}}_t$, which then serve as the proxy values for the unidentified $\boldsymbol{x}_t$ at time t

In order to use our proxy methodology to estimate risk factors, we first need to establish the soundness of our imposed condition. This paper gives insight into this by providing answers to the following questions:

> *What are the implications on the overall fit of a pricing model, which potentially yields arbitrage possibilities, when a no-arbitrage condition is imposed?*

We use Matlab to test the effect on the Nelson Siegel model and the Principal Component Model in combination with US constant maturity treasuries based yield data and the Stochastic Alpha Beta Rho model on S&P 500 index option prices. We first test the overall fit of the pricing models on the full data set, and sequentially we add our proxy penalties with distinct weights on the RMSE and the proxy error. Sequentially we strip tenors or strikes from the data set to determine the effect of our proxy in market with missing data.

Based on the observed results we conclude that our generic proxy has a slightly preferred status over using set pricing model directly based on overall RMSE but mostly on out-of-sample RMSE in scenarios where data is stripped from the original set. Our proxy also does not require additional data or expert judgements, which eases implementation process in existing pricing platforms.

III. Table of figures

Table of figures

IV. Table of tables

Contents

1. Introduction

In this paper we determine the impact of a no-arbitrage condition on parameter estimation using models which otherwise could yield arbitrage opportunities. Our aim is to test the appropriateness of using the generic proxy to fill in blanks in historical data sets.

We perform this study at Triple A Risk-Finance, a financial risk consultancy company based in Amsterdam. This study is part of a bigger research and development venture the banking team is undertaking, to establish itself as an expert party within the industry.

With this proxy methodology we aim to assist banks in their dual goal to implement a proxy methodology which generates sufficiently realistic market risk scenarios, and to gain the approval of the regulator to apply the methodology (BCBS, 2016) (BCBS, 2018).

Under new Basel regulations, banks are required to provide databases containing risk factors which influence their market book. These risk factors need to be based on historical data and give a realistic view of several market shocks. However, in practice these databases contain gaps where data is missing. These gaps come from illiquid markets, characterized by big bid-ask spreads, which are often linked to stressed markets, and therefore certainly relevant for these databases.

Regulations do allow for proxy methodologies, however these are submitted to regulatory review. Regulators want to see proxies which show a good track record for the actual positions. (BCBS, 2013). The methodologies themselves also need to be straightforward and based on models that are eligible to enjoy regulatory approval based on previous amendments.

In order to use our proxy methodology to estimate risk factors, we first need to establish the soundness of our imposed condition. This paper gives insight into this by providing an answer to the main question:

> *What are the implications on the overall fit of a pricing model, which potentially yields arbitrage possibilities, when a no-arbitrage condition is imposed?*

We accompany this question with the following sub questions:

> *What is arbitrage?*

We give a brief introduction on arbitrage followed by an example, demonstrating an arbitrage opportunity in a pricing model. To conclude we discuss the existence of arbitrage in affine term-structure models.

> *Which pricing models are non-arbitrage free by definition?*

We provide an overview of models that contain arbitrage and make an exposition of when models are generally arbitrage free.

> *How do we introduce a no-arbitrage condition to pricing models?*

Introduction

We explain our working method; we set out our methodology and prove its soundness.

How does a proxy effect the out-of-sample fit of a pricing model?

Concluding we come back to our main question by answering the last sub question. We test our methodology by implementing it in different pricing models. We compare estimated parameters of the pricing model with and without our imposed condition. Sequentially, we test our methodology for its soundness to backfill missing data in historical databases of risk factors.

In order to derive the result necessary to answer our main question, we also deliver three Matlab models which can be used to fit a Nelson Siegel model on yield data, a SABR model on volatility curves and a PCA model on yield data. These models can be fitted directly or by addition of our proxy, where different weights can be put on the proxy.

Our paper is organized as follows: in section 2 we discuss arbitrage and implicitly answer on sub question 2 and 3, followed by section 3 in which we explain our working method and so answer our last sub question. In section 4 to 6 we discuss pricing models and elaborate on our findings. We conclude, discuss and suggest further research in the last sections.

2. Arbitrage

A portfolio of financial instruments (portfolio) is represented by a vector of values, either positive or negative, where each point in the vector represent the position in an instrument. One is long in a position if the value is positive and short in the instrument if the value is negative. We define a portfolio u as an arbitrage portfolio if:

- The portfolio strategy is self-financing;
- $V_0 = 0$;
- $V_T \geq 0, P - almost\ surely$;
- $P(V_T > 0) > 0$.

Where V_t is the value of the portfolio at time t and P is a probability measure on some space Ω and a filtration i.e. an increasing family of σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$. The portfolio is self-financing if the value process satisfies:

$$dV_t = \sum_{i=0}^{N} h_t^i dS_t^i,$$

where S_t is the price of a stock at time t and is assumed to be a P-martingale[1] and portfolio h_t^i is the number of units of asset i at time t.

Arbitrage can be created by longing a set of financial products and/or derivatives at a certain prices and shorting another set of financial products and/or derivatives at a higher price, and vice versa, while making sure no exposure to the underlying is created.

We introduce one plain example; consider the stock Arcelor Mittal, which is traded on both the Euronext Amsterdam (MT) as well as the Bolsa de Madrid (MTS). Since it concerns the exact same stock type, in a perfect world[2], the prices of both instruments should be the same. When MT trades at a higher price than MTS, one could short MT and long MTS making an instantaneous profit while adding no exposure to Arcelor Mittal to his portfolio.

An extension of this example can be found by making the arbitrage trade triangular. Arcelor Mittal does not only trade on the Euronext Amsterdam (AMS: MT) and Bolsa de Madrid, but also on the New York Stock Exchange (NYSE: MT). If AMS: MT trades at a higher price as the NYSE: MT divided by the current EUR/USD exchange rate, one shorts AMS: MT, longs NYSE: MT and exchanges the euros it received for the short AMS: MT position into dollars, leaving the trader with an excess euro amount for its arbitrage trade while no exposure to Arcelor Mittal is added to his portfolio.

One of the main pillars supporting the modern theory of Mathematical Finance is the first Fundamental Theorem of Asset Pricing (FTAP). FTAP finds its origin dating back to work of Black

[1] Since V is self-financing and it consists only of martingales, itself is also a martingale.
[2] E.g. no transaction costs, no tax benefits, no costs for shorting a stock.

and Scholes (1973) and Merton (1973). In this framework one changes the underlying P-measure into a risk-neutral or martingale measure Q.

As an example, we prove that the underlying process in the Black and Scholes model is free of arbitrage, taken from lecture notes from Björk (2010), let us consider the Black and Scholes model. We define a process for an equity underlying S as:

$$\mathrm{d}S_t = \alpha S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t,$$

where α is a factor that projects an average return, σ is the volatility of the underlying and W is an Geometric Brownian Motion (GBM). Next we define a process of a risk free bond B by:

$$\mathrm{d}B_t = r_t B_t \mathrm{d}t,$$

where r_t is the (stochastic) risk free rate.

We introduce Z and set it to $Z = S/B$. This results in the process of Z of:

$$\mathrm{d}Z_t = Z_t(\alpha - r_t)\mathrm{d}t + Z_t \sigma \mathrm{d}W_t. \tag{2.1}$$

Next, we need to formulate a measure to transform real world probabilities P into risk neutral probabilities Q. We apply a Girsanov transformation, which is used to transform a physical measure into a risk-neutral measure, on $[0, T]$:

$$\begin{cases} \mathrm{d}L_t = L_t \varphi_t \mathrm{d}W_t, \\ L_0 = 1, \end{cases}$$

$$\mathrm{d}Q = L_T \mathrm{d}P, \text{on } \mathcal{F}_T,$$

where φ_t is a process that we can define later and L is a random variable denotes as:

$$L = \frac{dQ}{dP}, on\ \mathcal{F}.$$

From the Girsanov Theorem we have:

$$W_t^Q = W_t - \int_0^t \varphi_s \mathrm{d}s, \text{or}$$

$$\mathrm{d}W_t = \varphi \mathrm{d}t + \mathrm{d}W_t^Q, \tag{2.2}$$

where W^Q is a Q-Wiener process.. A Wiener process is used to represent the integral of a white noise Gaussian process.

By using (2. 2) and substituting this into the P-dynamics given in (2. 1) we obtain the Q-dynamics for Z:

$$\mathrm{d}Z_t = Z_t[\alpha - r_t + \sigma\varphi_t]\mathrm{d}t + Z_t \sigma \mathrm{d}W_t^Q.$$

Arbitrage

We find a unique martingale measure Q, with Girsanov kernel given by:

$$\varphi_t = \frac{r_t - \alpha}{\sigma}.$$

From here we induce the Q-dynamics of S, which are given by:

$$\mathrm{d}S_t = r_t S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t^Q.$$

Since we showed the existence of a unique martingale measure for S, we have proven that the underlying process in the Black and Scholes model is free of arbitrage (Björk T. , 2009).

3. Working method

In this section, we set out the working method as developed by Jakob Bosma (2018). We provide an example throughout the theory which is in italics. The example consist of a set of 4 call options with a price of c_t at time t, where our options have a tenor of 1, 2, 3 and 4 years respectively. The options are priced using the Black and Scholes method, for which's theoretical outset we refer to section 5.1.

"We consider a vector of m risk factors $\boldsymbol{x}_t$ derived at time t from n observable instruments with prices denoted by $p_t(\tau)$ for various time-to-maturity tenors τ. These risk factors could for instance be zero-coupon rates and plain interest-rate swaps as associated instruments; or forward volatilities derived from interest-rate caps and floors for a particular strike level.[3]

The bank maintains an in-house pricing system to value the underlying instrument with the valuation function $F(\boldsymbol{x}_t, \boldsymbol{\tau}_t)$. We gather all n prices observed at time t in a vector $\boldsymbol{p}_t$, all corresponding maturities at time t in $\boldsymbol{\tau}_t := (T_1 - t, T_2 - t, \ldots, T_n - t)$ and the associated valuation functions in

$$\boldsymbol{F}(\boldsymbol{x}_t, \boldsymbol{\tau}_t) = \big(F(\boldsymbol{x}_t, T_1 - t), F(\boldsymbol{x}_t, T_2 - t), \ldots, F(\boldsymbol{x}_t, T_n - t)\big)'.$$

Note that $t = 0$ lies in the past and $min(\boldsymbol{T}_{\boldsymbol{x}} - t_x) > 0$ otherwise the instrument would have been terminated. Typically, implied risk factors are inferred by solving the following system for $\boldsymbol{x}_t$ a process also referred to as bootstrapping or stripping:[4]

$$\boldsymbol{p}_t = \boldsymbol{F}(\boldsymbol{x}_t, \boldsymbol{\tau}_t) \qquad (3.1)$$

For the construction of a complete historical database of risk factors, (3. 1) solves the set of risk factors corresponding to the historical date. If prices (or quotes) for the underlying instrument are observed for all relevant historical dates there is no need for a proxy methodology.

However, due to changing appetite for instruments or periods of illiquidity in which insufficient market prices are observed that qualify as observable market data, not all risk factors for a historical date can be inferred from the prices and the historical database features missing risk factor values. To alleviate this issue, we propose a generic proxy methodology that incorporates the banks' in-house pricing functionalities."

---

[3] Although the approach can be straightforwardly extended to include for instance heterogeneous strike levels. To maintain a clear outline we safe these extensions for Section 3.3 on examples.

[4] In some cases, *m* < *n*, this system is not solved as such, but used to calibrate a limited number of parameter to infer an underlying relation between the risk factors, e.g. in the case of implied volatility smile models.

In the case of our call options, the pricing function F_t is defined by the Black and Scholes (BS) pricing function. The vector containing prices is denoted by:

$$\boldsymbol{F}(\boldsymbol{x}_t, \boldsymbol{\tau}_t) = (BS(\boldsymbol{x}_t, 1), BS(\boldsymbol{x}_t, 2), BS(\boldsymbol{x}_t, 3), BS(\boldsymbol{x}_t, 4))\,'.$$

The risk factors in $\boldsymbol{x}_t$ are: the price of the underlying, linked to delta (Δ); the underlying risk free rate, linked to rho (ρ) and the underlying Black and Scholes volatility σ_{BS}, linked to vega (v) of the underlying options.

"In addition to the valuation functionalities, the second item the method requires from a bank's pricing system are the Delta sensitivities with respect to each risk factor in x_t for all n instruments considered. We gather these Deltas in the following n-by-m matrix, where each row $\delta F(\boldsymbol{x}_t, \boldsymbol{\tau}_i)/\delta \boldsymbol{x}$ denotes all Delta sensitivities associated with instrument i and with tenor τ_i:

$$\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'} := \left(\frac{\delta F(\boldsymbol{x}_t,\ T_1 - t)}{\delta \boldsymbol{x}}, \frac{\delta F(\boldsymbol{x}_t,\ T_2 - t)}{\delta \boldsymbol{x}}, \dots, \frac{\delta F(\boldsymbol{x}_t,\ T_n - t)}{\delta \boldsymbol{x}}\right)'. \tag{3.2}$$

"

The matrix containing all risk sensitivities is given by (note that this matrix is not transposed):

$$\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}} := \begin{pmatrix} \Delta_1 & \rho_1 & v_1 \\ \Delta_2 & \rho_2 & v_2 \\ \Delta_3 & \rho_3 & v_3 \\ \Delta_4 & \rho_4 & v_4 \end{pmatrix}$$

"Third, we require from the in-house pricing system the Theta sensitivities for each instrument with respect to the remaining time-to-maturity and allocate these in a n dimensional vector, implicitly defined as:

$$diag\left\{\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{\tau}'}\right\} := \left(\frac{\delta F(x_t, T_1 - t)}{\delta \tau}, \frac{\delta F(x_t, T_2 - t)}{\delta \tau}, \dots, \frac{\delta F(x_t, T_n - t)}{\delta \tau}\right)'.$$

"

The diagonal of theta's (Θ) is described by:

$$diag\left\{\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{\tau}'}\right\} := (\Theta_1, \Theta_2, \Theta_3, \Theta_4)'$$

"The next step is to use (3. 2) to calculate an n-by-n projection matrix:

$$\boldsymbol{M}(\boldsymbol{x}_t, \boldsymbol{\tau}_t) = I_n - \frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\left(\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)'\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)\right)^{-1}\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)',$$

where $\boldsymbol{I}_n$ denotes a n-by-n identity matrix. In subsequent notation we suppress the arguments for $\boldsymbol{M}_t$."

In $\boldsymbol{M}_t$, each row (or column due to symmetry) can be interpreted as a portfolio vector containing weights of each of the 4 call options, which together sum up to 1.

"With the three components, which merely comprise a valuation function and two first-order Greeks, we are now in a position to define the generic proxy methodology.

Definition 1. Generic proxy methodology – *In the event we observe at a given historical date* t *insufficient market prices to infer all risk factor values* x_t *we can derive a proxy value* $\hat{x}_t$ *from risk factors and a historical overnight risk-free rate* r *observed at* $t-1$ *with the following expression:*

$$\boldsymbol{M}_{t-1} * \left(F(\widehat{\boldsymbol{x}}_t, \boldsymbol{\tau}_t) - \boldsymbol{F}_{t-1} * (1 + r_{t-1}) + diag \left\{ \frac{\delta \boldsymbol{F}_{t-1}}{\delta \boldsymbol{\tau}'} \right\} \right) = 0. \tag{3.3}$$

With historical data observed at date $t-1$ *we can solve the linear system (3. 3) for the instruments' theoretical prices at time* t *and subsequently solve for the risk factor values* $\widehat{\boldsymbol{x}}_t$*, which then serve as the proxy values for the unidentified* $\boldsymbol{x}_t$ *at time* t*. A full derivation of (3. 3) can be found in Section A in the Appendix.*

From (3. 3), we can regard each row (or column) in $\boldsymbol{M}_t$ *in (A. 5) as a hedging portfolio with asset weights associated with the considered derivatives at time t. These portfolios are instantaneously unaffected by changes in the underlying risk factors* $\boldsymbol{x}_t$*."*

Using our "hedging portfolio" $\boldsymbol{M}_{t-1}$ we are able to reduce the difference between $\boldsymbol{F}(\widehat{\boldsymbol{x}}_t, \boldsymbol{\tau}_t)$ and $\boldsymbol{F}_{t-1}$ to the risk free rate r and the time effect Θ.

The theorem described in definition 1 describes that the return of a vector of instruments, corrected for the return of the risk free rate and theta, multiplied with the hedging portfolio M_{t-1} is equal to zero. Using this knowledge and by determining $\boldsymbol{M}_{t-1}, \boldsymbol{F}_{t-1}, r_{t-1}$ and $diag \left\{ \frac{\delta F_{t-1}}{\delta \tau'} \right\}$ we are able to derive a value for $F_t(\widehat{\boldsymbol{x}}_t, \boldsymbol{\tau}_t)$. Then using the valuation function F we can deduce a proxy value for all risk factors $\widehat{\boldsymbol{x}}_t$.

4. Nelson Siegel model

In this, and the following two chapters, we apply our working method to a set of models to determine the effect the working method has on the estimated parameters. These models are chosen given that they are relatively simple and widely used within the industry. We start with the Nelson Siegel model.

4.1. Introduction

Nelson and Siegel (Nelson & Siegel, 1987) introduced a yield curve model with the purpose to be simple, parsimonious and flexible enough to represent the range of shapes generally associated with yield curves; monotonic, humped and S shaped. This yield structure could be used to determine the values of bonds. The Nelson Siegel model is widely used by central banks (BIS, c2005) (Gurkaynak, Sack, & Wright, 2006).

In their paper, Nelson and Siegel found a high correlation between the present value of a long-term bond implied by the fitter curves and the actual reported prices of the bond. This confirms their assumption that their model reflects the basic shape of the term structure and not just a local approximation.

Their proposed function is described by the function:

$$y(m) = \beta_0 + \beta_1 \frac{1 - e^{-m/\tau}}{m/\tau} + \beta_2 \left(\frac{1 - e^{-m/\tau}}{m/\tau} - e^{-m/\tau}\right)$$

Where m denotes the maturity and $y(m)$ the yield of the curve at maturity m. $\beta_{0,}, \beta_1, \beta_2$ and τ are parameters that need to be fitted via a least-squares or similar statistical technique. β_0 is interpreted as the long run level of interest rates; β_1 is the short term component; β_2 is the medium term component and τ is the decay factor. The effect of both the short and medium term component β_1 respectively β_2 converge to zero over time, which leaves β_0 as the singular component to determine the yield on the long run.

A research performed by Diebold and Li (Diebold & Li, 2005), addresses a key-practical problem with studies performed so far. They make a novel twist of interpretation of the Nelson Siegel model and furthermore go in of out-of-sample of forecasting of yield curves with their model. They point out that β_0 can be interpreted as a level, $\beta_1$as the slope and β_2 as the curvature of the yield curve. Their research shows that these parameters can be interpreted as factors that may vary over time and further more shows that the models are consistent with a variety of stylized facts regarding the yield curve. They use their models to produce term-structure forecasts at both short and long horizons, with encouraging results.

4.2. Proxy

We examine the effect of our proxy condition in the case of the Nelson Siegel model. The parameters β_0, β_1, β_2 and λ, that are used in the Nelson Siegel model to define the yield curve, are estimated based on a statistical fit, like the Root Mean Squared Error (RMSE). First, β_0, β_1 and β_2 are estimated per time unit, using a given λ. The fit error at time t is minimized using the following two equations:

$$\boldsymbol{y}(\beta_t^0,\beta_t^1,\beta_t^2,\lambda,\boldsymbol{\tau}_t) = \beta_t^0 + \beta_t^1\frac{1-e^{-\lambda\boldsymbol{\tau}_t}}{\lambda\boldsymbol{\tau}_t} + \beta_t^2\left(\frac{1-e^{-\lambda\boldsymbol{\tau}_t}}{\lambda\boldsymbol{\tau}_t} - e^{-\lambda\boldsymbol{\tau}_t}\right),$$

$$fit\ error(\hat{\beta}_t^0,\hat{\beta}_t^1,\hat{\beta}_t^2,\hat{\lambda},\boldsymbol{\tau_t}) = \sqrt{\sum_\tau\left(y(\tau_t) - y(\hat{\beta}_t^0,\hat{\beta}_t^1,\hat{\beta}_t^2,\hat{\lambda},\tau_t)\right)^2}.$$

Where $\boldsymbol{y}(\tau)$ presents the observed yield at time t for each of the maturities in $\boldsymbol{\tau}_t := (T_1 - t, T_2 - t, \dots, T_n - t)$. We name the β's that minimize the fit error at time t β^*. Next, λ^* is determined by altering λ in order to achieve the minimal final error[5]:

$$final\ error(\hat{\lambda},\boldsymbol{\tau}) = \sum_t fit\ error(\beta_t^{0*},\beta_t^{1*},\beta_t^{2*},\hat{\lambda},\tau_t).$$

In our example we look at yield data of Constant Maturity Treasuries (CMT) on a daily basis. Since an increment of a day does not lead to the shortening of the maturity with one day (e.g. we look at a bond with a maturity of 1 year and the next day we look at the next bond with a maturity of 1 year), in our case, maturity is defined as: $\boldsymbol{\tau}_t := (T_1, T_2, \dots, T_n)$. Hence, from now on we suppress t for $\boldsymbol{\tau}_t$.

The first component we require to implement our proxy is the valuation function. The value of a bond is determined by the price function of a bond:

$$B = e^{-\lambda T},$$

Where we replace the general yield λ with our Nelson-Siegel based yield:

$$\boldsymbol{F}(\beta_t^0,\beta_t^1,\beta_t^2,\lambda,\boldsymbol{\tau}) = \begin{pmatrix} e^{-T_1\left(\beta_t^0 + \beta_t^1\frac{(1-e^{-\lambda T_1})}{\lambda T_1} + \beta_t^2\left(\frac{(1-e^{-\lambda T_1})}{\lambda T_1} - e^{-\lambda T_1}\right)\right)} \\ e^{-T_2\left(\beta_t^0 + \beta_t^1\frac{(1-e^{-\lambda T_2})}{\lambda T_2} + \beta_t^2\left(\frac{(1-e^{-\lambda T_2})}{\lambda T_2} - e^{-\lambda T_2}\right)\right)} \\ \dots \\ e^{-T_n\left(\beta_t^0 + \beta_t^1\frac{(1-e^{-\lambda T_n})}{\lambda T_n} + \beta_t^2\left(\frac{(1-e^{-\lambda T_n})}{\lambda T_n} - e^{-\lambda T_n}\right)\right)} \end{pmatrix}.$$

This function determines the price of a bond given maturity T with the corresponding yield at time t, which follows from the Nelson Siegel equation. Next we introduce the delta sensitivities to each underlying of the three β's at time t (we suppress the arguments of $\boldsymbol{F}_t$ from now on):

[5] Note that for each try for lambda all $\beta's$ are determined again by minimizing the total error.

$$\frac{\delta \boldsymbol{F}}{\delta \beta_t^{0'}} := \begin{pmatrix} -T_1 e^{-T_1\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_1})}{\lambda T_1} + \beta_t^2\left(\frac{(1-e^{-\lambda T_1})}{\lambda T_1} - e^{-\lambda T_1}\right)\right)} \\ -T_2 e^{-T_2\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_2})}{\lambda T_2} + \beta_t^2\left(\frac{(1-e^{-\lambda T_2})}{\lambda T_2} - e^{-\lambda T_2}\right)\right)} \\ \dots \\ -T_n e^{-T_n\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_n})}{\lambda T_n} + \beta_t^2\left(\frac{(1-e^{-\lambda T_n})}{\lambda T_n} - e^{-\lambda T_n}\right)\right)} \end{pmatrix},$$

$$\frac{\delta \boldsymbol{F}}{\delta \beta_t^{1'}} := \begin{pmatrix} -T_1\left(\frac{(1-e^{-\lambda T_1})}{\lambda T_1}\right) e^{-T_1\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_1})}{\lambda T_1} + \beta_t^2\left(\frac{(1-e^{-\lambda T_1})}{\lambda T_1} - e^{-\lambda T_1}\right)\right)} \\ -T_2\left(\frac{(1-e^{-\lambda T_2})}{\lambda T_2}\right) e^{-T_2\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_2})}{\lambda T_2} + \beta_t^2\left(\frac{(1-e^{-\lambda T_2})}{\lambda T_2} - e^{-\lambda T_2}\right)\right)} \\ \dots \\ -T_n\left(\frac{(1-e^{-\lambda T_n})}{\lambda T_n}\right) e^{-T_n\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_n})}{\lambda T_n} + \beta_t^2\left(\frac{(1-e^{-\lambda T_n})}{\lambda T_n} - e^{-\lambda T_n}\right)\right)} \end{pmatrix},$$

$$\frac{\delta \boldsymbol{F}}{\delta \beta_t^{2'}} := \begin{pmatrix} -T_1\left(\frac{(1-e^{-\lambda T_1})}{\lambda T_1} - e^{-\lambda T_1}\right) e^{-T_1\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_1})}{\lambda T_1} + \beta_t^2\left(\frac{(1-e^{-\lambda T_1})}{\lambda T_1} - e^{-\lambda T_1}\right)\right)} \\ -T_2\left(\frac{(1-e^{-\lambda T_2})}{\lambda T_2} - e^{-\lambda T_2}\right) e^{-T_2\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_2})}{\lambda T_2} + \beta_t^2\left(\frac{(1-e^{-\lambda T_2})}{\lambda T_2} - e^{-\lambda T_2}\right)\right)} \\ \dots \\ -T_n\left(\frac{(1-e^{-\lambda T_n})}{\lambda T_n} - e^{-\lambda T_n}\right) e^{-T_n\left(\beta_t^0 + \beta_t^1 \frac{(1-e^{-\lambda T_n})}{\lambda T_n} + \beta_t^2\left(\frac{(1-e^{-\lambda T_n})}{\lambda T_n} - e^{-\lambda T_n}\right)\right)} \end{pmatrix}.$$

We combine these sensitivities into our Delta sensitivity matrix for a given time t:

$$\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'} := \left(\frac{\delta \boldsymbol{F}}{\delta \beta_t^{0'}}, \frac{\delta \boldsymbol{F}}{\delta \beta_t^{1'}}, \frac{\delta \boldsymbol{F}}{\delta \beta_t^{2'}}\right). \tag{4}$$

Since, in our dataset, there is no effect of time decay on our bond prices, we neglect the effect Theta would otherwise have. We are now able to form a projection matrix based on (4) and a n-by-n identity matrix $\boldsymbol{I}_n$:

$$\boldsymbol{M}(\boldsymbol{x}_t, \boldsymbol{\tau}) = \boldsymbol{I}_n - \frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\left(\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)'\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)\right)^{-1}\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)'$$

The *proxy error* at a given time t can now be calculated by using the following equation where $F(\mathbf{x}_{t-1}, \tau)$ represents the prices of the bonds with the observed betas from the day prior.

$$proxy\ error\left(\hat{\beta}_t^0, \hat{\beta}_t^1, \hat{\beta}_t^2, \hat{\lambda}, \boldsymbol{\tau}\right) = \boldsymbol{M}(\boldsymbol{x}_{t-1}, \boldsymbol{\tau}_t) * \left(\boldsymbol{F}_t(\widehat{\boldsymbol{x}}_t, \boldsymbol{\tau}) - \boldsymbol{F}_{t-1}(\boldsymbol{x}_{t-1}, \boldsymbol{\tau})\right)$$

Next the fit error at time t is calculated in the same way as it is calculated for the regular Nelson Siegel fit.

