Bid-Ask Spread Modelling in the South African Bond Market

Matthew Shaw

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

September 24, 2018

MPhil in Mathematical Finance,
University of Cape Town.

Declaration

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

Signed by candidate

September 24, 2018

Abstract

Pitsillis and Taylor (2014) calculate bid-ask spread estimates of South African government bonds over a single year, using the models of De Jong and Rindi (2009) and Huang and Stoll (1997). This dissertation tests the effectiveness of both models by comparing the modelled equity spread estimates against the actual equity spread estimates. Furthermore, this dissertation investigates the stability of the De Jong and Rindi (2009) and Huang and Stoll (1997) models in the bond market by extending the spread estimate dataset to run annually over 5 years. The final section of this dissertation proposes a new method of estimating the bond spread through the use of a Kalman filter, as it can be used to leverage information from an on-screen market (albeit a different market) to imply bid-ask spread estimates in an off-screen market. The results indicate that the Huang and Stoll (1997) model consistently outperforms the De Jong and Rindi (2009) model. Furthermore, the yield estimate results of Pitsillis and Taylor (2014) align with the results obtained in this dissertation. The spread estimate results are stable over the 5 year period, indicating a strong provision of liquidity by the Primary Dealers.

Acknowledgements

I would like to express my gratitude to Obeid Mohamed and David Taylor for their supervision and continued support throughout the dissertation. A huge vote of thanks needs to go to Cameron Hutchison, Dirk Van Heeswijk, Angela Millat and Fergus Wegener for proof reading my dissertation. I have to acknowledge the unconditional support and motivation from Cameron, Dirk, Chris Baker and Rowan Dalton. Finally, I have to thank my family for their love and belief in me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

National Treasury is in the process of implementing an electronic trading platform for the secondary market of South African government bonds (Ndzamela, 2015). The current process requires Primary Dealers (five domestic banks and three international banks) to bid every week for government bonds at National Treasury auctions. Once the bonds have been divided amongst the Primary Dealers, Primary Dealers trade the government bonds bilaterally in the secondary market through a request-for-quote driven system. Treasury's (2014) reasons for moving to an electronic trading platform are twofold. National treasury wishes to monitor the activity of the Primary Dealers more closely, in addition to increasing transparency in the South African government bond market. Investors (Brand, 2014) also believe that a shift in market regime to an electronic trading platform will increase liquidity.

Lybek and Sarr (2002) state that market participants perceive a liquid asset, as one that can be quickly sold in large volumes, without negatively affecting the price. By this definition, a liquid financial asset will have a small bid-ask spread, will settle quickly and can be traded in large volumes without drastically affecting the price. As such, Lybek and Sarr (2002) provide evidence of many different financial market liquidity concepts. Market participants will characterise asset liquidity in stable markets by low transaction costs and will characterise asset liquidity in stressed markets by the ability of prices to adjust to their new equilibrium. Since there is not one accepted definition of market liquidity, Lybek and Sarr (2002) place market liquidity into the following five categories : (i) tightness: a tight market will have low transaction costs, i.e. low bid-ask spreads in a quote-driven market, (ii) immediacy: an immediate market will execute orders quickly with fast settlement, (iii) depth: a deep market will have an abundance of buy and sell orders on either side of the traded price, (iv) breadth: a broad market will have large volumes either side of the traded price, resulting in minor changes to traded prices, and (v) resilience: a resilient market is one where the traded prices recover quickly after order imbalances occur.

Treasury (2016), in the 2015/2016 Debt Management Report, comments that liquidity in the South African bond market is resilient because of an adequate turnover ratio and an increase in trade volumes. Treasury (2016) uses the turnover ratio as a measure of bond market liquidity. The turnover ratio is a measure of the secondary market activity against the total outstanding amount. The turnover ratio and larger trade volumes are reflective of Lybek and Sarr's (2002) depth and breadth liquidity characteristic as they are volume-based measures. However, the report is unable to capture the tightness characteristic of the bond market, as there is no historical data of the bid-ask spread.

Clients selling or buying bonds request quotes from the Primary Dealers through different channels (telephone, Bloomberg chat rooms and proprietary client trading screens). Once a trade is executed, the Johannesburg Stock Exchange records only the post-trade information, which is published daily. Hence information pertaining to buy and sell orders is not recorded. In order to analyse liquidity from a market tightness perspective, it is necessary to create a model to measure the bid-ask spread.

A report on SA government bond liquidity and transparency by Pitsillis and Taylor (2014) analyses the bid-ask spread of the South African government bond market using two different methods. The report covers nominal bonds, in addition to inflation-linked bonds. The spread models were tested over one year of JSE trade level data, but the results from the models were not tested against actual spread estimates in a different market. Per se, the aim of this dissertation is to:

1. Adapt the spread models, used in the report of Pitsillis and Taylor (2014), to generate bid-ask spread estimates for equity data. Compare and analyse the modelled spread estimates of the adapted models against the actual spread estimates. Determine whether the models are suitable.

2. Replicate and apply the direct and indirect models of Pitsillis and Taylor (2014) to a larger bond yield dataset (applied over five one-year periods, rather than a single year) to confirm the stability of the spread.

3. Propose a new stochastic model of measuring the bid-ask spread through a Kalman filter, without the need to classify trades as bid quotes or ask quotes.

Chapter 2 details the method used to classify trade direction. Classifying each trade, as either a quoted bid yield or a quoted ask yield, is required for applying the direct and indirect models of Pitsillis and Taylor (2014) to bond yield data. Chapter

3 details and applies the model of De Jong and Rindi (2009) to equity data, in addition to applying an adapted version to the bond yield dataset. Chapter 4 details and applies the Huang and Stoll (1997) model to the equity dataset. Futhermore, Pitsillis and Taylor (2014) adapt the model of Huang and Stoll (1997) for suitable application to the bond yield dataset. Chapter 5 proposes a new method of estimating the spread through the use of a Kalman filter. Chapter 6 concludes the dissertation.

Chapter 2

Trade Classification

The South African bond market is traded over-the-counter (OTC) with eight primary dealers required to post two-sided quotes in a request-for-quote driven market. The trades, recorded by the JSE, contain information pertaining to trade time, trade size, trade yield, trade price and the trade counterparty. These trades are published daily but, unlike electronic books, there is an absence of historical data on sell orders and buy orders.

To measure the bid-ask spread in the South African bond market, Pitsillis and Taylor (2014) use the insights of Lee and Ready (1991) to classify intraday trades as either buyer-initiated trades or as seller-initiated trades. Two approaches are used to infer trade direction. The first approach is to classify trades according to the relationship between traded prices and quoted bid prices and ask prices. The second approach, termed the tick test by Lee and Ready (1991), is to classify trades by comparing trade prices to preceding trade prices. All trades can be categorised as a bid trade or an ask trade using the tick test. In practice, however, the tick test has some shortcomings, as certain trades are unclassifiable because they are either recorded out of sequence, or they are sold with special conditions (e.g. trades can be sold at a premium for early settlement). An alternative to the tick test, is the reverse tick test, which compares trade prices to subsequent trade prices.

In the tick test, the current trade is specified as a buyer-initiated trade (an ask), provided that the current price is above the preceding trade price. In the South African bond market, a trade will be specified, under the tick test, as a seller-initiated trade (a bid), provided that the current trade yield is above the preceding trade yield. Unlike the forward tick test, the reverse tick test compares the current trade price to the subsequent trade price. The current trade is specified as a seller-initiated trade (a bid), given that the current price is below the subsequent trade price. A trade will be specified as a buyer-initiated trade (an ask) in the South African bond market, provided that the current trade yield is above the subsequent trade yield. These two methods of trade classification agree when share prices and bond yields oscillate during trades, but they conflict when price movements occur consecutively in a single direction. For instance, a run of asks in the same direction are correctly classified by the tick test, when trading on prices, but they are incorrectly classified by the reverse tick test. The inverse is true for a run of bid prices. Lee and Ready (1991) clarify which method produces the correct classification by factoring in changes in quotes. Since a record of historical quoted bid prices and quoted ask prices do not exist in the South African bond market, Pitsillis and Taylor (2014) extend the tick methods of Lee and Ready (1991) to form a consolidated tick test. The result is an amalgamation of the tick test and reverse tick test. The consolidated tick test seeks to deal with the problem of incorrect trade classification, by considering the size of the yield change on either side of the traded yield.

The consolidated tick test requires the tick test and the reverse tick test to be applied to bond yield data, acquired from the JSE. The consolidated tick test subsequently selects which test method to use, based on the change in yield size, either side of the traded yield. A larger change in yield between the current yield (at time t) and preceding yield (at time $t-1$), compared to the change in yield between the current and subsequent yield (at time $t+1$), infers a greater market impact from trade time $t-1$ to t. Consequently, the tick test classification is chosen. A greater change in yield from the current yield to the subsequent yield implies a larger market impact after the current trade, therefore, the reverse tick test is chosen. The historical yield data contains small yield changes attributed to bond market microstructure noises. Part of this microstructure noise is due to multiple dealers quoting different clients simultaneously, such that there is no prevailing bid or ask quote at any given time. Pitsillis and Taylor (2014) consulted with the appropriate traders, where it was decided that a one basis point change in yield was sufficient to distinguish between a bid quote and an ask quote. The algorithm that follows highlights how the method of classification is chosen.

Algorithm 1 The Consolidated Tick Algorithm in pseudo-code (Pitsillis and Taylor, 2014)

1) Implement forward tick test classification Q_t^F, store result
2) Implement reverse tick test classification Q_t^R, store result
3) Iterate forward in time to classify trades , $\bar{Q}_t$
$Q_1 = Q_1^R$
$Q_{end} = Q_{end}^F$

if $|y_t - y_{t-1}| < 0.01$ **then**
    Use classification of previous trade, $\bar{Q}_t = \bar{Q}_{t-1}$
else
    if $Q_t^F = Q_t^R$ **then**
        $\bar{Q}_t = Q_t^F$
    else
        if $|y_t - y_{t-1}| < |y_{t+1} - y_t|$ **then**
            $\bar{Q}_t = Q_t^R$
        else
            $\bar{Q}_t = Q_t^F$
        end if
    end if
end if

Where y_t is the yield at time t and $\bar{Q}_t$ is the trade indicator specified below.

$$\bar{Q}_t = \begin{cases} +1 & \text{for a seller-initiated trade (traded against a bid quote)} \\ -1 & \text{for a buyer-initiated trade (traded against an ask quote),} \end{cases}$$

Q_t^R is the indicator for the reverse tick test and Q_t^F is the trade indicator for the tick test. The tick test cannot classify the first trade of a sequence as the first trade needs a prior trade to compare to. As a consequence, the reverse tick test is used to classify the direction of the first trade. The last trade is classified by the tick test, as the reverse tick test cannot classify the final trade direction, due to a lack of a succeeding trade.

Chapter 3

De Jong and Rindi (2009) Spread Estimates

De Jong and Rindi (2009) highlight numerous different spread measures, in addition to their shortfalls. Firstly, they specify the quoted spread as the average difference between the liquidity provider's best prevailing quoted ask prices and quoted bid prices. The quoted spread estimate generally represents the costs to enter and exit a trade.

$$S^Q = \frac{1}{T}\sum_{i=1}^{T}(A_i - B_i) \tag{3.1}$$

The time step, i, represents the time step of a quote change, sourced from either a change in a bid quote or an ask quote. There are several pertinent shortfalls with this measure of spread estimation. Quoted spreads are generally not preferred, since the quotes may only be valid for small volumes and also may not be binding.

De Jong and Rindi (2009) define another method of spread called the effective spread.

$$S^E = \frac{1}{T}\sum_{t=1}^{T}2Q_t(P_t - M_t) \tag{3.2}$$

where time step, t, represents the time step at which a transaction occurs, rather than the time of a change in quote. M_t is the mid-price, P_t is the transaction price and Q_t is the trade indicator. A history of bid quotes, ask quotes and traded prices are recorded by Bloomberg for JSE listed equities. For this reason, trade indicators are known at transaction times. Historical yield data differs from equity data because there are no recorded bid quotes and ask quotes. Pitsillis and Taylor (2014) use the consolidated tick test to classify these traded yields as either trades against a prevailing bid quote or trades against a prevailing ask quote. After the yield data is classified, trade indicators for both equity data and yield data are known. It can be shown that equation 3.2 can be simplified into the form of equation 3.3 below, when trading on prices.

$$S^E = \frac{1}{T}\sum_{t=1}^{T}(A_t - B_t) \tag{3.3}$$

The spread estimator is not unlike equation 3.4, when trading on yields.

$$S^E = \frac{1}{T}\sum_{t=1}^{T}(B_t^{(y)} - A_t^{(y)}) \tag{3.4}$$

Equations 3.3 and 3.4 require a running series of bid quotes and ask quotes, per transaction time step. The effective spread is chosen for this paper over the quoted spread because of the aforementioned problems pertaining to the quote sizes and prices. Running bid quote and running ask quote data for equities is readily available from Bloomberg and hence the spread estimate, for equation 3.3, is easily attainable. To attain the spread estimate for bond yields using equation 3.4, the traded yields need to be separated into a series of running bid yield quotes and running ask yield quotes. With the trade indicators from the output of the consolidated tick test and the traded yields, algorithm 1 is used to create a series of running bid quotes and ask quotes.

Algorithm 2 Decomposing traded yields into a series of bid quotes and ask quotes

if $\bar{Q}_t = +1$ (a seller initiated trade, traded against a bid quote) **then**
    $B_t = y_t$
    $A_t = A_{t-1}$
else
    if $\bar{Q}_t = -1$ (a buyer initiated trade, traded against an ask quote) **then**
        $B_t = B_{t-1}$
        $A_t = y_t$
    end if
end if

Where $\bar{Q}_t$ is the trade indicator, defined by the consolidated tick test of Pitsillis and Taylor (2014). B_t is the running bid yield quote and A_t is the running ask yield quote. A similar algorithm is run on equity data to compare the actual spread estimate to the spread estimate created with the trade classification and traded prices. Algorithm 3 below creates a series of running bid quotes and running ask quotes from traded prices and trade classification.

Algorithm 3 Decomposing traded prices into a series of bid quotes and ask quotes

if $Q_t = +1$ (a buyer initiated trade, traded against an ask quote) **then**
  $B_t = B_{t-1}$
  $A_t = y_t$
else
  if $Q_t = -1$ (a seller initiated trade, traded against a bid quote) **then**
    $B_t = y_t$
    $A_t = A_{t-1}$
  end if
end if

Where Q_t is the trade indicator, given by the Bloomberg historical data. B_t is the running quoted bid price and A_t is the running quoted ask price.

3.1 De Jong and Rindi (2009) Equity Spread Estimate Results

The spread estimate results for equities, measured over a month in 2017, are shown in table 3.1. The actual spread estimate, S^E_{act}, is attained using equation 3.3, where A_t and B_t are the respective quoted ask prices and quoted bid prices at each transaction time. The modelled spread estimate, S^E_{mod}, is also calculated using equation 3.3, but A_t and B_t are the respective running ask levels and running bid levels at each transaction time attained through algorithm 3. The purpose of modelling the De Jong and Rindi (2009) spread estimates is to test how modelled spread estimates differ from the actual spread estimates, in the absence of recorded bid quotes and ask quotes. A small percentage difference in the two values implies the modelled spread estimate is in the correct spread estimate range.

Share Tickers	Observations	S^E_{act} (cents)	S^E_{mod}(cents)	% Difference
BIL	55,133	10.46	13.65	-30.526%
BTI	52,836	48.19	52.10	-8.109%
CFR	18,789	4.98	6.40	-28.427%
FSR	45,628	2.48	2.87	-15.939%
MTN	92,152	6.46	7.51	-16.309%
NBK	36,233	33.10	24.99	24.504%
NPN	16,411	257.95	227.48	11.813%
SBK	125,485	7.61	8.92	-17.190%
SNH	83,982	3.23	3.94	-21.831%
SOL	99,482	22.21	24.08	-8.404%
VOD	83,270	8.04	9.52	-18.388%

Tab. 3.1: Actual equity spread estimates against modelled equity spread estimates.

In general the results show a significant difference between the modelled spread estimates and the actual spread estimates, with 9 out of the 11 modelled spread estimates displaying a difference greater than 10%. The modelled spread estimates of BHP Billiton (BIL) and Compagnie Financiere Richemont SA (CFR) exhibit the greatest deviation from the actual spread estimate, while the modelled spread estimates of Nedbank (NBK) and Naspers (NPN) display a difference in the opposite direction. The difference between the modelled spread estimates and the actual spread estimates originates from two sources, both of which can be identified by examining algorithm 3. Firstly, when a seller-initiated transaction occurs, it is assumed that the quoted ask price remains at the same level as the quoted ask price from the previous trade. Similarly, when a buyer-initiated transaction occurs, it is assumed that the quoted bid price remains at the same level as the quoted bid price from the previous trade. Empirically, quotes from previous transaction times are not always stationary at the subsequent transaction times, regardless of whether the trade is a buyer-initiated transaction or a seller-initiated transaction. This is demonstrated through Figure 3.1 below. The graph on the left hand side of figure 3.1 plots the traded prices, quoted bid prices and quoted ask prices. The graph on the right-hand side plots the traded prices and the series of running bid quotes and ask quotes, attained through the use of algorithm 3.

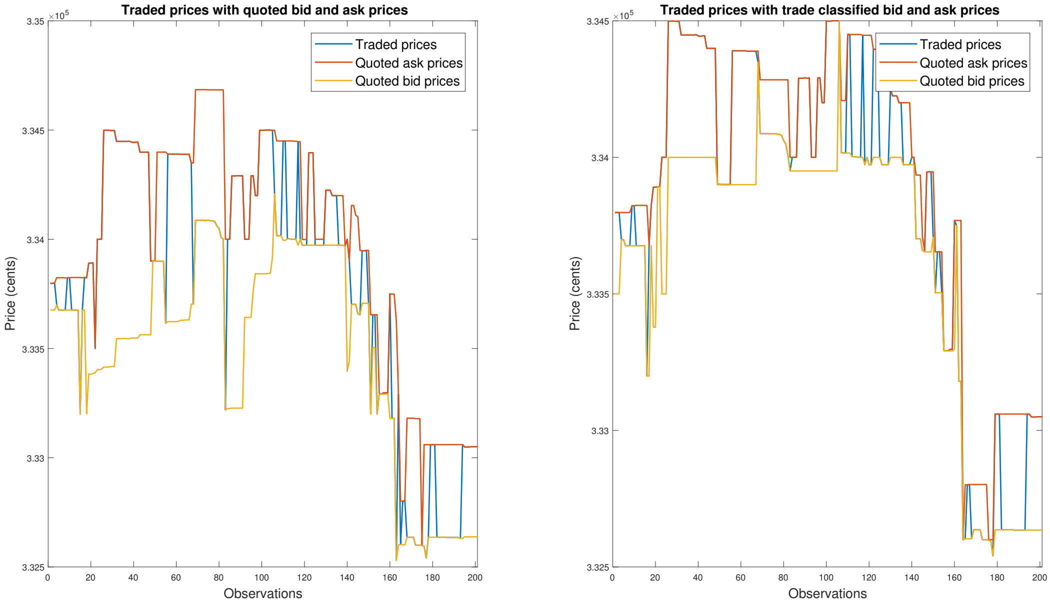


Fig. 3.1: Comparison of the running bid and ask series using quoted spreads and modelled spreads at transaction times.

