The impact of news on South African sovereign bond yields

Elizabeth-Ann van der Westhuizen, Leon Marx Brümmer and Cornelis Hendrik van Schalkwyk
Department of Financial Management, University of Pretoria, Pretoria, South Africa

*CONTACT Elizabeth-Ann van der Westhuizen. Email: ann.vanderwesthuizen@up.ac.za

Abstract

A reverse event study approach was used to investigate how the South African sovereign bond yield curve reacts to headline news. The change in daily yields, calculated as the difference between the natural log of zero-coupon yields on consecutive business days, were used in the analysis. Dates of abnormal daily yield changes were identified using GARCH models. News items for the sample period were classified into categories using supervised machine learning. A regression model was fitted to determine the link between the abnormal yield changes and the news categories. The results indicated that, for abnormal increases in yield (negative news), political news had an impact on all nodes. For abnormal decreases in yield (positive news), economic news had the greatest impact on the 10-year and political news on the 15- and 20-year nodes of the yield curve.

Keywords: Reverse event study; headline news; news classification; GARCH models; machine-learning; South African sovereign bond yield curve

JEL CODES: A12; E43; E44; G12; G14; N37

1. Introduction

At a time where the world is subjected to vast amounts of information daily, government bond investors and issuers face challenges in understanding how, when, and to what extent the release of new information affects bond yields. The South African sovereign bond market, which is a well-established and liquid bond market, is influenced by the investment risk appetite of a great number of domestic and international investors with a vested interest in the impact of news on the bond market (Brand, 2018). The sovereign yield curve is a collection of all the current information available to market participants, as well as the sum of all the risks relating to the sovereign yield curve, such as economic risk, fiscal policy risk, political risk, credit rating changes, and debt (Ferrucci, 2003). Changes in the government bond yields quoted in the market, and consequently in the yield curve, can be linked to specific news events. Changes in the South African sovereign bond yields are investigated by analysing the daily changes of specific nodes in the zero-coupon yield curve derived from the South African government bond yields. The changes in the zero-coupon yield curve on consecutive trading days, as observed at each node, are used as the daily yield change series in the analysis.

Different types of headline news represent new information that impact the bond market and thus drive investment decisions. The efficient market hypothesis (EMH) states that the total market comprised of rational investors will ensure that securities fully reflect all available information, to such an extent that all new information will in a very short time be reflected in a new price for the security (Fama, 1965). Merton (1976) confirmed that the arrival of new key

information has a significant impact on asset prices in accordance with the expected market efficiency.

Several event studies on bond yields and macroeconomic news show pronounced yield movements around the release times of news (Altavilla et al., 2014; Swanson, 2011; Urbschat & Watzka, 2020). News relating to political variables can also have a significant impact on bond yields (Gade et al., 2013). The intensity of the bond yield changes depends on the types of announcements and the surprise element contained in the announcement (Hördahl et al., 2020).

In a reverse event study performed by Bech and Lengwiler (2012) on the US Treasury yield curve, the dates of interest were firstly identified; thereafter news released during the event window was linked to the abnormal changes in the constant maturity yields on the identified dates. The reverse event study method enables the researcher first to determine abnormal yield event dates and then be able to associate news relevant to the event. Accordingly, the reverse event study approach was used in this study.

In this study, three generalised autoregressive conditional heteroskedasticity (GARCH) models were used to estimate the volatility of the daily yield changes embedded in the South African zero-coupon yield curve over the sample period of 2010 to 2020.

Abnormal changes in yields are caused by some event(s) or news release(s) at the time around the event window of the change. A substantial volume of headline news is released across many different sources and the process of hand-labelling or manually categorising each item according to topics or themes can be an onerous task. To deal with the processing of this corpus, the study applied supervised machine-learning algorithms for the text classification. This method efficiently and systematically classifies and examines texts in a way that humans cannot do (Jurka et al., 2013; Takahashi et al., 2007). A sorting algorithm was used to pre-classify headline news data, after which the data were classified into five broad categories using three supervised machine-learning algorithms. A stepwise process was followed to analyse the headline news using text analysis.

By linking news released within a specified news event window to the abnormal daily yield change date for the event window, the bond market reaction to specific news could be established and quantified to achieve the main objective of the study, namely, to determine the impact of news on South African sovereign bond yields. The results from this research provided insight into which news categories drive changes in the South African sovereign bond yields.

2. Literature review and hypotheses development

Three major aspects of the literature that support the study, reverse event studies, the South African sovereign bond market and yield curve, and the influence of news on the yield curve, are discussed, followed by the development of the research hypotheses.

2.1 Reverse event studies

The event study methodology has an extensive history and was used as early as 1933 when Dolley (1933) examined stock price reactions to stock splits. The two formative papers by Ball and Brown (1968) and Fama et al. (1969) gave rise to the modern event study methodology that is known and used today (Corrado, 2011). This method is regularly used as an analytical

tool (Fama, 1991) and has many variations and applications in financial research (Peterson, 1989). In their event study of stock splits Fama et al. (1969) investigated the process by which stock prices adjust to information, thus paving the way for many successive event studies with the same focus (Bar-Yosef & Brown, 1977; Burnwal & Rakshit, 2019; Eckbo, 2008; Fleming & Remolona, 1999a; Lamoureux & Poon, 1987).