Nelson Siegel model

$$\boldsymbol{y}(\beta_t^0, \beta_t^1, \beta_t^2, \lambda, \boldsymbol{\tau}) = \beta_t^0 + \beta_t^1 \frac{1 - e^{-\lambda\tau}}{\lambda\boldsymbol{\tau}} + \beta_t^2 (\frac{1 - e^{-\lambda\tau}}{\lambda\boldsymbol{\tau}} - e^{-\lambda\tau})$$

$$fit\ error(\hat{\beta}_t^0, \hat{\beta}_t^1, \hat{\beta}_t^2, \hat{\lambda}, \boldsymbol{\tau}) = \sqrt{\sum_\tau \left(y(\tau_t) - y(\hat{\beta}_t^0, \hat{\beta}_t^1, \hat{\beta}_t^2, \hat{\lambda}, \tau)\right)^2}$$

The final error for time t can now be determined. We add a weight p to the $proxy\ error$ in order to asses to impact on the parameter estimation for different scenarios. The final error is minimized by changing β_0, β_1 and β_2.

$$final\ error(\hat{\beta}_t^0, \hat{\beta}_t^1, \hat{\beta}_t^2, \hat{\lambda}, \boldsymbol{\tau}) = \sum_\tau \left(fit\ error(\hat{\beta}_t^0, \hat{\beta}_t^1, \hat{\beta}_t^2, \hat{\lambda}, \tau) + p * proxy\ error(\hat{\beta}_t^0, \hat{\beta}_t^1, \hat{\beta}_t^2, \hat{\lambda}, \tau)\right).$$

When we have the optimal betas β_o^*, β_1^* and β_2^*, we can determine the final error, and we minimize it by altering lambda to achieve the optimal value (λ^*).

$$final\ error(\hat{\lambda}, \boldsymbol{\tau}) = \sum_t final\ error(\beta_t^{0*}, \beta_t^{1*}, \beta_t^{2*}, \hat{\lambda}, \boldsymbol{\tau})$$

Due to the fact our minimizing algorithm does not behave properly (e.g. it returned local minima), we introduce boundaries for the values that the β's can take. For β_t^0 which reflects the level (e.g. long term yield) of the yield, we take $(\beta_{t-1}^0 \pm 2 * (y_t^{long} - y_{t-1}^{long}))$, where y^{long} represents the longest available tenor. For β_t^1 which reflects the slope of the yield (e.g. long yield minus the short yield), we take $\left(\beta_{t-1}^0 \pm 2 * \left(\left(y_t^{long} - y_t^{short}\right) - \left(y_{t-1}^{long} - y_{t-1}^{short}\right)\right)\right)$, where y^{short} represents the shortest available tenor. Finally, for β_t^2 which represents the curvature of the curve, we take $(\beta_t^2 \pm 0.01)$.

4.3. Results

We introduce our proxy to a dataset containing US Constant Maturity Treasuries (CMT) quotes retrieved from the Federal reserve of Louisiana website. The dataset ranges from the first trading day of 2006 (January 3rd) to the last trading day of 2015 (December 30th), as to include a financial crisis, namely the credit crisis of 2007/2008. Tenors include the 1, 3 and 6 months, and the 1, 2, 3, 5, 7, 10, 20 and 30 years.

We fit a yield curve for each day, using the method we described above for the p factor 0, 10 and 100. In this model, choosing $p = 10$, makes the size of the fit and proxy error about equal. $p = 0$ shows the fit of only using the Nelson Siegel model, and $p = 100$ puts a heavy weight, with an approximate ratio 1:10, on the fit and proxy error respectively. We test for a dataset which contains the full spectrum of tenors, and sequentially remove the short (1, 3, 6 month), medium (2, 3, 5, 7 year) and long (10, 20, 30 year) tenors. This will give us an insight in the information the other part of the term structure contains about the removed part.

At first, we determine the mean residual difference between the yield, as presented by the fitted yield curve for that day, and the actual yield from our data set, to create insight in the overall level of our fitted yield curves. We also determine the Root Mean Squared Error (RMSE), to show how large the deviance is. We advise the reader to focus on the RMSE in fit as well as out of fit and the total RMSE. The lower the RMSE the better the fit of the model. The results can be seen in Table 1 through

SUM LONG	***0***	***5.5000 E -06***	***3.2000 E -06***	***4.7481 E – 03***	***4.9020 E -03***	***4.8971 E -03***
SUM OTHER	*-3.2897 e -03*	*-1.4089 e-03*	*-1.2270 e -03*	*1.3788 e -02*	*1.2622 e -02*	*1.3417 e-02*

Table 4 and Figure 1, where we show what happens when we remove the short, medium and long tenors. The development of the β's over time can be observed in appendix B.

What becomes evident in our results is that while the sum of the RMSE of each tenor for p = 10 is higher for the full data set, it is lower out-of-sample for all datasets where some tenors are missing, indicating that our proxy version can model yield curves that are better fitting to the actual data when we leave out some data. This encourages us to believe that when data cannot be observed in the market, out proxy is a meaningful addition to the regular Nelson Siegel estimation process.

We also observe that the total fit (sum RMSE) drastically reduces when tenors are removed from the dataset. The scenarios in which data is removed have a total RMSE which is about 50 to 100% higher. This is however expected since these are out of sample results.

It is particularly noticeable that the proxy reduces the RMSE the most, in comparison the regular Nelson Siegel, in the regions where yield data has been left out.

FULL DATA	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 10***	***P = 100***	***P = 0***	***P = 10***	***P = 100***
1M	*-4.5678 e-04*	*-4.9019 e-04*	*-4.8769 e-04*	*1.2128 e-03*	*1.2180 e-03*	*1.2189 e-03*
3M	*-1.6203 e-04*	*-1.7476 e-04*	*-1.7265 e-04*	*5.2235 e-04*	*5.3055 e-04*	*5.3967 e-04*
6M	*5.4720 e-04*	*5.5774 e-04*	*5.5928 e-04*	*9.7023 e-04*	*9.7463 e-04*	*9.9451 e-04*
1Y	*3.9840 e-04*	*4.3424 e-04*	*4.3480 e-04*	*1.1168 e-03*	*1.1244 e-03*	*1.1357 e-03*
2Y	*-7.0983 e-06*	*3.0426 e-05*	*2.9772 e-05*	*7.9862 e-04*	*7.9830 e-04*	*8.2505 e-04*
3Y	*-3.0276 e- 04*	*-2.8834 e-04*	*-2.8924 e-04*	*5.8561 e-04*	*5.7776 e-04*	*6.3341 e-04*
5Y	*-2.0876 e -05*	*-2.3801 e-04*	*-2.3741 e-04*	*7.5279 e-04*	*7.7038 e-04*	*7.8699 e-04*
7Y	*3.9699 e -05*	*-5.3295 e-06*	*-2.0889 e-06*	*9.8897 e-04*	*9.8970 e-04*	*1.0001 e-03*
10Y	*5.9369 e -05*	*2.3262 e-05*	*3.0429 e-05*	*5.7161 e-04*	*5.6770 e-04*	*6.1283 e-04*
20Y	*-3.2958 e -04*	*3.4887 e-04*	*3.6326 e-04*	*9.7683 e-04*	*1.0003 e-03*	*1.0858 e-03*
30Y	*-2.3952 e -04*	*-1.9464 e-04*	*-1.7722 e-04*	*9.7694 e-04*	*9.7627 e-04*	*1.0009 e-03*
SUM	*-2.7068e -06*	*3.2620 e-06*	*5.1233 e-05*	*9.4735 e-03*	*9.5280 e-03*	*9.8338 e-03*

Table 1 Statistics of the full data set

MISSING SHORT	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 10***	***P = 100***	***P = 0***	***P = 10***	***P = 100***
1M	*-2.0622 e -03*	*-1.7378 e -03*	*-1.5320 e -03*	*4.6523 e -03*	*4.4657 e -03*	*4.3773 e -03*
3M	*-1.4725 e -03*	*-1.2228 e -03*	*-1.0610 e -03*	*3.4287 e -03*	*3.2841 e -03*	*3.2133 e -03*
6M	*-3.9433 e -04*	*-2.3478 e -04*	*-1.2661 e -04*	*1.8653 e -03*	*1.8009 e -03*	*1.7935 e -03*
1Y	*-1.6267 e -05*	*2.4079 e -05*	*5.9790 e -05*	*4.4176 e -04*	*4.5741 e -04*	*5.1153 e -04*
2Y	*8.4439 e -05*	*3.1628 e -05*	*1.3688 e -05*	*5.9512 e -04*	*5.8167 e -04*	*5.8043 e -04*
3Y	*-8.0713 e -05*	*-1.2670 e -04*	*-1.5266 e -04*	*5.5593 e -04*	*5.6896 e -04*	*5.9289 e -04*
5Y	*-6.9272 e -05*	*-5.7716 e -05*	*-5.4078 e -05*	*3.4192 e -04*	*3.3907 e -04*	*3.5325 e -04*
7Y	*5.8092 e -05*	*1.0306 e -04*	*1.2651 e -04*	*6.1612 e -04*	*6.1589 e -04*	*6.1516 e -04*

10Y	*-2.3790 e -06*	*4.3879 e -05*	*7.0278 e -05*	*4.7532 e -04*	*4.9102 e -04*	*5.1177 e -04*
20Y	*2.8002 e -04*	*2.6951 e -04*	*2.6621 e -04*	*8.7022 e -04*	*8.7585 e -04*	*9.2256 e -04*
30Y	*-2.5713 e -04*	*-2.9500 e -04*	*-3.1328 e -04*	*9.0715 e -04*	*9.0919 e -04*	*9.5204 e -04*
SUM	*-3.9322 e -03*	*-3.1981 e-03*	*-2.7032 e -03*	*1.4750 e -02*	*1.4390 e -02*	*1.4424 e -02*
SUM SHORT	*-3.9290 e -03*	*-3.1953 e -03*	*-2.7196 e -03*	*9.9463 e -03*	*9.5507 e -03*	*9.3841 e -03*
SUM OTHER	*-3.2100 e -07*	*-7.2600 e -07*	*1.6460 e -06*	*4.8035 e -03*	*4.8391 e -03*	*5.0396 e -03*

Table 2 Statistics of the data set without short tenors

MISSING MEDIUM	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 10***	***P = 100***	***P = 0***	***P = 10***	***P = 100***
1M	*-2.7773 e-04*	*-2.3273 e-04*	*2.4587 e-04*	*7.7855 e-04*	*8.1150 e-04*	*1.3750 e-03*
3M	*-1.0392 e-04*	*-9.2451 e-05*	*8.5703 e-05*	*4.3074 e-04*	*4.3606 e-04*	*5.6032 e-04*
6M	*4.4305 e-04*	*4.1810 e-04*	*2.3861 e-04*	*7.7677 e-04*	*7.8250 e-04*	*8.4529 e-04*
1Y	*3.1160 e-05*	*-2.8980 e-05*	*-6.5679 e-04*	*5.7328 e-04*	*6.1508 e-04*	*1.3704 e-03*
2Y	*-7.0531 e-04*	*-7.5143 e-04*	*-1.5947 e-03*	*1.4826 e-03*	*1.4818 e-03*	*2.3100 e-03*
3Y	*-1.1499 e-03*	*-1.1428 e-03*	*-1.7701 e-03*	*2.1364 e-03*	*2.1004 e-03*	*2.7207 e-03*
5Y	*-1.0622 e-03*	*-9.6629 e-04*	*-9.5263 e-04*	*2.2453 e-03*	*2.1698 e-03*	*2.1764 e-03*
7Y	*-6.6703 e-04*	*-5.4212 e-04*	*-1.6113 e-04*	*2.0343 e-03*	*1.9839 e-03*	*1.6601 e-03*
10Y	*-3.9630 e-04*	*-2.9234 e-04*	*1.0074 e-04*	*1.0437 e-03*	*1.1064 e-03*	*7.1634 e-04*
20Y	*3.4716 e-04*	*3.3795 e-04*	*4.4277 e-05*	*9.2826 e-04*	*9.3364 e-04*	*7.8729 e-04*
30Y	*-4.4145 e-05*	*-1.0421 e-04*	*-5.7351 e-05*	*8.1950 e-04*	*8.0812 e-04*	*4.2828 e-04*
SUM	*-3.5851 e-03*	*-3.3973 e-03*	*-4.4775 e-03*	*1.3249 e-02*	*1.3139 e-02*	*1.4950 e-02*
SUM MEDIUM	*-3.5844 e -03*	*-3.4026 e -03*	*4.4786 e -03*	*7.8986 e -03*	*7.7359 e -03*	*8.8672 e -03*
SUM OTHER	*7.0000 e- 07*	*5.3000 e -06*	*1.1000 e -06*	*5.3508 e -03*	*5.4933 e -03*	*6.0829 e -03*

Table 3 Statistics of the data set without medium tenors

MISSING LONG	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 10***	***P = 100***	***P = 0***	***P = 10***	***P = 100***
1M	*-3.4785 e-04*	*-3.6539 e-04*	*-3.4986 e-04*	*9.1704 e-04*	*9.3096 e-04*	*9.4489 e-04*
3M	*-1.3514 e-04*	*-1.3528 e-04*	*-1.3429 e-04*	*4.3513 e-04*	*4.4491 e-04*	*4.4459 e-04*
6M	*4.8735 e-04*	*5.0426 e-04*	*4.9093 e-04*	*9.1548 e-04*	*9.2184 e-04*	*9.1963 e-04*
1Y	*2.6208 e-04*	*2.8953 e-04*	*2.6708 e-04*	*7.4762 e-04*	*7.4948 e-04*	*7.5720 e-04*
2Y	*-8.9261 e-05*	*-8.6851 e-05*	*-8.9274 e-05*	*3.3883 e-04*	*3.7920 e-04*	*3.7627 e-04*
3Y	*-2.5534 e-04*	*-2.8663 e-04*	*-2.6158 e-04*	*5.9716 e-04*	*6.1774 e-04*	*6.0154 e-04*
5Y	*-3.2401 e-05*	*-6.8677 e-05*	*-3.9194 e-05*	*4.0954 e-04*	*4.1756 e-04*	*4.2228 e-04*
7Y	*1.1056 e-04*	*1.5458 e-04*	*1.1934 e-04*	*3.8726 e-04*	*4.4034 e-04*	*4.3071 e-04*
10Y	*-2.3820 e-04*	*-5.9969 e-06*	*-1.9290 e-04*	*2.2263 e-03*	*2.1559 e-03*	*2.2034 e-03*
20Y	*-9.8671 e-04*	*-2.7286 e-04*	*-8.4623 e-04*	*5.2636 e-03*	*4.8418 e-03*	*5.1313 e-03*
30Y	*-2.0648 e-03*	*-1.1300 e-03*	*-1.8791 e-04*	*6.2987 e-03*	*5.6245 e-03*	*6.0823 e-03*
SUM	*-3.2897 e-03*	*-1.4033 e-03*	*-1.2239 e-03*	*1.8537 e-02*	*1.7524 e-02*	*1.8314 e-02*
SUM LONG	*0*	*5.5000 e -06*	*3.2000 e -06*	*4.7481 e – 03*	*4.9020 e -03*	*4.8971 e -03*
SUM OTHER	*-3.2897 e -03*	*-1.4089 e-03*	*-1.2270 e -03*	*1.3788 e -02*	*1.2622 e -02*	*1.3417 e -02*

Nelson Siegel model

Table 4 Statistics of the data set without long tenors

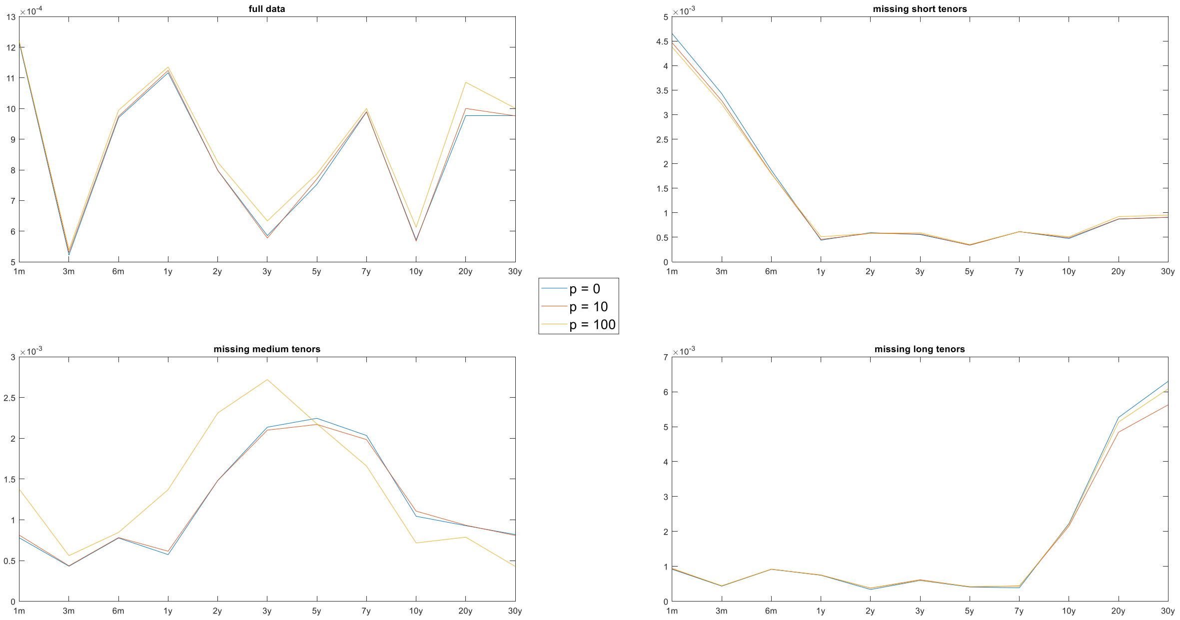

Figure 1 RMSE per tenor for all datasets

5. SABR model

Next we apply our working method to the Stochastic Alpha Beta Rho (SABR) model.

5.1. Introduction

The Black and Scholes (1973) and Black-76 (1975) model, developed to determine the value of an option in an arbitrage free or risk neutral world, are the industry standards for valuing vanilla options on equity and futures respectively. They can be extended to account for complex options like Bermudans and Asian options as well.

In the Black and Scholes model the value of a call option c_{BS} can be determined by the following function:

$$c_{BS} = S\,\mathcal{N}(d_1) - K\mathcal{N}(d_2)e^{-rt},$$

where

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma_{BS}^2}{2}\right) * (T - t)}{\sigma_{BS}\sqrt{T - t}},$$
$$d_2 = d_1 - \sigma_{BS}\sqrt{T - t},$$

and S is the spot price of an equity, $\mathcal{N}$ is the cumulative normal distribution function, K is the strike price of the option, r represents the risk-free rate, σ_{BS} is the Black and Scholes implied volatility of the underlying, t is the current time and T is the expiration date.

In Black-76's model the value of a call option c_{B-76} is represented by the following equations:

$$c_{B-76} = e^{-rt} * \left(F\mathcal{N}(d_1) - K\mathcal{N}(d_2)\right), \tag{5.1}$$

where,

$$d_1 = \frac{\ln\left(\frac{F}{K}\right) + \left(\frac{\sigma_{B-76}^2}{2}\right) * (T - t)}{\sigma_{B-76}\sqrt{T - t}},$$
$$d_2 = d_1 - \sigma_{B-76}\sqrt{T - t},$$

and F is the futures prices, σ_{B-76} is the Black-76 implied volatility and all others are the same as in the Black and Scholes model.

The Black and Scholes and Black-76 model assume a constant implied volatility σ over all strikes K. However when calibrated to market data, a volatility skew or smile is often observed. In Figure 2 we can see the implied volatility of a put option on the iShares Russel 2000 ETF versus the strike price. The star indicates the spot price.

The Stochastic Alpha Beta Rho (SABR) model, was developed by Hagan et al. (2002) in order to deal with the observed market skew in implied volatilities observed for different strikes. In a volatility skew, or volatility smile in jargon, one observes higher implied volatilities for In The Money (ITM) and Out of The Money (OTM) options as for At The money (ATM) options. Market skews started occurring in option pricing after the 1987 stock market crash. They were not present in the U.S. markets prior, indicating a market structure more in line with what the Black-Scholes model predicts. After 1987, traders realized that extreme events could happen, hence raising the prices (and implicitly the implied volatility) for ITM and OTM options.

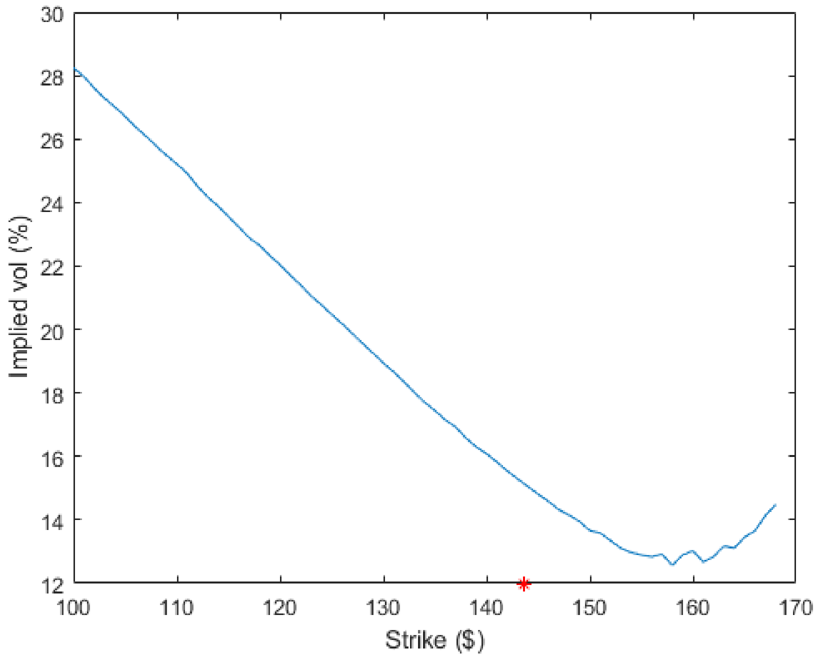

Figure 2 implied volatility vs. strike retrieved from a put (03/16/2018) on 09/21/2017 with iShares Russell 2000 ETF as underlying

The SABR model introduces stochastic volatility and lets the forward value satisfy:

$$\begin{aligned} \mathrm{d}\hat{F} &= \hat{\alpha}\hat{F}^{\beta}\mathrm{d}W_1, \qquad && \hat{F}(t=0) = f \\ \mathrm{d}\hat{\alpha} &= \nu\hat{\alpha}\mathrm{d}W_2, \qquad && \hat{\alpha}(t=0) = \alpha \end{aligned}$$

where t denotes time and in which the forward $\hat{F}$ and volatility $\hat{\alpha}$ are correlated by the following process:

$$\mathrm{d}W_1\mathrm{d}W_2 = \rho\mathrm{d}t.$$

A singular perturbation technique is used to obtain the prices of European options. The prices of the European options on a forward are given by (5. 1), where the implied volatility σ_{B-76} is given by the stochastic $\sigma_{B-76}(K,f)$ which can be derived from:

$$\begin{aligned} \sigma_{B-76}(K,f) = & \frac{\alpha}{(fK)^{\frac{1-\beta}{2}}\left\{1 + \frac{(1-\beta)^2}{24}\ln^2\left(\frac{f}{K}\right) + \frac{(1-\beta)^4}{1920}\ln^2\left(\frac{f}{K}\right) + \cdots\right\}} * \left(\frac{z}{x(z)}\right) \\ & * \left\{1 + \left[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4}\frac{\rho\beta\nu\alpha}{(fK)^{\frac{1-\beta}{2}}} + \frac{2-3\rho^2}{24}\nu^2\right](T-t) + \cdots,\right. \end{aligned}$$

where,

$$z = \frac{\nu}{\alpha}(fK)^{\frac{1-\beta}{2}} \ln\frac{f}{K},$$
$$x(z) = \ln\left\{\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right\}.$$

α, β, ρ and ν are factors that determine the shape of the volatility skew. All other variables are as in (5. 1). α can be interpreted as the level of the volatility. β, often referred to as the backbone, turns the curve from diagonal to horizontal. When $\beta = 0$ a stochastic Gaussian model is assumed, giving a steeply downward sloping backbone and when $\beta = 1$ is assumed the lognormal model is assumed. ν determines the steepness of the curve. ρ has the same effect on the curve as β does, hence it is common practice to choose a level of β and fitting ρ respectively. The effect of these Greeks can be seen in Figure 33.

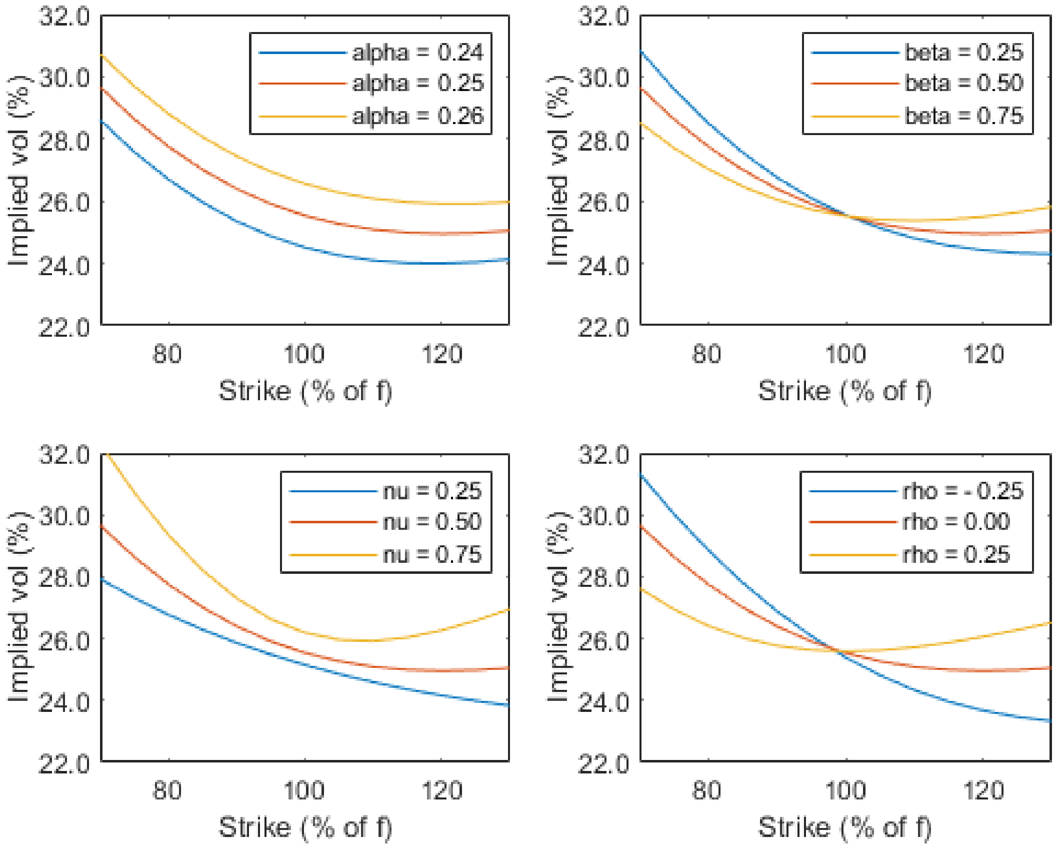

Figure 3 Effect of Greeks on volatility skew

5.2. Proxy

The SABR model parameters can be deduced by fitting the theoretical call prices, using the SABR volatility as input, to the observed market prices. First we choose a β that represents our view on the market. We assume lognormal shocks in the underlying prices, so we take $\beta = 1$.

We use the RMSE as a function for the fit error at time t that can be calculated using the following formula:

$$fit\ error(\hat{\alpha}_t, \beta, \hat{v}_t, \hat{\rho}_t, \tau_t, \boldsymbol{K}, f_t) = \sqrt{\sum_K \left(c(\tau_t, K, f_t) - c_{B-76}(\hat{\alpha}_t, \beta, \hat{v}_t, \hat{\rho}_t, \tau_t, K, f_t)\right)^2},$$

Where $c(\tau_t, K, f_t)$ denotes the observed market price of a call option with tenor τ, strike K and futures price f, all at time t.

If we minimize the fit error at time t by changing $\hat{\alpha}_t, \hat{v}_t$ and $\hat{\rho}_t$, we end up with the best fit represented by α_t^*, v_t^* and ρ_t^*.