The graph on the left-hand side depicts the non-static nature of both quotes at transaction times. This is most evident in the middle of the graph where a widening of the spread occurs. The modelled spreads trade tighter in this scenario, as a result of creating the running bid quote and ask quote series using algorithm 3. The actual quoted bid prices do not increase to the same extent as the modelled bid prices do, nor do the actual quoted ask prices stay at a lower level. Figure 3.2 below plots the actual spread and the modelled spread against time and subsequently shows the difference between the two methods.

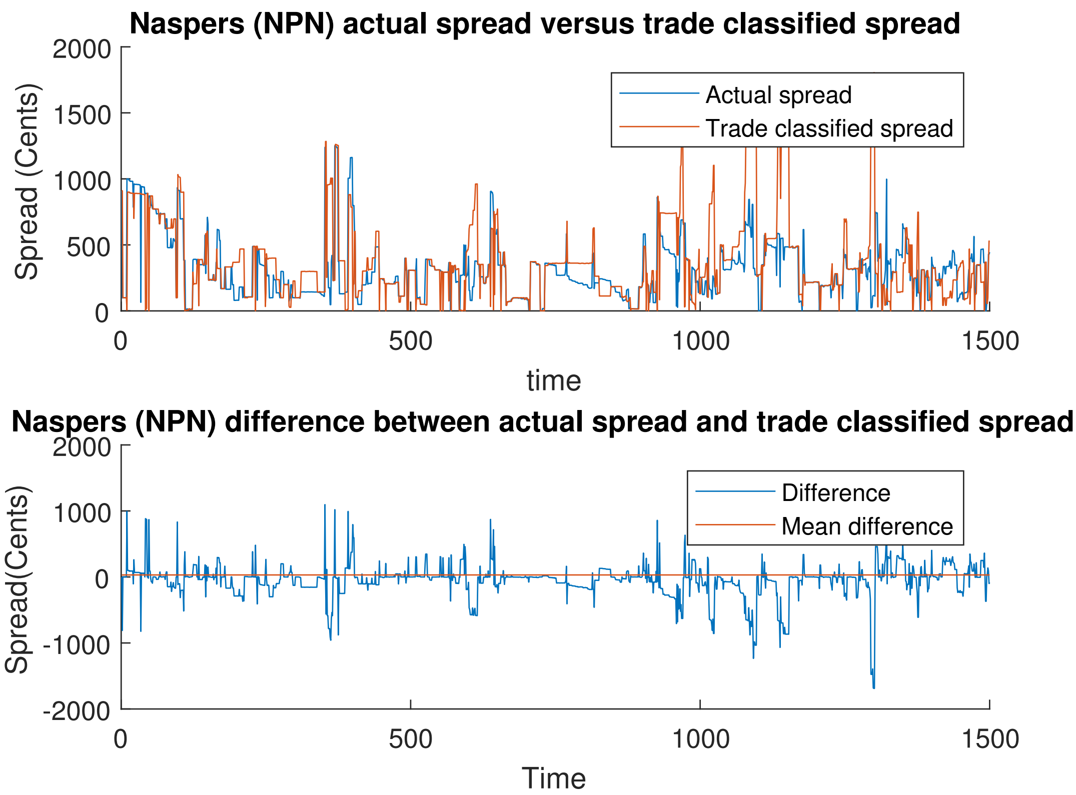


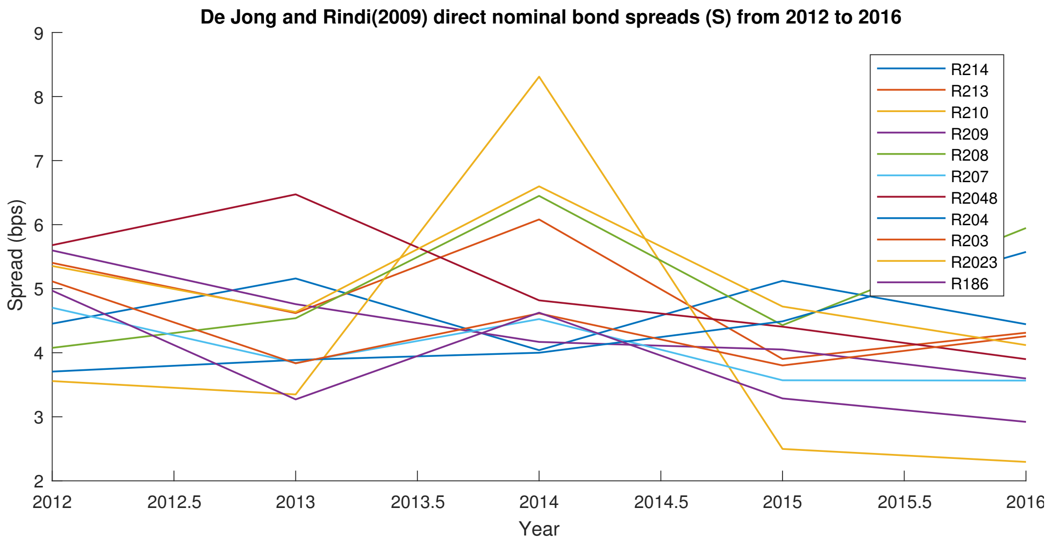
Fig. 3.2: Top: Actual spread of Naspers (NPN) against the modelled spread acquired using algorithm 3. Bottom: Time-series difference between actual spreads and modelled spreads.

Figure 3.2 provides evidence that moving quoted bid prices and moving quoted ask prices cause the difference between the actual spread estimate and the modelled spread estimate. Most of the modelled equity spread estimates overestimated the actual spread estimate. The equity spread estimates with a positive mean spread difference (modelled estimate minus actual estimate) indicate a combination of two movements. When buyer-initiated trades occur (trades against a quoted ask price), the modelled quoted bid price drops more than the actual quoted bid price on average. In addition to this, when seller-initiated trades occur (trades against a quoted bid price), the modelled ask quote increases more than the actual ask quote on average, also widening the spread. The opposite holds true for negative mean differences.

The mean difference between the Naspers actual spread estimates and the modelled spread estimates, shown in figure 3.2, is 30.47 cents. This amounts to 0.009437% of the last traded share price in the series of measurements. The modelled spread as a percentage of the last traded share price is 0.0705%, compared to the actual spread as a percentage of the last traded share price which amounts to 0.0799%. Although the percentage difference of modelled spread estimates to the actual spread estimates are relatively large, the spread estimates from the De Jong and Rindi (2009) method are in the correct range.

3.2 De Jong and Rindi (2009) Bond Spread Estimate Results

The consolidated tick test of Pitsillis and Taylor (2014) is applied to 19 different South African government bonds, both nominal bonds and inflation-linked bonds, to classify trade directions. Since the bond market trades on yield, algorithm 2 is used to generate a series of running quoted bid yields and running quoted ask yields. The yield spread estimate is calculated using equation 3.4. The spread estimates are calculated on a per annum time scale over 5 years to form an indication of spread stability. The nominal bonds and inflation-linked bonds, summarised in table 3.2, are plotted in figure 3.3. The top graph plots the nominal bond spread estimates, while the bottom graph plots the inflation-linked bond spread estimates. Bonds maturing before 2016 are omitted for the 5 year plot. The behaviour of the spreads is made easier to visualise by splitting nominal bonds and inflation-linked bonds in two. The more frequently traded nominal bonds display a stable spread between 3 and 6 basis points over the 5-year period. Most of the nominal bond spread estimates seem to display a tightening of the spread over time, with the notable exception of a spike in 2014 for most nominal bonds. This spike does not apprear evident in the inflation-linked bond spread estimates. The benchmark R186 bond, the most frequently traded bond, displays the lowest average nominal bond spread for the 5 years. In the inflation-linked market, the directional behavior of the spread estimates are pronounced. The 2012 spread estimates are banded between a 2.5 and 8 basis point range. A noticeable tightening of the spread estimates occurs on all of the inflation-linked bonds over the 5-year period. The 2016 spread estimates sit in a 2 basis point to 2.75 basis point range. In the report of Pitsillis and Taylor (2014), the appropriate traders agree on an approximate spread between 4 and 6 basis points, and note the smaller than expected inflation-linked bond spread.



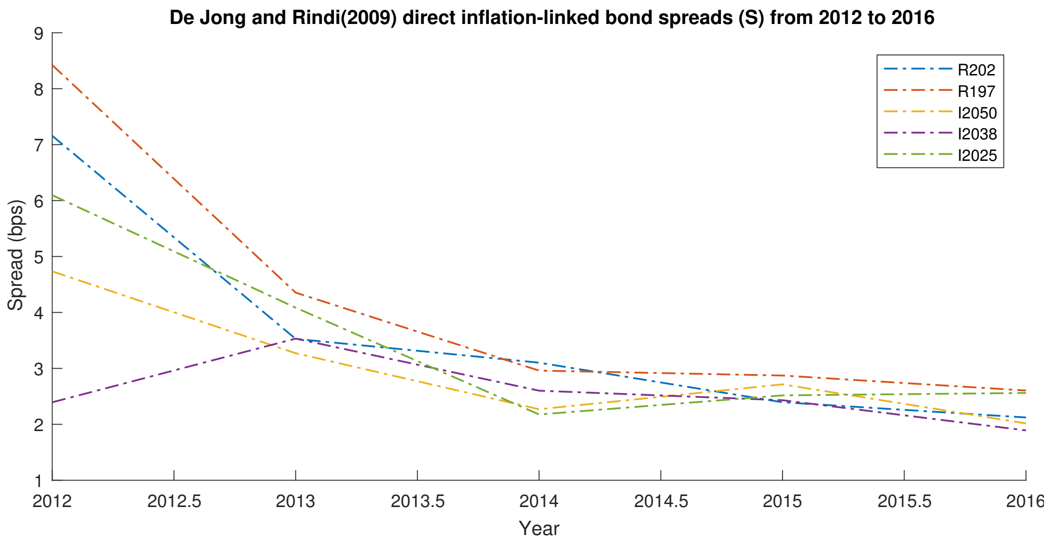


Fig. 3.3: Top: Nominal bond spread estimates over 5 years. Bottom: Inflation-linked bond spread estimates over 5 years.

	2012		2013		2014		2015		2016	
Bond List	Obs.	S(bps)	Obs.	S(bps)	Obs.	S(bps)	Obs.	S(bps)	Obs.	S(bps)
R214	9388	4.4533	7641	5.1590	18424	4.0397	6839	5.1222	5878	4.444
R213	10923	5.4027	8860	4.6186	19087	6.0807	7471	3.9022	6652	4.3120
R210	2302	3.5572	2198	3.348	2412	8.3094	1921	2.4977	1285	2.2951
R209	16366	5.5970	14416	4.7583	22854	4.1686	12123	4.0499	9951	3.5979
R208	19657	4.0761	10962	4.5388	21937	6.4487	8149	4.4296	7236	5.9465
R207	15140	4.7025	8083	3.8416	20393	4.5250	7596	3.5695	6352	3.5650
R206	5075	5.8269	781	4.1525	*		*		*	
R2048	3891	5.6793	5284	6.4722	15014	4.8172	9212	4.4063	10614	3.899
R204	14668	3.7054	6996	3.8871	16677	3.9998	6739	4.4896	6077	5.5732
R203	14400	5.1134	7789	3.8343	18364	4.6139	7172	3.801	3694	4.2577
R2023	6416	5.3541	8040	4.6350	19173	6.5977	7055	4.7212	7546	4.1188
R202	3089	7.1576	1640	3.5276	4959	3.1021	2487	2.3945	2355	2.1224
R201	5207	5.9366	1576	5.085	1565	9.1719	*		*	
R197	2428	8.4220	1630	4.3543	3835	2.9606	1905	2.8720	2099	2.6042
R186	43932	4.9683	57476	3.2712	80034	4.6251	67255	3.2860	65880	2.9203
R157	41068	5.7171	19660	3.2217	*		*		*	
I2050	481	4.7346	1124	3.2710	2181	2.2698	2125	2.7148	1760	2.0165
I2038	635	2.3940	1262	3.5324	3071	2.6014	1854	2.4309	1624	1.8911
I2025	822	3.0963	1370	4.0853	3435	2.1759	2393	2.5176	1825	2.5607

Tab. 3.2: Direct spread estimates for all standard spot trades from 2012 to 2016
Bonds marked with * have expired and hence no measurement was taken.

Chapter 4

Huang and Stoll (1997) and Pitsillis and Taylor (2014) Indirect Spread Models

This section of the dissertation explains the derivation and implementation of the Huang and Stoll (1997) indicator variable model to estimate the bid-ask spread. This section also details the adaptation of the Huang and Stoll (1997) indicator variable model, by Pitsillis and Taylor (2014), to suit the South African bond market, as South African bonds trade on yields rather than prices. Huang and Stoll's (1997) indicator variable model is a time-series model, incorporating historical transaction prices, the bid-ask spread and quoted prices. The modelled bid-ask spread is fragmented, by Huang and Stoll (1997), into three components, namely adverse selection (asymmetric information) costs, inventory holding costs and order processing costs. Adverse selection costs are described by Kyle (1985), as the costs imposed due to an information asymmetry between market makers and traders. Kyle (1985) believes that adverse selection costs emerge from traders having greater knowledge of the value of a share, over the providers of liquidity. Market makers have costs implicit in holding a portfolio of stocks and bonds. Ho and Stoll (1983) posit that these holding costs arise from trades in other bonds or stocks. When an order is executed for a specific bond or stock, market makers adjust their quotes in order to hedge their inventory. The final component of the bid-ask spread is the order processing cost. The bid-ask spread contains an order processing component during trade execution, comprising several obvious costs and a few hidden costs. Obvious costs include brokerage fees, clearing costs, back office costs, tax costs and the dealers margin, whilst some of the hidden costs arise from missed opportunities, such as inability to execute an order at a favorable time. The Huang and Stoll (1997) model and the adaptation to this model, by Pitsillis and Taylor (2014), are trade indicator models with the benefit of simplicity, since the models do not contain complicated lag structures. Additionally, the models are also robust against the innumerable statistical issues such as heteroskedasticity, serial correlation and measurement errors.

4.1 Model Derivation

The Huang and Stoll (1997) indicator variable model begins with an unobservable fair share value, V_t, and in the case of Pitsillis and Taylor (2014), an unobservable fair bond yield, y_t^*, before the placement of bid-ask quotes:

$$V_t = V_{t-1} + \alpha \frac{S}{2} Q_{t-1} + \varepsilon_t \tag{4.1}$$

$$y_t^* = y_{t-1}^* + \alpha \frac{S}{2} Q_{t-1} + \varepsilon_t \tag{4.2}$$

where V_t (y_t^*) is the fundamental unobservable share value (bond yield) at time t, absent of transaction costs. The percentage of half spread, attributable to adverse selection, is represented by α and the spread is denoted by S. Q_{t-1} is a buy-sell trade indicator, where the $\alpha \frac{S}{2} Q_{t-1}$ term represents the private information revealed from the previous trade. The serially uncorrelated shock from public information is represented by ε_t. When the α term is equal to 0, the fundamental share (bond) value changes exclusively from the random flow of public information. A mid-price equivalent to the fundamental value and the α term equal to 1, implies that the trader buying/selling the instrument from/to the market maker knows the true fundamental share (bond) value, in the absence of other costs. When α is greater than 1, it implies that the trader has greater knowledge of the fundamental instrument value than the market maker. The buy-sell trade indicators are opposite for the Huang and Stoll (1997) model and the Pitsillis and Taylor (2014) adaptation of the model, since bids are below asks when trading on prices, whilst the inverse holds true when trading on yields. The trade indicator for the Huang and Stoll (1997) model is shown below

$$Q_t = \begin{cases} +1 & \text{when the transaction initiated by a buyer} \\ -1 & \text{when the transaction is initiated by a seller} \end{cases}$$

Whilst the trade indicator for the Pitsillis and Taylor (2014) adaptation is the opposite.

$$\bar{Q}_t = \begin{cases} -1 & \text{when the transaction initiated by a buyer} \\ +1 & \text{when the transaction is initiated by a seller} \end{cases}$$

The fundamental and unobservable value of the share (V_t) and the yield (y_t^*) is directly dependent on its previous value, the private information of more informed traders from the last trade and the current shock from public information.