Evidence from the body of literature over several decades clearly indicates that shares (investigating price behaviour) were the most popular asset class for event studies (Armitage, 1995; Binder, 1998). Other asset classes that have been investigated using the event study methodology include: fixed income or bonds (Maul & Schiereck, 2016; Rossi, 1998), commodities (Roache & Rossi, 2010) and foreign exchange (Anderson et al., 2003; Ehrmann & Fratzscher, 2005; Galati & Ho, 2003). All these studies examined certain specific events that may have had an impact on the performance of the underlying instruments.

The premise of the event study methodology is that it is an empirical process that measures the impact of new information on the price of an asset. The methodology requires that the event be identified *a priori* to test for the significance of the event on the price of the chosen asset (Dutta, 2014; Peterson, 1989). In contrast, a reverse event study is based on the detection of the abnormal daily yield change dates before identifying the news items or news categories that can then be linked to the abnormal yield change.

Kugler and Weder di Mauro (2009) used a reverse event study approach in the currency market, Bech and Lengwiler (2012) examined the US Treasury yield curve with respect to the financial crisis during 1998 to 2011, and Ellison and Mullin (2001) studied the effect of President Clinton's health care reform proposal on pharmaceutical stock prices.

The reverse event study methodology, as applied by Janner and Schmidt (2015), is the research most closely related to this analysis. Their study tested the association between bond prices and corporate news announcements of firms listed in the prime segment of the German stock market. The researchers firstly determined the economically significant returns (ESARs) and then matched the bond data with the news data. Janner and Schmidt (2015) described the quantitative news content analysis as a linguistic classification algorithm that assigned each announcement to a specific news category. This approach allowed the researchers to draw comparisons with respect to the size and time of impact.

2.2 The South African sovereign bond market and yield curve

The structure of the South African government fixed coupon bond market is such that there are relatively few maturities with large outstanding nominal amounts (South African Reserve Bank, 2021). This, in turn, leads to the problem where there are fewer available data points to observe over time with irregularly spaced maturities from which to infer yield changes. The solution to this is to construct a yield curve from the known data points, and then infer the unknown rates.

Choudhry (2019) states that different types of yield curves exist, namely: the yield-to-maturity yield curve, the par yield curve, zero-coupon rate (spot) yield curve and the forward yield curve, which is derived from the zero-coupon (spot) yield curve (or vice versa). The zero-coupon rate yield curve (spot rate curve, or spot curve) is derived from the observed prices of coupon paying bonds and produces pure discount rates that can be used to value financial contracts accurately (Gzyl & Mayoral, 2016). The zero-coupon curve has the advantage of not having to rely on

any assumptions of the distribution of interest rates. It has no requirement to take mean-reversion into account, nor does the determination of the zero-coupon curve require that a universe of bonds exists with continuous maturities to enable the accurate reflection of observed market prices (Heath et al., 1990).

The zero-coupon yield curve is regarded as the true term structure of interest rates, as the rates exclude any reinvestment risk, and it is an important information source for central banks and investors (Collin, 2007). The true price of a bond is given as the present value (discounted at the spot rate for the tenor of the cash flow) of all future cash flows. Using no-arbitrage pricing principles, the price of the bond traded in the market should be equal to the true price of the bond, regardless of the structure used to determine the present value of the future cash flows (Moraleda & Pelsser, 2000). The zero-coupon yield curve is derived from the traded bonds through a bootstrap process. This bootstrap method of deriving the zero-coupon yield curve assumes that the true price of a bond is equal to the sum of all the future cash flows of the bond discounted at the calculated zero-coupon rate for each cash flow. The yield curve is then derived by using the shortest known rate and discounting each successive cash flow at the calculated rate. This process does sometimes require that interest rates are interpolated, as the dates of the cash flows seldom correspond exactly (Hagan & West, 2007).

2.3 The influence of news and the yield curve

News is a catalyst for asset price changes and the influence of news can be far-reaching. French and Roll (1986) found that specific types of macroeconomic news are consistently considered as important, namely: unemployment data, the producer price index (PPI) and the consumer price index (CPI). They concluded that a high degree of uncertainty in the interest rate markets can explain the magnitude of the market response. Some of the biggest increases and decreases in US bond yields took place after macroeconomic news releases (Fleming, 1997). Bond prices react instantly and significantly to major macroeconomic news announcements and market bid-offer spreads widen (Fleming & Remolona, 1999a). The intermediate maturities (1 to 5 years) of the US yield curve showed strong reactions to macroeconomic news announcements (Balduzzi et al., 2001; Faust et al., 2007; Fleming & Remolona, 1999b). Hördahl et al. (2020) studied the surprise element of macroeconomic news announcements along the yield curve by using six maturities and observed that the yield curve reacted to announcement surprises with the impact being stronger for long-term maturities than for short-term maturities. Lombardi et al. (2019) examined the reaction of sovereign bond yields to news events and central bank communications. Their study found that the longer-end of the yield curve is more sensitive to shocks contained in the content of central bank communications. De Santis (2020) studied the impact on long-term yields using news data from Bloomberg containing keywords relating to the Asset Purchase Programme (APP) and found that a positive market response to the news announcements had a significant impact on the Euro area's long-term sovereign yields.