Before we can introduce our proxy error, we first need to define our valuation function. This valuation function is equal to the Black-76 function:

$$\boldsymbol{F}(\alpha_t, \beta_t, v_t, \rho_t, \tau_t, \boldsymbol{K}, f_t) = \begin{pmatrix} c_{B-76}(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_1, f_t) \\ c_{B-76}(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_2, f_t) \\ \dots \\ c_{B-76}(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_n, f_t) \end{pmatrix}.$$

Next we determine the delta sensitivities of F_t to the underlying risk factors. Since a closed form solution of $\frac{\boldsymbol{F}}{\delta\,\alpha}, \frac{\boldsymbol{F}}{\delta\,v}$ and $\frac{\boldsymbol{F}}{\delta\,\boldsymbol{\rho}}$ is extremely complex, we derive an analytical alternative. We shock the risk factor of which we want to know its influence on the price by 1 basis point, subtract it from the regular price and multiply it by 10 thousand. In formula (we suppress the arguments of $\boldsymbol{F}_t$ from now on):

$$\frac{\delta\boldsymbol{F}}{\delta\,\alpha_t'} \approx \begin{pmatrix} (F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_1, f_t) - F(\alpha_t - 0.0001, \beta, v_t, \rho_t, \tau_t, K_1, f_t)) * 10000 \\ (F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_2, f_t) - F(\alpha_t - 0.0001, \beta, v_t, \rho_t, \tau_t, K_2, f_t)) * 10000 \\ \dots \\ (F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_n, f_t) - F(\alpha_t - 0.0001, \beta, v_t, \rho_t, \tau_t, K_n, f_t)) * 10000 \end{pmatrix}, \quad (5.2.1)$$

$$\frac{\delta\boldsymbol{F}}{\delta\,v_t'} \approx \begin{pmatrix} (F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_1, f_t) - F(\alpha_t, \beta, v_t - 0.0001, \rho_t, \tau_t, K_1, f_t)) * 10000 \\ (F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_2, f_t) - F(\alpha_t, \beta, v_t - 0.0001, \rho_t, \tau_t, K_2, f_t)) * 10000 \\ \dots \\ (F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_n, f_t) - F(\alpha_t, \beta, v_t - 0.0001, \rho_t, \tau_t, K_n, f_t)) * 10000 \end{pmatrix}, \quad (5.2.2)$$

$$\frac{\delta\boldsymbol{F}}{\delta\,\rho_t'} \approx \begin{pmatrix} \left(F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_1, f_t) - F(\alpha_t, \beta, v_t, \rho_t - 0.0001, \tau_t, K_1, f_t)\right) * 10000 \\ \left(F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_2, f_t) - F(\alpha_t, \beta, v_t, \rho_t - 0.0001, \tau_t, K_2, f_t)\right) * 10000 \\ \dots \\ \left(F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_n, f_t) - F(\alpha_t, \beta, v_t, \rho_t - 0.0001, \tau_t, K_n, f_t)\right) * 10000 \end{pmatrix}, \quad (5.2.3)$$

$$\frac{\delta\boldsymbol{F}}{\delta\,f_t'} \approx \begin{pmatrix} \left(F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_1, f_t) - F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_1, f_t - 0.0001)\right) * 10000 \\ \left(F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_2, f_t) - F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_2, f_t - 0.0001)\right) * 10000 \\ \dots \\ \left(F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_n, f_t) - F(\alpha_t, \beta, v_t, \rho_t, \tau_t, K_n, f_t - 0.0001)\right) * 10000 \end{pmatrix}. \quad (5.2.4)$$

Note that we are not interested in the effect of beta on the price since we keep this factor constant.

Using the risk sensitivities (5. 2. 1) through (5. 2. 4) in

$$\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'} := \left(\frac{\delta \boldsymbol{F}}{\delta \alpha_t'}, \frac{\delta \boldsymbol{F}}{\delta \nu_t'}, \frac{\delta \boldsymbol{F}}{\delta \rho_t'}, \frac{\delta \boldsymbol{F}}{\delta f_t'}\right),$$

enables us to form a projection matrix using an n-by-n identity matrix $\boldsymbol{I}_n$:

$$\boldsymbol{M}(\boldsymbol{x}_t, \boldsymbol{\tau}) = \boldsymbol{I}_n - \frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\left(\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)'\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)\right)^{-1}\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)'.$$

We define theta as:

$$\frac{\delta \boldsymbol{F}}{\delta \tau_t'} \approx \begin{pmatrix} \left(F(\alpha_t, \beta, \nu_t, \rho_t, \tau_t, K_1, f_t) - F(\alpha_t, \beta, \nu_t, \rho_t, \tau_t - 0.0001, K_1, f_t)\right) * 10000 \\ \left(F(\alpha_t, \beta, \nu_t, \rho_t, \tau_t, K_2, f_t) - F(\alpha_t, \beta, \nu_t, \rho_t, \tau_t - 0.0001, K_2, f_t)\right) * 10000 \\ \dots \\ \left(F(\alpha_t, \beta, \nu_t, \rho_t, \tau_t, K_n, f_t) - F(\alpha_t, \beta, \nu_t, \rho_t, \tau_t - 0.0001, K_n, f_t)\right) * 10000 \end{pmatrix}.$$

We introduce our proxy by adding an additional term to the fit error at time t, namely the proxy error. Which is given by:

$$\begin{aligned} &proxy\ error(\hat{\alpha}_t, \beta, \hat{\nu}_t, \hat{\rho}_t, \tau_t, \boldsymbol{K}, f_t) \\ &\quad = \boldsymbol{M}(\boldsymbol{x}_{t-1}, \boldsymbol{\tau}_t) \\ &\quad * \left(\boldsymbol{F}(\hat{\boldsymbol{x}}_t, \tau_t) - \boldsymbol{F}(\boldsymbol{x}_{t-1}, \tau_{t-1}) * (1+r)^{\tau_t - \tau_{t-1}} + diag\left\{\frac{\delta \boldsymbol{F}}{\delta \tau'}(\boldsymbol{x}_{t-1}, \tau_{t-1})\right\}\right). \end{aligned}$$

Note that for the SABR model, unlike for the Nelson Siegel model, we do account for the passing of time by use of the theta term and risk free return over the price at $t-1$.

Finally we determine the final error, by multiplying the proxy error at time t with a weight p and add the fit error at time t:

$$\begin{aligned} &final\ error(\hat{\alpha}_t, \beta, \hat{\nu}_t, \hat{\rho}_t, \tau_t, \boldsymbol{K}, f_t) \\ &\quad = \sum_K \left(fit\ error(\hat{\alpha}_t, \beta, \hat{\nu}_t, \hat{\rho}_t, \tau_t, \boldsymbol{K}, f_t) + p * proxy\ error(\hat{\alpha}_t, \beta, \hat{\nu}_t, \hat{\rho}_t, \tau_t, \boldsymbol{K}, f_t)\right). \end{aligned}$$

Minimizing the final error at each time t gives us the optimal values for α_t, ν_t and $\rho_t \sim \alpha_t^*, \nu_t^*$ and ρ_t^*.

5.3. Results

We introduce our proxy to a dataset containing data on S&P 500 options maturing at jan-2020, with strikes ranging from 2100 to 3300, ranging from 09-10-2018 to 13-03-2019, retrieved from a Bloomberg Terminal. In addition we use S&P 500 price data over the same time span, retrieved from Yahoo Finance. Finally, we use 1 week Libor (interest) rates, retrieved from Federal reserve of Louisiana website.

SABR model

We fit a SABR curve for each day, using the method we described above for the p factor 0, 100 and 1000. In this model, choosing $p = 100$, makes the size of the fit and proxy error about equal. $p = 0$ shows the fit of only using the SABR model, and $p = 1000$ puts a heavy weight, with an approximate ratio 1:10, on the fit and proxy error respectively.

At first, we determine the mean residual difference between the observed price, as presented by price based on the fitted SABR curve for that day, and the actual price from our data set, to create insight in the overall level of our fitted SABR curves. We also determine the Root Mean Squared Error (RMSE), to show how large the deviance is. The results can be seen in Table 5 through

SUM HIGH	***1.0297 E-02***	***1.0293 E - 02***	***1.0255 E - 02***	***2.0560 E - 02***	***2.0561 E - 02***	***2.0577 E - 02***
SUM OTHER	*-1.2343 e-02*	*-9.0316 e-.03*	*-9.4793 e -03*	*5.2399 e -03*	*5.2398 e -03*	*5.2429 e -03*

Table 8 and Figure 4, where we show what happens when we remove the low (2100-2450), medium (2475-2825) and high (2850-3300) strikes. The development of the α, β, ν and ρ over time can be observed in appendix C.

Our results show that our proxy only reduces the RMSE for the data set without the medium strikes. However, the reduction is only minimal. Therefore we cannot conclude that our proxy has material benefits, which contradicts our observations for the Nelson Siegel model. We propose three possible reasons for this:

- We tested our proxy on an equity option. Equity markets may show less of a functional relationship with parameters (in this case the α, β, ρ and ν) as interest rate markets. This may allow consistency between two consecutive periods to be minimal. Perhaps we might see distinct results when we apply it on an interest rate related option.
- Our dataset is quite small, which can lead to distorted results.
- Perhaps the term structure of the SABR model at a specific part does not contain information about the other parts of the volatility curve.

We also observe that the total fit (sum RMSE) drastically reduces when tenors are removed from the dataset. The scenarios in which data is removed have a total RMSE which is about 20 to 180% higher. Where 20% is linked to the scenario in which the medium strikes are removed, from which we can derive that these strikes have little influence on the shape of the SABR curve in comparison to the low and high strikes.

FULL DATA	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 100***	***P = 1000***	***P = 0***	***P = 100***	***P = 1000***
0.729	*-9.7652 e -05*	*-9.8315 e -05*	*-1.0448 e -04*	*3.0415 e -04*	*3.0398 e -04*	*3.0482 e -04*
0.737	*-1.1070 e -04*	*-1.1134 e -04*	*-1.1726 e -04*	*2.5564 e -04*	*2.5558 e -04*	*2.5781 e -04*
0.746	*-1.2911 e -04*	*-1.2972 e -04*	*-1.3538 e -04*	*2.4125 e -04*	*2.4123 e -04*	*2.4362 e -04*
0.755	*-1.3035 e -04*	*-1.3093 e -04*	*-1.3631 e -04*	*2.2156 e -04*	*2.2166 e -04*	*2.2550 e -04*
0.763	*-1.3385 e -04*	*-1.3440 e -04*	*-1.3947 e -04*	*2.1513 e -04*	*2.1552 e -04*	*2.2104 e -04*
0.772	*-1.2347 e -04*	*-1.2399 e -04*	*-1.2872 e -04*	*2.1636 e -04*	*2.1652 e -04*	*2.1960 e -04*
0.781	*-8.5628 e -05*	*-8.6102 e -05*	*-9.0478 e -05*	*2.1771 e -04*	*2.1801 e -04*	*2.2230 e -04*
0.789	*-4.3518 e -05*	*-4.3950 e -06*	*-4.7945 e -05*	*1.7150 e -04*	*1.7196 e -04*	*1.7771 e -04*
0.798	*-3.1937 e -05*	*-3.2326 e -05*	*-3.5916 e -05*	*1.7937 e -04*	*1.7961 e -04*	*1.8325 e -04*
0.807	*-4.1092 e -05*	*-4.1435 e -05*	*-4.4596 e -05*	*1.5465 e -04*	*1.5477 e -04*	*1.5729 e -04*
0.815	*-7.7359 e -06*	*-8.0301 e -06*	*-1.0740 e -06*	*1.7104 e -04*	*1.7130 e -04*	*1.7467 e -04*
0.824	*-3.8680 e -07*	*-4.1112 e -06*	*-6.3485 e -06*	*1.7109 e -04*	*1.7117 e -04*	*1.7267 e -04*
0.833	*9.6320 e -07*	*9.4421 e -06*	*7.6981 e -06*	*1.8853 e -04*	*1.8859 e -04*	*1.8948 e -04*
0.841	*3.0904 e -05*	*3.0770 e -06*	*2.9538 e -05*	*1.6935 e -04*	*1.6930 e -04*	*1.6906 e -04*
0.850	*3.8379 e -05*	*3.8302 e -05*	*3.7598 e -05*	*1.8194 e -04*	*1.8194 e -04*	*1.8229 e -04*
0.859	*7.5289 e -05*	*7.5271 e -05*	*7.5108 e -05*	*2.5811 e -04*	*2.5816 e -04*	*2.5889 e -04*
0.867	*8.4055 e -05*	*8.4096 e -05*	*8.4484 e -05*	*1.8840 e -04*	*1.8831 e -04*	*1.8775 e -04*
0.876	*1.0890 e -04*	*1.0900 e -04*	*1.0995 e -04*	*1.9857 e -04*	*1.9864 e -04*	*1.9922 e -04*
0.885	*1.2675 e -04*	*1.2691 e -04*	*1.2841 e -04*	*2.2340 e -04*	*2.2304 e -04*	*2.2347 e -04*
0.893	*1.3084 e -04*	*1.3106 e -04*	*1.3312 e -04*	*2.2701 e -04*	*2.2731 e -04*	*2.2982 e -04*
0.902	*1.675 e -04*	*1.6778 e -04*	*1.7038 e -04*	*3.1191 e -04*	*3.1246 e -04*	*3.1763 e -04*
0.911	*1.4402 e -04*	*1.4436 e -04*	*1.4748 e -04*	*2.4309 e -04*	*2.4361 e -04*	*2.4856 e -04*
0.920	*1.5607 e -04*	*1.5646 e -04*	*1.6008 e -04*	*3.8381 e -04*	*3.8345 e -04*	*3.8186 e -04*
0.928	*1.1737 e -04*	*1.1781 e -04*	*1.2188 e -04*	*2.3659 e -04*	*2.3699 e -04*	*2.4169 e -04*
0.937	*1.5049 e -04*	*1.5098 e -04*	*1.5545 e -04*	*4.2752 e -04*	*4.2835 e -04*	*4.3490 e -04*
0.946	*7.4335 e -05*	*7.4860 e -06*	*7.9680 e -06*	*1.7969 e -04*	*1.8005 e -04*	*1.8589 e -04*
0.954	*5.4082 e -05*	*5.4638 e -05*	*5.9728 e -05*	*1.9943 e -04*	*1.9983 e -04*	*2.0617 e -04*
0.963	*-5.2179 e -06*	*-4.6416 e -06*	*0.6304 e -07*	*1.7273 e -04*	*1.7241 e -04*	*1.7392 e -04*
0.972	*-3.3312 e -05*	*-3.2725 e -05*	*-2.7374 e -05*	*3.2084 e -04*	*3.2106 e -04*	*3.2543 e -04*
0.980	*-5.8938 e -05*	*-5.8355 e -05*	*-5.3041 e -05*	*1.4317 e -04*	*1.4281 e -04*	*1.4440 e -04*
0.989	*-1.1915 e -04*	*-1.1858 e -04*	*-1.1344 e -04*	*2.0744 e -04*	*2.0749 e -04*	*2.1010 e -04*
0.998	*-1.3104 e -04*	*-1.3050 e -04*	*-1.2566 e -04*	*2.1753 e -04*	*2.1756 e -04*	*2.1987 e -04*
1.006	*-1.3560 e -04*	*-1.3511 e -04*	*-1.3073 e -04*	*2.4495 e -04*	*2.4452 e -04*	*2.4227 e -04*
1.015	*-1.6940 e -04*	*-1.6898 e -04*	*-1.6521 e -04*	*2.3024 e -04*	*2.2994 e -04*	*2.2875 e -04*
1.024	*-1.4937 e -04*	*-1.4902 e -04*	*-1.4601 e -04*	*2.4971 e -04*	*2.4960 e -04*	*2.4951 e -04*
1.032	*-1.6855 e -04*	*-1.6830 e -04*	*-1.6619 e -04*	*2.3457 e -04*	*2.3430 e -04*	*2.3281 e -04*
1.041	*-1.2109 e -04*	*-1.2095 e -04*	*-1.1987 e -04*	*2.2886 e -04*	*2.2867 e -04*	*2.2766 e -04*
1.050	*-1.1556 e -04*	*-1.1553 e -04*	*-1.1558 e -04*	*1.7558 e -04*	*1.7549 e -04*	*1.7547 e -04*
1.058	*-9.6572 e -05*	*-9.6676 e -05*	*-9.7901 e -05*	*1.9190 e -04*	*1.9230 e -04*	*1.9691 e -04*
1.067	*-3.9647 e -05*	*-3.9882 e -05*	*-4.2309 e -05*	*1.0564 e -04*	*1.0589 e -04*	*1.1133 e -04*
1.076	*1.1966 e -06*	*8.3315 e -07*	*-2.7719 e -06*	*1.0977 e -04*	*1.1022 e -04*	*1.1875 e -04*
1.084	*6.0428 e -05*	*5.9942 e -05*	*5.5227 e -05*	*1.2171 e -04*	*1.2191 e -04*	*1.2975 e -04*
1.093	*7.9630 e -06*	*7.9033 e -05*	*7.3312 e -05*	*1.4515 e -04*	*1.4542 e -04*	*1.5466 e -04*
1.102	*1.1979 e -04*	*1.1910 e -04*	*1.1251 e -04*	*1.6025 e -04*	*1.5978 e -04*	*1.6387 e -04*
1.111	*1.7351 e -04*	*1.7274 e -04*	*1.6542 e -04*	*2.5152 e -04*	*2.5112 e -04*	*2.5341 e -04*
1.119	*1.7090 e -04*	*1.7006 e -04*	*1.6219 e -04*	*2.4750 e -04*	*2.4703 e -04*	*2.4956 e -04*
1.128	*1.9552 e -04*	*1.9463 e -04*	*1.8636 e -04*	*2.7588 e -04*	*2.7529 e -04*	*2.7629 e -04*
1.145	*2.2210 e -04*	*2.2118 e -04*	*2.1252 e -04*	*3.2339 e -04*	*3.2252 e -04*	*3.2036 e -04*
SUM	*2.0935 e -04*	*2.0537 e -04*	*1.6504 e -04*	*1.0495 e -02*	*1.0497 e -02*	*1.0622 e -02*

Table 5 Statistics for full data set

MISSING LOW	***MEAN RESIDUALS***			***RMSE***		
	P = 0	*P = 100*	*P = 1000*	*P = 0*	*P = 100*	*P = 1000*
0.729	*-5.9401 e -04*	*-5.9430 e -04*	*-5.9696 e -04*	*9.3191 e -04*	*9.3259 e -04*	*9.3930 e -04*
0.737	*-5.9999 e -04*	*-6.0028 e -04*	*-6.0295 e -04*	*8.9003 e -04*	*8.9063 e -04*	*8.9668 e -04*
0.746	*-6.1010 e -04*	*-6.1039 e -04*	*-6.1306 e -04*	*8.6013 e -04*	*8.6070 e -04*	*8.6636 e -04*
0.755	*-6.0176 e -04*	*-6.0205 e -04*	*-6.0471 e -04*	*8.3718 e -04*	*8.3777 e -04*	*8.4357 e -04*
0.763	*-5.9437 e -04*	*-5.9466 e -04*	*-5.9731 e -04*	*8.0097 e -04*	*8.0141 e -04*	*8.0590 e -04*
0.772	*-5.7175 e -04*	*-5.7204 e -04*	*-5.7467 e -04*	*7.7139 e -04*	*7.7184 e -04*	*7.7633 e -04*
0.781	*-5.2030 e -04*	*-5.2058 e -04*	*-5.2319 e -04*	*7.2635 e -04*	*7.2670 e -04*	*7.3028 e -04*
0.789	*-4.6319 e -04*	*-4.6347 e -04*	*-4.6604 e -04*	*6.5036 e -04*	*6.5089 e -04*	*6.5605 e -04*
0.798	*-4.3522 e -04*	*-4.3550 e -04*	*-4.3802 e -04*	*6.1353 e -04*	*6.1403 e -04*	*6.1895 e -04*
0.807	*-4.2659 e -04*	*-4.2686 e -04*	*-4.2932 e -04*	*5.7382 e -04*	*5.7417 e -04*	*5.7769 e -04*
0.815	*-3.7408 e -04*	*-3.7434 e -04*	*-3.7673 e -04*	*5.0715 e -04*	*5.0755 e -04*	*5.1150 e -04*
0.824	*-3.4969 e -04*	*-3.4995 e -04*	*-3.5225 e -04*	*4.6832 e -04*	*4.6864 e -04*	*4.7179 e -04*
0.833	*-3.1437 e -04*	*-3.1461 e -04*	*-3.1681 e -04*	*4.2519 e -04*	*4.2563 e -04*	*4.2969 e -04*
0.841	*-2.7003 e -04*	*-2.7026 e -04*	*-2.7234 e -04*	*3.7123 e -04*	*3.7172 e -04*	*3.7615 e -04*
0.850	*-2.3831 e -04*	*-2.3853 e -04*	*-2.4047 e -04*	*3.3686 e -04*	*3.3728 e -04*	*3.4110 e -04*
0.859	*-1.7601 e -04*	*-1.7629 e -04*	*-1.7809 e -04*	*2.9485 e -04*	*2.9512 e -04*	*2.9757 e -04*
0.867	*-1.4108 e -04*	*-1.4126 e -04*	*-1.4289 e -04*	*2.1392 e -04*	*2.1418 e -04*	*2.1653 e -04*
0.876	*-8.9203 e -05*	*-8.9362 e -05*	*-9.0804 e -05*	*1.4249 e -04*	*1.4266 e -04*	*1.4431 e -04*
0.885	*-4.3715 e -05*	*-4.3852 e -05*	*-4.5091 e -05*	*1.1975 e -04*	*1.1987 e -04*	*1.2109 e -04*
0.893	*-1.1585 e -05*	*-1.1698 e -05*	*-1.2721 e -05*	*1.0199 e -04*	*1.0194 e -04*	*1.0171 e -04*
0.902	*5.3258 e -05*	*5.3170 e -06*	*5.2376 e -05*	*2.2457 e -04*	*2.2455 e -04*	*2.2445 e -04*
0.911	*5.7833 e -05*	*5.7772 e -05*	*5.7218 e -05*	*1.3383 e -04*	*1.3380 e -04*	*1.3362 e -04*
0.920	*9.7496 e -05*	*9.7462 e -05*	*9.7154 e -05*	*3.3559 e -04*	*3.3561 e -04*	*3.3585 e -04*
0.928	*8.5619 e -05*	*8.5613 e -05*	*8.5554 e -05*	*1.8443 e -04*	*1.8442 e -04*	*1.8438 e -04*
0.937	*1.4439 e -04*	*1.4441 e -04*	*1.4460 e -04*	*4.1646 e -04*	*4.1655 e -04*	*4.1729 e -04*
0.946	*9.2296 e -05*	*9.2343 e -05*	*9.2769 e -05*	*1.5916 e -04*	*1.5924 e -04*	*1.6005 e -04*
0.954	*9.4085 e -05*	*9.4157 e -05*	*9.4809 e -05*	*2.0241 e -04*	*2.0253 e -04*	*2.0371 e -04*
0.963	*5.4351 e -05*	*5.4446 e -05*	*5.5302 e -05*	*1.6166 e -04*	*1.6165 e -04*	*1.6176 e -04*
0.972	*4.2886 e -05*	*4.3000 e-05*	*4.4035 e -05*	*2.9801 e -04*	*2.9781 e -04*	*2.9597 e -04*
0.980	*3.0500 e-05*	*3.0631 e -05*	*3.1812 e -05*	*9.4310 e -05*	*9.4164 e -05*	*9.3226 e -05*
0.989	*-2.0280 e -06*	*-2.0137 e -05*	*-1.8847 e -05*	*1.2652 e -04*	*1.2650 e -04*	*1.2632 e -04*
0.998	*-2.6904 e -05*	*-2.6753 e -05*	*-2.5397 e -05*	*1.2909 e -04*	*1.2907 e -04*	*1.2898 e -04*
1.006	*-3.0648 e -05*	*-3.0496 e -05*	*-2.9118 e -05*	*1.4958 e -04*	*1.4957 e -04*	*1.4964 e -04*
1.015	*-6.8237 e -05*	*-6.8087 e -05*	*-6.6734 e -05*	*1.0335 e -04*	*1.0328 e -04*	*1.0297 e -04*
1.024	*-5.6564 e -05*	*-5.6422 e -05*	*-5.5140 e -06*	*1.6132 e -04*	*1.6143 e -04*	*1.6255 e -04*
1.032	*-8.8477 e -05*	*-8.8347 e -05*	*-8.7179 e -05*	*1.2849 e -04*	*1.2846 e -04*	*1.2841 e -04*
1.041	*-5.7665 e -05*	*-5.7553 e -05*	*-5.6535 e -05*	*1.6157 e -04*	*1.6160 e -04*	*1.6204 e -04*
1.050	*-7.2008 e -05*	*-7.1916 e -05*	*-7.1080 e -06*	*1.0913 e -04*	*1.0924 e -04*	*1.1050 e -04*
1.058	*-7.5207 e -05*	*-7.5138 e -05*	*-7.4505 e -05*	*1.6015 e -04*	*1.5970 e -04*	*1.5563 e -04*
1.067	*-4.1679 e -05*	*-4.1634 e -05*	*-4.1216 e -05*	*8.5311 e -05*	*8.5594 e -05*	*8.8572 e -05*
1.076	*-2.4266 e -05*	*-2.4246 e -05*	*-2.4044 e -05*	*9.5739 e -05*	*9.6128 e -05*	*1.0001 e -04*
1.084	*1.2642 e -05*	*1.2639 e -05*	*1.2633 e -05*	*9.7277 e -05*	*9.7644 e -05*	*1.0144 e -04*
1.093	*1.1609 e -05*	*1.1583 e -05*	*1.1383 e -05*	*8.8284 e -05*	*8.7795 e -05*	*8.3803 e -05*
1.102	*3.434 e -06*	*3.4295 e -05*	*3.3922 e -05*	*7.8774 e -05*	*7.9062 e -05*	*8.2608 e -05*
1.111	*7.3839 e -05*	*7.3777 e -05*	*7.3257 e -05*	*1.5624 e -04*	*1.5643 e -04*	*1.5866 e -04*
1.119	*6.0324 e -05*	*6.0248 e -05*	*5.9606 e -05*	*1.2573 e -04*	*1.2507 e -04*	*1.1968 e -04*
1.128	*7.7207 e -05*	*7.7120 e -06*	*7.6383 e -05*	*1.5170 e -04*	*1.5123 e -04*	*1.4750 e -04*
1.145	*9.6488 e -05*	*9.6388 e -05*	*9.5529 e -05*	*1.7461 e -04*	*1.7434 e -04*	*1.7252 e -04*
SUM	*-6.8682 e -03*	*-6.872 e -03*	*-6.9059 e -03*	*1.5131 e -02*	*1.5138 e -02*	*1.5215 e -02*
SUM LOW	*-7.1399 e-03*	*-7.1441 e -03*	*-7.1829 e -03*	*1.0059 e -02*	*1.0067 e -02*	*1.0139 e-02*
SUM OTHER	*2.5899 e-04*	*1.5189 e -04*	*3.9064 e -04*	*4.9866 e -03*	*5.0711 e -03*	*5.0758 e -03*