The observable components of the Huang and Stoll (1997) model are the mid-price (M_t) and the traded price (P_t). The observable components of the Pitsillis and Taylor (2014) adaptation to the model are, similarly, the mid-yield ($M_t^{(y)}$) and the traded yield (y_t). According to inventory theory literature by Stoll (1978), providers of liquidity change the mid-price (M_t) relative to the fundamental value (V_t) based on the total inventory holding. The intent of the fundamental value adjustment is to invoke trades that are inventory equilibrating. The relationship between the mid-quote and the fundamental value for the Huang and Stoll (1997) model and the Pitsillis and Taylor (2014) model is shown by equation 4.3 and equation 4.4 respectively.

$$M_t = V_t + \beta \frac{S}{2} \sum_{i=1}^{t-1} Q_i \tag{4.3}$$

$$M_t^{(y)} = y_t^* + \beta \frac{S}{2} \sum_{i=1}^{t-1} \bar{Q}_i \tag{4.4}$$

The mid-price (mid-yield) is related to the inventory holding costs (β), as a proportion of half of the spread, and the accumulated inventory ($\sum Q_i$). Inventory accumulation occurs from the time that the market opens at Q_1 (the inventory that was previously accumulated) up until the last trade at Q_{t-1}. In the absence of inventory holdings, the mid-price (mid-yield) would be equivalent to the fair value. However, market makers bear the cost of holding inventory and this cost is consequently reflected in the quoted prices. The more expensive it is for the market maker to hold inventory (i.e. the higher the β), the more exaggerated the effect on the mid-price. A run of buyer-initiated transactions results in a positive accumulated inventory ($\sum Q_i$) term, consequently increasing the mid-price relative to the fundamental value. Market makers would have a net short position as a result of selling this inventory. The higher mid-price makes it more expensive for traders to buy inventory, preventing a larger net short position. Inversely, a run of seller-initiated transactions results in a negative accumulated inventory ($\sum Q_i$) term, consequently decreasing the mid-price relative to the fundamental value. Market makers would have a net long position as a result of buying inventory. The lower mid-prices result in market makers realising higher losses on accumulated inventory, further preventing an accumulation of inventory. An expression for the changes in mid-prices and mid-yields is attained by taking the difference between the current and previous time step of equation 4.3 and equation 4.4 respectively.

$$\Delta M_t = M_t - M_{t-1} = \Big(V_t + \beta\frac{S}{2}\sum_{i=1}^{t-1} Q_i\Big) - \Big(V_{t-1} + \beta\frac{S}{2}\sum_{i=1}^{t-2} Q_i\Big) = \Delta V_t + \beta\frac{S}{2}Q_{t-1} \quad (4.5)$$

$$\Delta M_t^{(y)} = \Delta y_t + \beta\frac{S}{2}\bar{Q}_{t-1} \quad (4.6)$$

By substituting in equation 4.1 into equation 4.5 and equation 4.2 into equation 4.6, the following expressions are acquired.

$$\Delta M_t = (\alpha + \beta)\frac{S}{2}Q_{t-1} + \varepsilon_t = (\lambda)\frac{S}{2}Q_{t-1} + \varepsilon_t \quad (4.7)$$

$$\Delta M_t^{(y)} = (\alpha + \beta)\frac{S}{2}\bar{Q}_{t-1} + \varepsilon_t = (\lambda)\frac{S}{2}\bar{Q}_{t-1} + \varepsilon_t \quad (4.8)$$

where the dealers' cost of liquidity (λ), excluding order processing costs, is defined as the sum of the inventory holding cost as a percentage of half spread (β) and the adverse selection cost as a percentage of half spread (α). This implies that the change in the mid-price and mid-yield is driven by the revealed information asymmetry and the inventory holding cost from the previous trade. Huang and Stoll (1997) denote the observed traded price by P_t below.

$$P_t = M_t + \frac{S}{2}Q_t + \eta_t \quad (4.9)$$

In the case of Pitsillis and Taylor (2014), the traded yield (y_t) is denoted similarly.

$$y_t = M_t^{(y)} + \frac{S}{2}\bar{Q}_t + \eta_t \quad (4.10)$$

Equation 4.9 above relates the mid-price, trade direction and the spread to the traded price. Similarly, equation 4.10 relates the mid-yield, trade direction and the spread to the traded yield. The spread term in equation 4.9 and equation 4.10 is not time dependent, consequently specifying a constant spread assumption. Since traded prices and traded yields do not sequentially change by a constant half spread amount each time step, an error term is introduced. This error term, η_t, performs two functions. The first function is to account for the deviation in spread per time step, while the second function is to allow for price discreteness in traded prices and traded yields. Huang and Stoll (1997) combine equation 4.9 and 4.7, while Pitsillis and Taylor (2014) combine equation 4.10 and 4.8 to acquire a basic regression model for changes in traded prices and traded yields respectively.

$$\Delta P_t = \frac{S}{2}(Q_t - Q_{t-1}) + \lambda\frac{S}{2}Q_{t-1} + e_t \quad (4.11)$$

$$\Delta y_t = \frac{S}{2}(\bar{Q}_t - \bar{Q}_{t-1}) + \lambda\frac{S}{2}\bar{Q}_{t-1} + e_t \tag{4.12}$$

Both Huang and Stoll (1997) and Pitsillis and Taylor (2014) represent the error term, e_t, as the sum of the error term from equation 4.7 and 4.8, ε_t, and the change in error term from equation 4.9 and 4.10, $\Delta\eta_t$. The regression models, equation 4.11 and equation 4.12, are non-linear trade indicator models which provide estimates for the traded spread (S) and the adjustments to the spread due to adverse selection and inventory holding costs. The two individual spread components cannot be separated out using equation 4.11 and 4.12. However, the residual component of the spread (the order-processing cost as a percentage of half spread), can be calculated as $1 - \lambda$. Labour costs, equipment costs and the dealer's margin are built into the order processing component of the spread.

4.2 Parameter Estimation of the Regression Models

The parameters for equation 4.11 and equation 4.12 can be estimated using a least squares procedure or a maximum likelihood procedure, both of which have strong distributional assumptions. There are practical complications, highlighted in the paper of Huang and Stoll (1997), when estimating parameters of a model which requires discrete prices. For this reason, Huang and Stoll (1997) adopt the Generalised Method of Moments (GMM) approach, as it is lighter on distributional assumptions. This is important for the error term containing rounding errors, in addition to accounting for conditional heteroskedasticity. Huang and Stoll (1997) and Pitsillis and Taylor (2014) begin the estimation procedure of equation 4.11 and equation 4.12 respectively, by defining a vector function.

$$f\left(x_t, \omega\right) = \begin{bmatrix} e_t Q_t \\ e_t Q_{t-1} \end{bmatrix}$$

$$f^{(y)}\left(x_t^{(y)}, \omega\right) = \begin{bmatrix} e_t \bar{Q}_t \\ e_t \bar{Q}_{t-1} \end{bmatrix}$$

where e_t corresponds to noise term in equation 4.11 and equation 4.12.

$$f\left(x_t, \omega\right) = \begin{bmatrix} (\Delta P_t - S/2(Q_t - Q_{t-1}) - \lambda S/2 Q_{t-1})Q_t \\ (\Delta P_t - S/2(Q_t - Q_{t-1}) - \lambda S/2 Q_{t-1})Q_{t-1} \end{bmatrix}$$

$$f^{(y)}\left(x_t^{(y)}, \omega\right) = \begin{bmatrix} (\Delta y_t - S/2(\bar{Q}_t - \bar{Q}_{t-1}) - \lambda S/2\bar{Q}_{t-1})\bar{Q}_t \\ (\Delta y_t - S/2(\bar{Q}_t - \bar{Q}_{t-1}) - \lambda S/2\bar{Q}_{t-1})\bar{Q}_{t-1} \end{bmatrix}$$

Where the row vectors for the Huang and Stoll (1997) model and the Pitsillis and Taylor (2014) model are shown below.

$$x_t = [\Delta P_t, \Delta Q_t, Q_t]$$

$$x_t^{(y)} = [\Delta y_t, \Delta \bar{Q}_t, \bar{Q}_t]$$

These are the observed variables, where ΔP_t is the change in equity price, sourced from Bloomberg for multiple different stocks and Δy_t is the change in traded yield for multiple South African government bonds, sourced from the JSE's historical trade level data. Historical yields are acquired for both nominal bonds and inflation-linked bonds. The trade direction for equity data, Q_t, is also sourced from Bloomberg, as a record of seller-initiated trades and buyer-initiated trades exists. The trade direction for the yield data, $\bar{Q}_t$, is assumed to be observed. Since trade-level data from the JSE only contains historically traded yields, a proxy measurement is defined using the consolidated tick test trade classification from Chapter 2.

Huang and Stoll (1997) exploit the assumption, used in the paper of Hansen (1982), that functions of an observed set of variables are orthogonal to the noise in an econometric model. The GMM model used by Huang and Stoll (1997) consequently implies the following moment condition to the regression equations.

$$\mathbb{E}[f(x_t, \omega)] = 0$$

$$\mathbb{E}[f^{(y)}(x_t^{(y)}, \omega)] = 0$$

Parameter estimates, λ and S, are then chosen to minimize the following norm function.

$$J_T(\omega) = g_T(\omega)' S_T g_T(\omega)$$

Where S_T is the symmetric observation weighting identity matrix and $g_T(\omega)$, shown below, is the sample mean of $f(x_t, \omega)$ or $f(x_t^{(y)}, \omega)$.

$$g_T(\omega) = \frac{1}{T} \sum_{i=1}^{T} [f(x_i, \omega)]$$

$$g_T(\omega) = \frac{1}{T} \sum_{i=1}^{T} [f(x_i^{(y)}, \omega)]$$

Hansen (1982) proves that the GMM estimator, (ω_t), is steady for a basic trade indicator model under weak regulatory circumstances and

$$\sqrt{T}(\hat{\omega}_T - \omega_0) \to N(0, \Omega))$$

where Ω is defined as

$$\Omega = (D_0' S_0^{-1} D_0)^{-1}$$

where D_0 and S_0 are as follows:

$$D_0 = \mathbb{E}\left[\frac{\partial f}{\partial S}, \frac{\partial f}{\partial \lambda}\right]$$

$$S_0 = \mathbb{E}[f(x_t, \omega) f(x_t, \omega)]$$

$$S_0 = \mathbb{E}[f(x_t^{(y)}, \omega) f(x_t^{(y)}, \omega)]$$

Standard errors are estimated by using the sample means that correspond to D_0 and S_0.

4.3 Huang and Stoll (1997) Equity Spread Results

The intention of testing the Huang and Stoll (1997) model on equity data is to investigate how far out the spread estimates could be. Testing this model on equity data also enables a comparison between the Huang and Stoll (1997) method and the De Jong and Rindi (2009) method. The results for the equity spread estimates, using the method of Huang and Stoll (1997), are shown in table 4.1. Similary to chapter 3, actual spread estimates are compared to modelled (Huang and Stoll (1997)) spread estimates. Hence, the models are applied to the same set of JSE listed shares. The efficacy of the Huang and Stoll (1997) method is determined by the percentage difference of the modelled spread estimate to the actual spread estimate. Availability of the historical quoted bid prices and quoted ask prices allows for the comparison of spread estimates, but it does not confirm the dealers' cost of liquidity (λ).

Share Ticker	Obs.	S_{act} (cents)	S (cents)	Std. error	λ	Std. error	% Difference
BIL	55,133	10.46	11.47	0.1346	0.291	0.0045	9.61%
BTI	52,836	48.19	48.95	0.5418	0.204	0.0040	1.58%
CFR	18,789	4.98	5.73	0.0901	0.226	0.0059	15.02%
FSR	45,628	2.48	2.47	0.0221	0.187	0.0032	-0.36%
MTN	92,152	6.46	6.41	0.0762	0.198	0.0029	-0.75%
NED	38,038	33.10	33.21	0.39	0.15	0.0041	0.33%
NPN	56,832	257.95	234.01	2.60	0.10	0.0025	-9.28%
SBK	125,485	7.61	7.73	0.0572	0.162	0.0019	1.59%
SNH	83,982	3.23	3.55	0.0324	0.182	0.0025	9.95%
SOL	99,482	22.21	23.67	0.1733	0.131	0.0019	6.59%
VOD	83,270	8.04	8.37	0.0998	0.140	0.0024	4.03%

Tab. 4.1: Actual spread estimates against Huang and Stoll (1997) spread estimates, along with the cost liquidity (λ) as a percentage of spread and their standard errors.

The spreads estimates using the Huang and Stoll (1997) model are overestimated in 8 out of the 11 shares tested. The greatest percentage difference of the Huang and Stoll (1997) spread estimates is the Compagnie Richemont (CFR) share at 15.02%, compared to the 30.53% of the De Jong and Rindi (2009) spread estimates. Furthermore, the mean percentage difference using the Huang and Stoll (1997) method is 3.48%, compared to the 11.71% using the De Jong and Rindi (2009) method. The spread estimate standard errors of all the measured shares are small relative to the spread estimate, indicating an accurate estimation procedure. Although the model will not be able to get the exact spread estimate, because of the absence of bid and ask quotes, having a model that is on average -3.5% out still allows for a deduction of the spread of bond yields. When using the percentage difference as a measure of model accuracy, the equity spread estimates attained through the Huang and Stoll (1997) method perform better than those attained through the De Jong and Rindi (2009) method.

4.4 Pitsillis and Taylor (2014) Bond Yield Spread Results

Since the aim of this dissertation is to model the bid-ask spread in the South African bond market, the adapted Huang and Stoll (1997) spread model carried out by Pitsillis and Taylor (2014) is implemented to measure the spread over a 5-year time horizon. The data in the report of Pitsillis and Taylor (2014) was run over one year. Extending the data set to 5 years confirms the stability of the bid-ask spread.

All Trade Results

The spread estimate results for all trade types align with the De Jong and Rindi (2009) spread estimate results from chapter 3. The results of the bond spreads and dealers' cost of liquidity for all trade types are summarised in table 4.2 and illustrated in figure 4.1. Figure 4.1 splits the inflation-linked bonds and the nominal bonds into two separate graphs, as the inflation-linked bond spread estimates appear to trade at tighter levels than the nominal bond spreads. The top graph, in figure 4.1, displays the annual spread estimates of the South African nominal bonds. These spreads trade mostly in a $3 - 4.5$ basis point range. The Huang and Stoll (1997) spread estimates trade tighter than the De Jong and Rindi (2009) spread estimates, where both methods have contrasting results in 2014. The De Jong and Rindi (2009) results show a spike in 2014, unlike the Huang and Stoll (1997) results which show an annual spread reduction. A similar result is painted for both methods in the inflation-linked bond market. Both methods show a reduction in spread estimates and a convergence to a $1.8 - 2.6$ basis point range. The dealers' cost of liquidity for the majority of all 19 bonds sits around $0.2\% - 0.4\%$ in 2012 and gradually reduces to a $10\% - 20\%$ range in 2016. As a result, the order processing costs $(1 - \lambda)$ increase over the 5 year time period. The variability in the order processing costs is due to the dealers' margin, given that the trading infrastructure costs are closer to fixed costs. A decreasing cost of liquidity implies lower inventory holding costs as a percent and a higher dealers' margin, given a stable spread level. The R186 bond spread decreases from 2.95 to 2.7 basis points from 2012 to 2016. Over the same time, the dealers' cost of liquidity decreases from 40.39% to 27.83%. This could be as a results of decreased inventory holding costs from increased liquidity and an increased dealers' margin to maintain similar profit levels to previous years. The inflation-linked bond trades and nominal bond trades display similar spread levels to the Pitsillis and Taylor (2014) report, where emphasis must be placed on the market practitioners' view that the inflation-linked bond results are smaller than expected.

	2012			2013			2014			2015			2016		
Bond	Trades	S (bps)	λ	Trades	S (bps)	λ	Trades	S (bps)	λ	Trades	S (bps)	λ	Trades	S (bps)	λ
R186	26794	2.9495	0.40395	61128	3.3652	0.37340	52304	2.3861	0.35306	67737	3.0947	0.33018	66488	2.7068	0.27835
R157	24972	2.6189	0.50797	20617	3.4550	0.40132	*						*		
R208	8507	3.6138	0.40774	11529	4.1448	0.41153	9669	3.5002	0.36191	8168	4.1866	0.34663	7273	5.9035	0.27838
R209	8163	2.9532	0.31412	15128	3.9070	0.26631	12146	3.2315	0.25811	12155	3.4117	0.30832	9992	3.0675	0.31316
R203	5764	4.2773	0.42410	8155	4.0055	0.35641	7763	3.4886	0.36830	7204	3.4999	0.35422	3720	3.7648	0.33735
R207	5685	4.3946	0.29249	8502	3.6043	0.38731	7257	3.5717	0.26670	7621	3.5759	0.31278	6367	3.2913	0.26157
R204	4725	3.5584	0.39243	7321	4.1617	0.38239	6694	3.5533	0.43517	6757	4.1972	0.44146	6087	5.3196	0.35091
R213	4404	3.4088	0.36363	9344	4.1247	0.32160	8256	3.9107	0.42330	7507	3.5783	0.37731	6660	3.7904	0.28172
R214	3370	3.5758	0.42315	8098	4.5730	0.29972	7960	3.3226	0.29619	6865	4.4388	0.36890	5921	3.8611	0.26952
R2023	3311	3.6077	0.27231	8432	4.2090	0.28336	9323	3.0536	0.29458	7080	3.6391	0.28873	7569	3.6231	0.24195
R201	2042	4.9951	0.39520	1611	4.4711	0.45562	233	3.1817	0.47496				*		
R2048	1398	4.2180	0.29853	5586	5.2998	0.29159	1378	2.8320	0.21998	9238	3.9741	0.26226	10640	3.6174	0.18820
R210	1295	3.5166	0.15684	2320	4.5281	0.11932	7335	2.9081	0.24346	1934	2.3892	0.20411	1296	2.1367	0.16071
R206	1266	8.1694	0.39917	813	4.5316	0.45013	*						*		
R197	1213	3.0574	0.18676	1699	3.9626	0.25236	1693	2.6580	0.21005	1923	2.5729	0.18721	2109	2.4947	0.09910
R202	977	3.0363	0.28705	1731	3.3758	0.26369	1691	2.2504	0.21926	2504	2.1346	0.15773	2364	2.1326	0.09614
I2025	547	2.5778	0.01626	1453	3.4341	0.21921	1588	2.3019	0.21370	2407	2.4365	0.21035	1829	2.6907	0.15047
I2038	412	2.6575	0.20981	1332	3.1810	0.18288	1643	1.9738	0.19971	1871	2.3161	0.15704	1632	1.9661	0.13308
I2050	360	2.5084	0.29137	1186	3.3142	0.16976	1262	1.8808	0.22668	2138	2.2155	0.20255	1774	2.0240	0.14851

Tab. 4.2: Spread estimates and cost of liquidity (λ) estimates for all standard spot trades from 2012 to 2016.