In recent years the bond markets of developed countries as well as emerging market countries have experienced turmoil and uncertainty (Ahmad et al., 2018). Communications by government officials and politicians have an impact on the financial markets, and even more so in turbulent times (Gade et al., 2013). In general, the yield curve is affected by macroeconomic announcements (Balduzzi et al., 1996), fiscal news (Falagiarda & Gregori, 2015), ratings changes (Andritzky et al., 2007), and liquidity constraints. A study by Mohl and Sondermann (2013) further suggests that political communication is a variable that impacts bond spreads. Piazzesi (2001) investigated release dates of different macroeconomic variables, Federal Reserve (Fed) policy events, and Federal Open Market Committee (FOMC)

announcements as jumps in developing an arbitrage-fee time series model of yields that can incorporate these jumps. The rationale provided for this approach was that the information released on these dates caused strong reactions in the yield curve.

South Africa is one of a handful of emerging economies that is able to borrow in local currency for long maturities at fixed or floating rates, both from the domestic market and international markets (Eichengreen et al., 2003).

A limited amount of research has been conducted on emerging markets and news variables, and no research has been conducted on the influence of different news categories on the South African yield curve, which this study seeks to address.

2.4 Research hypotheses

The sovereign bond yield curve is regarded as the domestic interest rate risk-free curve. As such, it is used as input for the pricing, valuation and risk determination of financial instruments, or derivatives thereof, that are dependent on a risk-free interest rate. As increased volatility is generally an indication of increased risks, the change in the daily yields is used to analyse the yield curve. The volatility in these daily yields can provide key information on the efficiency and transparency of the bond market. This leads to the first research hypothesis:

> H1: Increased volatility in the sample period can be associated with abnormal daily yield changes.

The sovereign bond zero-coupon yield curve represents the collective sum of all information available relating to the sovereign issuer. Any changes in this body of knowledge are reflected as level changes on the nodes of the sovereign bond zero-coupon yield curve. The types of information that lead to bigger changes in the sovereign bond zero-coupon yield curve are studied and analysed. This forms the basis of the second research hypothesis:

> H2: Abnormal daily yield changes of the zero-coupon yield curve can be associated with specific news categories.

Not all news elicits equal responses from investors across all maturities. The extent of the change in the sovereign zero-coupon yield curve and, indeed the specific tenor that is affected by the news, is influenced by the type of news being released. This leads to the third research hypothesis:

> H3: There is a difference between the influences of different headline news categories on the abnormal zero-coupon daily yield changes.

3. Methodology and data

The broad outline of the reverse event study methodology framework was followed to achieve the main objective of this study. The analyses proceeded as follows:

1. The abnormal daily yield event dates were identified from the change in daily yields using GARCH models.
2. News was classified into five categories using supervised machine-learning algorithms.
3. The identified abnormal event dates were matched to the classified event window news.

4. The impact of the news on the rates was quantified.

3.1 Data and sample

The sample of government bonds used was selected from all listed fixed-coupon South African government bonds during the period of 2010 to 2020. The sample bonds included were all constituents of the FTSE/JSE government bond index (GOVI), as these bonds are viewed by market participants as the domestic benchmark bonds. Some of the bonds initially selected for the short-dated yield curve points matured during the period under review. Such bonds were excluded on maturity and new bonds were included in the sample as the GOVI index rebalanced based on the underlying constituents. The raw bond data collected was the end-of-day closing yield-to-maturity (YTM) rate for each of the bonds. The static term structure model of Hagan and West (2008), namely the unameliorated monotone convex bootstrap method, was used to construct the zero-coupon yield curve for each given set of bond yields for all trading dates in the sample period.

The news data consisted of daily news announcements relating to South Africa from 1 January 2010 to 31 December 2019. The news was sourced using the search functionality of Eikon newswire services provided by Refinitiv, owned by Thomson Reuters. News headlines were filtered based on relevancy using the data provider's news monitor function. All available news providers and sources that produced relevant news headlines were included, as it was imperative that all headline news announcements relating to South Africa, both domestic and international, were included in the sample of news during the event window. The total sample of headline news announcements comprised 84 341 items published by 683 sources. Each headline announcement in the dataset was assigned to either the trading date on which the announcement was made, or the following trading date if the announcement was released on a non-trading date.

The final dataset consisted of the abnormal yield event dates, daily changes in zero-coupon yield change for the abnormal yield event dates and the categorised event window headline news. The units of observation relating to the government bonds were the daily changes of the zero-coupon yields for the 2-, 5-, 10-, 15- and 20-year nodes. Each node comprised 2 498 observations over the period from 1 January 2010 to 31 December 2019. For the units of observation, only the headline news announcements released during the event window were used to identify the news categories that related to the abnormal yield change event dates.

3.2 Determination of abnormal yield changes and abnormal yield change event dates

The daily yield change series used in the GARCH analysis is calculated as the difference between the natural logarithms of the zero-coupon yields on consecutive trading days. Several recent event studies address the issue of autoregressive conditional heteroskedastic effects (ARCH) of the residuals persisting after models have been fitted to financial time series data and have applied GARCH models to address this problem (Brockett et al., 1999; Cam & Ramiah, 2012; Sabet et al., 2012).