Table 6 Statistics for data set without low strikes

MISSING MEDIUM	***MEAN RESIDUALS***			***RMSE***		
	P = 0	*P = 100*	*P = 1000*	*P = 0*	*P = 100*	*P = 1000*
0.729	*-8.1864 e -05*	*-8.2386 e -05*	*-8.6722 e -05*	*2.8356 e -04*	*2.8325 e -04*	*2.8262 e -04*
0.737	*-8.9858 e -05*	*-9.0328 e -05*	*-9.4235 e -05*	*2.2812 e -04*	*2.2820 e -04*	*2.3070 e -04*
0.746	*-1.0295 e -04*	*-1.0337 e -04*	*-1.0682 e -04*	*2.1241 e -04*	*2.1237 e -04*	*2.1355 e -04*
0.755	*-9.8604 e -05*	*-9.8958 e -05*	*-1.0193 e -04*	*1.8071 e -04*	*1.8072 e -04*	*1.8213 e -04*
0.763	*-9.6248 e -05*	*-9.6540 e -06*	*-9.9011 e -05*	*1.7230 e -04*	*1.7245 e -04*	*1.7455 e -04*
0.772	*-7.9747 e -05*	*-7.9974 e -05*	*-8.1921 e -05*	*1.7615 e -04*	*1.7591 e -04*	*1.7434 e -04*
0.781	*-3.5518 e -05*	*-3.5677 e -05*	*-3.7080 e -06*	*1.8379 e -04*	*1.8386 e -04*	*1.8464 e -04*
0.789	*1.3228 e -05*	*1.3138 e -05*	*1.2296 e -05*	*1.3925 e -04*	*1.3928 e -04*	*1.3958 e -04*
0.798	*3.1684 e -05*	*3.1664 e -05*	*3.1397 e -05*	*1.5918 e -04*	*1.5919 e -04*	*1.5928 e -04*
0.807	*2.9628 e -05*	*2.9680 e -06*	*2.9998 e -05*	*1.4379 e -04*	*1.4378 e -04*	*1.4389 e -04*
0.815	*7.0285 e -05*	*7.0409 e -05*	*7.1319 e -05*	*1.9217 e -04*	*1.9210 e -04*	*1.9180 e -04*
0.824	*8.1631 e -05*	*8.1828 e -05*	*8.3331 e -05*	*2.0327 e -04*	*2.0326 e -04*	*2.0389 e -04*
0.833	*1.0275 e -04*	*1.0302 e -04*	*1.0511 e -04*	*2.3201 e -04*	*2.3196 e -04*	*2.3253 e -04*
0.841	*1.3175 e -04*	*1.3209 e -04*	*1.3477 e -04*	*2.4498 e -04*	*2.4473 e -04*	*2.4383 e -04*
0.850	*1.4701 e -04*	*1.4742 e -04*	*1.5066 e -04*	*2.6562 e -04*	*2.6552 e -04*	*2.6645 e -04*
0.859	*1.9170 e -04*	*1.9217 e -04*	*1.9596 e -04*	*3.5600 e -04*	*3.5596 e -04*	*3.5731 e -04*
0.867	*2.0817 e -04*	*2.0871 e -04*	*2.1301 e -04*	*3.3069 e -04*	*3.3049 e -04*	*3.3097 e -04*
0.876	*2.4058 e -04*	*2.4117 e -04*	*2.4595 e -04*	*3.6785 e -04*	*3.6760 e -04*	*3.6774 e -04*
0.885	*2.6574 e -04*	*2.6638 e -04*	*2.7160 e -04*	*4.0631 e -04*	*4.0596 e -04*	*4.0543 e -04*
0.893	*2.7680 e -04*	*2.7749 e -04*	*2.8310 e -04*	*4.2809 e -04*	*4.2778 e -04*	*4.2778 e -04*
0.902	*3.1995 e -04*	*3.2068 e -04*	*3.2661 e -04*	*5.0216 e -04*	*5.0180 e -04*	*5.0129 e -04*
0.911	*3.0234 e -04*	*3.0311 e -04*	*3.0931 e -04*	*4.5940 e -04*	*4.5938 e -04*	*4.6191 e -04*
0.920	*3.1951 e -04*	*3.2029 e -04*	*3.2669 e -04*	*5.6113 e -04*	*5.6089 e -04*	*5.6105 e -04*
0.928	*2.8498 e -04*	*2.8578 e -04*	*2.9230 e -04*	*4.5024 e -04*	*4.4999 e -04*	*4.5092 e -04*
0.937	*3.2116 e -04*	*3.2196 e -04*	*3.2852 e -04*	*5.9683 e -04*	*5.9682 e -04*	*5.9891 e -04*
0.946	*2.4674 e -04*	*2.4753 e -04*	*2.5404 e -04*	*4.0314 e -04*	*4.0292 e -04*	*4.0426 e -04*
0.954	*2.2670 e -04*	*2.2747 e -04*	*2.3386 e -04*	*3.9669 e -04*	*3.9631 e -04*	*3.9635 e -04*
0.963	*1.6587 e -04*	*1.6662 e -04*	*1.7279 e -04*	*3.3966 e -04*	*3.3917 e -04*	*3.3839 e -04*
0.972	*1.3433 e -04*	*1.3504 e -04*	*1.4091 e -04*	*3.6990 e -04*	*3.6997 e -04*	*3.7294 e -04*
0.980	*1.0313 e -04*	*1.0379 e -04*	*1.0929 e -04*	*2.3073 e -04*	*2.3079 e -04*	*2.3448 e -04*
0.989	*3.5073 e -05*	*3.5681 e -05*	*4.0724 e -05*	*2.0856 e -04*	*2.0803 e -04*	*2.0679 e -04*
0.998	*1.2969 e -05*	*1.3514 e -05*	*1.8037 e -05*	*1.5338 e -04*	*1.5409 e -04*	*1.6284 e -04*
1.006	*-4.2174 e -06*	*-3.7441 e -06*	*2.0056 e -07*	*1.7058 e -04*	*1.7079 e -04*	*1.7414 e -04*
1.015	*-5.3003 e -05*	*-5.2606 e -05*	*-4.9289 e -05*	*1.1307 e -04*	*1.1337 e -04*	*1.1775 e -04*
1.024	*-5.0124 e -05*	*-4.9808 e -05*	*-4.7156 e -05*	*1.5195 e -04*	*1.5241 e -04*	*1.5710 e -04*
1.032	*-8.8338 e -05*	*-8.8105 e -05*	*-8.6142 e -05*	*1.2839 e -04*	*1.2835 e -04*	*1.2863 e -04*
1.041	*-6.1329 e -05*	*-6.1181 e -05*	*-5.9918 e -05*	*1.689 e -04*	*1.6888 e -04*	*1.6893 e -04*
1.050	*-7.7081 e -05*	*-7.7015 e -05*	*-7.6450 e -06*	*1.1942 e -04*	*1.1942 e -04*	*1.1946 e -04*
1.058	*-7.9510 e -06*	*-7.9525 e -05*	*-7.9639 e -05*	*1.6963 e -04*	*1.6986 e -04*	*1.7161 e -04*
1.067	*-4.3364 e -05*	*-4.3457 e -05*	*-4.4218 e -05*	*9.6145 e -05*	*9.5870 e -05*	*9.4105 e -05*
1.076	*-2.1911 e -05*	*-2.2075 e -05*	*-2.3439 e -05*	*1.0827 e -04*	*1.0797 e -04*	*1.0637 e -04*
1.084	*1.9977 e -05*	*1.9748 e -05*	*1.7838 e -05*	*1.0881 e -04*	*1.0825 e -04*	*1.0496 e -04*
1.093	*2.4370 e -06*	*2.4085 e -05*	*2.1693 e -05*	*1.0514 e -04*	*1.0598 e -04*	*1.1354 e -04*
1.102	*5.2533 e -05*	*5.2199 e -05*	*4.9395 e -05*	*9.3633 e -05*	*9.2631 e -05*	*8.7059 e -05*
1.111	*9.7110 e -06*	*9.6736 e -05*	*9.3594 e -05*	*1.8218 e -04*	*1.8169 e -04*	*1.7943 e -04*
1.119	*8.8074 e -05*	*8.7669 e -05*	*8.4258 e -05*	*1.4975 e -04*	*1.5073 e -04*	*1.5965 e -04*
1.128	*1.0870 e -04*	*1.0827 e -04*	*1.0466 e -04*	*1.7010 e -04*	*1.7081 e -04*	*1.7782 e -04*
1.145	*1.3317 e -04*	*1.3271 e -04*	*1.2888 e -04*	*2.1552 e -04*	*2.1578 e -04*	*2.1921 e -04*
SUM	*3.7240 e -03*	*3.7333 e -03*	*3.8081 e -03*	*1.1830 e -02*	*1.1827 e -02*	*1.1883 e -02*
SUM MEDIUM	*3.4640 e -03*	*3.4752 e -03*	*3.5667 e -03*	*6.2048 e -03*	*6.2020 e -03*	*6.2221 e -03*
SUM OTHER	*2.2216 e -04*	*3.1826 e -04*	*3.4357 e -04*	*5.6248 e -03*	*5.5390 e -03*	*5.6609 e -03*

Table 7 Statistics of the data set without medium strikes

SABR model

MISSING HIGH	***MEAN RESIDUALS***			***RMSE***		
	P = 0	*P = 100*	*P = 1000*	*P = 0*	*P = 100*	*P = 1000*
0.729	*7,4022 e -05*	*7,3897 e -05*	*7,2694 e -05*	*1.9585 e-04*	*1.9576 e-04*	*1.9525 e-04*
0.737	*4,6508 e -05*	*4,6397 e -05*	*4,5334 e -05*	*1.3903 e-04*	*1.3911 e-04*	*1.4009 e-04*
0.746	*1,2818 e -05*	*1,2721 e -05*	*1,1803 e -05*	*1.4203 e-04*	*1.4206 e-04*	*1.4259 e-04*
0.755	*-4,4557e -06*	*-4,5364 e -06*	*-5,3072 e -06*	*1.0145 e-04*	*1.0148 e-04*	*1.0196 e-04*
0.763	*-2,4689 e -05*	*-2,4754 e -05*	*-2,5375 e -05*	*1.2609 e-04*	*1.2619 e-04*	*1.2719 e-04*
0.772	*-3,1665 e -05*	*-3,1714 e -05*	*-3,2183 e -05*	*1.5022 e-04*	*1.5020 e-04*	*1.5002 e-04*
0.781	*-1,1703 e -05*	*-1,1737 e -05*	*-1,2056 e -05*	*1.7564 e-04*	*1.7565 e-04*	*1.7579 e-04*
0.789	*1,2099 e -05*	*1,2081 e -05*	*1,1911 e -05*	*1.4039 e-04*	*1.4039 e-04*	*1.4038 e-04*
0.798	*5,0743 e -06*	*5,0710 e -03*	*5,0461 e -06*	*1.5516 e-04*	*1.5516 e-04*	*1.5521 e-04*
0.807	*-2,2835 e -05*	*-2,2824 e -05*	*-2,2710 e -06*	*1.3043 e-04*	*1.3041 e-04*	*1.3022 e-04*
0.815	*-8,2141 e -06*	*-8,1892 e -06*	*-7,9432 e -06*	*1.4691 e-04*	*1.4691 e-04*	*1.4692 e-04*
0.824	*-2,2867 e -05*	*-2,2829 e -05*	*-2,2462 e -05*	*1.4472 e-04*	*1.4473 e-04*	*1.4485 e-04*
0.833	*-2,7455 e -05*	*-2,7406 e -05*	*-2,6928 e -05*	*1.5840 e-04*	*1.5850 e-04*	*1.5946 e-04*
0.841	*-2,3590 e -06*	*-2,3531 e -05*	*-2,2959 e -05*	*1.3407 e-04*	*1.3416 e-04*	*1.3504 e-04*
0.850	*-3,2568 e -05*	*-3.2500 e-05*	*-3,1850 e -06*	*1.4621 e-04*	*1.4634 e-04*	*1.4765 e-04*
0.859	*-1,0850 e -06*	*-1,0776 e -05*	*-1,0067 e -05*	*2.0597 e-04*	*2.0611 e-04*	*2.0741 e-04*
0.867	*-1,5680 e -06*	*-1,5603 e -05*	*-1,4855 e -05*	*1.1659 e-04*	*1.1662 e-04*	*1.1703 e-04*
0.876	*-2,4681 e -06*	*-2,3891 e -06*	*-1,6270 e -03*	*8,1196 e -05*	*8,1186 e -05*	*8,1340 e -06*
0.885	*6,1064 e -06*	*6,1845 e -06*	*6,9362 e -06*	*1.1671 e-04*	*1.1668 e-04*	*1.1658 e-04*
0.893	*3,7153 e -06*	*3,7897 e -06*	*4,5047 e -06*	*1.2079 e-04*	*1.2069 e-04*	*1.1994 e-04*
0.902	*3,7128 e -05*	*3,7196 e -05*	*3,7847 e -05*	*2.3264 e-04*	*2.3258 e-04*	*2.3206 e-04*
0.911	*1,4118 e -05*	*1,4177 e -05*	*1,4736 e -05*	*1.2945 e-04*	*1.2948 e-04*	*1.2986 e-04*
0.920	*3,0847 e -05*	*3,0893 e -05*	*3,1332 e -05*	*3.2967 e-04*	*3.2963 e-04*	*3.2933 e-04*
0.928	*1,5354 e -06*	*1,5664 e -06*	*1,8583 e -06*	*1.6861 e-04*	*1.6858 e-04*	*1.6832 e-04*
0.937	*4,9249 e -05*	*4,9262 e -05*	*4,9381 e -05*	*3.7272 e-04*	*3.7273 e-04*	*3.7288 e-04*
0.946	*-6,6271 e -06*	*-6,6342 e -06*	*-6,7122 e -06*	*1.0829 e-04*	*1.0831 e-04*	*1.0846 e-04*
0.954	*-4,7015 e-03*	*-4,9987 e-03*	*-7,9582 e -03*	*1.5467 e-04*	*1.5468 e-04*	*1.5487 e-04*
0.963	*-2,6850 e -06*	*-2,6904 e -05*	*-2,7436 e -05*	*1.3509 e-04*	*1.3515 e-04*	*1.3579 e-04*
0.972	*-1,5225 e -05*	*-1,5306 e -05*	*-1,6086 e -05*	*2.8327 e-04*	*2.8309 e-04*	*2.8146 e-04*
0.980	*5,8227 e -06*	*5,7156 e -06*	*4,6767 e -06*	*1.1278 e-04*	*1.1252 e-04*	*1.1056 e-04*
0.989	*-7,7702 e -03*	*-9,1149 e-03*	*-2,2124 e -06*	*1.3514 e-04*	*1.3531 e-04*	*1.3749 e-04*
0.998	*4,7642 e -05*	*4,7480 e -06*	*4,5919 e -05*	*2.4977 e-04*	*2.4945 e-04*	*2.4691 e-04*
1.006	*1.0954 e-04*	*1.0936 e-04*	*1.0754 e-04*	*3.1895 e-04*	*3.1877 e-04*	*3.1753 e-04*
1.015	*1.4747 e-04*	*1.4726 e-04*	*1.4520 e-04*	*3.9601 e-04*	*3.9588 e-04*	*3.9510 e-04*
1.024	*2.4322 e-04*	*2.4298 e-04*	*2.4071 e-04*	*5.5978 e-04*	*5.5959 e-04*	*5.5828 e-04*
1.032	*3.0195 e-04*	*3.0169 e-04*	*2.9922 e-04*	*6.5981 e-04*	*6.5975 e-04*	*6.5963 e-04*
1.041	*4.2727 e-04*	*4.2700 e-04*	*4.2437 e-04*	*8.3114 e-04*	*8.3109 e-04*	*8.3107 e-04*
1.050	*5.0786 e-04*	*5.0757 e-04*	*5.0482 e-04*	*9.8549 e-04*	*9.8544 e-04*	*9.8537 e-04*
1.058	*5.9594 e-04*	*5.9564 e-04*	*5.9281 e-04*	*1.1699 e-03*	*1.1701 e-03*	*1.1728 e-03*
1.067	*7.1257 e-04*	*7.1227 e-04*	*7.0942 e-04*	*1.3310 e-03*	*1.3310 e-03*	*1.3311 e-03*
1.076	*8.0037 e-04*	*8.0007 e-04*	*7.9724 e-04*	*1.4970 e-03*	*1.4970 e-03*	*1.4972 e-03*
1.084	*8.9081 e-04*	*8.9052 e-04*	*8.8777 e-04*	*1.6343 e-03*	*1.6343 e-03*	*1.6348 e-03*
1.093	*9.2333 e-04*	*9.2306 e-04*	*9.2044 e-04*	*1.7496 e-03*	*1.7498 e-03*	*1.7519 e-03*
1.102	*9.5804 e-04*	*9.5778 e-04*	*9.5534 e-04*	*1.8241 e-03*	*1.8243 e-03*	*1.8257 e-03*
1.111	*9.8820 e-04*	*9.8796 e-04*	*9.8572 e-04*	*1.9115 e-03*	*1.9117 e-03*	*1.9135 e-03*
1.119	*9.4622 e-04*	*9.4601 e-04*	*9.4398 e-04*	*1.9102 e-03*	*1.9106 e-03*	*1.9140 e-03*
1.128	*9.1940 e-04*	*9.1921 e-04*	*9.1739 e-04*	*1.8991 e-03*	*1.8994 e-03*	*1.9030 e-03*
1.145	*8.2516 e-04*	*8.2502 e-04*	*8.2356 e-04*	*1.8824 e-03*	*1.8828 e-03*	*1.8863 e-03*
SUM	*-2.0461 e-03*	*1.2618 e-03*	*7.7627 e -04*	*2.5800 e-02*	*2.5801 e-02*	*2.5820 e-02*
SUM HIGH	*1.0297 e-02*	*1.0293 e -02*	*1.0255 e -02*	*2.0560 e -02*	*2.0561 e -02*	*2.0577 e -02*
SUM OTHER	*-1.2343 e-02*	*-9.0316 e-.03*	*-9.4793 e -03*	*5.2399 e -03*	*5.2398 e -03*	*5.2429 e -03*

Table 8 Statistics of the data set without high strikes

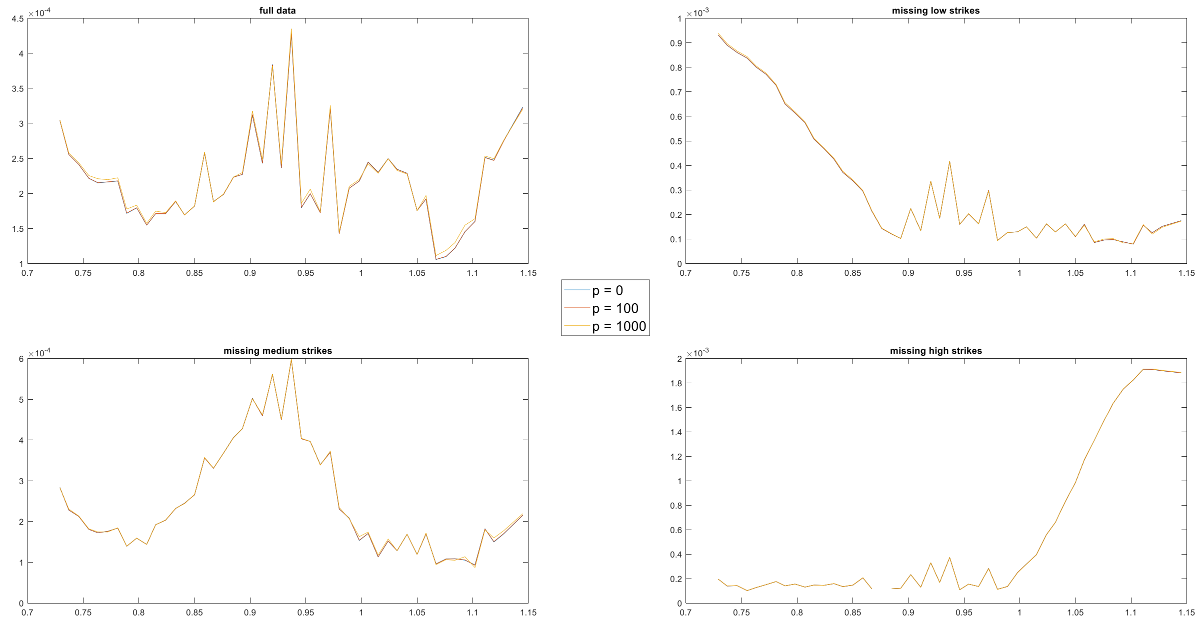

Figure 4 RMSE per strike for all datasets

6. Principal component model

The final model on which we apply our theorem is the Principal Component model.

6.1. Introduction

In recent years, data science has played an increasingly bigger role in the financial world. One of the techniques that has gotten a lot of attention from financial institutions is the Principal Component Analysis (PCA). PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. Where the first component explains the most of the variance and each next component explains most of the remaining variance. The main advantage of PCA is that it does not require some predefined functional form.

PCA was invented by Karl Pearson (1901) and later independently named and developed by Harold Hotelling (1933). Besides its use in finance, it is also widely used in chemistry, geology, engineering and other fields where big data streams need to be analyzed.

We use a Singular Value Decomposition (SVD) variant of PCA. We define M as in m-by-n dimensional matrix whose entries come from the filed K, which is either the field of real numbers or the field of complex numbers. Then the singular value decomposition of M exists and is a factorization of the form

$$M = U\Sigma V',$$

Where

- U is an m-by-m-dimensional unitary matrix over K (if $K = \mathbb{R}$, unitary matrices are orthogonal matrices),
- Σ is a diagonal m-by-n-dimensional matrix with non-negative real numbers of the diagonal,
- V is an $n \; x \; n$ unitary matrix over K, and V' is the conjugate transpose V.

A visualization of PCA can be seen in Figure 5.

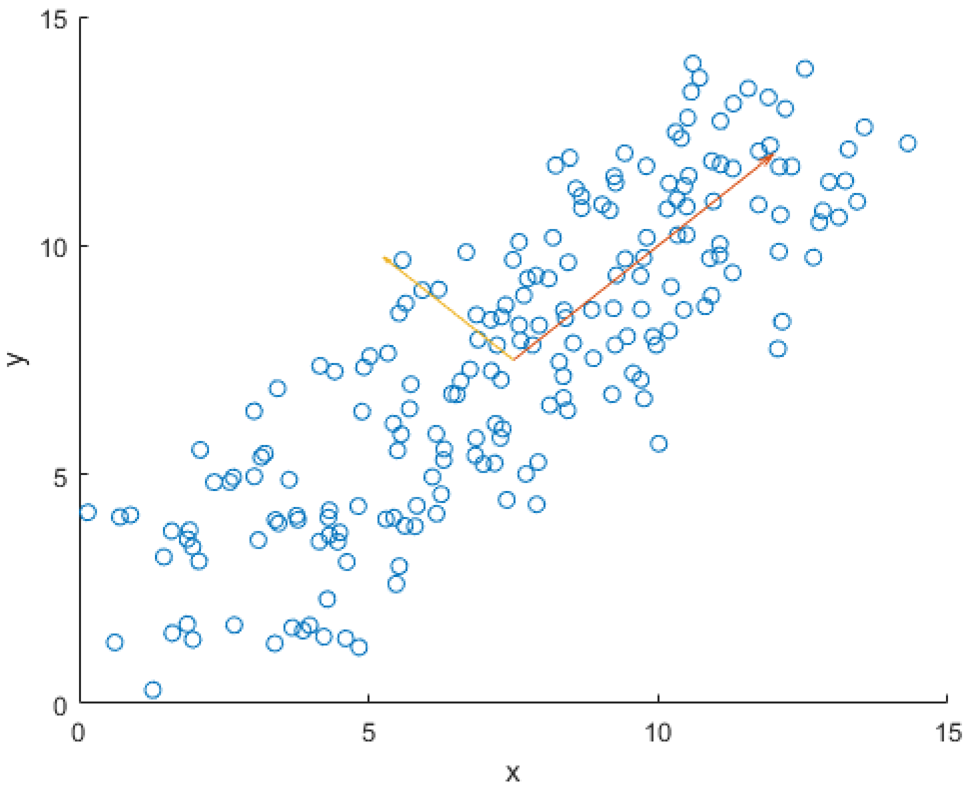

Figure 5 visualization of PCA

This model differs from the two examples we have treated before. Whereas the Nelson Siegel and SABR model are described by a functional form, the PCA is not bounded. The principal components can take any shape. However, when applying the PCA to yield data, one does observe a linked with the Nelson Siegel model. Where in the Nelson Siegel model β_0 can be interpreted as variable linked to the level, β_1 to the slope and β_2 to the curvature. When applying PCA, the first PC is also linked to level, the second also to slope and the third also to the curvature. We visualize this phenomenon in the figure below.

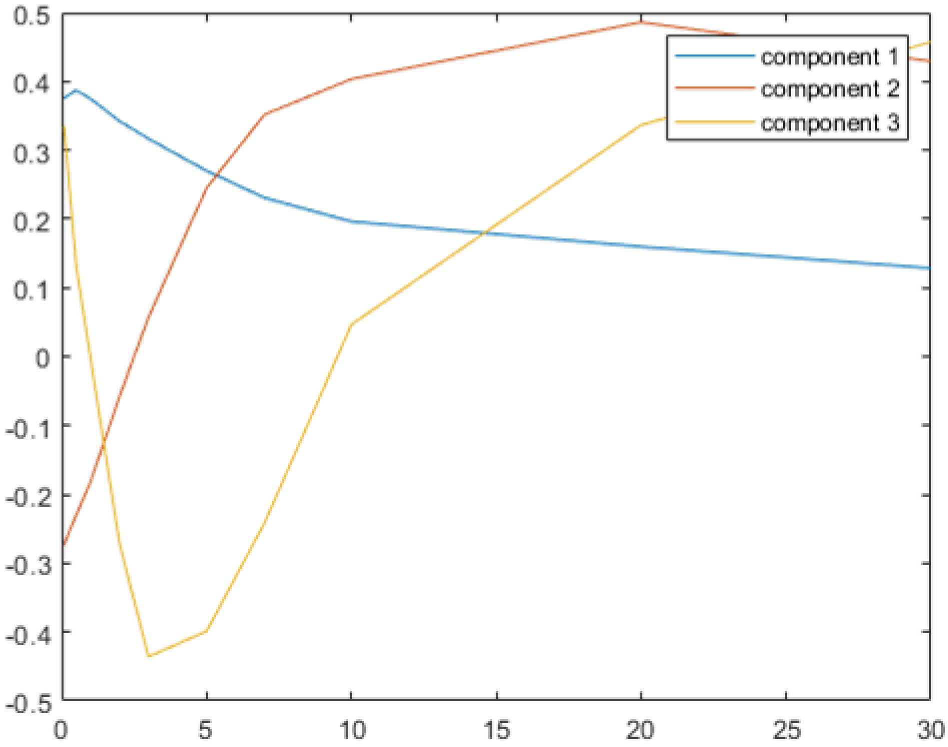

Figure 6 principle components of yield changes

6.2. Proxy

The dataset we use for the PCA, is equal to the data set used to test the Nelson-Siegel model. This has the added benefit that we are able to compare the results, not only between the model with and without the proxy itself, but also between the models both using the proxy or not using the proxy.

We set out how the scores can be fit in two scenarios; with and without our proxy. The first step in both scenarios is to determine the principal components. We determine the principal components using the function Matlab provides, which is based on singular value decomposition. Using the first three components we are able to determine the yield using the following formula:

$$y(s1_t, s2_t, s3_t, \boldsymbol{pc1}, \boldsymbol{pc2}, \boldsymbol{pc3},, \boldsymbol{\tau}_t) = s1_t * \boldsymbol{pc1} + s2_t * \boldsymbol{pc2} + s3_t * \boldsymbol{pc3},$$

where sx_t denotes the score of component x at time t and $\boldsymbol{pcx}$ is a vector denoting the principal component x. Note the lack of the τ parameter, which implies that the yield is not (directly) depended on the time. Due to the distinct values in the $\boldsymbol{pc}$ vector $(pc_1, pc_2, \dots, pc_n)$ (each value corresponding to one maturity) we do have an indirect effect of τ.

The fit error is defined by:

$$fit\ error(\widehat{s1}_t, \widehat{s2}_t, \widehat{s3}_t, \boldsymbol{pc1}, \boldsymbol{pc2}, \boldsymbol{pc3},, \boldsymbol{\tau}_t) = \sqrt{\sum_{\tau} \left(y_t(\tau_t) - y_t\left(\widehat{s1}_t, \widehat{s2}_t, \widehat{s3}_t, pc1, pc2, pc3, \tau_t\right)\right)^2}.$$

Where $\boldsymbol{y}_t(\tau)$ presents the observed yield at time t for each of the maturities in $\boldsymbol{\tau}_t := (T_1 - t, T_2 - t, \dots, T_n - t)$. We name the scores that minimize the fit error at time t sx^*.