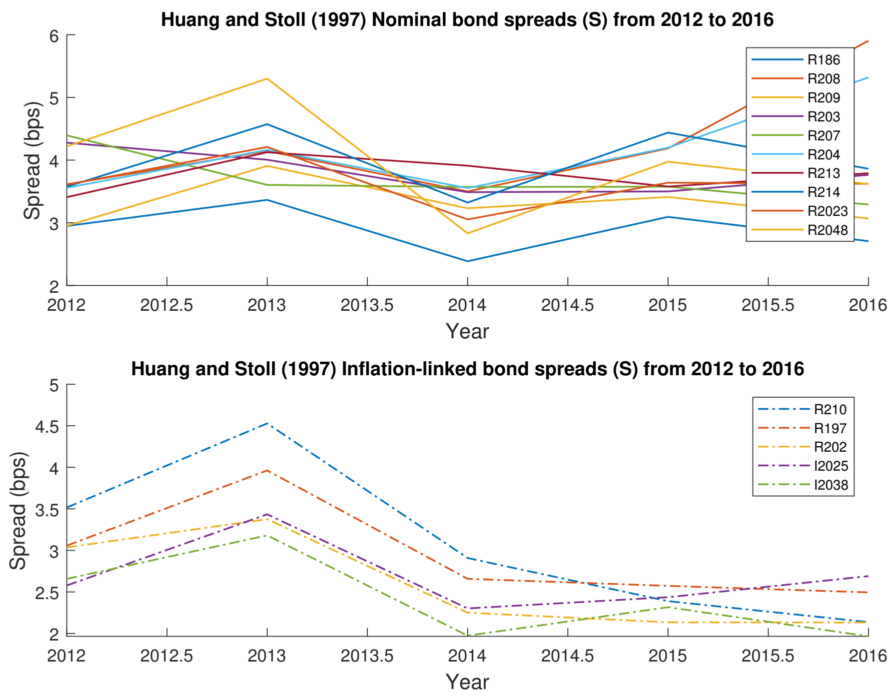


Fig. 4.1: Top: series of Huang and Stoll (1997) bid-ask spreads for SA nominal bonds. Bottom: series of Huang and Stoll (1997) bid-ask spreads for SA inflation-linked bonds.

Trade Bucketing Results

Pitsillis and Taylor (2014) identify a U-shaped phenomenon when spread estimates are split into trade sizes. Small trades and huge trades exhibit a higher spread than medium and large trades, see Pitsillis and Taylor (2014) for trade bucket sizing. This U-shape is consistent for each year that the spreads were modelled. With few exceptions, the effect is most pronounced in the more frequently traded bonds, see Appendix A. A squeeze in spread in the large and mid-range sizes may be as a result of increased competition. The spread expansion in huge buckets may be because of dealers having to take on excess liquidity costs which have a consequential effect on the spread. Pitsillis and Taylor (2014) reported a similar phenomenon with the trade bucket sizing and, despite the pattern of a trade-size linked spread, the difference between the buckets is still smaller than 1 basis point. Figure 4.2 plots the spread against the dealers' cost of liquidity as a percentage of spread. The first detectable pattern is the consistently lower dealers' cost of liquidity as a percent for small trades relative to all other trade sizes. The smaller percentage may result from an increased percentage in the dealers' margin, since there is less competition in the smaller trade sizes. It is hard to recognise a pattern in the other buckets, as the dealers' cost of liquidity is highly variable. Figure 4.2 does give an indication of spread stability. The medium and large trades are the most frequently traded out of all the bucket sizes and display the least variation in spread levels.

RFQ vs IDB Results

Figure 4.3 illustrates the cost of liquidity (λ) as a percentage of spread against the bid-ask spread of inter-dealer broker (IDB) trades and request-for-quote (RFQ) trades. The results of the IDB against RFQ trades are shown in appendix A. The standard errors for most of the IDB and RFQ trades are small, indicating a successful estimation procedure. Certain less frequently traded bonds have high standard errors, but are still recorded for reporting purposes. The spread levels for the nominal bonds remains steady, generally trading in the 3-5 basis point band. The inflation-linked bond spreads narrowed over the 5-year period from a 2-5 basis point band to a 1.8-2.5 basis point band. The more frequently traded bonds, especially the R186 and R209, consistently display a greater spread in the RFQ market, compared to the IDB market. While the spread levels are similar for the other IDB trades and RFQ trades, the dealers' cost of liquidity for the RFQ trades is prominantly lower across all bonds. This pattern is evident in figure 4.3 and is consistent for the 5-year period over which it was measured. A higher dealers' cost of liquidity is evident for trades that are traded less frequently when IDB and RFQ trades are viewed independently. A higher cost of liquidity could imply that dealers narrow their margins for the bonds that are traded more frequently and widen them for the less frequently traded bonds. The IDB and RFQ trade results in this paper align with those of Pitsillis and Taylor (2014). The more frequently traded bonds display a higher spread in the RFQ market and the dealers' cost of liquidity is uniformly higher in the IDB market for the 5 year period.

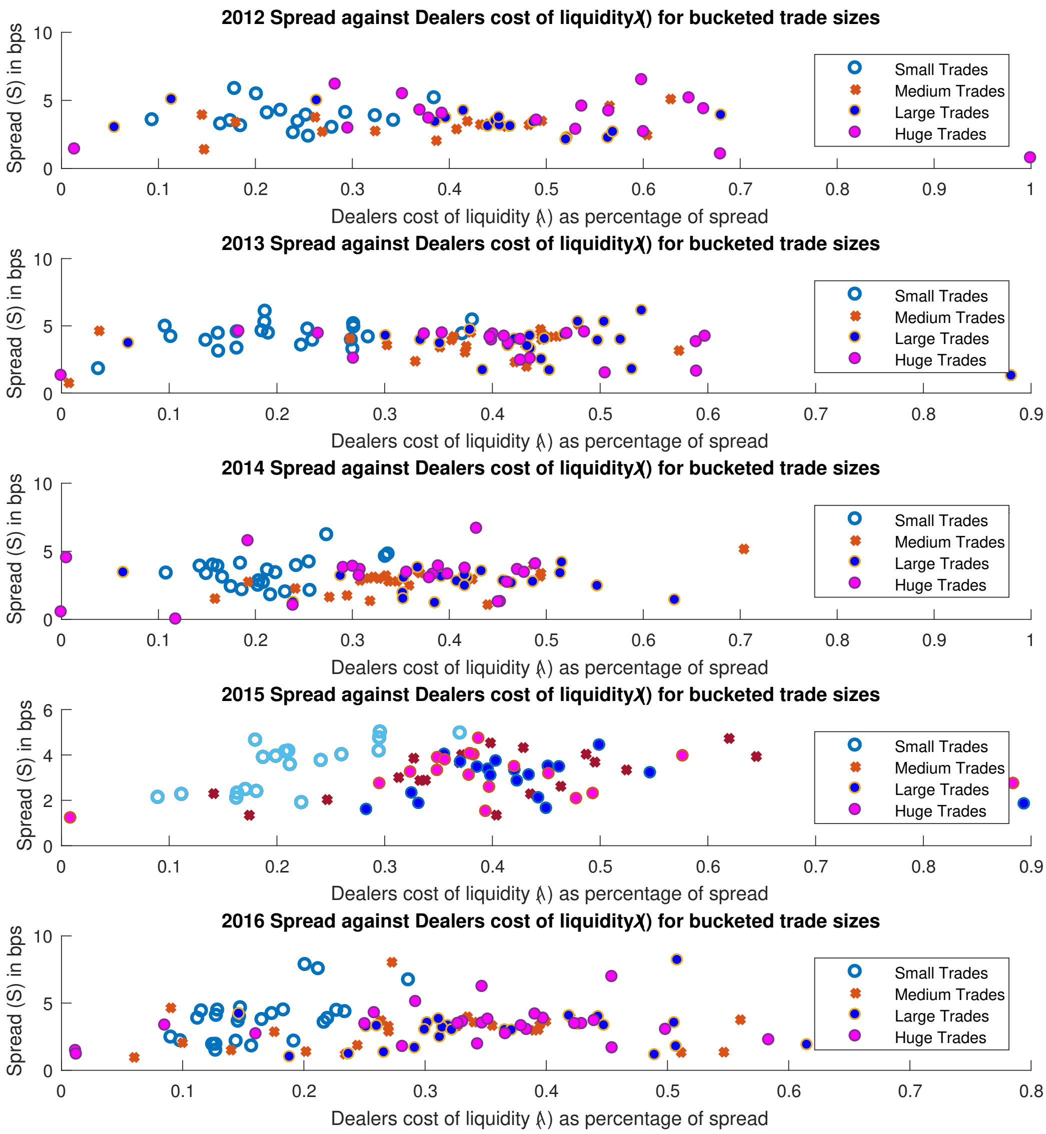


Fig. 4.2: Spread against dealers' cost of liquidity (λ) with trade size bucketing, over a 5 year period.

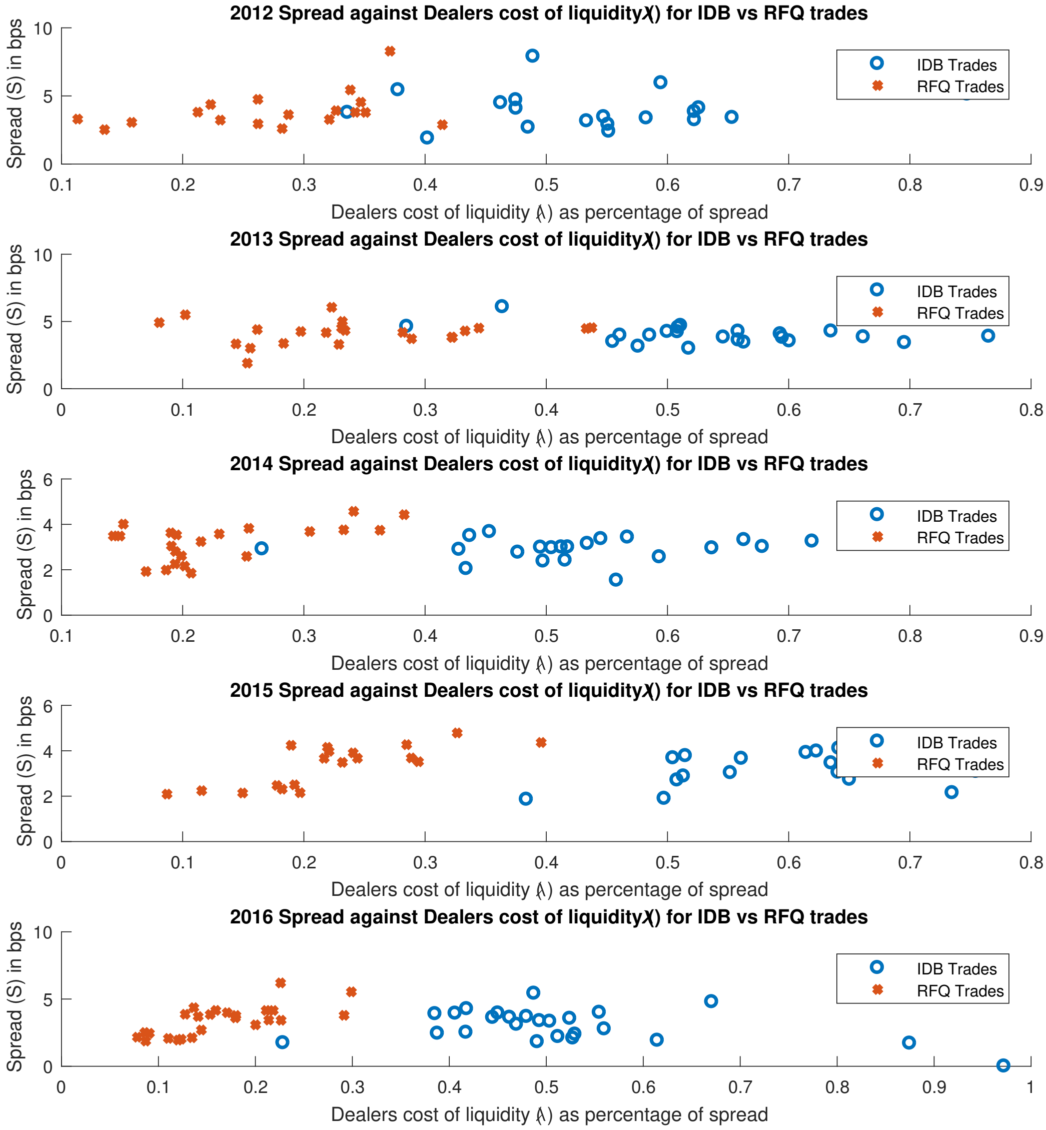


Fig. 4.3: Spread against dealers' cost of liquidity (λ) for IDB and RFQ trades, over a 5 year period.

Chapter 5

Spread Estimation through the Kalman Filter

This section of the dissertation proposes a new method of estimating the bid-ask spread, in addition to demonstrating its implementation. The Kalman filter is chosen as it can be used to leverage information from an on-screen market (the equity market) to imply bid-ask spread estimates in an off-screen market (the bond market). Traded prices ($z_t^{(eq)}$), quoted bid prices (B_t) and quoted ask prices (A_t) can all be obtained in an on-screen equity market. An equity spread estimate ($S_{act}^{(eq)}$) can be obtained using these quoted equity bid prices and quoted equity ask prices through equation 3.3. An equity spread estimate ($S_{kf}^{(eq)}$) can also be obtained through the use of a Kalman filter, using traded prices and a specified noise term (R). The Kalman filter takes an observable process and extracts an unobservable process. In this dissertation, the observable process is a time-series of traded equity prices ($z_t^{(eq)}$). The unobservable process is a time-series of equity mid-prices ($m_t^{(eq)}$). A specified noise term, in the Kalman filter, controls the relationship between these processes. An equity mid-price process can be extracted for a given arbitrary noise term. This extracted mid-price process is used alongside the observable traded equity prices to create a spread estimate ($S_{kf}^{(eq)}$), assuming that the difference between the traded equity price and the equity mid-price is half of the spread. It is possible to then calibrate this noise term ($\hat{R}$) by setting the equity spread estimates ($S_{act}^{(eq)}$), using on-screen quoted data, equal to the spread estimates ($S_{kf}^{(eq)}$), attained using the extracted mid-prices from the Kalman filter.

Calculating spread estimates in the South African government bond market isn't feasible using the former method, due to lack of quoted ask prices and quoted bid prices. The latter method, using the Kalman filter, is possible. Traded bond prices ($z_t^{(b)}$) and a calibrated equity noise term ($\hat{R}$) can be used to extract the bond mid-prices ($m_t^{(b)}$). The modelled bond spread estimate ($S_{kf}^{(b)}$) can subsequently be attained.

Two conditions need to be met when calculating bond spread estimates using a calibrated equity noise term. Firstly, the equity traded prices and the bond traded prices need to be re-based to the same price level and measured over the same transaction times. This is important because the spread is calculated as a currency amount, rather than a percentage. Secondly, the re-based equity share prices must have a similar standard deviation on returns to the bond prices. Once the appropriate equity is chosen, the bond mid-prices are extracted from traded bond prices using the Kalman filter with the calibrated equity noise term. The bond mid-prices are converted back to their original price levels. The BESA (2005) yield-from-price algorithm is used to convert the bond mid-prices into bond mid-yields. The bond mid-yields and the traded bond yields are used in calculating the bond spread estimate, using equation 3.4. The steps below summarise the proposed method of attaining the spread of a South African government bond.

1. Re-base traded equity prices ($z_t^{(eq)}$), equity quotes (A_t and B_t) and bond prices ($z_t^{(b)}$) to the same price level.

2. Align equity transaction times to bond transaction times.

3. Calculate actual equity spread estimate ($S_{act}^{(eq)}$) using re-based equity bid quotes (B_t) and re-based equity ask quotes (A_t).

4. Calculate modelled equity spread estimate ($S_{kf}^{(eq)}$) using the Kalman filter and an arbitrary equity noise term (R).

5. Calibrate equity noise term (R) by setting ($S_{act}^{(eq)}$)=($S_{kf}^{(eq)}$).

6. Use calibrated equity noise ($\hat{R}$) term to extract underlying bond mid-price process ($\hat{m}_t^{(b)}$).

7. Re-base the bond mid-prices ($\hat{m}_t^{(b)}$) to their original level.

8. Convert the bond mid-prices into bond mid-yields ($\hat{y}_t^{mid}$) using the BESA (2005) yield-from-price algorithm.

9. Calculate the modelled bond spread estimate ($s_{kf}^{(b)}$) using bond traded yields (y_t^{traded}) and extracted bond mid-yields ($\hat{y}_t^{mid}$).

The aim of this chapter is to present a potential approach of distilling bond mid-prices from traded prices using the Kalman Filter. Much more empirical work is required to find equity prices that match the dynamics of bond clean prices. As an illustration of the approach, a bond spread estimate of the South African benchmark bond is acquired from 20 business days worth of data. The following sections detail the state-space formulation of the Kalman filter, the calibration of the equity noise term and the calculation of the spread estimate.

5.1 Kalman Filter and State-Space Formulation

The Kalman filter is a recursive procedure used to extract an unobservable process from a set of noisy observations. The observations are linked to the underlying process via some measurement equation. When the system is cast in state-space form, the Kalman filter provides a recursive mechanism for estimating the evolution of the underlying process through time. Furthermore, under this formulation, a likelihood function is attainable. This is useful for the estimation of parameters, given the state-space formulation.

5.1.1 State-Space Formulation

This dissertation uses a linear measurement equation and a linear transition equation, as required in a typical state-space formulation. The linear equations are formulated with an approximation of three different stochastic processes using the Euler-Maruyama method. The measurement equation links the traded price of a share (bond) to the mid-price of a share (bond).

Measurement Equation

The observed variables (the bond clean prices and share prices) are given by

$$z_t = Hm_t + v_t \tag{5.1}$$

where z_t is the traded share price ($z_t^{(eq)}$) or the traded clean bond price ($z_t^{(b)}$) and m_t is the unobservable mid-price of the bond ($m_t^{(b)}$) or share ($m_t^{(eq)}$), at time step t. H is the constant that links the mid-price to the traded price. It is assumed, in this paper, that the traded price, z_t, is a noisy observation of the mid-price and hence the constant, H, is set to one. The measurement error, v_t, in the traded price is assumed to be normally distributed and additive.

$$v_t \sim N(0, R)$$

where R is the variance in the measurement error, as the measurement equation is one-dimensional. The error term, v_t, allows for the perturbation of the mid-price to the traded price and is interpreted economically as half of the bid-ask spread. For instance, if the measurement equation is based in rand terms and the error term has a standard deviation, $\sqrt{R}$, of $R10$, then we expect the traded price to trade within $R10$ of the mid-price 68.27% of the time.

Transition Equation

The transition equation describes the evolution of the unobservable mid-price process, m_{t+1} (either $m_{t+1}^{(eq)}$ or $m_{t+1}^{(b)}$), as a linear function of the previous state variable, m_t. The transition equation is expressed in the following form.

$$m_{t+1} = Fm_t + E + Gw_{t+1} \tag{5.2}$$

where F, E and G are constants attained using three different stochastic processes namely: geometric Brownian motion (GBM), arithmetic Brownian motion (ABM) and an Ornstein-Uhlenbeck (OU) process, summarised in table 5.1. The error term in the transition equation is also assumed to be normally distributed and additive.

$$w_{t+1} \sim N(0, Q)$$

where Q is the variance of the transition equation.

This dissertation uses three different stochastic processes to model the transition equation. The stochastic differential equations (SDE) for each of these processes are approximated using the Euler-Maruyama method to attain a numerical solution of linear form.