The three GARCH models used were the standard generalised autoregressive conditional heteroskedasticity model (GARCH) as introduced by Engle (1982), the exponential GARCH model (EGARCH) by Nelson and Cao (1992) and the Glosten, Jagannathan and Runkle GARCH model (GJR) (Glosten et al., 1993).

Abnormal daily yield changes can be readily identified visually as spikes in the plots of the GARCH models. As a non-parametric approach, the Chebyshev inequality was used to identify abnormal daily yield change observations that fall outside of two standard deviations from the mean for each of the nodes for all the GARCH models (Gaur et al., 2019). These points were then used as the abnormal daily yield observations, and the dates on which the abnormal change in yield were observed were used as the abnormal event dates. The data were split into dates on which the abnormal yield changes were positive (abnormal increase event dates) and dates on which the abnormal yield changes were negative (abnormal decrease event dates).

3.3 Text analysis and news classification

News items from various sources were categorised into five distinct news topics using supervised machine-learning. For the supervised machine-learning training set a Naïve Bayes classification model was used to allocate headline news items to the news categories. The headline news categories used for the analysis were economic, political, credit, international and emerging market news.

The importance and relevance of each category were:

1. South African economic headline news (Eco)

   Words used for the categorisation of economic headline news included those relating to economic variables and concepts, inter alia, CPI, PPI and gross domestic product (GDP).

2. South African political headline news (Pol)

   Words relating to statements made by political figures and politically related words, inter alia, minister, strikes and protests were used for this categorisation.

3. South African sovereign credit rating news (Cre)

   This categorisation employed words relating to credit rating headline news announcements made by rating agencies, inter alia, Moody's, Fitch, credit, and credit rating.

4. International news references to South Africa (Int)

   Words relating to international headline news that referenced South Africa and other international occurrences that related to events in South Africa, inter alia global markets, United States, US, United Nations and UN were employed in this categorisation.

5. Emerging market news (EMM).

Words used for this categorisation related to emerging market headline news that were relevant to, or had an impact on, South Africa, inter alia, emerging markets, African markets and emerging markets. Three machine-learning algorithms were applied for the categorisation: the support vector machine (SVM) (Meyer et al., 2014), the scale linear discriminant analysis (SLDA) (Peters et al., 2012) and boosting algorithm (Logitboost) (Jurka et al., 2013).

In Table 1 the performance of each of the algorithms across every category is provided. For the three algorithms that were used, the best precision results (precision scores) and overall performance (F-score) were obtained by the SVM algorithm. The emerging market category had the lowest performance when comparing the recall and F-score percentages.

Table 1. Algorithm summary.

	SVM			SLDA			Logitboost		
	Precision	Recall	F score	Precision	Recall	F score	Precision	Recall	F score
Cre	0.99	0.91	0.95	0.98	0.89	0.93	0.99	0.92	0.95
Eco	0.93	0.78	0.85	0.89	0.7	0.78	0.86	0.8	0.83
EMM	0.87	0.26	0.4	0.53	0.29	0.37	0.89	0.25	0.39
Int	0.79	0.71	0.75	0.55	0.73	0.63	0.78	0.7	0.74
Pol	0.87	0.66	0.75	0.83	0.62	0.71	0.88	0.64	0.74

In each instance of headline news, the consensus code, as allocated by the algorithms, was used as input for the robustness and reliability check performed to verify the results.

3.4 Linking abnormal yield change to news categories

In order to link the abnormal daily yield changes on a specific date to news, a count of the number of each category's headline news items during the event window was done. In this study a short event window was used for news released around the event date. A study by Fleming and Remolona (1999a) motivated that news, and particularly surprise news announcements, have an immediate impact on treasury yields and that news released closest to the event date has the greatest influence. In this study a short event window was used for news released around the event date.

News that was released after closing time, or even very close to closing time, was only reflected in the daily change on the following trading day. The event window must therefore be at least two days. This study implemented a three-day event window (T_{-1}, T_0, T_{+1}) of South African trading dates. All headline news released during the event window was analysed and categorised. The three-day event window comprised the trading day before (T_{-1}) the event date (T_0), with (T_{+1}) being the trading day after the event date. To include news that was released on non-trading days (over weekends or during public holidays) these announcements were allocated to the first trading day after the release date. The determination of the event time window was consistent with other bond market studies. More specifically, it was the same time period as that used in the study by Ederington et al. (2015) on bond markets using bond yields.

The final data sample comprised the abnormal event dates, the zero-coupon abnormal daily yield change, and the news categories with the different frequencies of the classifications for each of the event windows. In this analysis a multiple linear regression model was fitted using forward selection, with the dependent variable, the abnormal daily yield change in each of the nodes, and the independent variables, the counts of headline news categories. The relationship could then be expressed as shown in Equationequation 1.

$$Y_t = \beta_0 + \sum_{i=1} \beta_i X_{it} + \epsilon_t \quad (1)$$

where:

Yt = the value of the dependent variable at time *t*;

β0 = the intercept or constant term;

βi = the coefficient for independent variable *i*;

*X*it = the value of independent variable *i* at time *t*; and

ϵt = the residual or error term at time *t*.