In our example we look at yield data of Constant Maturity Treasuries (CMT) on a daily basis. Since in increment of a day does not lead to the shortening of the maturity with one day (e.g. we look at a bond with a maturity of 1 year and the next day we look at the next bond with a maturity of 1 year), in our case, maturity is defined as: $\boldsymbol{\tau}_t := (T_1, T_2, \dots, T_n)$. Hence, from now on we suppress t for $\boldsymbol{\tau}_t$.

The first component we require to implement our proxy is the valuation function:

$$\boldsymbol{F}(s1_t, s2_t, s3_t, \boldsymbol{pc1}, \boldsymbol{pc2}, \boldsymbol{pc3}, \boldsymbol{\tau}) = \begin{pmatrix} e^{-T_1(s1_t * pc1_1 + s2_t * pc2_1 + s3_t * pc3_1)} \\ e^{-T_2(s1_t * pc1_2 + s2_t * pc2_2 + s3_t * pc3_2)} \\ \dots \\ e^{-T_n(s1_t * pc1_n + s2_t * pc2_n + s3_t * pc3_n)} \end{pmatrix}.$$

This function determines the price of a bond given maturity T with the corresponding yield at time t, which follows from the PCA equation. Next we introduce the delta sensitivities to each underlying of the three sx's at time t. Due to the fact that there is not functional form in the PCA, we use a numerical approximation to find the deltas. From now on we suppress the arguments of $\boldsymbol{F}$:

$$\frac{\delta \boldsymbol{F}}{\delta s1_t'} \approx \begin{pmatrix} \left(\boldsymbol{F}(s1_t, s2_t, s3_t, pc1_1, pc2_1, pc3_1, \tau_1) - \boldsymbol{F}(s1_t - 0.0001, s2_t, s3_t, pc1_1, pc2_1, pc3_1, \tau_1)\right) * 10000 \\ \left(\boldsymbol{F}(s1_t, s2_t, s3_t, pc1_2, pc2_2, pc3_2, \tau_2) - \boldsymbol{F}(s1_t - 0.0001, s2_t, s3_t, pc1_2, pc2_2, pc3_2, \tau_2)\right) * 10000 \\ \dots \\ \left(\boldsymbol{F}(s1_t, s2_t, s3_t, pc1_n, pc2_n, pc3_n, \tau_n) - \boldsymbol{F}(s1_t - 0.0001, s2_t, s3_t, pc1_n, pc2_n, pc3_n, \tau_n)\right) * 10000 \end{pmatrix}$$

$$\frac{\delta \boldsymbol{F}}{\delta s2_t'} \approx \begin{pmatrix} \left(\boldsymbol{F}(s1_t, s2_t, s3_t, pc1_1, pc2_1, pc3_1, \tau_1) - \boldsymbol{F}(s1_t, s2_t - 0.0001, s3_t, pc1_1, pc2_1, pc3_1, \tau_1)\right) * 10000 \\ \left(\boldsymbol{F}(s1_t, s2_t, s3_t, pc1_2, pc2_2, pc3_2, \tau_2) - \boldsymbol{F}(s1_t, s2_t - 0.0001, s3_t, pc1_2, pc2_2, pc3_2, \tau_2)\right) * 10000 \\ \dots \\ \left(\boldsymbol{F}(s1_t, s2_t, s3_t, pc1_n, pc2_n, pc3_n, \tau_n) - \boldsymbol{F}(s1_t, s2_t - 0.0001, s3_t, pc1_n, pc2_n, pc3_n, \tau_n)\right) * 10000 \end{pmatrix}$$

$$\frac{\delta \boldsymbol{F}_t}{\delta s3_t'} \approx \begin{pmatrix} \left(\boldsymbol{F}_t(s1_t, s2_t, s3_t, pc1_1, pc2_1, pc3_1, \tau_1) - \boldsymbol{F}_t(s1_t, s2_t, s3_t - 0.0001, pc1_1, pc2_1, pc3_1, \tau_1)\right) * 10000 \\ \left(\boldsymbol{F}_t(s1_t, s2_t, s3_t, pc1_2, pc2_2, pc3_2, \tau_2) - \boldsymbol{F}_t(s1_t, s2_t, s3_t - 0.0001, pc1_2, pc2_2, pc3_2, \tau_2)\right) * 10000 \\ \dots \\ \left(\boldsymbol{F}_t(s1_t, s2_t, s3_t, pc1_n, pc2_n, pc3_n, \tau_n) - \boldsymbol{F}_t(s1_t, s2_t, s3_t - 0.0001, pc1_n, pc2_n, pc3_n, \tau_n)\right) * 10000 \end{pmatrix}$$

We combine these sensitivities into our Delta sensitivity matrix for a given time t:

$$\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'} := \left(\frac{\delta \boldsymbol{F}}{\delta s1_t{}'}, \frac{\delta \boldsymbol{F}}{\delta s2_t{}'}, \frac{\delta \boldsymbol{F}}{\delta s3_t{}'}\right).$$

Since, in our dataset, there is no effect of time decay on our bond prices, we neglect the effect Theta would otherwise have. We are now able to form a projection matrix based on the equation above and a n-by-n identity matrix $\boldsymbol{I}_n$:

$$\boldsymbol{M}(\boldsymbol{x}_t, \boldsymbol{\tau}_t) = \boldsymbol{I}_n - \frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\left(\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)'\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)\right)^{-1}\left(\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}'}\right)'.$$

The *proxy error* at a given time t can now be calculated by using the following equation where $F(\mathbf{x}_{t-1}, \tau)$ represents the prices of the bonds with the observed betas from the day prior.

$$proxy\ error(s1_t, s2_t, s3_t, \boldsymbol{\tau}) = \boldsymbol{M}(\boldsymbol{x}_{t-1}, \boldsymbol{\tau}_t) * \left(\boldsymbol{F}(\widehat{\boldsymbol{x}}_t, \boldsymbol{\tau}) - \boldsymbol{F}(\boldsymbol{x}_{t-1}, \boldsymbol{\tau})\right).$$

The final error for time t can now be determined. We add a weight p to the *proxy error* in order to asses to impact on the parameter estimation for different scenarios. The final error is minimized by changing $s1$, $s2$ and $s3$, in order to achieve $s1^*$, $s2^*$ and $s3^*$.

$$final\ error_t(s1_t, s2_t, s3_t, \boldsymbol{\tau}) = \sum_{\tau}\left(fit\ error_t(s1_t, s2_t, s3_t, \tau) + p * proxy\ error_t(s1_t, s2_t, s3_t, \tau)\right).$$

Due to the fact our minimizing algorithm does not behave properly (e.g. it returned local minima), we introduce boundaries for the values which the scores can take. We choose these boundaries to be $sx\ \pm 0.005$ based on observed daily score changes.

6.3. Results

We introduce our proxy to a dataset containing US Constant Maturity Treasuries (CMT) quotes retrieved from the Federal reserve of Louisiana website. The dataset ranges from the first trading day of 2006 (January 3rd) to the last trading day of 2015 (December 30th), as to include a financial crisis, namely the credit crisis of 2007/2008. Tenors include the 1, 3 and 6 months, and the 1, 2, 3, 5, 7, 10, 20 and 30 years.

We use Matlabs PCA function to determine our principal components. Note that we fit these on the full data set. This implicitly means that we are using an "in-sample" fit, in contrast to an "out-of-sample" fit where determines the principal components based on data and test its model on other data.

We fit a yield curve for each day, using the method we described above for the p factor 0, 10 and 100. In this model, choosing $p = 10$, makes the size of the fit and proxy error about equal. $P = 0$ shows the fit of only using the Nelson Siegel model, and $p = 100$ puts a heavy weight, with an approximate ratio 1:10, on the fit and proxy error respectively.

At first, we determine the mean residual difference between the yield, as presented by the fitted yield curve for that day, and the actual yield from our data set, to create insight in the overall level of our fitted yield curves. We also determine the Root Mean Squared Error (RMSE), to show how large the deviance is. The results can be seen in Table 10 through Table 12 and Figure 7, where we show what happens when we remove the short (1, 3, 6 month), medium (2, 3, 5, 7 year) and long (10, 20, 30 year) tenors. The development of the β's over time can be observed in appendix D.

What becomes evident in our results is that while the sum of the RMSE of each tenor for p = 10 is higher for the full data set, it is lower in all datasets where some tenors are missing, indicating that our proxy version can model yield curves that are better fitting to the actual data when we leave out some data. This encourages us to believe that when data cannot be observed in the market, out proxy is a meaningful addition to the regular PCA estimation process.

We also observe that the total fit (sum RMSE) drastically reduces when tenors are removed from the dataset. The scenarios in which data is removed have a total RMSE which is about 50% higher. This is however expected since these are out of sample results.

Also note that the fit of the PCA, while based on an "in-sample" method still performs far worse as the Nelson-Siegel model based on a much lower mean residual and lower RMSE as well for the Nelson-Siegel model.

FULL DATA	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 10***	***P = 100***	***P = 0***	***P = 10***	***P = 100***
1M	*-1.0041 e -03*	*-1.0017 e -03*	*-1.0649 e -03*	*1.5001 e -03*	*1.5016 e -03*	*2.0877 e -03*
3M	*-8.8380 e -04*	*-8.8201 e -04*	*-9.3938 e -04*	*1.0390 e -03*	*1.0381 e -03*	*1.5863 e -03*
6M	*-4.5998 e -06*	*-3.7974 e -06*	*-5.3200 e -06*	*6.2898 e -04*	*6.2796 e -04*	*1.0706 e -03*
1Y	*6.1582 e -04*	*6.1555 e -04*	*5.7962 e -04*	*1.0527 e -03*	*1.0526 e -03*	*1.1419 e -03*
2Y	*1.4959 e -03*	*1.4936 e -03*	*1.4907 e -03*	*1.6492 e -03*	*1.6470 e -03*	*1.7502 e -03*
3Y	*1.6317 e -03*	*1.6281 e -03*	*1.6545 e -03*	*1.6877 e -03*	*1.6836 e -03*	*2.1047 e -03*
5Y	*-5.6103 e -04*	*-5.6429 e -04*	*-4.9438 e -04*	*7.7174 e -04*	*7.7475 e -04*	*1.8545 e -03*
7Y	*-1.4743 e -03*	*-1.4762 e -03*	*-1.3840 e -03*	*1.6261 e -03*	*1.6282 e -03*	*2.3373 e -03*
10Y	*-1.5723 e -03*	*-1.5719 e -03*	*-1.4724 e -03*	*1.6204 e -03*	*1.6210 e -03*	*2.0668 e -03*
20Y	*-1.2274 e -03*	*-1.2245 e -03*	*-1.1104 e -03*	*1.4288 e -03*	*1.4286 e -03*	*1.8000 e -03*
30Y	*3.4671 e -03*	*3.4709 e -03*	*3.569 e -03*	*3.5188 e -03*	*3.5240 e -03*	*3.7118 e -03*
SUM	*4.8301 e -04*	*4.8356 e -04*	*7.7513 e -04*	*1.6524 e -02*	*1.6528 e -02*	*2.1512 e -02*

Table 9 Statistisc of the full data set

MISSING SHORT	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 10***	***P = 100***	***P = 0***	***P = 10***	***P = 100***
1M	*-4.0553 e -03*	*-4.0455 e -03*	*-3.9664 e -03*	*4.8171 e -03*	*4.8107 e -03*	*4.8272 e -03*
3M	*-3.6981 e -03*	*-3.6894 e -03*	*-3.6201 e -03*	*4.1436 e -03*	*4.1366 e -03*	*4.1527 e -03*
6M	*-2.4560 e -03*	*-2.4490 e -03*	*-2.3948 e -03*	*2.6519 e -03*	*2.6461 e -03*	*2.6795 e -03*
1Y	*-1.3043 e -03*	*-1.2996 e -03*	*-1.2643 e -03*	*1.3597 e -03*	*1.3560 e -03*	*1.4088 e -03*
2Y	*7.3162 e -04*	*7.3160 e -04*	*7.2778 e -04*	*8.2818 e -04*	*8.2791 e -04*	*8.3301 e -04*
3Y	*1.6982 e -03*	*1.6950 e -03*	*1.6653 e -03*	*1.7459 e -03*	*1.7430 e -03*	*1.7313 e -03*
5Y	*8.3392 e -05*	*7.9192 e -05*	*4.3022 e -05*	*2.6314 e -04*	*2.6818 e -04*	*5.3464 e -04*
7Y	*-8.1144 e -04*	*-8.1448 e -04*	*-8.3903 e -04*	*8.9701 e -04*	*9.0037 e -04*	*1.0395 e -03*
10Y	*-1.3961 e -03*	*-1.3958 e -03*	*-1.3908 e -03*	*1.4426 e -03*	*1.4425 e -03*	*1.4859 e -03*
20Y	*-1.4541 e -03*	*-1.4507 e -03*	*-1.4171 e -03*	*1.5932 e -03*	*1.5911 e -03*	*1.6172 e -03*
30Y	*2.8948 e -03*	*2.8999 e -03*	*2.9476 e -03*	*2.9593 e -03*	*2.9656 e -03*	*3.0431 e -03*
SUM	*-9.7672 e -03*	*-9.7388 e -03*	*-9.509 e -03*	*2.2702 e -02*	*2.2688 e -02*	*2.3353 e -02*
SUM SHORT	*-1.0209 e -02*	*-1.1084 e -02*	*-9.9813 e -03*	*1.1613 e -02*	*1.1593 e -02*	*1.1659 e -02*
SUM OTHER	*4.4207 e -04*	*4.4511 e -04*	*4.7247 e -04*	*1.1089 e-02*	*1.1095 e-02*	*1.1693 e -02*

Table 10 Statistics of the data set without short tenors

MISSING MEDIUM	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 10***	***P = 100***	***P = 0***	***P = 10***	***P = 100***
1M	*-1.2232 e -03*	*-1.2213 e -03*	*-1.2252 e -03*	*1.4520 e -03*	*1.4528 e -03*	*1.5901 e -03*
3M	*-7.7855 e -04*	*-7.7756 e -04*	*-7.8325 e -04*	*8.8354 e -04*	*8.8309 e -04*	*9.6440 e -04*
6M	*5.9244 e -04*	*5.9203 e -04*	*5.8335 e -04*	*7.9129 e -04*	*7.8996 e -04*	*7.9746 e -04*
1Y	*1.7079 e -03*	*1.7060 e -03*	*1.6955 e -03*	*1.8011 e -03*	*1.7990 e -03*	*1.8291 e -03*
2Y	*3.5588 e -03*	*3.5538 e -03*	*3.5419 e -03*	*3.7245 e -03*	*3.7184 e -03*	*3.8819 e -03*
3Y	*4.2985 e -03*	*4.2915 e -03*	*4.2817 e -03*	*4.6333 e -03*	*4.6247 e -03*	*4.9189 e -03*
5Y	*1.9414 e -03*	*1.9342 e -03*	*1.9408 e -03*	*2.7516 e -03*	*2.7432 e -03*	*3.3937 e -03*
7Y	*3.8782 e -04*	*3.8195 e -04*	*4.0337 e -04*	*1.8455 e -03*	*1.8411 e -03*	*2.6454 e -03*
10Y	*-8.5312 e -04*	*-8.5601 e -04*	*-8.1888 e -04*	*1.1283 e -03*	*1.1304 e -03*	*1.7979 e -03*
20Y	*-1.6465 e -03*	*-1.6467 e -03*	*-1.5910 e -03*	*1.7581 e -03*	*1.7582 e -03*	*2.0012 e -03*
30Y	*2.4617 e -03*	*2.4633 e -03*	*2.5189 e -03*	*2.5169 e -03*	*2.5183 e -03*	*2.6479 e -03*
SUM	*1.0447 e -02*	*1.0421 e -02*	*1.0547 e -02*	*2.3286 e -02*	*2.3259 e -02*	*2.6468 e -02*
SUM MEDIUM	*1.0187 e -02*	*1.0161 e -02*	*1.0168 e -02*	*1.2955 e -02*	*1.2927 e -02*	*1.4840 e -02*
SUM OTHER	*2.6067 e -04*	*2.5967 e -04*	*3.794 e -04*	*1.0331 e -02*	*1.0332 e -02*	*1.1628 e -02*

Table 11 Statistics of the data set without medium tenors

MISSING LONG	***MEAN RESIDUALS***			***RMSE***		
	P = 0	***P = 10***	***P = 100***	***P = 0***	***P = 10***	***P = 100***
1M	*-7.5025 e -06*	*-7.4817 e -06*	*-6.9683 e -06*	*8.9658 e -04*	*8.9653 e -04*	*8.9155 e -04*
3M	*-2.4179 e -04*	*-2.4177 e -04*	*-2.4097 e -04*	*5.1544 e -04*	*5.1543 e -04*	*5.1085 e -04*
6M	*5.4170 e -06*	*5.4180 e -06*	*5.5486 e -05*	*6.1629 e -04*	*6.1633 e -04*	*6.2402 e -04*
1Y	*1.4590 e -04*	*1.4591 e -04*	*1.4754 e -04*	*7.1311 e -04*	*7.1316 e -04*	*7.2283 e -04*
2Y	*1.7577 e -04*	*1.7576 e -04*	*1.7771 e -04*	*3.3665 e -04*	*3.3664 e -04*	*3.4012 e -04*
3Y	*7.0717 e -05*	*7.0704 e -05*	*7.2386 e -05*	*4.9138 e -04*	*4.9135 e -04*	*4.8646 e -04*
5Y	*-5.2855 e -04*	*-5.2855 e -04*	*-5.2936 e -04*	*6.5260 e -04*	*6.5266 e -04*	*6.5497 e -04*
7Y	*3.1554 e -04*	*3.1556 e -04*	*3.1244 e -04*	*4.8203 e -04*	*4.8194 e -04*	*4.8395 e -04*
10Y	*2.3968 e -03*	*2.3969 e -03*	*2.3913 e -03*	*2.9314 e -03*	*2.9313 e -03*	*2.9227 e -03*
20Y	*5.1756 e -03*	*5.1756 e -03*	*5.1671 e -03*	*6.1651 e -03*	*6.1651 e -03*	*6.1514 e -03*
30Y	*1.0130 e -02*	*1.0130 e -02*	*1.0121 e -02*	*1.0588 e -02*	*1.0588 e -02*	*1.0576 e -02*
SUM	*1.7686 e -02*	*1.7687 e -02*	*1.7668 e -02*	*2.4389 e -02*	*2.4388 e -02*	*2.4365 e -02*
SUM LONG	*1.7702 e -02*	*1.7703 e -02*	*1.7679 e -02*	*1.9685 e -02*	*1.9684 e -02*	*1.9650 e -02*
SUM OTHER	*-6.4499 e -05*	*-6.4450 e -05*	*-1.1736 e -05*	*4.7041 e -03*	*4.7040 e -03*	*4.7148 e -03*

Table 12 Statistics of the data set without long tenors

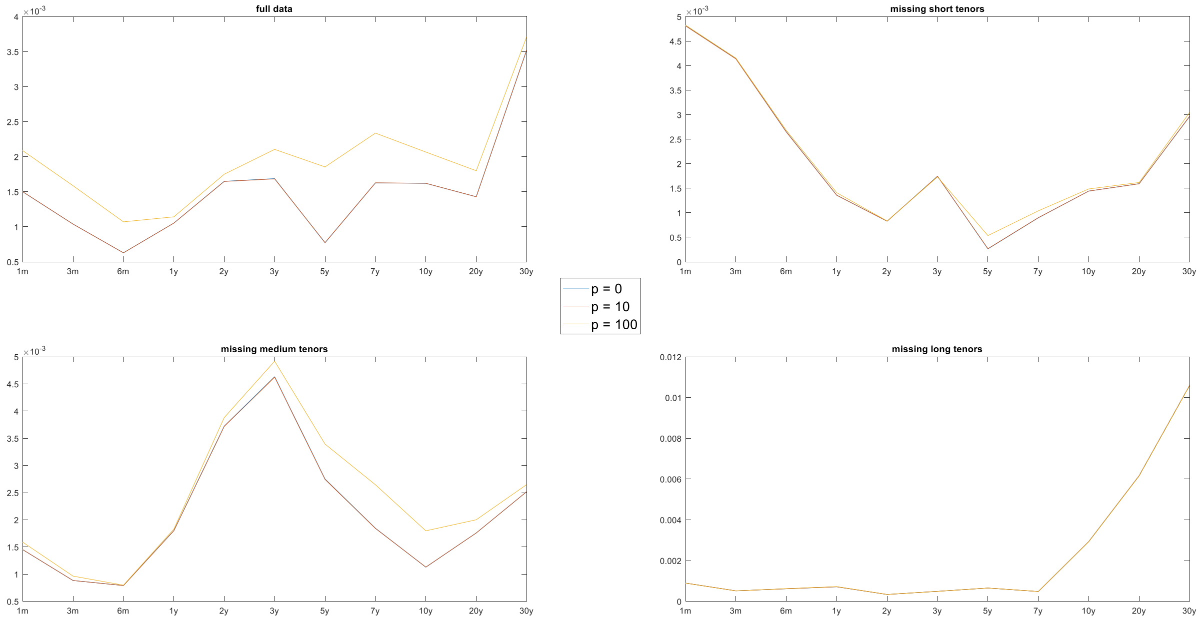

Figure 7 RMSE per tenor for all datasets

7. Discussion

In this chapter we discuss the analysis and results which we respectively performed and determined in the section above. To begin with, we point out that our results are influenced by the used datasets. These datasets however from reliable sources (Fed St. Louisiana and Bloomberg) are still subjected to manipulation (CMT yields are bootstrapped) or market limited liquidity (option data). Specifically in the case of the options used for the SABR model, we point out that the dataset is limited and only tested over distinct strikes, not maturities. The choice for the option chain used was mostly based on data limitations.

Secondly, we acknowledge that the results for the Nelson Siegel and PCA models are, however present, low in absolute terms. We point out that RMSE are lowered for out of sample maturities, but since the RMSE is higher overall than the in-sample errors, regardless of the chosen p value, the absolute effect of the error reduction might provide a skewed view. We did not perform a test for significance, which means that we are not sure of the statistical significance of our results, allowing our observed results to be solitarily based on chance. However, due to the large sample size used, we have implicitly reduced the chance component.

Finally we also point out that our proxy is only tested for a small sample of pricing models, related to vanilla instruments. The introduction of more complex models in order to price exotic products like caps or variance swaps might result into a lower improvement in out of sample performance of our proxy or worse.

8. Conclusion

Based on the observed results we conclude that our generic proxy has a slightly preferred status over using set pricing model directly based on overall RMSE but mostly on out-of-sample RMSE in scenarios where data is stripped from the original set. Our proxy also does not require additional data or expert judgements, which eases implementation process in existing pricing platforms. In the remaining of this chapter we discuss the conclusion on a model level.

What becomes evident in our results for the Nelson Siegel model is that while the sum of the RMSE of each tenor for p = 10 is higher for the full data set, it is lower out-of-sample in all datasets where some tenors are missing, indicating that our proxy version can model yield curves that are better fitting to the actual data when we leave out some data. This encourages us to believe that when data cannot be observed in the market, out proxy is a meaningful addition to the regular Nelson Siegel estimation process.

Our results for the SABR model show that our proxy only reduces the RMSE for the data set without the medium strikes. However, the reduction is only minimal. Therefore we cannot conclude that our proxy has material benefits.

In case of the PCA model our analysis show results similar to those of the Nelson Siegel model. While the sum of the RMSE of each tenor for p = 10 is higher for the full data set, it is lower out-of-sample in all datasets where some tenors are missing, indicating as with the Nelson Siegel model, that the proxy has value. We point out that the Nelson-Siegel model overall showed a lower RMSE for all data sets, and that the added value of the proxy was more significant for the Nelson Siegel as for the PCA. The latter can be subjected to the reduction in overall RMSE for the dataset where medium and high maturities are missing with the Nelson Siegel model, where in the case of the PCA our model fails to do so.

9. Further research

In this chapter we highlight some relevant topics which were not part of our thesis, but where we envision relevant findings to be made. Starting off with the use of our proxy on data sets where gaps in the datasets are artificially created. In our tests we left out complete tenors/strikes to test whether the functional forms of our selected pricing model contained information on the missing tenors. A quick analysis has shown that creating gaps and backfilling them with our proxy shows good results.

Furthermore, we suggest to test our proxy on pricing models which are used on more exotic marketable instruments. For example: floors and variance swaps, where the need for an alternative backfilling methodology may even be of greater use due to the significantly more stressed markets.

We also recommend to investigate the possibility of using our proxy outside the financial markets. Backfilling data is a general problem in the practical world. We suspect that underlying factors can be derived as long as they are linked to a functional form, which's derivative can be either derived or approximated.

Finally we suggest to extent the performed research on the PCA, by focusing on the phase in which we construct the components (perform the PCA). We performed regular PCA directly on the data, therefore the components do not contain any added information. We see room the combine our proxy with the PCA to get enhanced components.

10. Bibliography

BCBS. (2013). *Fundamental review of the trading book: A revised market risk framework.* Bank of International Settlements.

BCBS. (2016). *Standards - Minimum capital requirements for market risk. .* Bank of International Settlements.

BCBS. (2018). *Consultative Document - Revisions to the minimum requirements for market risk.* Bank of International Settlements.

BIS. (c2005). *Zero-coupon Yield Curves: Technical Documentation.* Basel: Bank of International Settlements.

Björk, T. (2009). *Arbitrage Theory in Continous Time.* Oxford University Press, Inc.

Björk, T. (2010). *Basic Arbitrage Theory (lecture notes).* KTH.

Black, F. (1975). The pricing of commodity contracts. *Journal of Financial Economics*, 167-179.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilties. *Journal of Political Economy*, 637-659.

Black, F., & Scholes, M. (1973). The pricing of Options and Corporate Liabilties. *The Journal of Political Economy*, 637-654.

Bosma, J. (2018). *FRTB Generic proxy method.* Amsterdam: Triple A - Risk Finance.

Diebold, F. X., & Li, C. (2005). *Forecasting the term structure of government bond yields.* Journal of Econometrics.

Gurkaynak, R. S., Sack, B., & Wright, J. H. (2006). *The U.S. Treasury Yield Curve: 1961 to the Present.* Washington D.C.: Federal Reserve Board.

Hagan, P. S., Kumar, D., Lesniewski, A. S., & Woodward, D. E. (2002). Managing smile risk. *Wilmott.*

Hotelling, H. (1933). Analysis of a complex statiscal variables into principal components. *Journal of Educational Psychology*, 417-441 and 498-520.

Merton, R. (1973). The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 141-183.

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious Modeling of Yield Curve. *Journal of Business*, vol. 60, no. 4, pp 473-489.

Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine*, 559-572.

Svensson, L. E. (1995). *Estimating forward interest rates with the extended Nelson & Siegel model.* Sveriges Riksbank Quarterly Review.