1. Arithmetic Brownian Motion
   The first stochastic process follows an arithmetic Brownian motion (ABM), with real-world dynamics that satisfy the following SDE.

   $$dm_t = \mu dt + \sigma dW_t \tag{5.3}$$

   where μ is the constant drift term and σ is the constant volatility term. W_t is a Brownian motion with $\mathbb{E}(W_t) = 0$ and $\mathbb{V}(W_t) = t$. The Euler-Maruyama approximation of the SDE of equation 5.3 is given by

   $$m_{t+1} = m_t + \mu\Delta t + \sigma\Delta W_t$$

   whereby the conditional mean and conditional variance are normally distributed under the real-world measure as shown below:

   $$\mathbb{E}[m_{t+1}] = m_t + \mu\Delta t$$

   $$\mathbb{V}[m_{t+1}] = \sigma^2\Delta t$$

2. Geometric Brownian Motion
   The second stochastic process follows a geometric Brownian motion (GBM), with real-world dynamics that satisfy the following SDE:

$$dm_t = \mu m_t dt + \sigma m_t dW_t$$

   where μ is the constant percentage drift term and σ is the constant percentage volatility term. W_t is a Brownian motion with $\mathbb{E}(W_t) = 0$ and $\mathbb{V}(W_t) = t$. However, since the transition equation needs to be of linear form, the log of the mid-prices is used and takes on the following form.

$$dln(m_t) = (\mu - \sigma^2/2)dt + \sigma dW_t \tag{5.4}$$

   where the extra term in the drift component arises from the quadratic variation of the SDE. The Euler-Maruyama approximation of the SDE of equation 5.4 is given by

$$ln(m_{t+1}) = ln(m_t) + (\mu - \sigma^2/2)\Delta t + \sigma \Delta W_t$$

   whereby the conditional mean and conditional variance are normally distributed under the real-world measure as shown below:

$$\mathbb{E}[ln(m_{t+1})] = ln(m_t) + (\mu - \sigma^2/2)\Delta t$$

$$\mathbb{V}[ln(m_{t+1})] = \sigma^2 \Delta t$$

3. Ornstein-Uhlenbeck process
   The final stochastic process used is the Ornstein-Uhlenbeck process, with real-world dynamics that satisfy the following SDE.

$$dm_t = \kappa(\theta - m_t)dt + \sigma dW_t \tag{5.5}$$

   where κ is the constant representing the speed of mean reversion, θ is the long term mean and σ is the constant volatility of the mid-price. The Euler-Maruyama approximation of equation 5.5 is given by

$$m_{t+1} = (1 - \kappa\Delta t)m_t + \kappa\theta\Delta t + \sigma \Delta W_t$$

   where the conditional mean and conditional variance are normally distributed under the real-world measure as shown below:

$$\mathbb{E}[m_{t+1}] = (1 - \kappa\Delta t)m_t + \kappa\theta\Delta t$$

$$\mathbb{V}[m_{t+1}] = \sigma^2\Delta t$$

The linear form of the approximation from above is summarised in a table format for the transition equation.

Model	F	E	G	Q
ABM	1	$\mu\Delta t$	1	$\sigma^2\Delta t$
GBM	1	$(\mu - \sigma^2/2)\Delta t$	1	$\sigma^2\Delta t$
Ornstein-Uhlenbeck	$(1 - \kappa\Delta t)$	$\kappa\theta\Delta t$	1	$\sigma^2\Delta t$

Tab. 5.1: System scalars of state-space form for ABM, GBM and Ornstein-Uhlenbeck processes.

In addition to modelling the mid-price to imply a bid-ask spread, comparing the three aforementioned models enables for the comparison of the fit of each model. The state-space formulation is represented by the transition equation (equation 5.2) and the measurement equation (equation 5.1).

5.1.2 Kalman Filter Procedure

The steps used in the Kalman filtering procedure are attained from Babbs and Nowman (1999) and Harvey (1990). The Kalman filter allows for the construction of a likelihood function of the state-space model. A likelihood based inference can then be made on the model's parameters. For a given set of parameters, δ, a likelihood function, $f(z_1, ..., z_T; \delta)$, associated with the state-space model above can be obtained. Bayes theorem allows for the factoring of the conditional density functions per time step, as shown below.

$$f(z_1, ..., z_T; \delta) = \prod_{t=1}^{T} f(z_t|z^{t-1}, \delta) \tag{5.6}$$

where $z^{t-1} \equiv \{z_0, ..., z_{t-1}\}$ is the history of observations. The Kalman filter uses a recursive procedure, at each time step, to derive the conditional densities and the state variables. This procedure is highlighted in the following four steps.

1. Initialisation step

   Harvey (1990) uses the unconditional mean and unconditional variance from the transition equation to initialise the state vector. The unconditional mean and unconditional variance, for each of the stochastic process, is highlighted below.

$$\mathbb{E}[m_1] = m_{1|0} = Fm_0 + E \tag{5.7}$$

$$\mathbb{V}[m_1] = \sigma_{1|0} = \sigma^2 \Delta t \tag{5.8}$$

where equation 5.7 and equation 5.8 are the expected values and variances of the OU, ABM or GBM processes respectively. The initial value m_0 is chosen to be $R100$ for reasons that will be described in the following section. The unconditional variances are chosen to be the variances from the transition equations. The mean and variance are updated using the following formulae.

$$m_{1|1} = m_{1|0} + \Sigma_{1|0} H (H' \Sigma_{1|0} H + R)^{-1} (z_1 - H' m_{1|0})$$

$$\Sigma_{1|1} = \Sigma_{1|0} - \Sigma_{1|0} H (H' \Sigma_{1|0} H + R)^{-1} H' \Sigma_{1|0}$$

2. Prediction step

The transition equation, equation 5.2, is then used to predict the conditional mean and conditional variance.

$$m_{t|t-1} = \mathbb{E}[m_t | z^{t-1}] = F m_{t-1|t-1} + E$$

$$\Sigma_{t|t-1} = \mathbb{E}[(m_t - m_{t|t-1})(m_t - m_{t|t-1})' | z^{t-1}] = F \Sigma_{t|t-1} F' + GQG'$$

A prediction for the next measurement is made, $z_{t|t-1} = H' m_{t|t-1}$ and the prediction error is calculated.

$$e_t = z_t - z_{t|t-1} = v_t + H'(m_t - m_{t|t-1})$$

where the error term is normally distributed, as z_t and $z_{t|t-1}$ are Gaussian and additive.

$$e_t \sim N(0, H \Sigma_{t|t-1} H' + R)$$

The density function $f(e_t, \delta) = f(z_t|z^{t-1})$ because $z_t = e_t + z_{t|t-1}$. This means that the density function, $f(z_t|z^{t-1}, \delta)$, at each point in time is

$$f(z_t|z^{t-1}, \delta) = \frac{1}{\sqrt{2\pi|H\Sigma_{t|t-1}H' + R|}} e^{-\frac{e_t'(H\Sigma_{t|t-1}H'+R)^{-1}e_t}{2}} \tag{5.9}$$

The state predictions for the expectation and variance are updated to $z_{t|t}$ and $\Sigma_{t|t}$, given the new information of z_t at time steps $t = 1, 2, ..., T$, to get the density function of the subsequent time step, $f(z_{t+1}|z^t, \delta)$.

3. Correction/Update step

Once the data is observed at time step t, the mid-price mean and variance predictions are updated to $m_{t|t}$ and $\Sigma_{t|t}$ to compute the subsequent density function, $f(z_{t+1}|z_t, \delta)$.

$$m_{t|t} = m_{t|t-1} + K_t(z_t - Hm_{t|t-1})$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t(H\Sigma_{t|t-1}H' + R)K_t'$$

where K_t, the Kalman gain is shown below.

$$K_t = \Sigma_{t|t-1}H'(H\Sigma_{t|t-1}H' + R)^{-1}$$

The update step serves to correct the prediction step, in addition to reducing the prediction error variance through the Kalman gain term.

4. Construction of the likelihood function

Steps 2 to 3 are repeated for each time step, where the Kalman filter recursively computes the density functions $f(z_t|z^{t-1}; \delta)$ and the underlying state variables. Once all the density functions are computed, the likelihood function is attained through the product of the individual conditional density functions, given that the prediction errors are normally distributed and independent.

$$L(z^T, \delta) = \prod_{t=1}^{T} f(z_t|z^{t-1}, \delta) \tag{5.10}$$

Parameter Estimation through Maximising the Likelihood Function

The likelihood function above provides the likelihood of the data, given a fixed set of parameters. The parameters of the state space model are $\delta = \{H, R, F, E, G, Q\}$. It is assumed that the parameters H and G are constant and equal to 1, and that R is known. The optimal parameters are the parameters $(\hat{F}, \hat{E}, \hat{Q}, \hat{R})$ that maximise the likelihood function. The log of the likelihood function is taken to estimate the optimal parameters $\delta = \{\hat{F}, \hat{E}, \hat{Q}\}$, and is shown below for the case where observations are one-dimensional.

$$lnL(z^T, \delta) = -\sum_{t=1}^{T} \left(\frac{ln(2\pi)}{2} + \frac{ln|H'\Sigma_{t|t-1}H + R|}{2} + \frac{e_t'(H\Sigma_{t|t-1}H' + R)^{-1}e_t}{2} \right) \tag{5.11}$$

In this paper, the parameters are estimated numerically by minimising the negative log-likelihood function, rather than maximising the log-likelihood function, using an unconstrained non-linear optimiser.

5.1.3 Parameter Estimation for Simulated ABM, GBM and OU Processes

A short general simulation study is conducted to test the implementation of the Kalman Filter. Bond clean prices are simulated for each of the stochastic processes, using a known set of parameters. The efficacy of the filter is determined by the ability to return the simulated parameters when the calibrated noise term, $\hat{R}$, is equal to 0, meaning the bond mid-prices are equivalent to the bond traded prices. This ensures that the returned parameters are the optimal parameters for the bond clean prices and the bond clean mid-prices, rather than solely the bond clean mid-prices. These optimal parameters are estimated through a combination of the Kalman filter and a non-linear optimiser. A more accurate estimation of optimal parameters is determined by simulating 10 different price paths and taking the average of those optimal parameters. The results of the parameter estimates are summarised in the table below.

Process	Parameter	True Value	Estimate mean
ABM	μ	0.05	0.0544
	σ	0.2	0.2005
GBM	μ	0.05	0.0529
	σ	0.2	0.2002
OU	κ	4	4.1768
	θ	0.05	0.0536
	σ	0.2	0.1991

Tab. 5.2: True parameter estimates against the returned parameter estimates from Monte Carlo simulations.

Table 5.2 shows that the parameters for each process are close to their true values and seem to be unbiased. A greater number of simulations should result in a more accurate estimation of the parameters. The choice of 10 simulations for each method was chosen to reduce computational time.

5.2 Noise Calibration and Spread Estimation

This section of the dissertation explains how the optimal measurement error term ($\hat{R}$), the noise term controlling the relationship between the traded price and the mid-price, is calibrated. Furthermore, this section details how this calibrated error term is applied to the bond market to infer information about the bid-ask spread. Numerous different noise terms are calibrated using JSE listed large-cap shares, mid-cap shares and small-cap shares, acquired using a Bloomberg terminal. Using shares of different market capitalisation enables the investigation of the spread behaviour under different liquidity conditions. The uses of the Kalman filter are twofold. The first is to calibrate the noise term ($\hat{R}$) for the South African equities by setting the actual equity spread estimate ($S_{act}^{(eq)}$) equal to the modelled spread estimate ($S_{kf}^{(eq)}$). The second is to extract the South African bond's underlying optimal process (the mid-price process, $\hat{m}_t^{(b)}$) using the calibrated equity noise term ($\hat{R}$). Finally, the modelled bond spread estimate ($S_{kf}^{(b)}$) is attained using the bond traded prices ($z_t^{(b)}$) and the bond traded mid-prices ($\hat{m}_t^{(b)}$). To compare the equity market to the bond market, it is necessary to compare returns on prices, rather than comparing returns on prices with returns on yields. Furthermore, bond clean prices are preferred over all-in-prices, as accrued interest induces artificial volatility in all-in prices. The following steps detail the procedure of acquiring the bid-ask spread for the bond market, using calibrated equity error terms.

1. Re-basing assets step

   Re-base all the share traded prices, quoted bid prices, quoted ask prices and bond clean prices to start at the same level. In this paper, the variables are re-based to a level of R100 for the arithmetic Brownian motion and Ornstein-Uhlenbeck processes. Variables using geometric Brownian motion for the underlying transition process are re-based to a level of log(R100). The traded share prices, quoted bid prices, quoted ask prices and the clean prices are evolved using the respective share returns and clean price returns. The equity transaction times are aligned to R186 bond transaction times, through linear interpolation, for calibration of the noise term R. Time step, t, is specified as a transaction time for the R186 bond. The graph below depicts the re-based share prices and bond clean prices, for transaction times t.

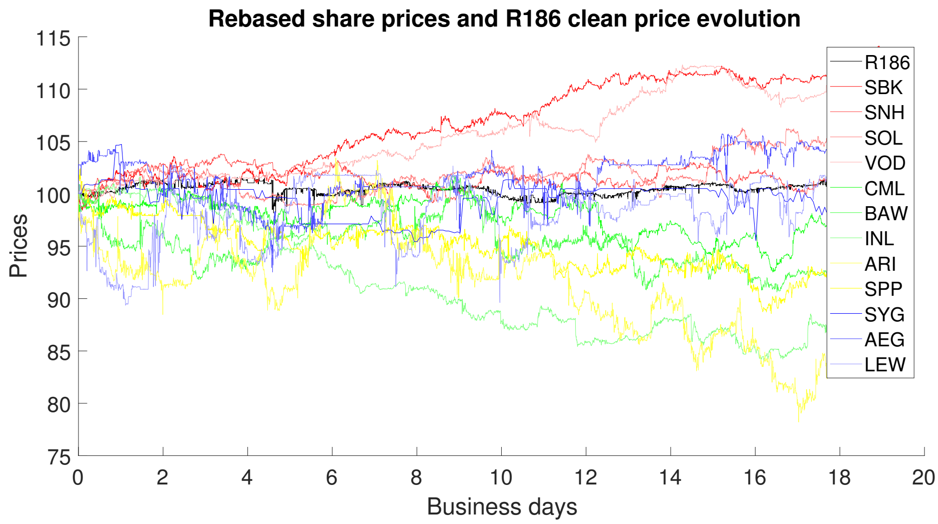


Fig. 5.1: Time-series plot of large-cap, mid-cap and small-cap re-based share prices, in addition to re-based R186 bond clean prices.

   The re-based R186 bond clean prices tend to exhibit smaller drift and noise components relative to the share prices. Whilst there may be price-based instruments akin to the behaviour of bond clean prices, share prices are used because of the availability of historical bid-ask quotes.

2. Error term ($\hat{R}$) calibration step

   The error term variance (R) of v_t, in equation 5.1, controls the size of the spread. A larger variance term implies a greater difference between the traded equity prices and the equity mid-prices. A smaller variance implies that the traded equity prices trade closer to the equity mid-prices. It is assumed that the underlying equity mid-price process has unique optimal transition parameters $(\hat{F}, \hat{E}, \hat{Q})$, but a calibrated equity measurement variance term $(\hat{R})$ can be used to infer information about the deviation of the traded price from the mid-price. To attain the optimal estimate $(\hat{R})$ of the measurement equation, an initial guess (R_0) is specified. The traded equity price observations, time between transaction times, the initial (R_0) term and initial transition parameter guesses F_0, E_0, Q_0 are passed through the Kalman filter to extract the underlying mid-price process $(m_t^{(eq)})$, in addition to a log-likelihood value. Optimal transition equation parameters $\hat{F}, \hat{E}, \hat{Q}$ are estimated numerically by minimising the negative log-likelihood function from equation 5.11. Once optimal parameters are attained, these optimal parameters are passed back through the filter to get an optimal mid-price process $(\hat{m}_t^{(eq)})$ for a given measurement variance (R_0). A modelled De Jong and Rindi (2009) equity spread estimate can subsequently be acquired using the formula below:

$$S_{kf}^{(eq)} = \frac{1}{T}\sum_{t=1}^{T} 2|z_t - \hat{m}_t^{(eq)}| \tag{5.12}$$

The actual De Jong and Rindi (2009) equity spread estimates can be attained from traded equity prices and actual equity mid-prices (m_t^{act}), acquired from quoted bid prices and quoted ask prices.

$$\hat{S}_{act}^{(eq)} = \frac{1}{T}\sum_{t=1}^{T} 2\left|z_t - \frac{A_t + B_t}{2}\right| = \frac{1}{T}\sum_{t=1}^{T} 2|z_t - m_t^{act}| \tag{5.13}$$

The $S_{kf}^{(eq)}$ and $\hat{S}_{act}^{(eq)}$ terms will differ in values because of the initial noise term (R_0). The optimal Kalman filter spread estimate, $\hat{S}_{kf}^{(eq)}$, will have the same spread estimate as the actual spread estimate, $\hat{S}_{act}^{(eq)}$, using the optimal noise term $\hat{R}$. The optimal noise term $\hat{R}$ is calibrated by minimising the squared difference between the actual spread estimate, $\hat{S}_{act}^{(eq)}$, and the modelled Kalman filter spread estimate, $\hat{S}_{kf}^{(eq)}$. The calibrated noise term $\hat{R}$ is subsequently used to estimate bond spreads.

3. Mid-price filtering step

In this dissertation, 12 different noise terms $(\hat{R})$ are calibrated, using three different stochastic processes (table 5.1) as the underlying transition equations for each of the large-cap, mid-cap and small-cap shares. Each of these noise terms are used independently to investigate the effect on the spread. The calibrated noise term, initial parameter guesses and the clean price observations ($z_t^{(b)}$) are combined with maximum likelihood estimation to acquire optimal transition parameters for the re-based R186 clean prices. These optimal parameters, along with the calibrated equity noise terms and clean price observations are run through the Kalman filter to acquire 12 different series of re-based R186 clean mid-prices ($\hat{m}_t^{(b)}$).