Using the data for each abnormal event date for each node, Equationequation 1 was used to generate a system of equations that were solved for βi using the ordinary least squares (OLS) method. In some cases, the assumption of normality for the residuals of the fitted model was violated. In these cases, the absolute value of the standardised residuals was greater than three. This was then corrected by removing some of the outlier values and refitting the stepwise selection multiple linear regression model.

4. Results

Table 2 shows that for all the nodes the economic news category had the most headline articles released during the event window for the abnormal increase event dates. In the 1-, 2-, 5- and 10-year nodes more than 60% of the abnormal increase event dates were classified as economic news. Only 49% of the dates classified as economic news appear in the 20-year node. Around 20% of the abnormal increase event dates were attributed to news articles categorised as political news. The third highest count of news was the emerging markets news (EMM) category.

Table 2. Aggregate contribution of news per node for daily abnormal increase event dates.

Node	Cre	Eco	EMM	Int	Pol	Total
2Y	0	23	4	3	7	37
	0%	62%	11%	8%	19%	
5Y	0	23	7	2	6	38
	0%	61%	18%	5%	16%	
10Y	0	28	3	2	9	42
	0%	67%	7%	5%	21%	
15Y	0	27	6	3	9	45
	0%	60%	13%	7%	20%	
20Y	0	21	8	4	10	43
	0%	49%	19%	9%	23%	

Note: The summary of headline news classifications for abnormal daily yield increase event dates shows the number of abnormal event dates for which the specific news category had the highest number of articles released during the event window. Abnormal increase event dates are the dates identified on which abnormal daily yield increases were observed.

For the abnormal decrease event dates (which represent news experienced as positive for the fixed income market) the results were similar to the abnormal increase event dates (negative news). Economic news had the highest count per event date. Table 3 shows the aggregate counts per news category for each node for the abnormal decrease event dates over the event windows. For the abnormal decrease event dates, there were some occurrences of news headlines classified as credit news.

Table 3. Aggregate contribution of news per node for daily abnormal decrease event dates.

Node	Cre	Eco	EMM	Int	Pol	Total
2Y	3	25	2	2	7	39
	8%	64%	5%	5%	18%	
5Y	2	25	5	2	6	40
	5%	63%	13%	5%	15%	
10Y	2	27	4	4	7	44
	5%	61%	9%	9%	16%	
15Y	1	29	3	4	9	46
	2%	63%	7%	9%	20%	
20Y	1	20	4	3	10	38
	3%	53%	11%	8%	26%	

Note: The summary of headline news classifications for abnormal daily yield decrease event dates shows the number of abnormal event dates for which the specific news category had the highest number of articles released during the event window. Abnormal decrease event dates are the dates identified on which abnormal daily yield decreases were observed.

Economic news was the headline category that had the highest count for both the abnormal increase and decrease event dates, followed by political news (except for the 5-year increase where emerging market news had a higher count) followed by emerging market, international and credit news. Multiple linear regression was used to examine the relationship between the news categories and the changes in yield on the zero-coupon curve nodes.

The results of the linear regression models are presented in Table 4 for abnormal increases in yield (daily yield increase) and Table 5 for abnormal decreases in yield (daily yield decrease). The final regression model for all nodes in the abnormal increases in yield had statistically significant F-test results of $p < .001$ and the 5-year node F test had $p < 0.05$, indicating that at least one beta coefficient differs from zero. However, in the 5-year node, the linear regression model did show that the coefficients of the selected dependent variables were not statistically significant, despite the absence of multi-collinearity, indicating that a linear model might not be suitable.

Table 4. Results of multiple linear regression for all nodes of the yield curve for abnormal increase in yields.

Node		Intercept	Cre	Eco	EMM	Int	Pol	Model
2-Year	Coeff	0.000					0.613	
	p-value	1.000					<.001	<.001
	R^2 Val							.376
	Adj R^2							.358
5-Year	Coeff	0.000				0.284	0.288	
	p-value	1.000				.087	.083	.045
	R^2 Val							.176
	Adj R^2							.124
10-Year	Coeff	0.000					0.614	
	p-value	1.000					<.001	<.001
	R^2 Val							.377
	Adj R^2							.361
15-Year	Coeff	0.000	0.299				0.573	
	p-value	1.000	.019				<.001	<.001
	R^2 Val							.618
	Adj R^2							.598
20-Year	Coeff	0.000					0.823	
	p-value	1.000					<.001	<.001
	R^2 Val							.678
	Adj R^2							.670

Note: Coefficients in table are the standardised or beta coefficients for all independent variables

Table 5. Results of multiple linear regression for all nodes of the yield curve for abnormal decrease in yields.

Node		Intercept	Cre	Eco	EMM	Int	Pol	Model
2-Year	Coeff	0.000	0.112	0.117	0.138	−0.121	−0.276	
	p-value	1.000	.523	.550	.460	.490	.153	.701
	R^2 Val							.130
	Adj R^2							−.066
5-Year	Coeff	0.000	0.130	0.274	−0.082	0.199	0.220	
	p-value	1.000	.427	.117	.607	.219	.189	.222
	R^2 Val							.179
	Adj R^2							.058
10-Year	Coeff	0.000		−0.428				
	p-value	1.000		.004				.004
	R^2 Val							.183
	Adj R^2							.164
15-Year	Coeff	0.000					−0.759	
	p-value	1.000					<.001	<.001
	R^2 Val							.576
	Adj R^2							.566
20-Year	Coeff	0.000					−0.650	
	p-value	1.000					<.001	<.001
	R^2 Val							.423
	Adj R^2							.407

Note: Coefficients in table are the standardised or beta coefficients for all independent variables

Political news was the only significant predictor of abnormal daily yield increases for the 2-, 10-, and 20-year nodes, explaining 35.8%, 36.1%, and 67% of the variances, respectively. However, in the 15-year node, both the political news and credit news categories were significant predictors of abnormal increases in yield. These two predictors explained 59,8% of the daily yield increases for the 15-year node.