11. Appendices

A. Derivation of risk factor proxy value

[This chapter is written by Jakob Bosma]

Let $P_t(\tau)$ denote the price of a derivate at time t with time-to-maturity $\tau := T - t$ and $T \geq t$. We assume that derivative prices are generated by the following continuous-time model:

$$P_t(\tau) := F_t(x_t, \tau);$$

$$\mathrm{d}\boldsymbol{x}_t = \boldsymbol{\mu}(\boldsymbol{x}_t, t)dt + \boldsymbol{\sigma}(\boldsymbol{x}_t, t)\mathrm{d}\boldsymbol{W}_t.$$

Where F_t is a strictly positive twice differentiable function, and $\boldsymbol{x}_t$ a m-dimensional vector of tradable factors. $\boldsymbol{\mu}(\boldsymbol{x}_t, t)$ and $\boldsymbol{\sigma}(\boldsymbol{x}_t, t)$ are adapted processes and $\boldsymbol{W}_t$ denotes a k-dimensional Brownian motion; all satisfying the usual regularity conditions. [6]

Using Itŏ's Lemma, we obtain the following specification for the dynamics of the derivative's price, where we suppress the arguments for F_t:

$$\mathrm{d}p_t = \left(\frac{\delta F_t}{\delta \boldsymbol{x}'}\boldsymbol{\mu}(\boldsymbol{x}_t, t) + \frac{1}{2}tr\left\{\frac{\delta^2 F_t}{\delta \boldsymbol{x}\boldsymbol{x}'}\boldsymbol{\sigma}(\boldsymbol{x}_t, t)\boldsymbol{\sigma}(\boldsymbol{x}_t, t)'\right\}\right)\mathrm{d}t + \frac{\delta F_t}{\delta \boldsymbol{x}'}(\boldsymbol{x}_t, t)\mathrm{d}\boldsymbol{W}_t. \qquad (\mathrm{A.1})$$

No-arbitrage requires that the derivative's instantaneous return equals the instantaneous short rate r_t in addition to the various market risk premia derived from the exposure to the factors $\boldsymbol{\lambda}(\boldsymbol{x}_t, t)$, including time-to-maturity decay. For a derivation of instantaneous returns under the risk natural measure we refer to Chapter 15 in Björk (2009). This chapter covers the multidimensional case with the number of risk sources k not necessarily equal to the number of factors m, as considered here. We arrive at the following no-arbitrage condition:

$$\frac{\delta F_t}{\delta \boldsymbol{x}'}\boldsymbol{\mu}(\boldsymbol{x}_t, t) + \frac{1}{2}tr\left\{\frac{\delta^2 F_t}{\delta \boldsymbol{x}\boldsymbol{x}'}\boldsymbol{\sigma}(\boldsymbol{x}_t, t)\boldsymbol{\sigma}(\boldsymbol{x}_t, t)'\right\} = F_t r_t + \frac{\delta F_t}{\delta \boldsymbol{x}'}\boldsymbol{\sigma}(\boldsymbol{x}_t, t)\boldsymbol{\lambda}(\boldsymbol{x}_t, t) + \frac{\delta F_t}{\delta \tau}. \qquad (A.2)$$

[6] Note that the dynamics of $\boldsymbol{x}_t$ are not necessarily incorporated in the derivation of the derivative's valuation function F_t. To illustrate this point, consider the Black Scholes call option price with constant volatility. In this context we consider the Vanilla pricing function of Black Scholes as such and regard it as a factor model in the above specification with the constant volatility *parameter* replaced with a dynamic volatility *factor.*

We now move on to the multi-dimensional case and consider n prices for the derivative $\boldsymbol{p}_t$ with time-to-maturities given by $\boldsymbol{\tau}_t := (T_1 - t, T_2 - t, \ldots, T_n - t)$, such that:

$$\boldsymbol{p}_t := \boldsymbol{F}_t(\boldsymbol{x}_t, \boldsymbol{\tau}_t) = \left(F_t(\boldsymbol{x}_t, T_1 - t), F_t(\boldsymbol{x}_t, T_2 - t), \ldots, F_t(\boldsymbol{x}_t, T_n - t)\right)'.$$

Substitution of the no-arbitrage condition (A. 1) into (A. 2) and considering the vector of derivative prices p_t then yields (we again suppress the arguments for F_t):

$$\mathrm{d}\boldsymbol{p}_t = \left(\boldsymbol{F}_t r_t + \frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{x}'}\boldsymbol{\sigma}(\boldsymbol{x}_t, t)\boldsymbol{\lambda}(\boldsymbol{x}_t, t) + diag\left\{\frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{\tau}'}\right\}\right)\mathrm{d}t + \frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{x}'}\boldsymbol{\sigma}(\boldsymbol{x}_t, t)\mathrm{d}\boldsymbol{W}_t \qquad (A.3)$$

We introduce the following projection matrix which features a collection of Delta price sensitivities of the derivative price with respect to the considered risk factors. This n-by-n dimensional matrix is defined as follows:

$$\boldsymbol{M}_t(\boldsymbol{x}_t, \boldsymbol{\tau}_t) = \boldsymbol{I}_n - \frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{x}'}\left(\left(\frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{x}'}\right)'\left(\frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{x}'}\right)\right)^{-1}\left(\frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{x}'}\right)', \qquad (A.4)$$

where $\boldsymbol{I}_n$ denotes a n-by-n identity matrix. This projection matrix eliminates the noise term and the unidentified risk premium in (A. 3) by premultiplying the dynamics (A. 3) with (A. 4). This step results in the following expression, where we suppress the arguments for $\boldsymbol{M}_t$:

$$\boldsymbol{M}_t\mathrm{d}\boldsymbol{p}_t = \boldsymbol{M}_t\left(\boldsymbol{F}_t r_t + diag\left\{\frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{\tau}'}\right\}\right)\mathrm{d}t. \qquad (A.5)$$

We may need to check whether a pseudo-inverse applies here to ensure invertability of the inner matrix term in the projection matrix, for instance the Moore-Penrose inverse matrix.

The rows (and columns due to symmetry) of $\boldsymbol{M}_t$ in (A. 5) can be regarded as portfolio weights associated with the considered derivatives at time t. These portfolios are instantaneously unaffected by changes in the underlying risk factors $\boldsymbol{x}_t$ and are therefore risk free. As this expressions follows the condition of absence of arbitrage, the instantaneous return comprises the short rate and the time-to-maturity decay premium resulting from a finite maturity tenor.

For a sufficiently small time interval Δt condition (A. 5) can be approximated by:

$$\boldsymbol{M}_t\Delta\boldsymbol{p}_t \approx \boldsymbol{M}_t\left(\boldsymbol{F}_t r_t + diag\left\{\frac{\delta \boldsymbol{F}_t}{\delta \boldsymbol{\tau}_t'}\right\}\right)\Delta t,$$

Where $\Delta\boldsymbol{p}_t := \boldsymbol{p}_{t+\Delta t} - \boldsymbol{p}_t$.

With $\Delta t = 1$ we have the following condition to infer the missing factor value $\widehat{\boldsymbol{x}}_t$:

$$\boldsymbol{M}_{t-1}\left(\boldsymbol{F}_t(\widehat{\boldsymbol{x}}_t, \boldsymbol{\tau}_t) - \boldsymbol{F}_{t-1}(\boldsymbol{x}_{t-1}, \boldsymbol{\tau}_{t-1})(1 + r_{t-1}) + diag\left\{\frac{\delta \boldsymbol{F}_{t-1}}{\delta \tau'}(\boldsymbol{x}_{t-1}, \boldsymbol{\tau}_{t-1})\right\}\right) = 0 \qquad (A.6)$$

Given realized prices and risk factor values at time $t-1$ we can identify with (A. 6) current prices $\boldsymbol{F}_t(\boldsymbol{x}_t . \boldsymbol{\tau}_t)$ of the considered derivative at time t, and estimate the missing factor by inferring the approximate value $\widehat{\boldsymbol{x}}_t$ for $\boldsymbol{x}_t$.