4. Yield conversion and spread estimation step

   The 12 different time-series of clean mid-prices are transformed back to their original level to get the correct level of spread. The clean mid-prices are converted back to all-in-prices using the South African bonds pricing formula (see BESA (2005)). The De Jong and Rindi (2009) spread estimates and the Huang and Stoll (1997) spread estimates of the R186 bond are expressed in terms of yields. For this reason, the new mid all-in-prices are converted into yields using the BESA (2005) yield-from-price algorithm. The 12 mid-yield time-series are used along with the traded yields to calculate the Kalman filter spread estimate using the equation below:

$$\hat{s}_{kf}^{(b)} = \frac{1}{T}\sum_{t=1}^{T} 2|y_t^{traded} - \hat{y}_t^{mid}| \tag{5.14}$$

   where the lower case s denotes the spread in terms of yield for comparison to section 3 and section 4.

5.3 Spread Results

When selecting a share to infer information about the bid-ask spread of a South African government bond, it is important to choose assets with similar characteristics. The standard deviation on returns, over similar time intervals, is an important metric for this. Table 5.3 lists the standard deviation of re-based large-cap, mid-cap and small-cap equities as well as the re-based benchmark South African bond.

Asset Name	Std. Dev. On Returns (x10^{-3})
R186	1.90
SBK	0.93
SNH	0.86
SOL	0.93
VOD	0.75
CML	1.45
BAW	2.59
INL	1.20
ARI	4.14
SPP	2.33
SYG	0.52
AEG	3.24
LEW	3.60

Tab. 5.3: Standard deviation on returns for rebased large-cap, mid-cap and small-cap shares, where $\Delta t = 0.032$ days.

It is evident from above that the shares with the closest standard deviation to the R186 bond are the mid-cap equities: Coronation Fund Managers Limited (CML), Spar (SPP) and Investec Limited (INL). These shares will be used as the best proxy for the R186 spread estimate. It would be more robust to calculate the calibrated noise terms for every share listed on the JSE-main board to find shares more comparable to the bonds. However this is outside the scope of this dissertation, as this is a proposal of a potential approach rather than a thorough empirical analysis.

Table 5.4 displays all of the spread results (represented in terms of yield) for the R186 benchmark bond. The Calibration asset column refers to the share name of calibrated noise term $\hat{R}$ used to acquire the R186 bond spread. The corresponding calibrated noise term, $\hat{R}$, is displayed on the right hand side of each transition process column. The transition equation parameters for the underlying mid-price process and the log-likelihood values are placed in the middle of the transition process columns and the spread estimates are represented by $\hat{s}^{(b)}_{kf}$.

The R186 spreads using the calibrated noise terms of the mid-cap shares lie between 1.16 basis points and 3.01 basis points for all of the transition processes considered. These were the equities with the closest standard deviation on returns to the R186 bond. The measurements for the Kalman filter spread estimates were taken in 2016. The spread estimates across all of the processes are consistent using calibrated noise terms from mid-cap and large-cap shares. The noise terms from small-cap shares display larger spread estimates, when using ABM and OU processes as the underlying transition equation. From table 5.4, it is evident that the calibrated noise term, $\hat{R}$, affects the size of the bond spread. The more liquidly traded large-cap stocks have smaller error terms, which consequently result in a smaller spread estimates of the R186 bond.

The results from each underlying transition equation perform differently for each market capitalisation size. GBM performs most consistently across market capitalisation sizes, compared to the ABM and OU processes. The GBM and ABM processes return a low μ parameter. This small drift component is evident from figure 5.1. The parameters returned from the OU process exhibit a low long-term mean parameter, θ, relative to the R100 starting point. It is hard to foresee any mean reversion effect as the model is filtering on prices, in addition to only running over a 20 business day period. This dissertation assumes that the transition equation used in calibrating the noise term (R) and extracting the bond mid-price ($m_t^{(b)}$) are the same. Based on figure 5.1, the bond clean prices and the share prices evolve very differently. An extension to this model may be to calibrate the equity noise term using GBM and then extracting the unobservable bond mid-prices using an OU process as the underlying transition equation.

Calibration Asset	GBM					ABM					OU					
	$\hat{s}_{kf}^{(b)}$ (bps)	μ	σ	nLL	$Noise(\hat{R})$	$\hat{s}_{kf}^{(b)}$ (bps)	μ	σ	nLL	$Noise(\hat{R})$	$\hat{s}_{kf}^{(b)}$ (bps)	κ	θ	σ	nLL	$Noise(\hat{R})$
SBK	0.32	0.00080	0.025	24383	0.00982	0.36	0.0456	2.4257	1426	1.035	0.37	-0.00064	0.858	2.417	1429	1.04
SNH	0.30	0.00077	0.025	24370	0.00953	0.29	0.0457	2.5372	1387	0.946	0.31	-0.00062	0.382	2.513	1396	0.97
SOL	0.50	0.00069	0.022	24479	0.01183	0.51	0.0456	2.1993	1506	1.199	0.53	-0.00066	-2.983	2.178	1513	1.21
VOD	0.33	0.00078	0.025	24390	0.00998	0.38	0.0457	2.3945	1437	1.059	0.38	-0.00064	-0.558	2.388	1440	1.06
CML	1.63	0.00046	0.011	24794	0.02082	1.58	0.0415	1.1624	1815	2.041	1.58	-0.00078	0.710	1.162	1815	2.04
BAW	2.35	0.00041	0.008	24457	0.03044	2.28	0.0389	0.8107	1555	2.918	2.29	-0.00076	0.576	0.804	1555	2.92
INL	1.40	0.00049	0.013	24777	0.01883	1.16	0.0437	1.4767	1749	1.706	1.17	-0.00074	-3.226	1.470	1751	1.71
ARI	3.01	0.00040	0.006	23324	0.04739	2.89	0.0383	0.6233	607	4.377	2.86	-0.00076	0.739	0.631	672	4.28
SPP	2.46	0.00042	0.007	24327	0.03260	2.37	0.0386	0.7774	1464	3.083	2.37	-0.00076	0.570	0.777	1464	3.08
SYG	1.43	0.00050	0.013	24781	0.01904	12.01	0.0275	0.2177	-15042	131.646	12.03	-0.00056	1.069	0.217	-15053	131.93
AEG	1.76	0.00047	0.011	24785	0.02210	4.41	0.0393	0.4475	-3805	13.119	4.41	-0.00079	0.679	0.447	-3809	13.13
LEW	2.00	0.00043	0.009	24717	0.02496	4.81	0.0394	0.4323	-4841	16.381	4.81	-0.00079	0.726	0.432	-4841	16.38

Tab. 5.4: R186 spread results using calibrated noise terms from large-cap, mid-cap and small-cap shares under GBM, ABM and OU transition equations.

Chapter 6

Conclusion

The Huang and Stoll (1997) method applied to equity data exhibited closer estimates to the actual spread estimates, than the De Jong and Rindi (2009) method. Out of the 11 modelled spread estimates, 9 were overestimated using the De Jong and Rindi (2009) method, where only two spread estimates showed a percentage difference to the actual spread estimate under 10%. The difference between the modelled spread estimates and the actual spread estimates originates from algorithm 3, where it is assumed that quoted prices stay constant from the previous trade if opposing quotes are triggered. Out of the 11 spread estimates, 8 were overestimated using the Huang and Stoll (1997) model, where only one spread estimate displayed a percentage difference greater than 10%. The equity spread estimates attained through the Huang and Stoll (1997) method perform better than the De Jong and Rindi (2009) method, when using the percentage difference as metric for model performance.

The Huang and Stoll (1997) nominal bond spread estimates trade tighter in a $3 - 4.5$ basis point range than the De Jong and Rindi (2009) nominal bond spread estimates which trade in a $3 - 6$ basis point range. Both methods have contrasting results in 2014. The De Jong and Rindi (2009) results show a spike in 2014, unlike the Huang and Stoll (1997) results which show an annual spread reduction. This spread estimate spike is not evident for the inflation-linked market using both methods. Both methods show a reduction in spread estimates in the inflation-linked bond market. A $2-3.5$ basis point range is observed for the Huang and Stoll (1997) method and a $2 - 8$ basis point range is observed for the De Jong and Rindi (2009) method in 2012. The gradual reduction of inflation-linked bond spread estimates converges to a $1.8-2.6$ basis point range for both methods in 2016. In the report of Pitsillis and Taylor (2014), the spread estimates from both methods lie within the range of this dissertation, where the spreads stretch mostly between $2 - 6$ basis points over the 5 years. The inflation-linked bond spread estimates in the report of Pitsillis and Taylor (2014) are at the lower end of the range, but they align with the levels in this dissertation. The inflation-linked bond levels are lower than the market practitioners expected. The traders agreed on a $4-6$ basis point range for 2014 in the paper of Pitsillis and Taylor (2014). The spread estimates in this dissertation for the entire period ranged between $2-6$ basis points. The smaller spreads were mostly observed in the later years for more frequently traded bonds, indicating a more liquid market. The stability of the spread is indicative of a strong provision of liquidity by Primary Dealers. The U-shaped spread phenomenon, identified in the Pitsillis and Taylor (2014) report, linking the size of the spread estimate to the trade size is observed each year and is most evident in the more frequently traded bonds. The medium and large trades are the most frequently traded and display the least variation in spread levels. In addition to this, the dealers' cost of liquidity as a percentage is lower in general for small trades relative to all other trade sizes. Certain frequently traded bonds, especially the R186 and R209, consistently display a greater spread in the RFQ market, compared to the IDB market. The dealers' cost of liquidity for the RFQ trades are uniformly lower than the IDB trades across all bonds, in agreement with the Pitsillis and Taylor (2014) paper.

The spread estimates, calculated with the new proposed Kalman filter method, on the R186 bond mostly range between $0.3-4.8$ basis points, across all transition equations and noises using different market capitalisations. The re-based Coronation Fund Managers Limited (CML) prices display the most similar characteristics to re-based bond clean prices, based on standard deviation of returns. This is followed by Barloworld Limited (BAW) and then Investec Limited (INL). A spread estimate of 1.63, 1.58 and 1.58 basis points is calculated for the yield spread estimates using GBM, ABM and OU processes respectively as the underlying transition equations. All of these spread estimates use the CML calibrated noise term. The R186 spread estimate range is between $1.16-3.01$ basis points for all of the transition processes considered for mid-cap shares. On average these are slightly lower than the chapter 3 and chapter 4 spread estimates. However, these values should not be compared as the modelled Kalman filter spread estimates were measured over a 20 day period in 2016. The calibrated noise terms of the liquid large-cap shares were small, consequently resulting in smaller R186 spread estimates. Conversely, the calibrated noise terms of the less liquid small-cap shares returned spread estimates that were large relative to the other chapters. The results from each underlying transition equation perform differently for each market capitalisation size. GBM performs most consistently across market capitalisation sizes, compared to the ABM and OU processes. The parameters returned by the OU process seemed unlikely as the mean reversion level was far off from the re-based R100.

Bibliography

Babbs, S. H. and Nowman, K. B. (1999). Kalman filtering of generalized vasicek term structure models, *Journal of Financial and Quantitative Analysis* **34**(1): 115–130.

BESA (2005). Bond pricing formula, *Technical report*, Bond Exchange of South Africa.

Brand, R. (2014). Electronic trade to boost local bonds. BLOOMBERG.

De Jong, F. and Rindi, B. (2009). *The microstructure of financial markets*, Cambridge University Press.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators, *Econometrica: Journal of the Econometric Society* pp. 1029–1054.

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*, Cambridge university press.

Ho, T. S. and Stoll, H. R. (1983). The dynamics of dealer markets under competition, *The Journal of finance* **38**(4): 1053–1074.

Huang, R. D. and Stoll, H. R. (1997). The components of the bid-ask spread: A general approach, *Review of Financial Studies* **10**(4): 995–1034.

Kyle, A. S. (1985). Continuous auctions and insider trading, *Econometrica: Journal of the Econometric Society* pp. 1315–1335.

Lee, C. and Ready, M. J. (1991). Inferring trade direction from intraday data, *The Journal of Finance* **46**(2): 733–746.

Lybek, M. T. and Sarr, M. A. (2002). *Measuring liquidity in financial markets*, number 2-232, International Monetary Fund.

Ndzamela, P. (2015). *New way to trade*, Financial Mail.

Pitsillis, Z. and Taylor, D. (2014). The IOSCO transparency principle and modelling the bid-ask spread, *Technical Report 1*, African Institute of Financial Markets and Risk Managment.

Stoll, H. R. (1978). The supply of dealer services in securities markets, *The Journal of Finance* **33**(4): 1133–1151.

Treasury, N. (2014). Electronic trading platform for rsa government bonds. Media Statement.

Treasury, N. (2016). 2015/2016 debt management report, Debt Management Report.

Appendix A

Huang and Stoll (1997) Bond Spread Results

A.1 Huang and Stoll (1997) Bucketed Bond Spread Results

		Small Trades					Medium Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R186	26794	11775	3.4278	0.00088195	0.2443	0.007631665	6502	2.8928	0.000591402	0.4075	0.00673432
R157	24972	6169	3.5109	0.003811033	0.3429	0.024330133	5262	2.7515	0.002995027	0.3236	0.025774036
R209	8163	4527	3.2479	0.001036237	0.1645	0.01017861	1176	3.0725	0.001751224	0.4588	0.020527781
R208	8507	3910	4.2519	0.001394275	0.2265	0.012130676	1580	3.4933	0.00162738	0.4954	0.018029058
R2023	3311	1620	3.4877	0.002603486	0.1746	0.022909927	608	3.2806	0.003072452	0.4406	0.034823295
R213	4404	2375	3.8969	0.0017099	0.2526	0.015944962	744	3.1983	0.002475106	0.4821	0.029176024
R203	5764	2560	5.8501	0.007041389	0.1789	0.034478059	1030	5.0966	0.009600004	0.6283	0.093725089
R207	5685	2535	5.4531	0.003782555	0.2014	0.020122487	1074	3.9507	0.004248195	0.1448	0.026174032
R204	4725	2356	4.0855	0.002092158	0.2934	0.013232617	764	3.2256	0.002315729	0.4320	0.024983894
R214	3370	1419	3.8526	0.00178962	0.3240	0.017499303	680	3.4747	0.00215288	0.4186	0.021016211
R2048	1398	545	4.0596	0.004269089	0.2126	0.024926498	295	3.7706	0.004425365	0.2615	0.031985425
R210	1295	1054	3.5532	0.004489065	0.0939	0.020541385	137	2.4564	0.004098721	0.6040	0.091305777
R202	977	722	2.9962	0.002300287	0.2793	0.036614225	133	2.7057	0.00558605	0.2690	0.043020904
R197	1213	1016	3.1158	0.002438491	0.1851	0.021296599	135	3.4008	0.003646464	0.1795	0.020881429
I2038	412	215	2.3472	0.003510067	0.2551	0.061371252	71	1.4067	0.003715633	0.1471	0.062202949
I2050	360	188	2.6029	0.003134915	0.2395	0.047732748	64	2.0411	0.00458195	0.3867	0.067207797
R201	2042	994	5.1656	0.003306922	0.3848	0.025890586	351	4.5938	0.005494636	0.5653	0.055459792
R206	1266	619	10.1662	0.023174363	0.3553	0.075713923	220	9.5874	0.043185435	0.4559	0.164018835

Tab. A.1: Spread estimate results for small and medium trades in 2012

		Large Trades					Huge Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R186	26794	6984	3.0785	0.000682813	0.4404	0.009598562	1533	3.6757	0.001380909	0.3792	0.01245151
R157	24972	8888	2.2248	0.000718437	0.5641	0.034889836	4653	2.8537	0.003069527	0.5308	0.038963886
R209	8163	1885	3.0797	0.001210351	0.4634	0.014229448	575	4.0308	0.002640705	0.3929	0.021937282
R208	8507	2322	3.4686	0.00180276	0.4481	0.021658256	695	4.5584	0.005684674	0.5370	0.058246701
R2023	3311	808	3.4028	0.002838667	0.3860	0.026850576	275	5.4683	0.008823963	0.3520	0.047399188
R213	4404	975	3.1491	0.001895771	0.4523	0.019736578	310	4.2619	0.0036346	0.3701	0.02711117
R203	5764	1626	3.3897	0.003852074	0.4874	0.033372885	548	3.1881	0.006964853	0.7890	0.158938953
R207	5685	1515	4.2184	0.004433784	0.4150	0.031110797	561	3.5240	0.004415326	0.4904	0.051944715
R204	4725	1245	3.7326	0.005297948	0.4517	0.047820315	360	2.9445	0.002447142	0.2957	0.031435616
R214	3370	1000	3.6984	0.001884501	0.3968	0.017342376	271	4.2038	0.004158532	0.5648	0.049613962
R2048	1398	380	4.2731	0.004255648	0.3698	0.029442603	178	4.3660	0.007044433	0.6627	0.090090576
R210	1295	79	3.9011	0.008912088	0.6806	0.141395895	25	5.1609	0.041168658	0.6476	0.504726708
R202	977	80	2.6600	0.004795707	0.5692	0.091685695	42	2.6808	0.008852175	0.6006	0.186598164
R197	1213	59	3.0000	0.004772449	0.0551	0.008836704	3	1.4083	0.010890728	0.0138	0.008919809
I2038	412	76	2.2411	0.007914822	0.5218	0.13552153	50	1.0472	0.006322247	0.6800	0.363491575
I2050	360	50	2.1015	0.010907686	0.5207	0.216831512	58	0.7572	0.007038398	0.9998	0.885522849
R201	2042	458	4.9769	0.005054318	0.2636	0.025747716	239	6.1714	0.009917667	0.2826	0.036310392
R206	1266	285	5.0517	0.009232586	0.1139	0.08094153	142	6.4951	0.047087789	0.5987	0.382933904