The stepped-forward selection method of fitting the multiple linear regression model did not converge to a feasible solution for all the nodes in the abnormal yield decreases. The F-test results for the final regression models in the 2-year and 5-year nodes showed no significant linear relationships between news categories and decreases in yield. For these nodes, a multiple regression analysis was performed, which included all the independent variables. The results of the multiple linear regression analysis for the decrease in yields are shown in Table 5.

The final regression model for the 15- and 20-year nodes in the abnormal decreases in yields had statistically significant ($p < .001$) F-test results (the 10-year node F test had $p < 0.05$), indicating that at least one of the beta coefficients differs from zero. In the 10-year node, the economic category proved to be the strongest predictor explaining 16.4% of the decreases in yield. For the 15- and 20-year nodes, the political news category was the only statistically significant predictor of the abnormal yield decreases. The political news category explained 56.6% of the variance in the decreases in yield for the 15-year node and 40.7% of the variance in the decreases in yield for the 20-year node.

The linear regression models showed a similar fit for abnormal yield decreases and increases. The longer-term nodes displayed a linear relationship between political news and abnormal yield changes.

5. Robustness checks

5.1 Confirming the accuracy of the zero-coupon yield curve

To verify the accuracy of the derived zero-coupon rates, the difference between the actual mark-to-market (MTM) price (at the closing YTM) and the derived price using the bootstrapped zero-coupon yield rates were compared. A Wilcoxon rank-sum statistical test was used, and the results indicated that there were no statistically significant differences between the MTM prices and the derived prices.

5.2 Significance test of abnormal yield changes

The Wilcoxon signed rank test was performed to confirm that the abnormal daily yield changes were indeed statistically different from the full sample of daily yield changes. The Wilcoxon signed ranked test results affirmed that the abnormal daily yield changes were statistically different from the corresponding data population.

5.3 Text analysis analytics and reliability check

For the text analysis a two-step process was followed to ensure a robust and reliable final data sample. Firstly, the analytics function used provided statistical measures of the accuracy of the consensus category code allocated by the machine-learning algorithms. Secondly, a manual relevance and reliability check was employed with respect to each category's classification. Each headline news item per category was reviewed with respect to its classification and it was either considered to be correct or it needed to be either excluded or recoded. After this process, the final sample for statistical analysis represented all news that was relevant during the event window for each abnormal date.

An example of this process for the political news category is provided in Table 6. As the word "president" and the names of presidential aides and ministers were included in the search terms for the political news category, the machine-learning algorithms included both news headlines in the table. During the manual check, the first news item was excluded from the event window news, as the news content has no bearing on the South African fixed income markets. The second headline item in Table 6 was accepted as being correctly classified as a political news category item by the machine-learning algorithm.

Table 6. Extract of news items excluded and included in the headline news category: Political news.

Item	Date	Result	Newswire	Headline news item
1	04 05 2010	Excluded	ECLPCM	South Africa : IFAD President urges African leaders to invest in agriculture
2	10 12 2015	Remain Pol	COUREP	President sacks the finance minister

Note: Extract of political category manual robustness check.

6. Conclusion

The reverse event study methodology applied in this study allowed the researchers to link identified abnormal yield changes in the South African sovereign bond market to news categories during the period 1 January 2010 to 31 December 2019. Consistent with the first research hypothesis, using the conditional variance as determined by the GARCH models and the Chebyshev inequality, abnormal daily yield change dates were identified from the volatility

in the zero-coupon rates for each of the nodes. As per the second hypothesis, supervised machine-learning algorithms were used to categorise the news, and by fitting a multiple linear regression model, specific news categories were linked to abnormal daily yield changes. Lastly, the empirical evidence supports the hypothesis that political news has a significant impact on the daily yield changes and consequently on the zero-coupon yields, and thus on the South African sovereign bond yield curve. For the stated research hypotheses, the null hypothesis in each case has been rejected.