B. Nelson Siegel parameter developments

B.1. Full data

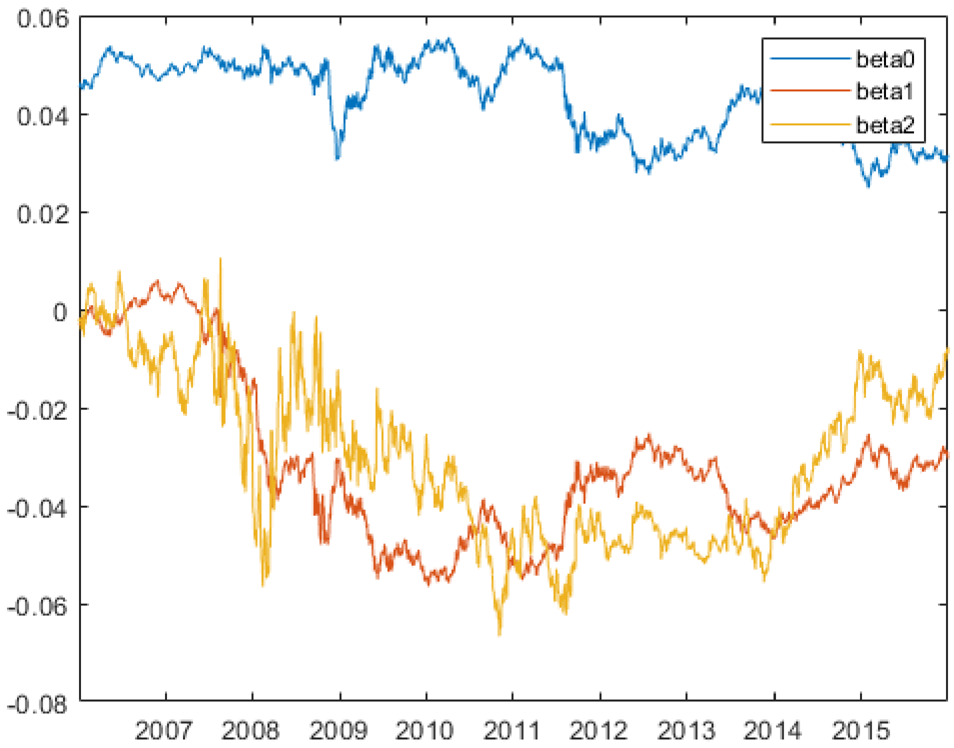

Figure 8 p = 0, full data

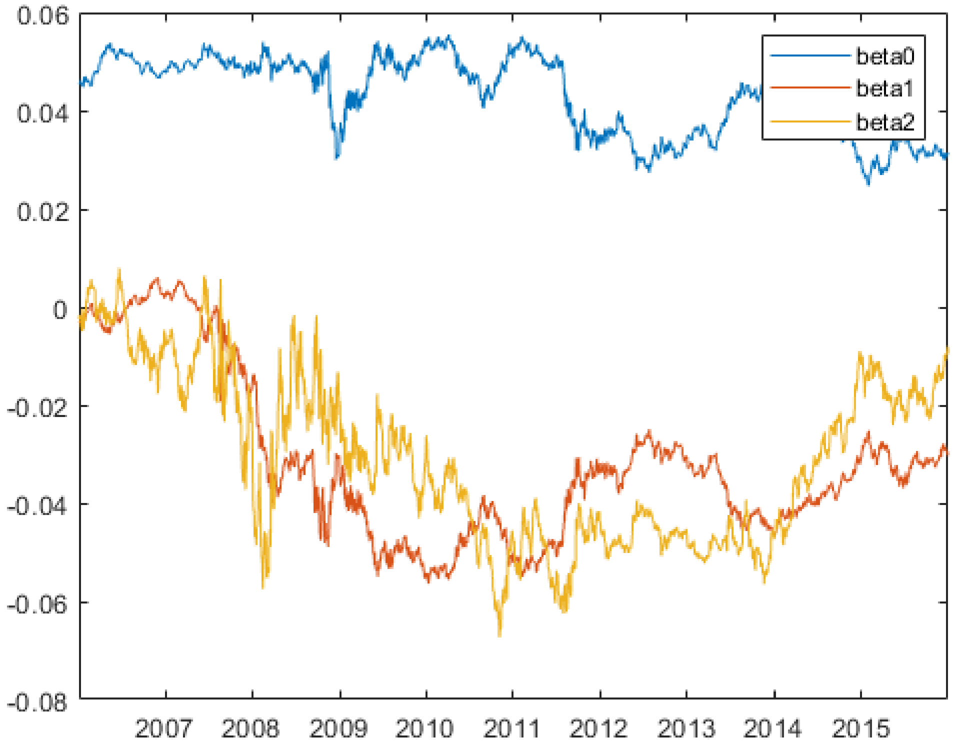

Figure 9 p = 10, full data

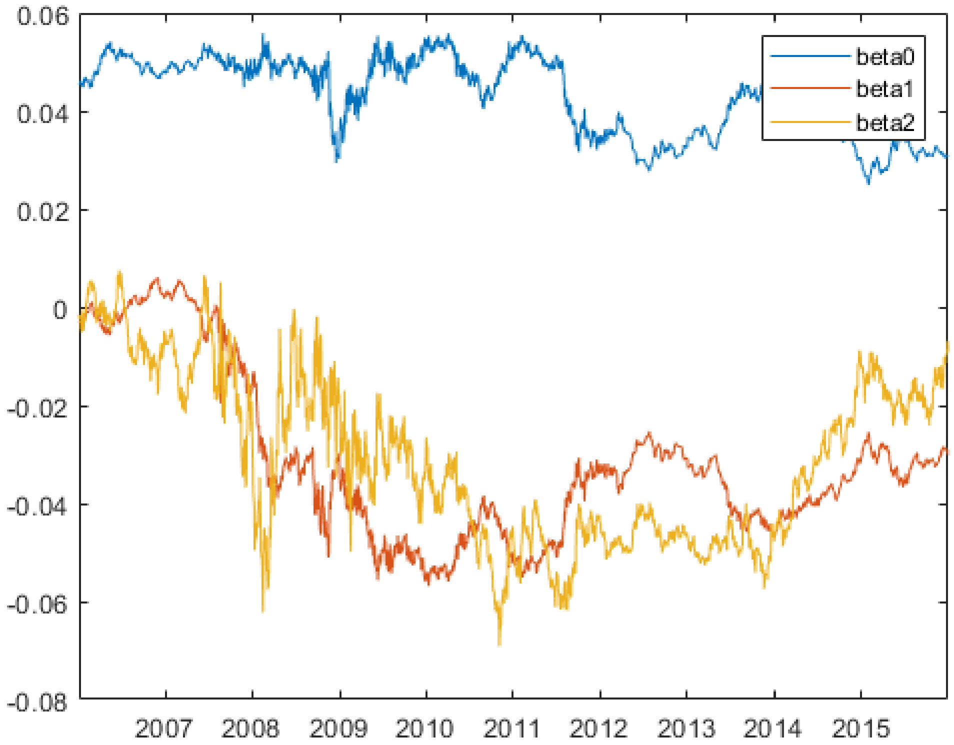

Figure 10 p =100, full data

B.2. Missing short tenors

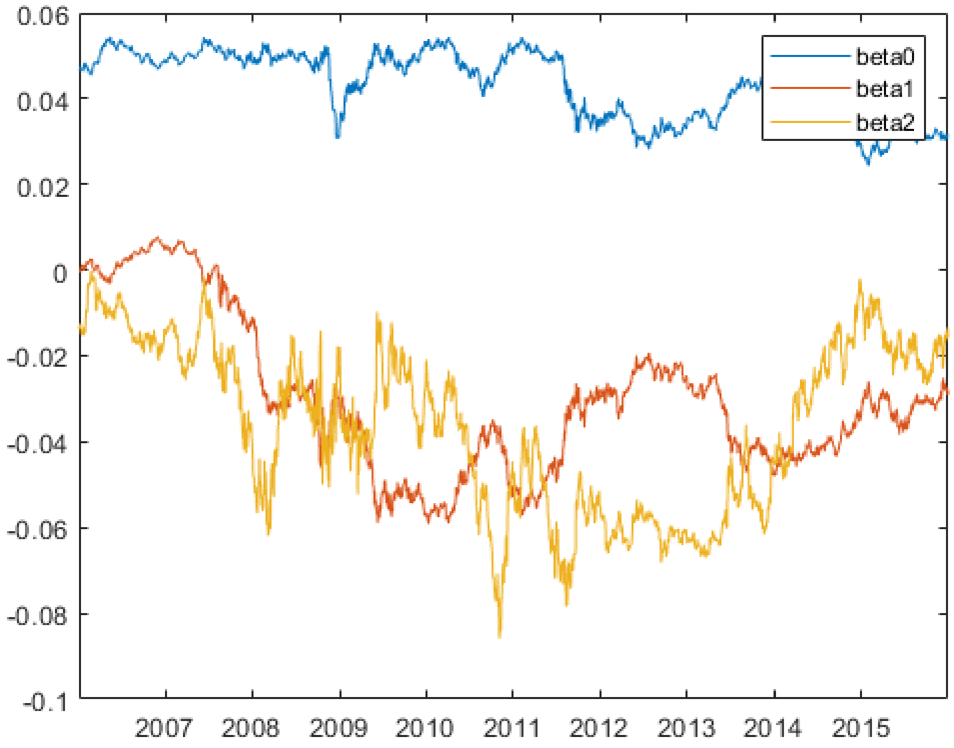

Figure 11 p = 0, missing short tenors

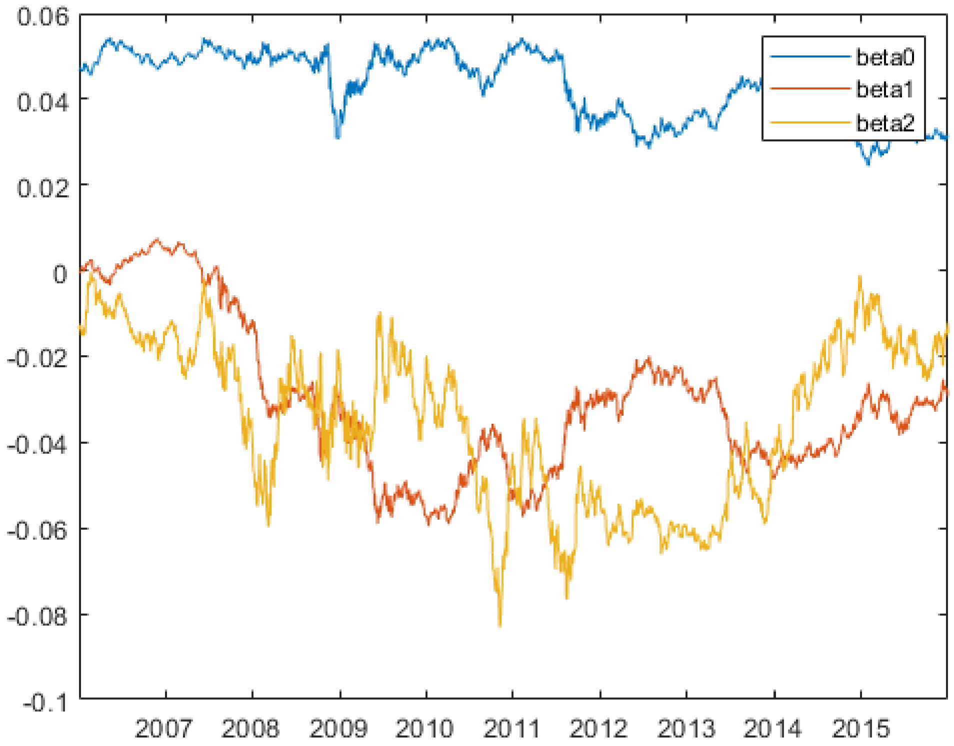

Figure 12 p = 10, missing short tenors

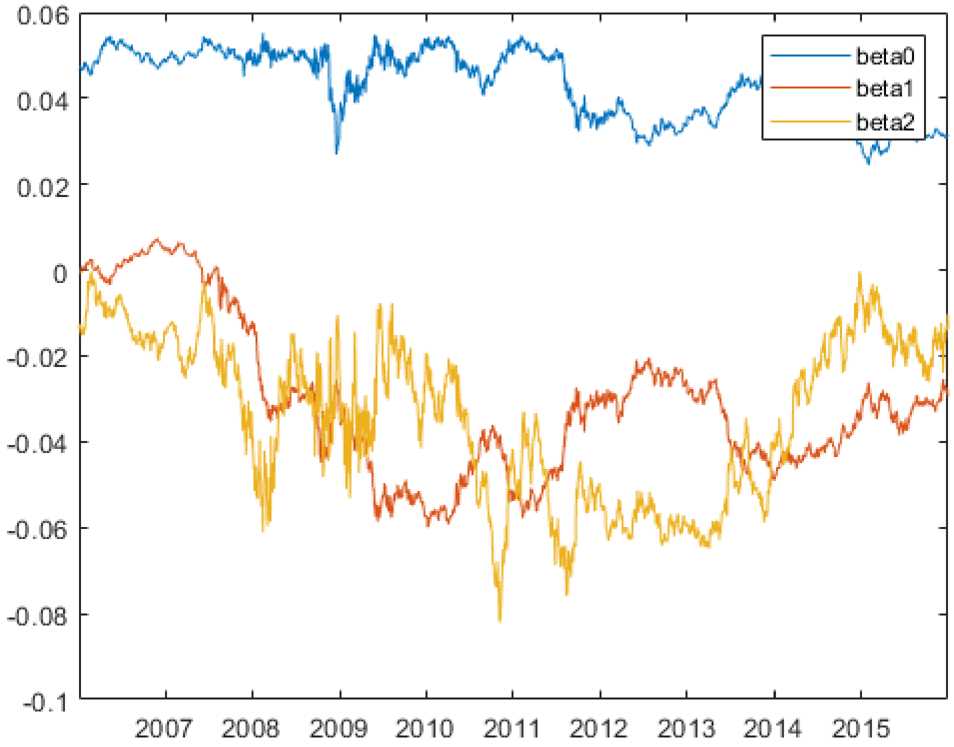

Figure 13 p = 100, missing short tenors

B.3. Missing medium tenors

Figure 14 p = 0, missing medium tenors

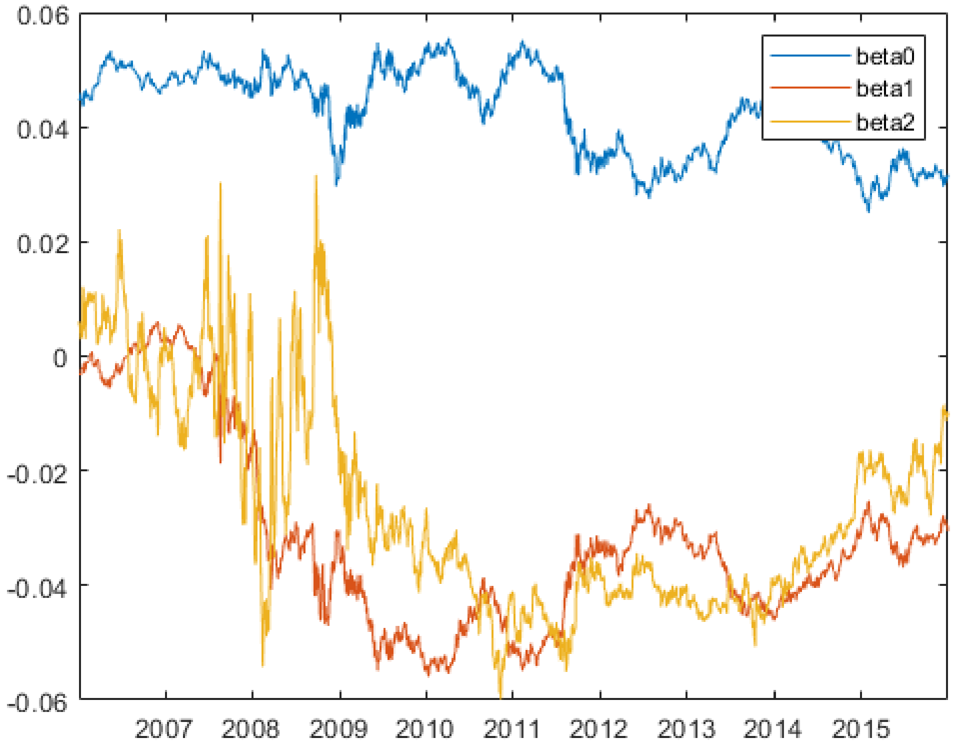

Figure 15 p = 10, missing medium tenors

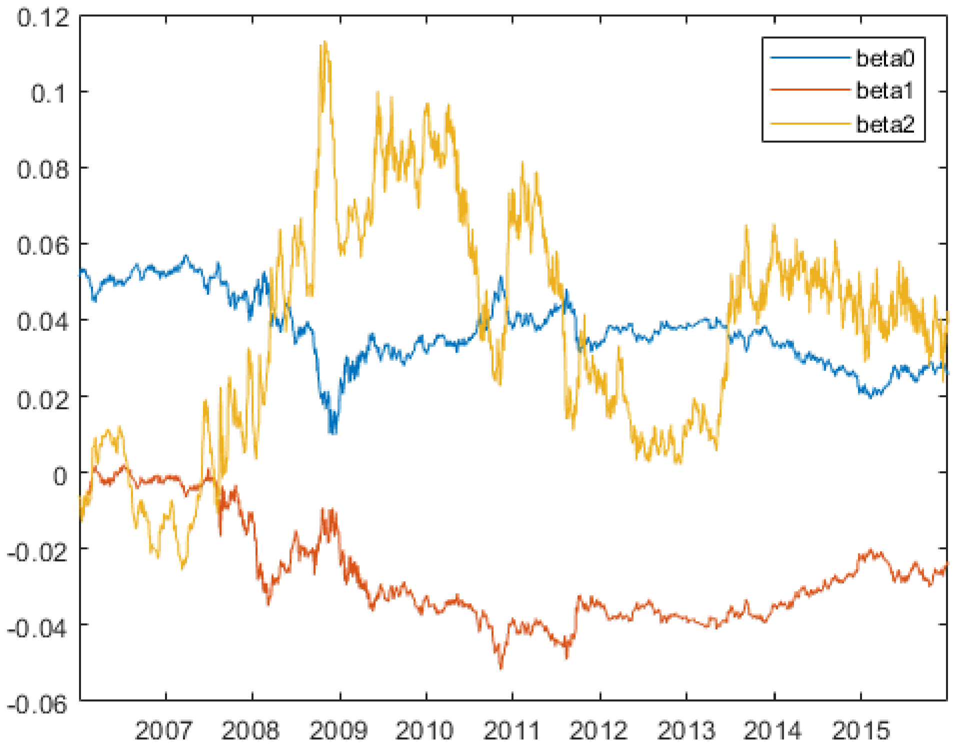

Figure 16 p = 100, missing medium tenors

B.4. Missing long tenors

Figure 17 p = 0, missing long tenors

Figure 18 p = 10, missing long tenors

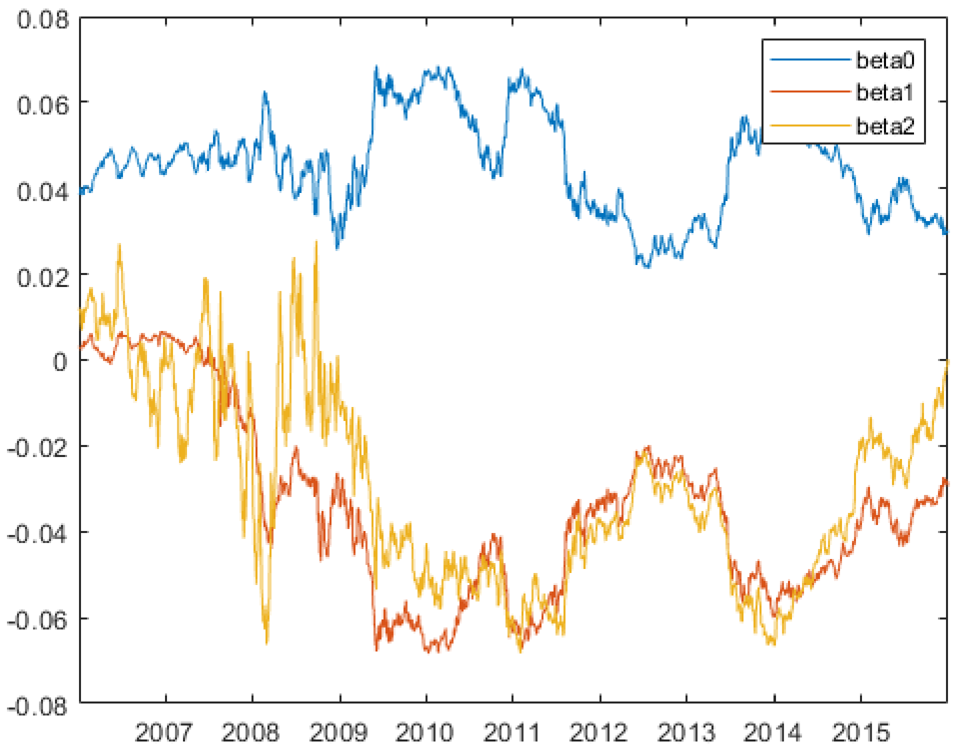

Figure 19 p = 100 missing long tenors

C. SABR parameter development

C.1. Full data

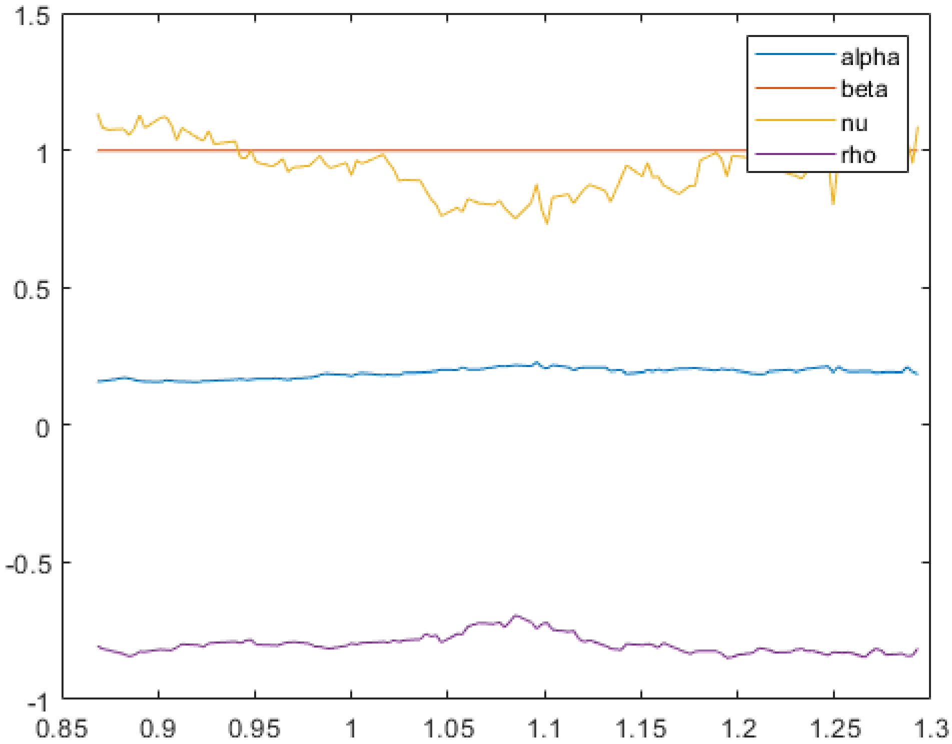

Figure 20 p =0, full data

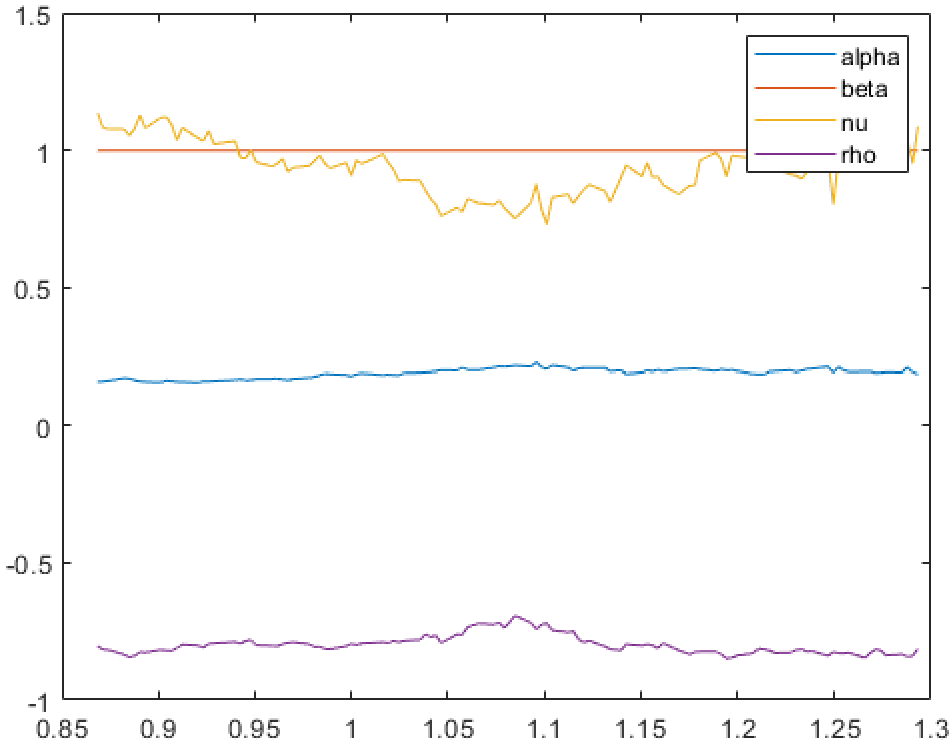

Figure 21 p = 100, full data

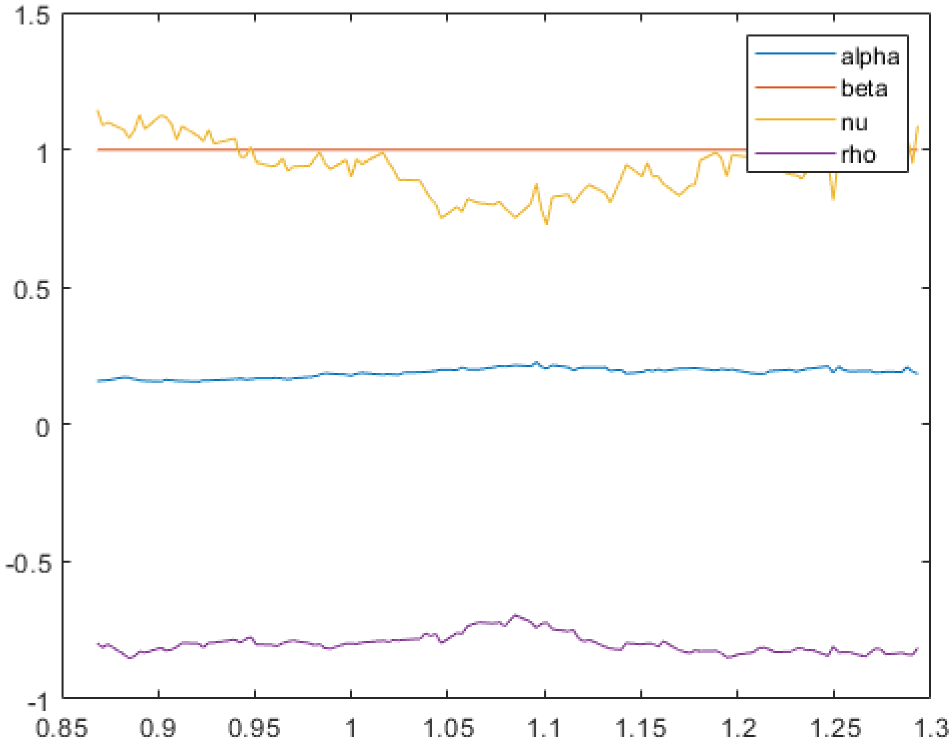

Figure 22 p = 1000, full data

C.2. Missing low strikes

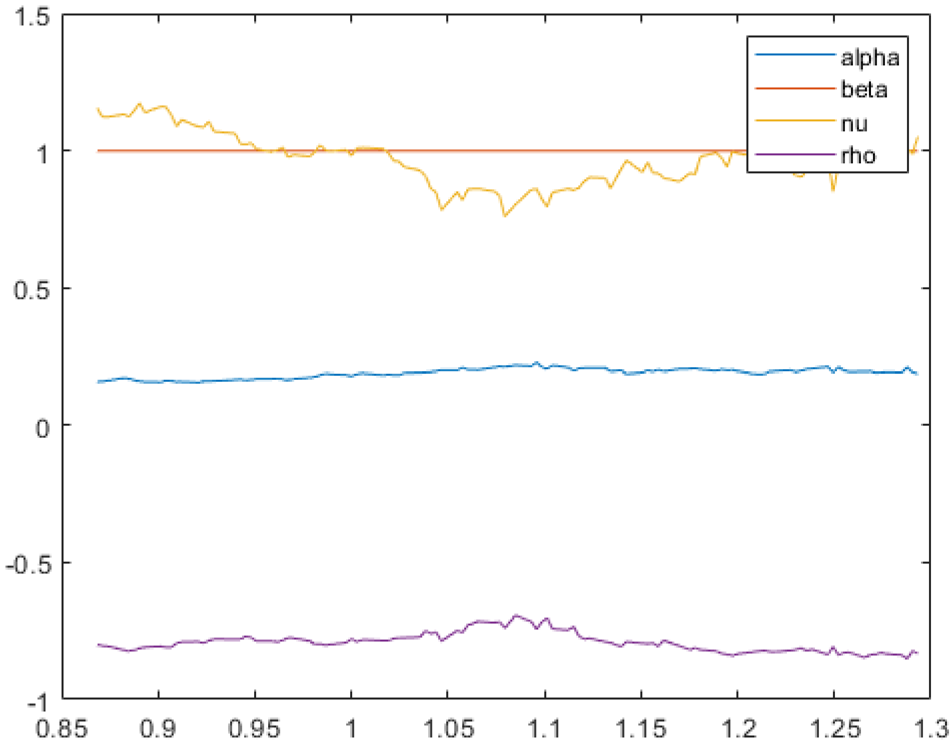

Figure 23 p = 0, missing low strikes

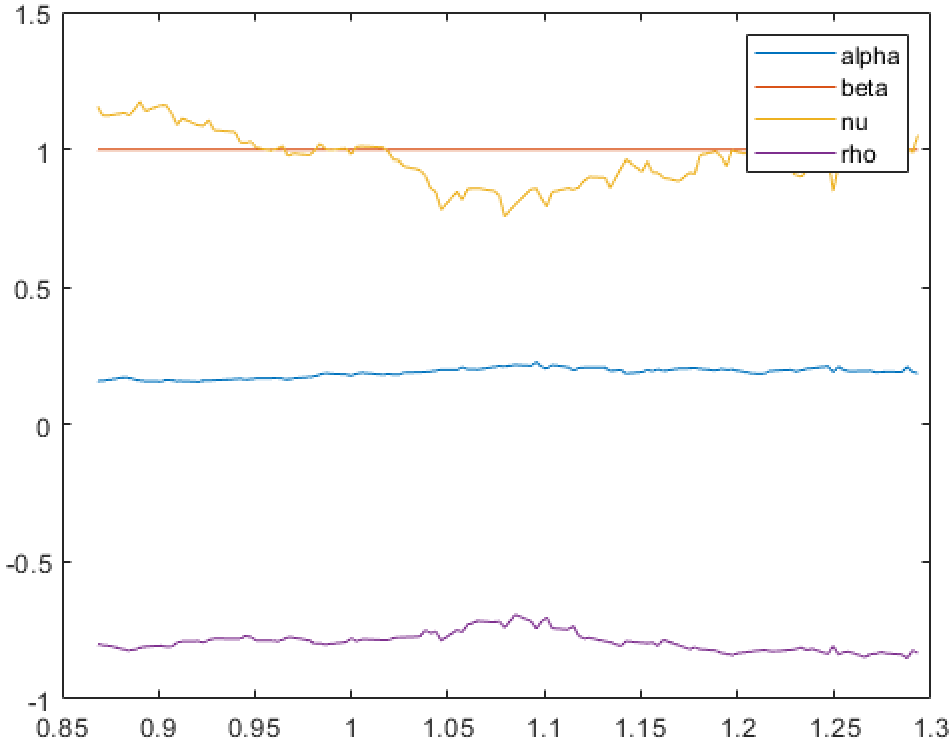

Figure 24 p = 100, missing low strikes

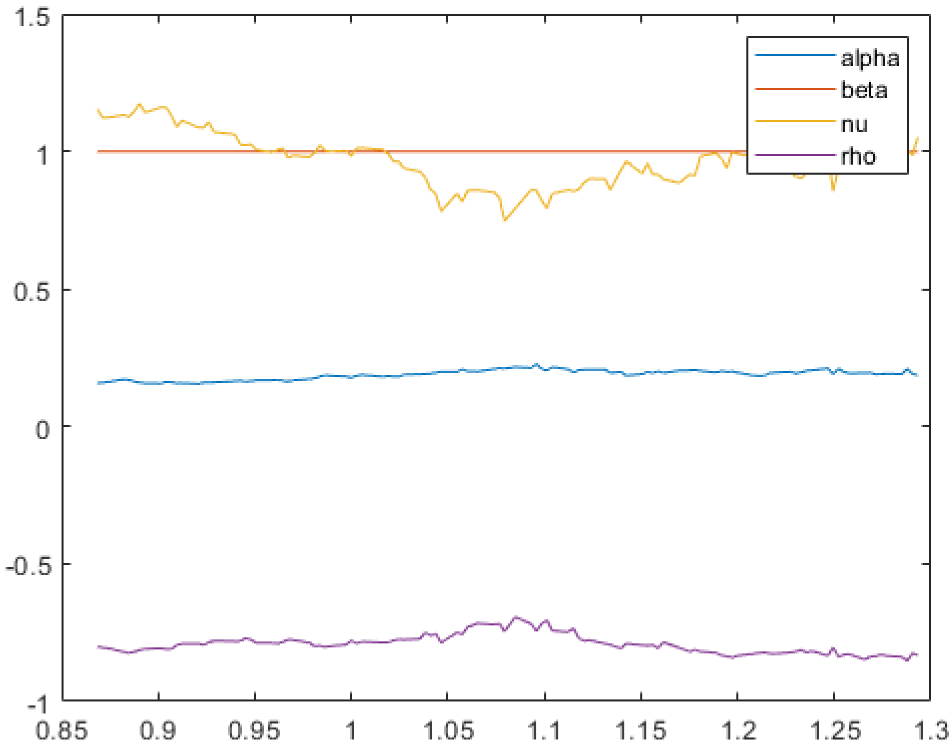

Figure 25 p = 1000, missing low strikes

C.3. Missing medium strikes

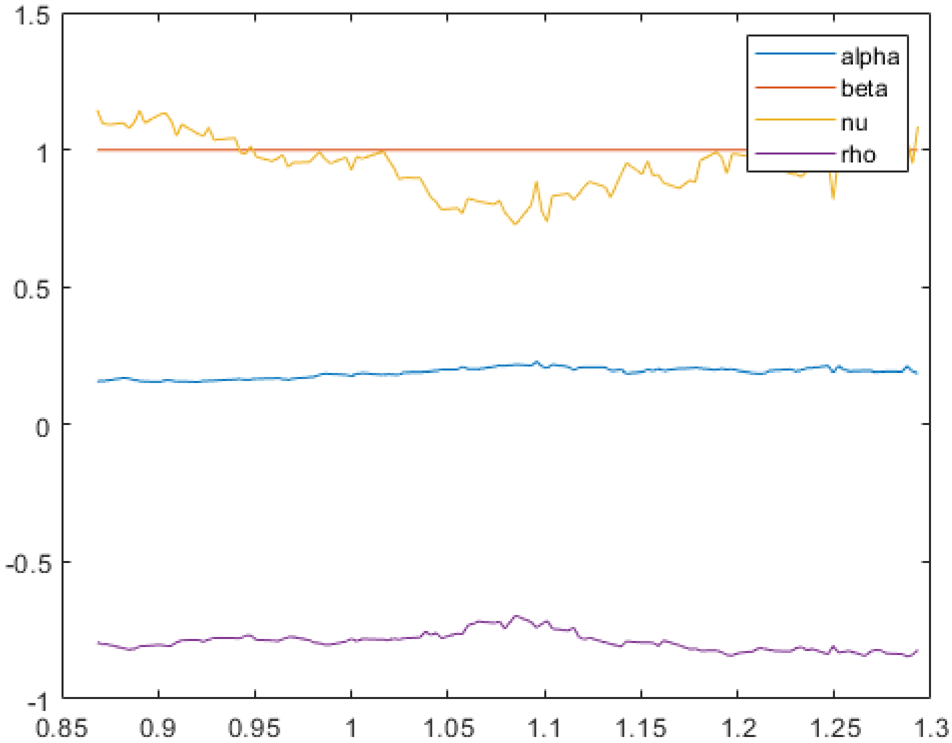