Tab. A.2: Spread estimate results for large and huge trades in 2012

		Small Trades					Medium Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R186	61128	27139	3.9064	0.000663154	0.2335	0.005711872	16200	3.4136	0.000550949	0.3513	0.004846971
R157	20617	5789	4.4360	0.002770708	0.1924	0.013459351	4235	3.5808	0.001355476	0.4306	0.011132832
R209	15128	9052	4.4204	0.001139601	0.1459	0.00809044	2616	4.2200	0.001768992	0.3642	0.011351809
R208	11529	5730	5.1393	0.003047515	0.2715	0.010282143	2229	3.7757	0.001507097	0.4138	0.013003096
R2023	8432	4422	4.5342	0.001744595	0.1634	0.01098679	1497	4.5106	0.002721349	0.3801	0.01760184
R213	9344	5064	4.6061	0.002284342	0.1865	0.013211526	1575	3.9859	0.002444913	0.4456	0.025186868
R203	8155	3892	4.7395	0.003773589	0.2290	0.025012579	1387	4.1841	0.004612576	0.4528	0.036200799
R207	8502	4064	4.1558	0.001390822	0.2851	0.011692033	1578	3.4837	0.00182751	0.3757	0.016570233
R204	7321	3611	4.8964	0.002035873	0.2717	0.013659593	1212	3.9457	0.002416438	0.3620	0.018503479
R214	8098	4191	5.2550	0.002690365	0.1891	0.015414095	1567	4.2156	0.002888192	0.4608	0.0232544
R2048	5586	3289	6.0657	0.005955172	0.1894	0.023698367	1028	4.7405	0.003691186	0.4446	0.041110614
R2037	1927	991	3.9022	0.003507456	0.1344	0.019488314	395	4.0602	0.004147045	0.2678	0.023318013
R2030	526	269	4.1801	0.010368365	0.1019	0.033108154	109	4.6264	0.010374825	0.0351	0.052525243
R210	2320	1893	4.9501	0.014668631	0.0966	0.024364687	219	2.2995	0.00384675	0.4208	0.059852697
R202	1731	1234	3.2503	0.001844846	0.2708	0.026356682	290	3.0383	0.004898001	0.3746	0.043618296
R197	1699	1405	3.9300	0.003521069	0.2694	0.046060342	184	3.5779	0.005449532	0.3021	0.037267572
I2025	1453	935	3.5434	0.003800068	0.2231	0.03724186	199	1.9804	0.004541899	0.4316	0.076677106
I2038	1332	921	3.0937	0.002853594	0.1462	0.029406939	157	3.1614	0.007466231	0.5732	0.113081299

Tab. A.3: Spread estimate result for small and medium trades in 2013

		Large Trades					Huge Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R186	61128	15087	3.5058	0.000631627	0.4149	0.005739715	2702	4.3779	0.00195299	0.3371	0.013299196
R157	20617	7010	3.2735	0.000965296	0.4357	0.011150849	3583	3.6980	0.001670744	0.4150	0.015401376
R209	15128	2875	4.0052	0.001586904	0.4489	0.013562172	585	4.2377	0.002948499	0.3987	0.025507068
R208	11529	2925	3.9514	0.004964799	0.5194	0.033976377	645	3.9296	0.002921626	0.3995	0.026537796
R2023	8432	1942	4.2345	0.002273471	0.4353	0.018192086	571	4.3723	0.0030303	0.4007	0.024659194
R213	9344	2111	4.3290	0.002248846	0.4673	0.017315772	594	4.2063	0.004098832	0.4113	0.035555290
R203	8155	2171	3.6729	0.001613964	0.3516	0.017317377	705	4.4147	0.006786607	0.4695	0.054951961
R207	8502	2287	3.4717	0.001698199	0.4327	0.016990589	573	4.1338	0.004118612	0.3981	0.033918040
R204	7321	1904	3.9832	0.002467569	0.4211	0.022365600	594	4.5227	0.003510843	0.4858	0.033068779
R214	8098	1871	4.6825	0.003239624	0.3796	0.020058843	469	4.4505	0.003827871	0.3537	0.026966066
R2048	5586	1037	5.2881	0.007843606	0.4799	0.048099488	232	3.9729	0.006136001	0.4263	0.060723224
R2037	1927	432	3.9318	0.003843677	0.3336	0.025771967	109	4.2051	0.010953156	0.5977	0.142852734
R2030	526	112	6.1162	0.012777813	0.5390	0.083026533	36	3.7956	0.01866294	0.5895	0.262808488
R210	2320	151	2.4861	0.004496483	0.4459	0.071182174	57	2.5499	0.009838364	0.4353	0.160657662
R202	1731	144	4.2435	0.006804665	0.3013	0.040116045	63	1.2903	0.003861062	0.0000	0.027854700
R197	1699	107	3.6944	0.006144842	0.0625	0.013688508	3	10.7831	0.070823012	0.0000	0.001531853
I2025	1453	231	1.7524	0.00357823	0.5299	0.080060422	88	2.4259	0.009149267	0.4264	0.143621879
I2038	1332	170	1.6796	0.00359102	0.3914	0.073583646	84	1.4800	0.005955299	0.5050	0.182602918
I2050	1186	140	1.6662	0.004237675	0.4534	0.094972231	85	2.5663	0.007063018	0.2713	0.062557016
R201	1611	289	5.2793	0.005607774	0.5042	0.042228017	70	4.4257	0.010628999	0.2386	0.050297984
I2046	288	67	1.2693	0.006519055	0.8821	0.365865913	57	1.6062	0.008724398	0.5898	0.248703000
R206	813	180	3.8816	0.005666087	0.4982	0.052784551	78	4.5866	0.007553387	0.1649	0.022001176

Tab. A.4: Spread estimate results for large and huge trades in 2013

		Small Trades					Medium Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R159	550	252	6.1973	0.007828266	0.2739	0.038489866	33	5.1702	0.016835515	0.7039	0.214691445
R197	1693	1418	2.4809	0.001195248	0.2032	0.025535598	187	2.8590	0.007168081	0.3080	0.055607462
R201	233	132	16.5836	0.018476991	0.4026	0.060829902	44	7.9703	0.02538656	0.9725	0.267991815
R202	1691	1253	2.1061	0.000996961	0.2567	0.022859867	274	1.7617	0.003390987	0.2945	0.043447784
R2023	9323	4915	3.4050	0.002537484	0.2214	0.020470534	1443	3.0003	0.002210961	0.3272	0.018756182
R203	7763	3349	3.9315	0.002595607	0.2427	0.020495676	1366	3.4068	0.002881465	0.3697	0.025318171
R2030	5115	2510	3.8853	0.002069024	0.1617	0.014353992	888	3.0226	0.002818901	0.3171	0.021622926
R2032	2854	1459	3.8905	0.00369494	0.1433	0.032686419	476	2.7568	0.002674885	0.1926	0.035391164
R2037	4809	2852	3.3429	0.00174517	0.1498	0.01311543	786	3.1248	0.002282979	0.3218	0.018755199
R204	6694	3167	4.6101	0.00204302	0.3344	0.015612355	973	3.5364	0.003610048	0.3676	0.029030449
R2044	1573	809	3.3794	0.003126026	0.1084	0.018293393	245	3.3866	0.005418361	0.4945	0.061944939
R2048	7335	4536	3.0789	0.000983215	0.1665	0.009279149	1195	2.7938	0.002364248	0.3460	0.021961464
R207	7257	3543	4.1008	0.004765473	0.1850	0.036009016	1277	4.0245	0.011518708	0.4865	0.100238409
R208	9669	5223	4.1997	0.00191961	0.2558	0.016101068	1559	3.1592	0.002164113	0.4937	0.02707628
R209	12146	7140	3.9684	0.00471761	0.1566	0.016242453	2004	2.9587	0.00185818	0.4237	0.023563664
R210	1378	1219	2.8138	0.00339755	0.2040	0.027954193	92	2.2728	0.004722805	0.2412	0.040775444
R213	8256	4563	4.7842	0.009459491	0.3372	0.091592351	1355	2.7969	0.002161658	0.3373	0.017344182
R214	7960	4307	3.6067	0.001557703	0.2128	0.017232364	1407	3.2222	0.00234735	0.3346	0.020521025

Tab. A.5: Spread estimate small and medium trades in 2014

		Large Trades					Huge Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R159	550	217	2.7346	0.003118764	0.4867	0.057371759	48	5.7533	0.012318671	0.1927	0.030474292
R197	1693	87	2.4358	0.004707005	0.5531	0.098867383	1	4.5066	6.56E-05	0.0056	7.73E-06
R201	233	49	9.4796	0.03172078	0.5629	0.147963887	8	4.4808	0.040177819	0.9683	0.840126317
R202	1691	109	1.9190	0.004849866	0.3525	0.079701112	55	0.5236	0.002439339	0.0000	0.043308502
R2023	9323	2152	2.7781	0.001330475	0.4083	0.016208289	813	3.6444	0.003372174	0.3083	0.023388296
R203	7763	2193	3.5244	0.00286808	0.4336	0.02724328	855	3.6408	0.004912909	0.4707	0.052315576
R2030	5115	1277	3.4585	0.002771809	0.3906	0.024338546	440	3.7894	0.005120978	0.2910	0.03260151
R2032	2854	644	3.7850	0.006127858	0.3680	0.034393994	275	3.4404	0.004878351	0.3564	0.041438195
R2037	4809	882	3.1819	0.002840968	0.2884	0.020382154	289	4.0614	0.005552259	0.4893	0.06053484
R204	6694	1829	3.0248	0.001444041	0.4191	0.020390356	725	3.3149	0.00267571	0.3987	0.02591724
R2044	1573	379	2.6412	0.003150112	0.4646	0.046294629	140	2.7028	0.006785228	0.4602	0.09944771
R2048	7335	1279	2.8077	0.001309162	0.4562	0.016774354	325	3.8775	0.003837994	0.3009	0.026464145
R207	7257	1774	3.4261	0.002807222	0.0642	0.04510477	663	3.4316	0.004182338	0.4777	0.048670425
R208	9669	2230	3.1232	0.002572174	0.3923	0.020893733	657	3.7391	0.002457284	0.4166	0.023940929
R209	12146	2479	3.0369	0.001788129	0.3536	0.019515983	523	3.2964	0.002626435	0.3835	0.027799346
R210	1378	61	3.3684	0.008629454	0.5150	0.123335957	6	6.6686	0.028569813	0.4284	0.181947053
R213	8256	1848	4.1618	0.011535671	0.5165	0.090173964	490	3.1871	0.00247985	0.3074	0.021154165
R214	7960	1745	3.2370	0.002390544	0.4164	0.01898745	501	3.9006	0.003588274	0.3891	0.031806286

Tab. A.6: Spread estimate results for large and huge trades in 2014

		Small Trades					Medium Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R186	67737	26353	3.8779	0.000990925	0.1878	0.006696241	15977	3.0048	0.000718272	0.3130	0.00651852
R209	12155	7175	3.5619	0.001466679	0.2127	0.013250593	1780	3.3403	0.002781848	0.5245	0.035017851
R208	8168	5041	4.7329	0.001877624	0.2957	0.016583758	957	4.0329	0.004972165	0.4870	0.051271153
R2023	7080	4128	4.1309	0.002493752	0.2084	0.018035106	1014	2.8752	0.00174445	0.3326	0.019578107
R213	7507	4459	4.1605	0.001778416	0.2952	0.016051981	926	2.8881	0.002357731	0.3379	0.02417115
R203	7204	3306	3.7429	0.001625484	0.2414	0.018726873	855	4.5303	0.005285628	0.3982	0.040371865
R207	7621	3833	3.9965	0.002090306	0.2606	0.019141917	1078	4.0120	0.004781318	0.3714	0.03533472
R204	6757	3754	4.9517	0.002114191	0.3704	0.020448052	803	4.7278	0.007573212	0.6198	0.082123316
R214	6865	4302	4.9976	0.006041822	0.2963	0.059767715	950	3.6581	0.004200726	0.3694	0.031776657
R2048	9238	5885	4.1604	0.001922377	0.2112	0.012727146	1362	3.8556	0.002552586	0.3270	0.018733562
R2037	6997	4409	4.6385	0.002985095	0.1805	0.01465404	885	3.6839	0.003345184	0.4952	0.038982161
R2030	7348	4174	3.9280	0.001811725	0.1996	0.013900417	1092	4.3222	0.004892464	0.4287	0.038256638
R210	1934	1726	2.3046	0.001788276	0.1638	0.028767742	118	3.9319	0.011076375	0.6450	0.166926824
R202	2504	2017	2.0835	0.001544164	0.1630	0.018686198	283	1.3431	0.00161147	0.1744	0.020918523
R197	1923	1756	2.4617	0.001363175	0.1714	0.020483834	116	2.6188	0.008194339	0.4635	0.128726507
I2025	2407	1906	2.3771	0.002562912	0.1816	0.02053299	217	2.2962	0.011111747	0.1413	0.050766978
I2038	1871	1399	2.2482	0.00167454	0.1120	0.021355964	231	2.2861	0.005277444	0.4353	0.079264976
I2050	2138	1634	1.8838	0.001671352	0.2233	0.029156416	219	2.0303	0.004036396	0.2467	0.042840407

Tab. A.7: Spread estimate results for small and medium trades in 2015

		Large Trades					Huge Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R186	67737	20758	2.8403	0.000626347	0.4234	0.007447654	4649	3.9965	0.002003213	0.3831	0.016449701
R209	12155	2509	3.3564	0.002649418	0.3964	0.022631673	691	4.7185	0.008939467	0.3876	0.064972276
R208	8168	1504	3.6777	0.003156045	0.3708	0.024782938	666	3.2332	0.003159303	0.3244	0.026903845
R2023	7080	1383	3.4835	0.004051476	0.4522	0.037562351	555	3.8684	0.005445436	0.3495	0.040798533
R213	7507	1660	3.2026	0.001949987	0.5469	0.02629807	462	3.7744	0.004589684	0.3565	0.037354613
R203	7204	1685	3.4601	0.003205321	0.4625	0.028294171	1358	3.3040	0.002813433	0.3492	0.023233264
R207	7621	1788	3.0824	0.002968255	0.3988	0.028520213	922	2.7281	0.001798084	0.2957	0.01645017
R204	6757	1362	3.1031	0.00234596	0.4344	0.027675636	838	3.1653	0.003832874	0.4527	0.044689453
R214	6865	1232	4.4211	0.010816922	0.4997	0.086288248	381	3.9472	0.003893902	0.5771	0.053418556
R2048	9238	1513	4.0131	0.002727993	0.3559	0.019052198	478	4.0441	0.00719612	0.3797	0.06059094
R2037	6997	1230	3.4507	0.003575937	0.3864	0.029075558	473	3.4715	0.004544705	0.4208	0.048985782
R2030	7348	1507	3.3166	0.00312169	0.4211	0.032551514	575	3.0975	0.004024449	0.3787	0.041840901
R210	1934	76	1.6339	0.003437457	0.4504	0.087783959	14	2.2857	0.009148084	0.4938	0.194885388
R202	2504	140	2.0950	0.005549834	0.4432	0.105942566	64	1.2095	0.003554559	0.0088	0.010173724
R197	1923	45	3.7194	0.007733793	0.4037	0.078522591	6	2.7273	0.017360699	0.8840	0.560883827
I2025	2407	225	1.8287	0.003947056	0.8940	0.188378892	59	2.5662	0.005890086	0.3976	0.086339085
I2038	1871	187	1.5809	0.00228631	0.2835	0.034778766	54	1.7495	0.014814598	0.7778	0.640048422
I2050	2138	198	2.3095	0.007369165	0.3255	0.082405722	87	2.0618	0.010874173	0.4783	0.234504295

Tab. A.8: Spread estimate results for large and huge trades in 2015

		Small Trades					Medium Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R159	352	125	13.6097	0.016920206	0.5699	0.064706457	30	8.0374	0.017944226	0.2727	0.053717779
R186	66488	21934	3.6264	0.000753997	0.1461	0.004289127	9953	2.9071	0.001332256	0.2698	0.008158699
R197	2109	1916	2.4485	0.001489983	0.0906	0.0093919	106	2.0538	0.003562884	0.0999	0.013356275
R2023	7569	4301	4.1989	0.002455536	0.1740	0.012034997	842	3.3117	0.002775089	0.3267	0.023253148
R203	3720	1771	4.3402	0.002500355	0.2342	0.019333618	308	4.6430	0.009166219	0.0903	0.014789147
R2030	6503	3655	4.4579	0.002579772	0.1292	0.010462612	726	3.7100	0.003244321	0.2637	0.01806183
R2032	4750	2595	4.6325	0.002589661	0.1480	0.012097468	508	3.2900	0.00443995	0.2696	0.028107683
R2035	5141	2781	3.9636	0.002342084	0.1472	0.012008275	592	3.3732	0.003903796	0.3945	0.037657673
R2037	7174	4259	3.7528	0.00179477	0.1657	0.011516777	817	3.0276	0.003293852	0.3934	0.034566713
R204	6087	3624	6.7088	0.00299071	0.2866	0.016752455	551	3.7641	0.004169514	0.5602	0.055625566
R2040	4466	2502	3.8532	0.002895658	0.1127	0.01145829	671	3.3310	0.003649907	0.3554	0.028928849
R2044	9203	6014	4.3998	0.002748392	0.1158	0.012312629	850	3.5832	0.003728821	0.3402	0.029498567
R2048	10640	6350	4.0621	0.001565147	0.1281	0.00766092	1067	3.2429	0.002468279	0.3246	0.020421307
R207	6367	3147	3.8367	0.002096413	0.2199	0.015895196	656	2.8701	0.002715773	0.1756	0.019229666
R208	7273	4220	7.5365	0.002860878	0.2122	0.011892381	665	3.9599	0.003686509	0.4197	0.033438666
R209	9992	4940	3.5486	0.001747115	0.2170	0.015802103	1260	2.9747	0.001953519	0.3899	0.020667065
R210	1296	1108	2.1478	0.001406615	0.1444	0.016362353	82	1.3312	0.003435725	0.5113	0.120210739
R211	890	703	7.8471	0.021775474	0.2013	0.078289688	38	4.2692	0.022605183	0.7286	0.371376415