From the exploratory analysis of the news that was categorised, the economic news category had the most articles and the highest count of articles per event date for all of the nodes for abnormal increases and decreases in yield changes. The economic news category was only selected as a predictor for the abnormal yield changes for one node in the multiple regression analysis. This means that, although economic news contributes to the abnormal yield changes, it is seldom the main cause for abnormal yield changes in the South African sovereign bond yield curve. Rather, in the sample period, political news was the main predictor of abnormal yield changes in the South African sovereign bond market.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Ahmad, W., Mishra, A. V., & Daly, K. J. (2018). Financial connectedness of BRICS and global sovereign bond markets. *Emerging Markets Review*, *37*, 1–16. https://doi.org/10.1016/j.ememar.2018.02.006

Altavilla, C., Giannone, D., & Lenza, M. (2014). The financial and macroeconomic effects of OMT Announcements. (ECB Working Paper No. 1707). *SSRN Electronic Journal*. Available at https://doi.org/10.2139/ssrn.2464118

Anderson, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *The American Economic Review*, *93*(1), 38–62. https://doi.org/10.1257/000282803321455151

Andritzky, J. R., Bannister, G. J., & Tamirisa, N. T. (2007). The impact of macroeconomic announcements on emerging market bonds. *Emerging Markets Review*, *8*(1), 20–37. https://doi.org/10.1016/j.ememar.2006.05.001

Armitage, S. (1995). Event study methods and evidence on their performance. *Journal of Economic Surveys*, *9*(1), 25–52. https://doi.org/10.1111/j.1467-6419.1995.tb00109.x

Balduzzi, P., Elton, E. J., & Green, T. C. (1996). *Economic news and the yield curve: Evidence from the US Treasury market*. Retrieved from https://pages.stern.nyu.edu/~eelton/working_papers/Economic_News_and_Yield_Curve.pdf

Balduzzi, P., Elton, E. J., & Green, T. C. (2001). Economic news and bond prices: evidence from the U.S. Treasury market. *Journal of Financial and Quantitative Analysis*, *36*(4), 523–543. https://doi.org/10.2307/2676223

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, *6*(2), 159–178. https://doi.org/10.2307/2490232

Bar-Yosef, S., & Brown, L. D. (1977). A reexamination of stock splits using moving betas. *The Journal of Finance*, *32*(4), 1069–1080. https://doi.org/10.1111/j.1540-6261.1977.tb03310.x

Bech, M. L., & Lengwiler, Y. (2012). The financial crisis and the changing dynamics of the yield curve. (BIS Paper (No. 650). Available at https://ssrn.com/abstract=2081972

Binder, J. (1998). The event study methodology since 1969. *Review of Quantitative Finance and Accounting*, *11*(2), 111–137. https://doi.org/10.1023/A:1008295500105

Brand, R. (2018). Foreigners can't resist South African debt. That's a problem. Bloomberg, May 19. https://www.bloomberg.com/news/articles/2018-05-10/foreigners-can-t-resist-south-african-bonds-that-s-a-problem

Brockett, P. L., Chen, H.-M., & Garven, J. R. (1999). A new stochastically flexible event methodology with application to Proposition 103. *Insurance, Mathematics & Economics*, *25*(2), 197–217. https://doi.org/10.1016/S0167-6687(98)00046-8

Burnwal, A., & Rakshit, D. (2019). Announcement effect of stock split on price behaviour and market liquidity of shares: A study. *Xi'an Jianzhu Keji Daxue Xuebao/Journal of Xi'an University of Architecture and Technology*, *10*(12), 659–671. https://www.researchgate.net/profile/Debdas-Rakshit/publication/349623404

Cam, M.-A., & Ramiah, V. (2012). The influence of systematic risk factors and econometric adjustments in catastrophic event studies. *Review of Quantitative Finance and Accounting*, *42*(2), 171–189. https://doi.org/10.1007/s11156-012-0338-4

Choudhry, M. (2019). *Analysing and interpreting the yield curve*. [Online]. Wiley. https://doi.org/10.1002/9781119444695

Collin, M. (2007). The flattening of the yield curve: Causes and economic policy implications. *Economic Review*, National Bank of Belgium, 1, 47–60, June. Handle: RePEc:nbb:ecrart:y:2007:m:june:i:i:p:47-60

Corrado, C. J. (2011). Event studies: A methodology review. *Accounting and Finance*, *51*(1), 207–234. https://doi.org/10.1111/j.1467-629X.2010.00375.x

De Santis, R. A. (2020). Impact of the Asset Purchase Programme on euro area government bond yields using market news. *Economic Modelling*, *86*, 192–209. https://doi.org/10.1016/j.econmod.2019.06.011

Dolley, J. C. (1933). Characteristics and procedure of common stock split-ups. *Harvard Business Review*, *11*(3), 316–326.

Dutta, A. (2014). Parametric and Nonparametric Event Study Tests: A Review. *International Business Research*, *7*(12), 136–142. https://doi.org/10.5539/ibr.v7n12p136

Eckbo, B. E. (Ed.). (2008). *Handbook of empirical corporate finance set*. Amsterdam & Oxford: Elsevier.

Ederington, L., Guan, W., & Yang, L. (2015). Bond market event study methods. *Journal of Banking & Finance*, *58*, 281–293. https://doi.org/10.1016/j.jbankfin.2015.03.013

Ehrmann, M., & Fratzscher, M. (2005). Exchange rates and fundamentals: New evidence from real-time data. *Journal of International Money and Finance*, *24*(2), 317–341. https://doi.org/10.1016/j.jimonfin.2004.12.010

Eichengreen, B., Hausmann, R., & Panizza, U. (2003). The pain of original sin. In B. Eichengreen, & R. Hausmann (Eds.), *Other people's money: Debt denomination and financial instability in emerging market economies* (pp. 13–47). Chicago: University of Chicago Press.