Figure 26 p = 0, missing medium strikes

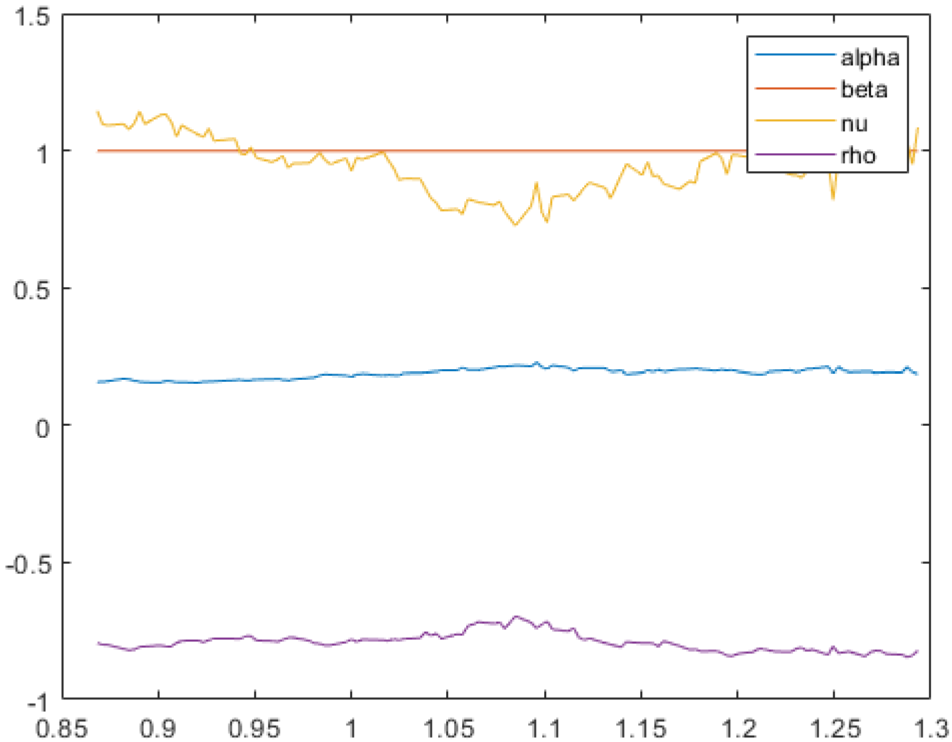

Figure 27 p = 100, missing medium strikes

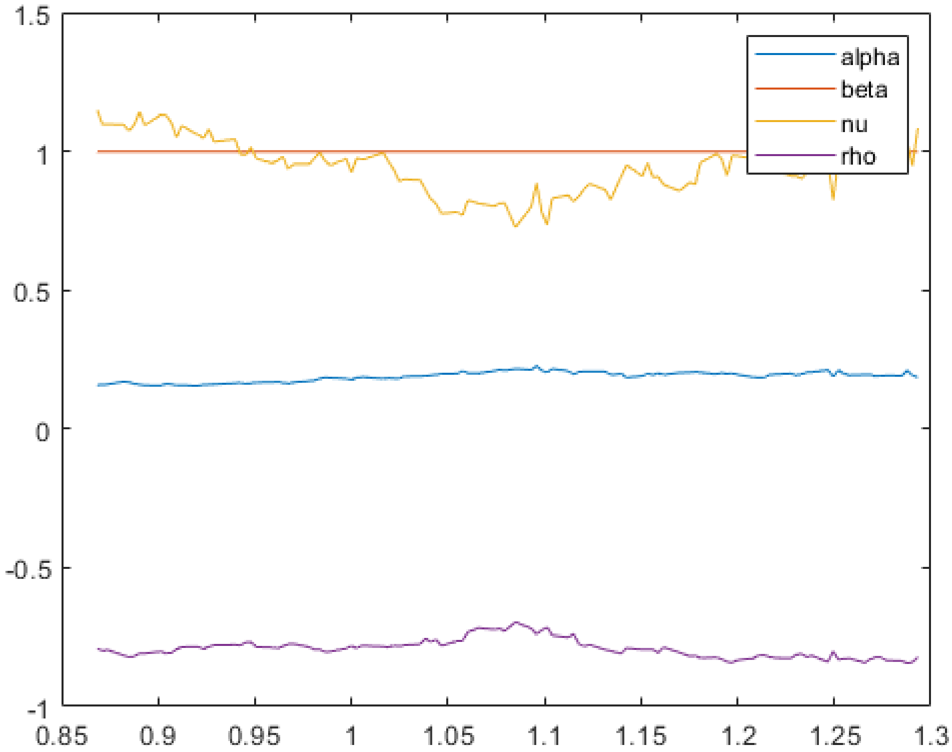

Figure 28 p = 1000, missing medium strikes

C.4. Missing high strikes

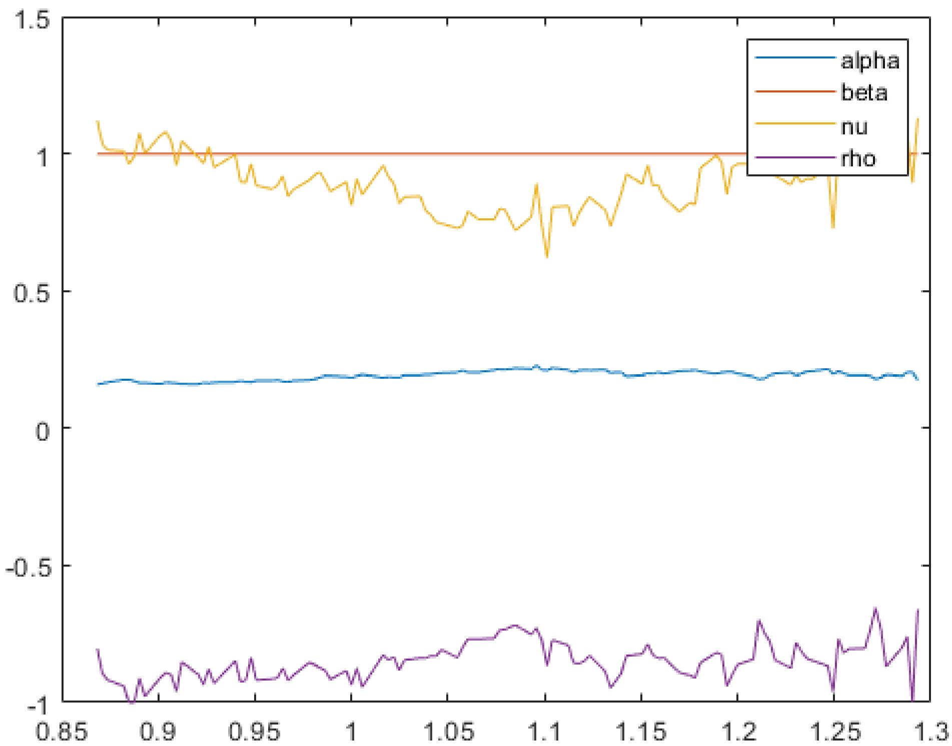

Figure 29 p = 0, missing high strikes

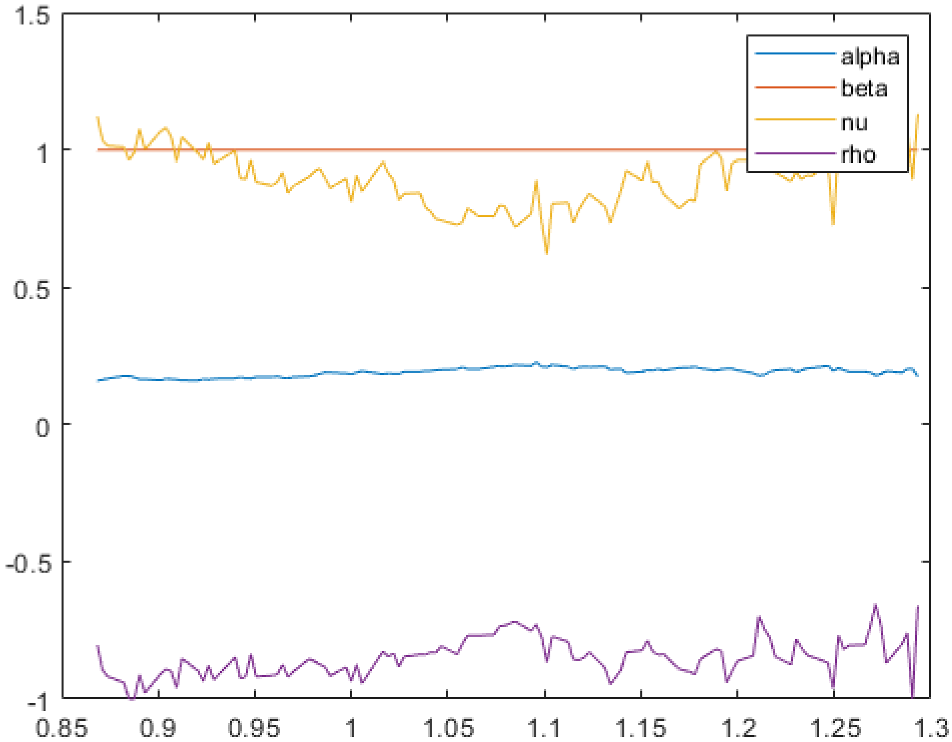

Figure 30 p = 100, missing high strikes

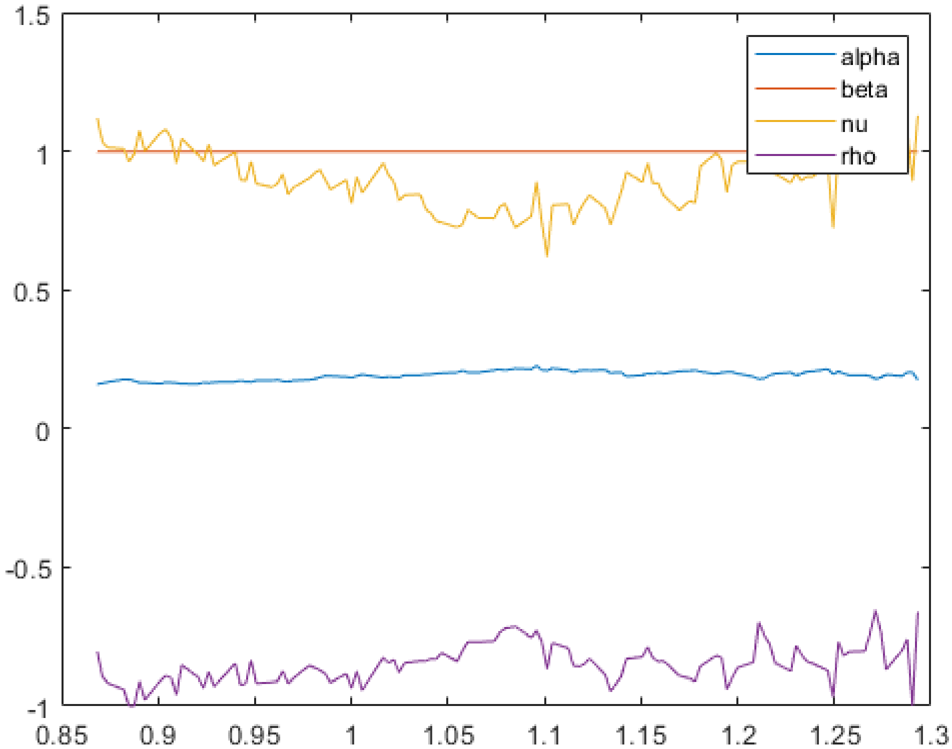

Figure 31 p = 1000, missing high strikes

D. PCA score developments

D.1. Full data

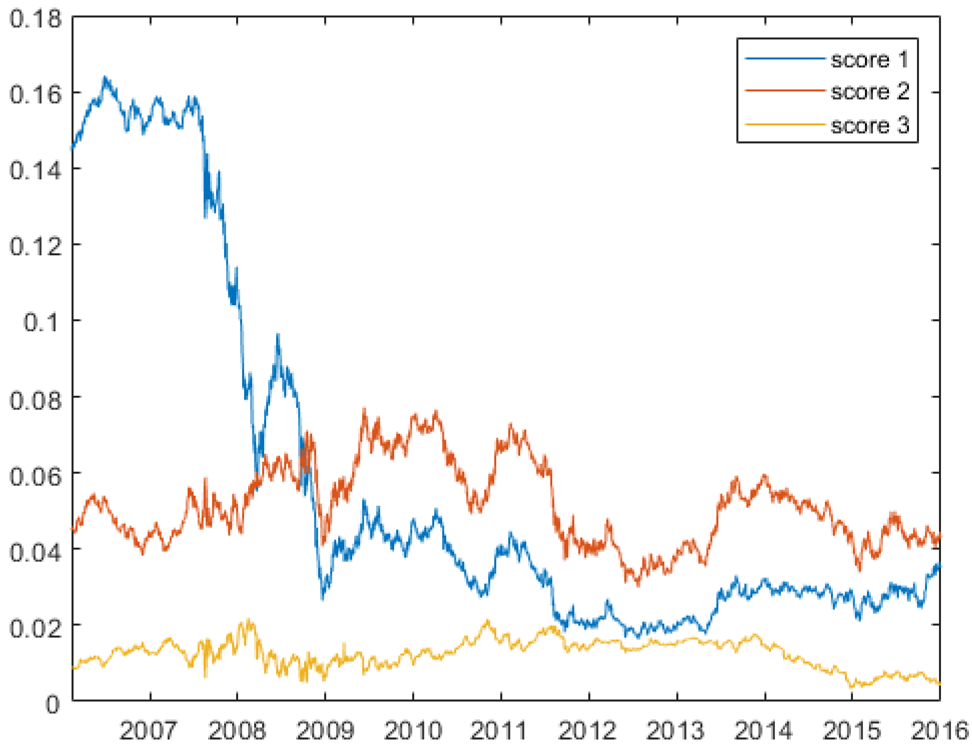

Figure 32 p = 0, full data

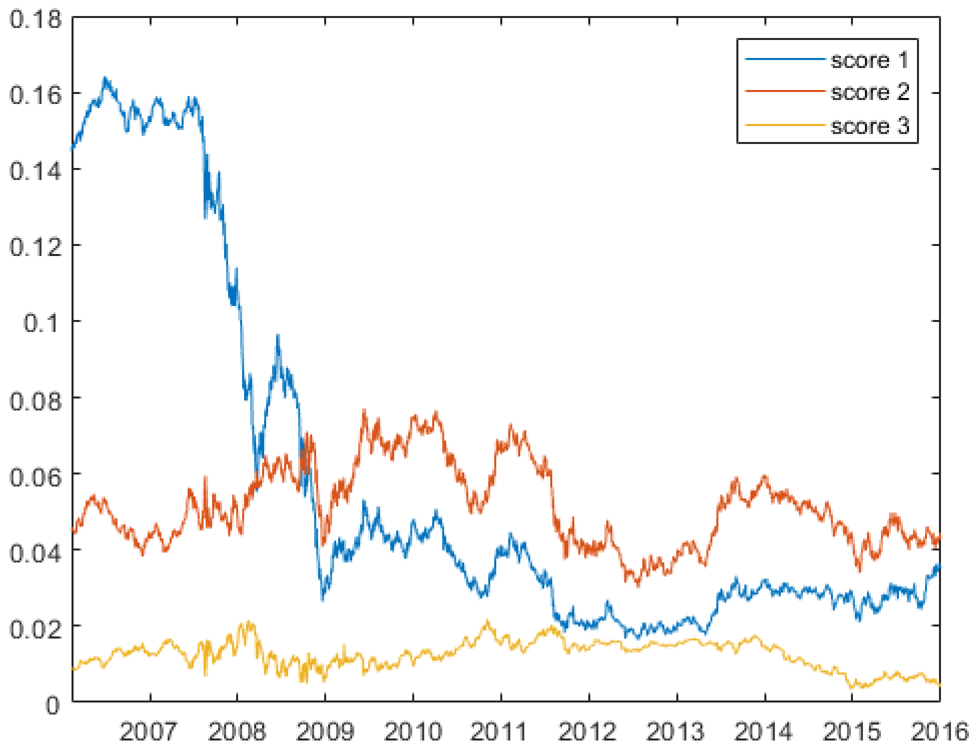

Figure 33 p = 10, full data

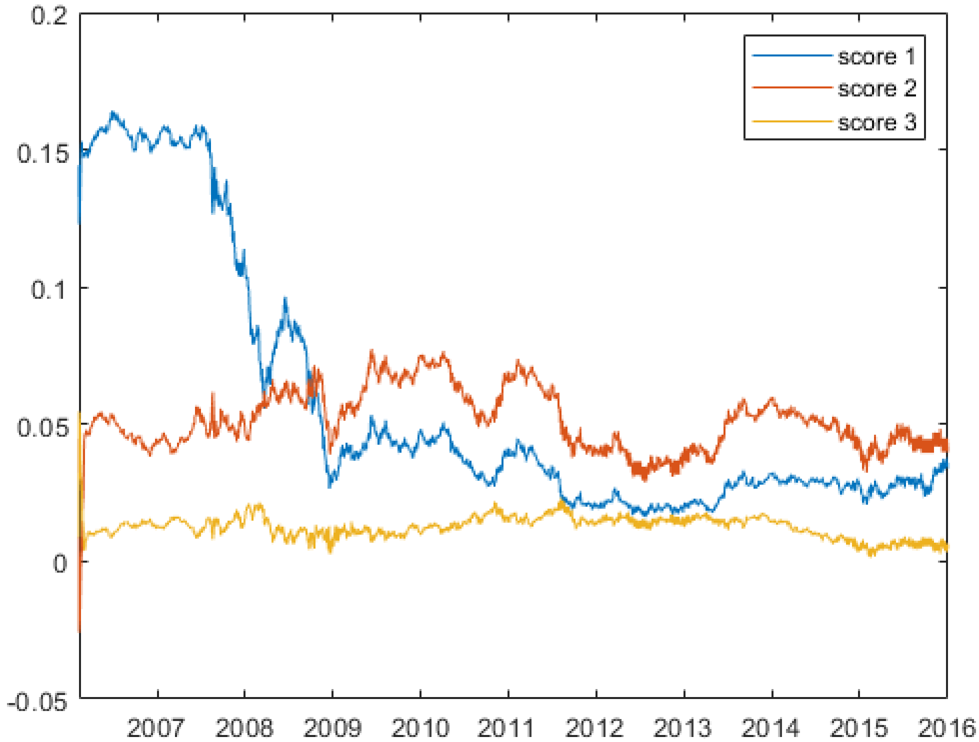

Figure 34 p =100, full data

D.2. Missing short tenors

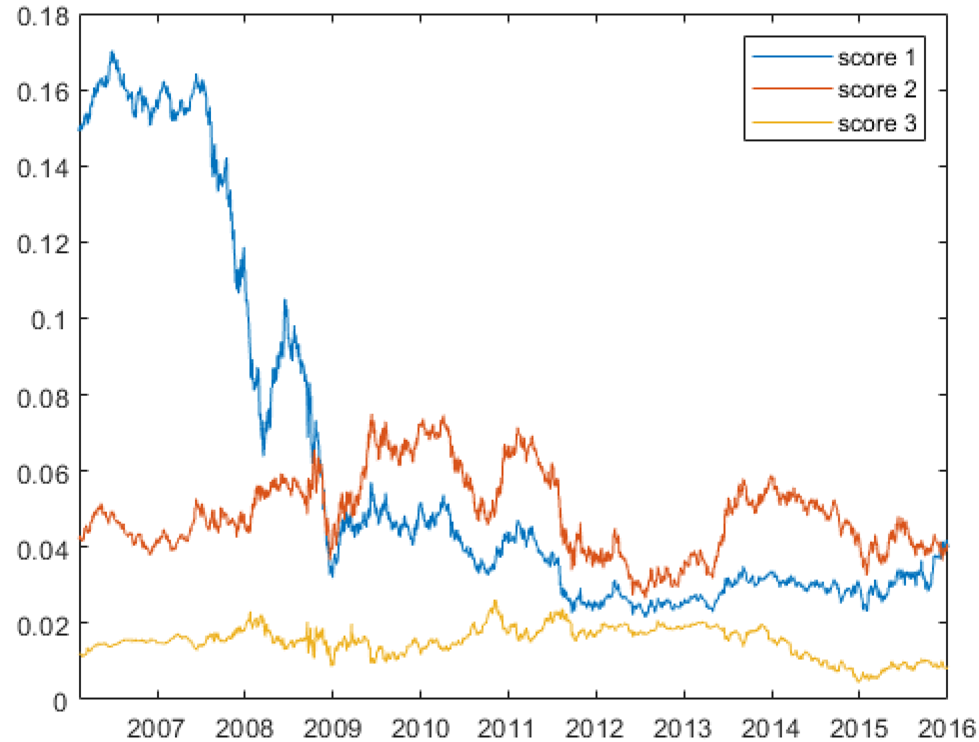

Figure 35 p = 0, missing short tenors

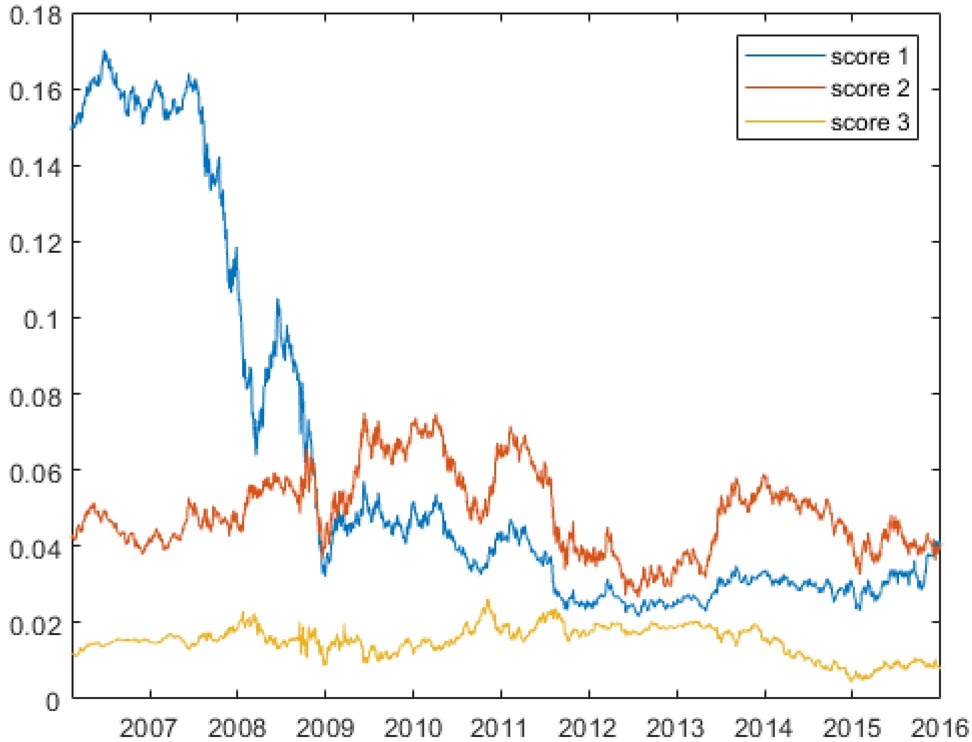

Figure 36 p = 10, missing short tenors

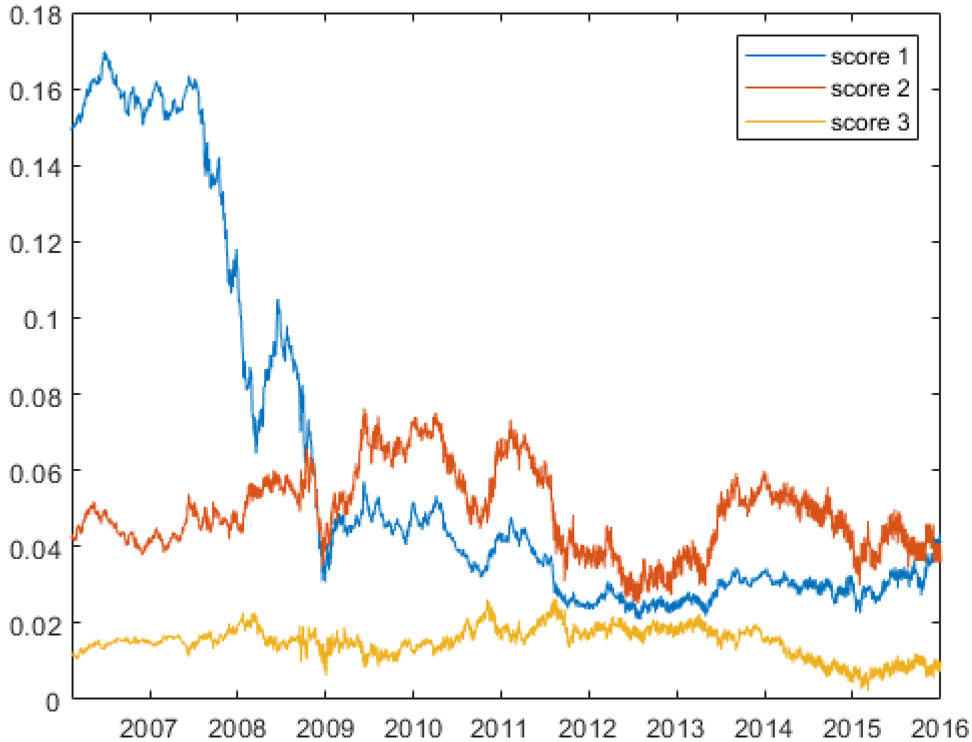

Figure 37 p = 100, missing short tenors

D.3. Missing medium tenors

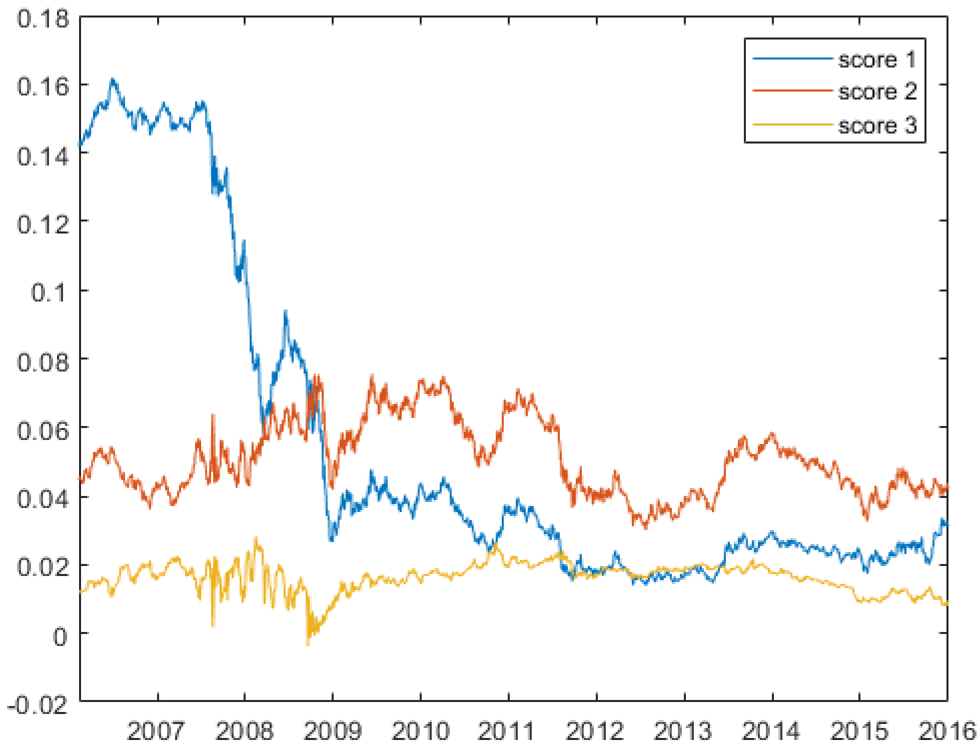

Figure 38 p = 0, missing medium tenors

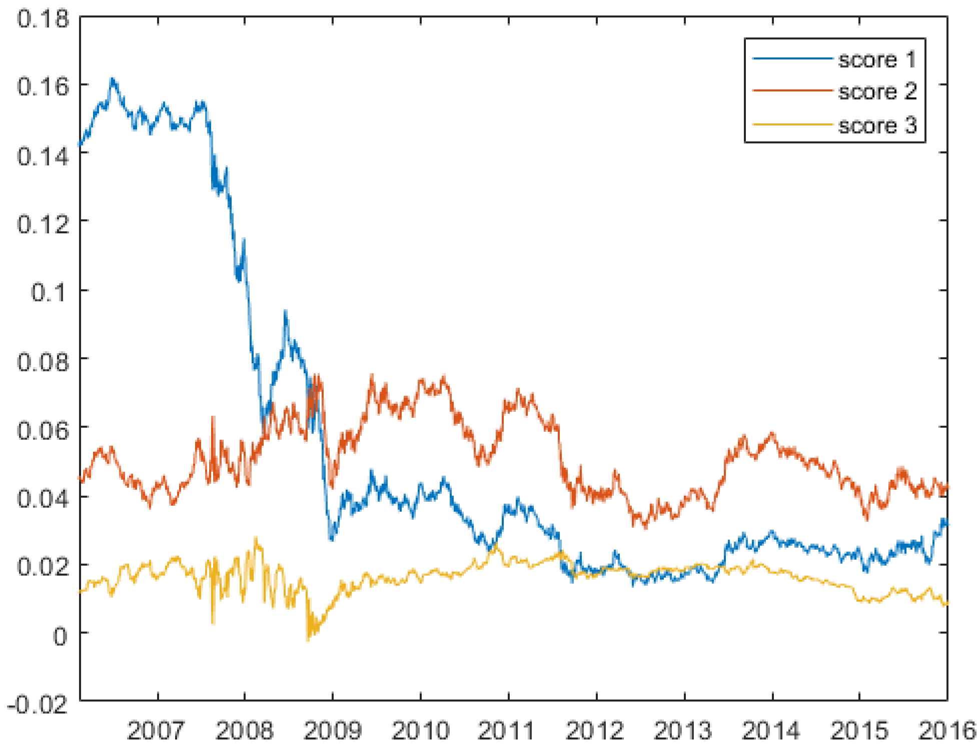

Figure 39 p = 10, missing medium tenors

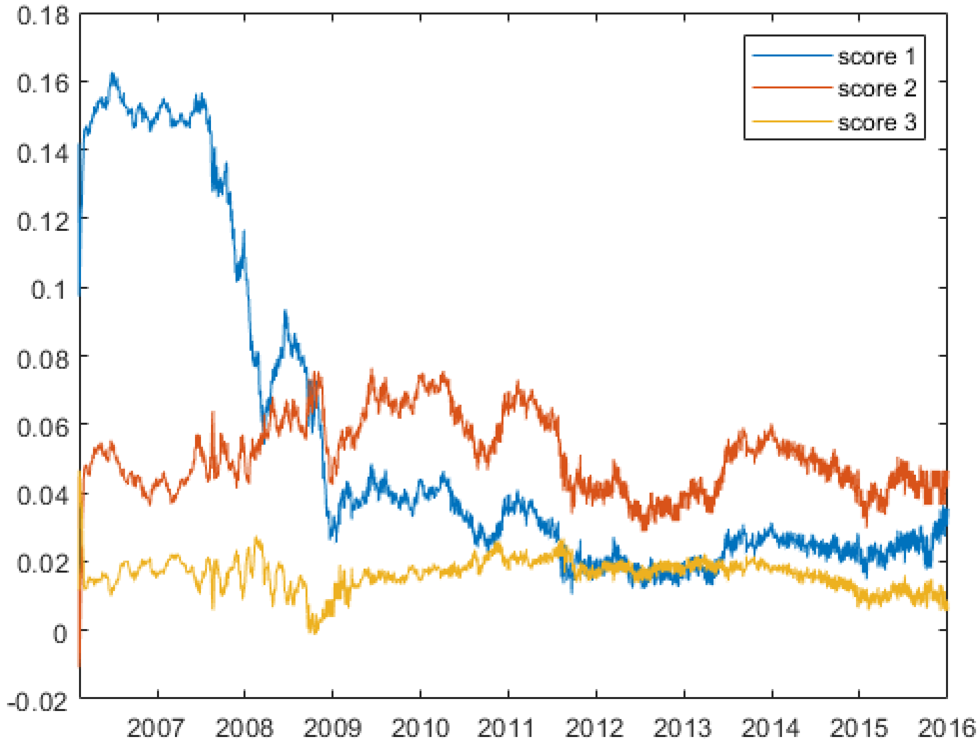

Figure 40 p = 100, missing medium tenors

D.4. Missing long tenors

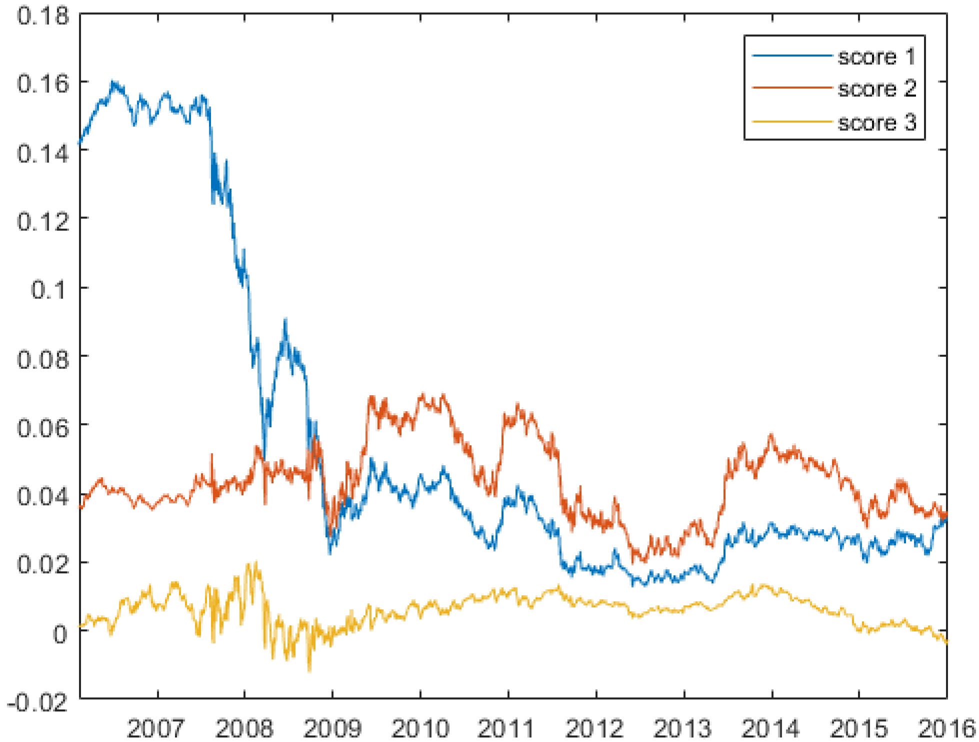

Figure 41 p = 0, missing long tenors

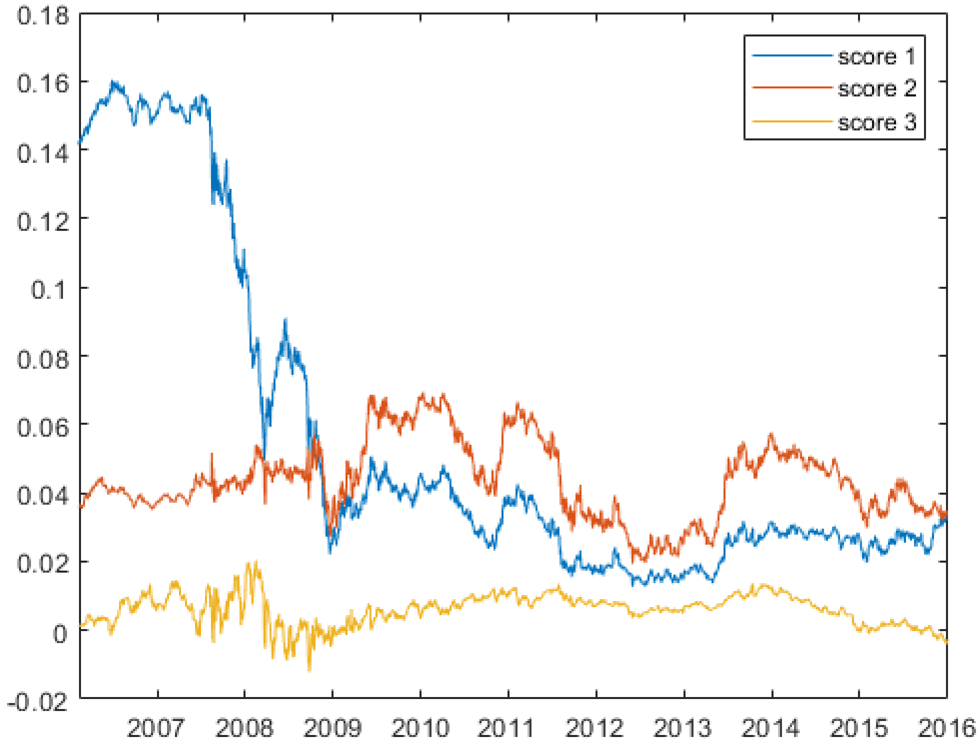

Figure 42 p = 10, missing long tenors

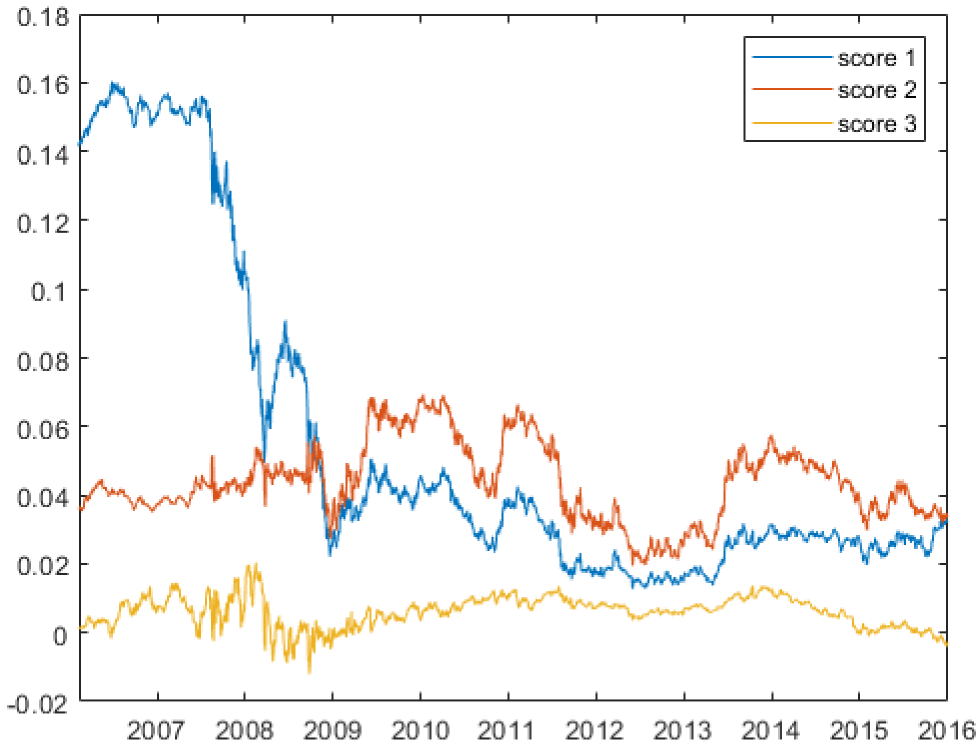

Figure 43 p = 100 missing long tenors