Tab. A.9: Spread estimate results for small and medium trades in 2016

		Large Trades					Huge Trades				
BondList	Total Trades	Trades	S(bps)	Std. error	Lambda	Std. error	Trades	S(bps)	Std. error	Lambda	Std. error
R159	352	118	8.1850	0.014683672	0.5084	0.060927418	79	6.2159	0.014370683	0.3473	0.061214912
R186	66488	27145	2.4494	0.000386534	0.3125	0.006126296	7456	3.0186	0.001188413	0.3841	0.012155668
R197	2109	81	1.8796	0.007148878	0.6154	0.223884476	6	3.3429	0.009379161	0.0855	0.023339806
R2023	7569	1676	2.9797	0.002347112	0.3662	0.024052282	750	3.4948	0.00400019	0.3474	0.032618548
R203	3720	847	3.5154	0.003082099	0.5059	0.034638919	794	3.4546	0.002782077	0.4295	0.026180009
R2030	6503	1471	3.4047	0.002663819	0.4266	0.026690266	651	3.0216	0.003341821	0.4987	0.048617032
R2032	4750	1197	3.3045	0.002356974	0.3193	0.021498601	450	3.6805	0.006257079	0.4398	0.063668957
R2035	5141	1241	3.2826	0.002384394	0.2605	0.018060341	527	3.8544	0.005608864	0.3977	0.048665773
R2037	7174	1481	3.5135	0.002365348	0.3022	0.015699149	617	3.5989	0.003652434	0.3312	0.028116811
R204	6087	1158	3.9682	0.003489211	0.4432	0.030196686	754	3.4529	0.003407951	0.4237	0.034324793
R2040	4466	907	3.1612	0.003337093	0.3146	0.028666855	386	3.4412	0.005215423	0.2506	0.031783676
R2044	9203	1638	3.8001	0.00306071	0.3116	0.018302894	701	4.1658	0.004367096	0.3910	0.036630348
R2048	10640	2265	3.2296	0.001682599	0.2517	0.011296498	958	3.7801	0.003419206	0.3521	0.026650734
R207	6367	1540	2.9827	0.002048756	0.3225	0.01770308	1024	2.7048	0.001984847	0.3666	0.021484966
R208	7273	1573	4.0375	0.002835219	0.4191	0.022612384	815	5.0923	0.009950693	0.2925	0.046863319
R209	9992	2907	2.9584	0.00169981	0.3717	0.01492274	885	3.2933	0.00456171	0.3796	0.043502688
R210	1296	91	1.1353	0.00426509	0.4897	0.167871284	15	2.2499	0.018645695	0.5838	0.477265839
R211	890	78	18.7403	0.141181142	0.5953	0.407794957	71	6.9515	0.022198353	0.4544	0.126725256
R212	2111	94	1.7583	0.004414816	0.5074	0.120443718	29	1.4485	0.004901241	0.0119	0.03616513
R213	6660	1653	3.3255	0.002619917	0.4479	0.025033617	643	3.4656	0.004576781	0.3274	0.037202768
R214	5921	1336	2.9991	0.002059646	0.3000	0.028674953	416	4.2634	0.011745413	0.2584	0.059181154
I2025	1829	144	4.1883	0.026293119	0.1468	0.068506701	31	2.6921	0.009451521	0.1607	0.052366413
I2033	1491	181	1.2027	0.002928219	0.2373	0.044521204	56	1.6500	0.006695479	0.4546	0.173604447
I2038	1632	175	0.9841	0.001899154	0.1885	0.029661895	30	1.1870	0.009256031	0.0125	0.006424744
I2046	1947	146	1.3160	0.003078357	0.2664	0.053243431	60	1.7528	0.005961538	0.2816	0.088967299

Tab. A.10: Spread estimate results for large and huge trades in 2016

A.2 Huang and Stoll (1997) IDB vs RFQ Bond Spread Results

		IDB Trades					RFQ Trades				
BondList	Total Trades	Trades	Spread(bps)	Std. Error	Lambda	Std. Error	Trades	Spread(bps)	Std. Error	Lambda	Std. Error
R186	26794	6992	2.6817	0.00057488	0.4853	0.008143599	19802	3.2725	0.000517369	0.3210	0.010533321
R157	24972	6642	2.3897	0.002067155	0.5518	0.026084379	18330	2.8815	0.001421684	0.4143	0.043074654
R209	8163	1824	2.8962	0.001031919	0.5511	0.014490162	6339	3.2242	0.00082575	0.2311	0.010975297
R208	8507	1690	3.2289	0.001803696	0.6225	0.025213798	6817	3.9181	0.000989954	0.3266	0.018432327
R2023	3311	503	3.3980	0.003376634	0.6536	0.053321567	2808	3.8064	0.001881862	0.2125	0.020643901
R213	4404	956	3.3660	0.001940082	0.5827	0.025653872	3448	3.6196	0.001183979	0.2873	0.017506213
R203	5764	1106	4.1058	0.00648361	0.6261	0.071539401	4658	4.5383	0.003650394	0.3470	0.05856286
R207	5685	1135	3.7704	0.00346383	0.3362	0.02457458	4550	4.7465	0.002619532	0.2621	0.024604917
R204	4725	871	3.1521	0.001760898	0.5335	0.030693146	3854	3.7945	0.002076706	0.3424	0.030547441
R214	3370	863	3.4539	0.001907916	0.5476	0.021826088	2507	3.7873	0.00116946	0.3509	0.01929145
R2048	1398	306	4.6992	0.005505605	0.4751	0.040349901	1092	4.3678	0.002712248	0.2231	0.027459434
R210	1295	101	5.9458	0.016163254	0.5949	0.142366133	1194	3.3006	0.003431026	0.1134	0.029892529
R202	977	83	3.8378	0.007685452	0.6224	0.109992506	894	2.9490	0.001904527	0.2622	0.034397423
R197	1213	100	4.0742	0.00579632	0.4753	0.057345567	1113	3.0578	0.00201451	0.1580	0.020857432
I2038	412	29	5.1142	0.019795519	0.8475	0.30419869	383	2.5231	0.003181444	0.1356	0.039872584
I2050	360	48	1.8854	0.004304081	0.4023	0.066948392	312	2.6018	0.002500313	0.2821	0.062310912
R201	2042	504	4.4843	0.004127956	0.4626	0.03599886	1538	5.4394	0.002829791	0.3383	0.027233778
R206	1266	242	7.8915	0.035852106	0.4893	0.157580737	1024	8.2857	0.015066764	0.3711	0.112257565
I2025	547	27	5.4320	0.016882696	0.3779	0.106276719	520	7.9275	0.243026481	0.0125	0.007638018

Tab. A.11: Spread estimate results for IDB vs RFQ trades in 2012

		IDB Trades					RFQ Trades				
BondList	Total Trades	Trades	Spread(bps)	Std. Error	Lambda	Std. Error	Trades	Spread(bps)	Std. Error	Lambda	Std. Error
R186	61128	15783	3.1419	0.000492253	0.4760	0.005759068	45345	3.7176	0.000425257	0.2889	0.00693892
R157	20617	4618	2.9912	0.000994079	0.5178	0.014186473	15999	3.8401	0.001051158	0.3220	0.013111946
R209	15128	3162	3.9658	0.00142765	0.4610	0.012428544	11966	4.2552	0.000901978	0.1975	0.008485009
R208	11529	2274	3.6116	0.001347129	0.5586	0.0159048	9255	4.5090	0.002286416	0.3444	0.036759289
R2023	8432	1391	3.9644	0.002268088	0.4856	0.022232826	7041	4.5157	0.001226055	0.2311	0.01190021
R213	9344	1913	4.2712	0.002196205	0.6352	0.026790356	7431	4.3271	0.001522972	0.2340	0.01393708
R203	8155	1474	3.8391	0.003057555	0.6618	0.041551649	6681	4.1836	0.002083879	0.2816	0.034490948
R207	8502	1619	3.4447	0.001605718	0.5633	0.020017425	6883	3.8146	0.000974773	0.3226	0.016142318
R204	7321	1383	4.2402	0.002802922	0.5000	0.027792165	5938	4.3004	0.001274172	0.3330	0.016219679
R214	8098	1719	4.2175	0.002223186	0.5084	0.020018685	6379	5.0031	0.001950986	0.2318	0.017434563
R2048	5586	1068	4.0652	0.002233412	0.5933	0.025437137	4518	6.0521	0.004651771	0.2230	0.027735235
R2037	1927	384	4.6175	0.004567383	0.2851	0.022637866	1543	4.3990	0.002736578	0.1616	0.018348999
R2030	526	111	6.0784	0.020664811	0.3640	0.093029437	415	5.4998	0.007859124	0.1022	0.034955324
R210	2320	242	3.8009	0.005574317	0.5950	0.074112457	2078	4.9157	0.01419211	0.0807	0.02008505
R202	1731	180	4.2633	0.005402054	0.5585	0.059448084	1551	3.2954	0.001693871	0.2289	0.026569242
R197	1699	124	3.4105	0.004813528	0.6958	0.088707732	1575	4.1710	0.003210577	0.2184	0.041347061
I2025	1453	117	3.8841	0.014436237	0.7653	0.257155757	1336	3.3715	0.00281886	0.1833	0.033585347
I2038	1332	127	3.8235	0.008173039	0.5463	0.101190145	1205	3.0138	0.002682647	0.1558	0.025957629
I2050	1186	76	3.5370	0.006392051	0.6007	0.098524118	1110	3.3349	0.004082346	0.1440	0.022624796
R201	1611	296	4.6843	0.004516669	0.5111	0.039798175	1315	4.4769	0.002338797	0.4329	0.035251302
I2046	288	33	3.4955	0.014800813	0.4551	0.159853104	255	1.9035	0.002599492	0.1534	0.059993234
R206	813	125	4.5474	0.009160067	0.5091	0.077416681	688	4.5397	0.002833644	0.4376	0.050188013

Tab. A.12: Spread estimate results for IDB vs RFQ trades in 2013

		IDB Trades					RFQ Trades				
BondList	Total Trades	Trades	Spread(bps)	Std. Error	Lambda	Std. Error	Trades	Spread(bps)	Std. Error	Lambda	Std. Error
R159	550	105	3.1515	0.004636897	0.8720	0.107904126	445	4.5713	0.004036412	0.3411	0.048230759
R197	1693	88	3.4946	0.006946018	0.4371	0.077777288	1605	2.6109	0.001471104	0.1991	0.020655287
R201	233	34	6.9835	0.018540754	0.8765	0.206520831	199	14.1486	0.013841188	0.4561	0.069925314
R202	1691	116	2.5579	0.004064453	0.5935	0.084424634	1575	2.2547	0.001035478	0.1937	0.021550798
R2023	9323	1700	3.0119	0.001809492	0.6785	0.034657035	7623	3.2408	0.001533965	0.2151	0.0145204
R203	7763	1343	2.7611	0.00126922	0.4768	0.023039737	6420	3.7579	0.001728844	0.3329	0.022436707
R2030	5115	1059	3.6687	0.002420914	0.4534	0.023802956	4056	3.6265	0.001754578	0.1902	0.014278545
R2032	2854	554	2.9580	0.002785322	0.5046	0.037019954	2300	4.0104	0.003452771	0.1511	0.030188561
R2037	4809	778	3.4272	0.002742097	0.5672	0.035718696	4031	3.4849	0.001419451	0.1483	0.011882244
R204	6694	1283	3.3153	0.001592783	0.6633	0.024626571	5411	3.7409	0.001309009	0.3628	0.021273817
R2044	1573	289	3.2502	0.004072934	0.7196	0.070504612	1284	3.4894	0.002421874	0.1428	0.021938826
R2048	7335	1233	2.9937	0.001321369	0.5179	0.018383532	6102	3.0468	0.000810954	0.1907	0.011878871
R207	7257	1342	2.9081	0.001232151	0.2658	0.053605575	5915	3.8232	0.003657521	0.2546	0.025903384
R208	9669	1777	3.3592	0.001867931	0.5453	0.026987802	7892	3.6845	0.001349591	0.3049	0.017178939
R209	12146	2497	2.9929	0.001355392	0.5126	0.01847393	9649	3.5346	0.002928433	0.1950	0.015485345
R210	1378	111	2.9517	0.006133788	0.6369	0.117237254	1267	2.8036	0.003170979	0.1940	0.027481448
R213	8256	1752	2.8886	0.001247494	0.4282	0.014084842	6504	4.4257	0.006991653	0.3828	0.089978523
R214	7960	1607	3.1481	0.00187642	0.5341	0.022146193	6353	3.5770	0.001334813	0.2302	0.017300642
I2025	1588	121	2.9844	0.008577352	0.4955	0.124729302	1467	2.1599	0.001808383	0.2017	0.036454245
I2038	1643	104	1.5288	0.00357994	0.5581	0.119910042	1539	1.9883	0.001004686	0.1864	0.022545522
I2046	1068	83	2.3720	0.006504862	0.4976	0.12072939	985	1.9236	0.001118693	0.1698	0.022219085
I2050	1262	131	2.0417	0.003199723	0.4341	0.057167159	1131	1.8491	0.000920656	0.2070	0.023775392
R186	52304	14483	2.4115	0.000410787	0.5158	0.006192082	37821	2.5904	0.000293856	0.2528	0.005945005

Tab. A.13: Spread estimate results for IDB vs RFQ trades in 2014

		IDB Trades					RFQ Trades				
BondList	Total Trades	Trades	Spread(bps)	Std. Error	Lambda	Std. Error	Trades	Spread(bps)	Std. Error	Lambda	Std. Error
R186	67737	18456	2.8768	0.000594804	0.5134	0.007490103	49281	3.4876	0.000595672	0.2318	0.007672471
R209	12155	2658	3.4567	0.001661987	0.6352	0.022730869	9497	3.6684	0.001523657	0.2168	0.016961448
R208	8168	1175	4.4759	0.004026817	0.6533	0.051282861	6993	4.2682	0.001531602	0.2848	0.016865792
R2023	7080	1196	3.0266	0.002484219	0.5521	0.035004591	5884	3.9070	0.001941677	0.2407	0.018344192
R213	7507	1465	3.6617	0.002464694	0.7293	0.038854947	6042	3.6833	0.001231161	0.2886	0.018522548
R203	7204	1029	3.9748	0.004211613	0.6232	0.053685715	6175	3.5181	0.001293706	0.2946	0.023011986
R207	7621	1150	3.8819	0.00382503	0.6480	0.050861589	6471	3.6712	0.001652649	0.2442	0.017465201
R204	6757	1075	3.9117	0.004143491	0.6146	0.050524723	5682	4.3645	0.001687037	0.3958	0.027349727
R214	6865	1257	3.7739	0.003611357	0.5150	0.035036335	5608	4.7853	0.005176344	0.3265	0.057966623
R2048	9238	1639	3.6780	0.001713869	0.5047	0.019612355	7599	4.1501	0.001650296	0.2194	0.01231102
R2037	6997	1164	4.1041	0.003622106	0.6416	0.045687773	5833	4.2394	0.002140992	0.1895	0.016601749
R2030	7348	1279	3.6552	0.003070664	0.5612	0.036158221	6069	3.9782	0.001779898	0.2208	0.015378472
R210	1934	126	3.0439	0.008499308	0.6410	0.162741045	1808	2.3096	0.001520343	0.1821	0.028329932
R202	2504	178	1.8523	0.003183596	0.3838	0.057807824	2326	2.1392	0.001187922	0.1495	0.018397978
R197	1923	108	2.7307	0.007487792	0.6504	0.163437647	1815	2.4711	0.001283181	0.1778	0.020653989
I2025	2407	109	2.1416	0.00275525	0.7351	0.089079913	2298	2.4986	0.002936389	0.1923	0.038162547
I2038	1871	134	3.0794	0.007230951	0.7547	0.161506194	1737	2.2417	0.001489805	0.1157	0.019692022
I2050	2138	184	1.8923	0.002474301	0.4975	0.057393983	1954	2.1518	0.001839181	0.1969	0.027860396
I2046	2009	142	2.7026	0.006999604	0.5081	0.117518308	1867	2.0904	0.001572714	0.0871	0.014898679

Tab. A.14: Spread estimate results for IDB vs RFQ trades in 2015

		IDB Trades					RFQ Trades				
BondList	Total Trades	Trades	Spread(bps)	Std. Error	Lambda	Std. Error	Trades	Spread(bps)	Std. Error	Lambda	Std. Error
R159	352	90	7.8045	0.01684903	0.7768	0.119838581	262	9.4511	0.008529563	0.5339	0.07544308
R186	66488	18715	2.5103	0.000581278	0.4175	0.007267655	47773	3.0811	0.000444944	0.2002	0.006069586
R197	2109	80	1.9144	0.003828381	0.6148	0.117157461	2029	2.4535	0.001352509	0.0907	0.009451665
R202	2364	118	2.1930	0.006135993	0.5123	0.132581307	2246	2.1638	0.001261281	0.0783	0.011040206
R2023	7569	1444	3.5448	0.002508274	0.5245	0.028953388	6125	3.7794	0.001755999	0.1799	0.01253438
R203	3720	577	3.9925	0.003666288	0.5550	0.040525074	3143	3.7895	0.001872403	0.2915	0.020611911
R2030	6503	1297	3.6868	0.002652031	0.4799	0.026513928	5206	3.9864	0.001697543	0.1709	0.014336466
R2032	4750	920	3.9436	0.003500391	0.4504	0.029826251	3830	4.1544	0.001930048	0.1592	0.013343324
R2035	5141	1065	3.6231	0.002878227	0.4625	0.027463356	4076	3.8497	0.001807546	0.1534	0.013418467
R2037	7174	1321	3.6147	0.002390169	0.4450	0.023356257	5853	3.6030	0.001285879	0.1794	0.012048511
R204	6087	893	4.7823	0.005015378	0.6709	0.058395653	5194	5.5344	0.002095907	0.2990	0.017782323
R2040	4466	684	3.9293	0.004478963	0.4063	0.036181897	3782	3.8674	0.002130831	0.1276	0.01292123
R2044	9203	1215	4.2636	0.003583007	0.4181	0.028082578	7988	4.3568	0.00188798	0.1365	0.012847533
R2048	10640	2062	3.8922	0.002060175	0.3855	0.015812898	8578	3.6995	0.001121744	0.1410	0.008732142
R207	6367	1186	3.1033	0.001913034	0.4696	0.022611759	5181	3.4373	0.001457816	0.2138	0.019153906
R208	7273	1385	5.4126	0.005036613	0.4875	0.034383609	5888	6.1988	0.002106337	0.2260	0.014578557
R209	9992	2621	2.7625	0.001333685	0.5601	0.021813871	7371	3.4234	0.001344476	0.2265	0.015817514
R210	1296	88	1.7003	0.004131828	0.8750	0.198250189	1208	2.1239	0.001349709	0.1347	0.016350343
R211	890	59	13.8748	0.080354249	0.9277	0.500795919	831	9.4566	0.020798761	0.2011	0.058759313
R212	2111	57	2.3699	0.007202808	0.5299	0.154731608	2054	2.5101	0.003222908	0.0859	0.011743332
R213	6660	1397	3.3297	0.002298589	0.5039	0.0257593	5263	4.1461	0.001617008	0.2184	0.015412689
R214	5921	1303	3.3736	0.002031996	0.4933	0.023094165	4618	4.1502	0.002217034	0.2111	0.014721941
I2025	1829	78	2.4397	0.005138297	0.3880	0.076576681	1751	2.6973	0.005558046	0.1443	0.053966105
I2033	1491	73	0.0000	0.007834875	0.9723	5.68E+18	1418	1.8754	0.001216619	0.0869	0.012940845
I2038	1632	102	1.8082	0.00320348	0.4910	0.078762945	1530	1.9460	0.001059497	0.1203	0.013847103
I2046	1947	170	1.7318	0.003387816	0.2287	0.042345984	1777	2.0721	0.001733198	0.1103	0.013780483
I2050	1774	141	2.0837	0.003460852	0.5277	0.078176048	1633	2.0179	0.001027559	0.1235	0.01469487

Tab. A.15: Spread estimate results for IDB vs RFQ trades in 2016