Ellison, S. F., & Mullin, W. P. (2001). Gradual incorporation of information: Pharmaceutical stocks and the evolution of President Clinton's health care reform. *The Journal of Law & Economics*, *44*(1), 89–129. https://doi.org/10.1086/320269

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987–1007. https://doi.org/10.2307/1912773

Falagiarda, M., & Gregori, W. D. (2015). The impact of fiscal policy announcements by the Italian government on the sovereign spread: A comparative analysis. *European Journal of Political Economy*, *39*, 288–304. https://doi.org/10.1016/j.ejpoleco.2015.07.002

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*(1), 34–105. https://doi.org/10.1086/294743

Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance*, *46*(5), 1575–1617. https://doi.org/10.1111/j.1540-6261.1991.tb04636.x

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, *10*(1), 1–21. https://doi.org/10.2307/2525569

Faust, J., Rogers, J. H., Wang, S.-Y. B., & Wright, J. H. (2007). The high-frequency response of exchange rates and interest rates to macroeconomic announcements. *Journal of Monetary Economics*, *54*(4), 1051–1068. https://doi.org/10.1016/j.jmoneco.2006.05.015

Ferrucci, G. (2003). Empirical determinants of emerging market economies' sovereign bond spreads. (Bank of England Working Paper No. 205). *SSRN Electronic Journal*. Available at https://doi.org/10.2139/ssrn.597422

Fleming, M. J. (1997). What moves the Bond Market? *Economic Policy Review*, *3*(4): December 1997. Available at https://ssrn.com/abstract=1016154

Fleming, M. J., & Remolona, E. M. (1999a). Price formation and liquidity in the U.S. Treasury market: The response to public information. *The Journal of Finance*, *54*(5), 1901–1915. https://doi.org/10.1111/0022-1082.00172

Fleming, M. J., & Remolona, E. M. (1999b). The term structure of announcement effects. (BIS Working Paper No. 71). *SSRN Electronic Journal*. Available at https://doi.org/10.2139/ssrn.166430

French, K. R., & Roll, R. (1986). Stock return variances. *Journal of Financial Economics*, *17*(1), 5–26. https://doi.org/10.1016/0304-405X(86)90004-8

Gade, T., Glöckler, G., Salines, M., & Strodthoff, S. (2013). 'Loose lips sinking markets?': The impact of political communication on sovereign bond spreads. (ECB Occasional Paper No. 150, July 2013). *SSRN Electronic Journal*. Available at https://doi.org/10.2139/ssrn.2291222

Galati, G., & Ho, C. (2003). Macroeconomic news and the euro/dollar exchange rate. *Economic Notes*, *32*(3), 371–398. https://doi.org/10.1111/1468-0300.00118

Gaur, M., Makonin, S., Bajic, I. V., & Majumdar, A. (2019). Performance evaluation of techniques for identifying abnormal energy consumption in buildings. *IEEE Access : Practical Innovations, Open Solutions*, *7*, 62721–62733. https://doi.org/10.1109/ACCESS.2019.2915641

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

Gzyl, H., & Mayoral, S. (2016). Determination of zero-coupon and spot rates from treasury data by maximum entropy methods. *Physica A*, *456*, 38–50. https://doi.org/10.1016/j.physa.2016.02.066

Hagan, P. S., & West, G. (2007). Interpolation methods for curve construction. *Applied Mathematical Finance*, *13*(2), 89–129. https://doi.org/10.1080/13504860500396032

Hagan, P. S., & West, G. (2008, May). Methods for constructing a yield curve. Wilmott Magazine. Retrieved from https://downloads.dxfeed.com/specifications/dxLibOptions/HaganWest.pdf.

Heath, D., Jarrow, R., & Morton, A. (1990). Bond pricing and the term structure of interest rates: A discrete time approximation. *Journal of Financial and Quantitative Analysis*, *25*(4), 419–440. https://doi.org/10.2307/2331009

Hördahl, P., Remolona, E. M., & Valente, G. (2020). Expectations and risk premia at 8:30 a.m.: Deciphering the responses of bond yields to macroeconomic announcements. *Journal of Business & Economic Statistics*, *38*(1), 27–42. https://doi.org/10.1080/07350015.2018.1429278

Janner, S., & Schmidt, D. (2015). Are economically significant bond returns explained by corporate news? An examination of the German corporate bond market. *Financial Markets and Portfolio Management*, *29*(3), 271–298. https://doi.org/10.1007/s11408-015-0253-5

Jurka, T. P., Collingwood, L., Boydstun, A. E., Grossman, E., & van Atteveldt, W. (2013). RTextTools: A Supervised Learning Package for Text Classification. *R Journal*, *5*(1), 6–12. https://doi.org/10.32614/RJ-2013-001

Kugler, P., & Weder di Mauro, B. (2009). *The demise of the Swiss interest rate puzzle*. (WWZ Working Paper 04/09 (B-093)). Wirtschaftswissenschaftliches Zentrum (WWZ) der Universität Basel. https://edoc.unibas.ch/61355/1/20180306170905_5a9ebd21ecfe9.pdf

Lamoureux, C. G., & Poon, P. (1987). The market reaction to stock splits. *The Journal of Finance*, *42*(5), 1347–1370. https://doi.org/10.1111/j.1540-6261.1987.tb04370.x

Lombardi, D., Siklos, P. L., & St. Amand, S. (2019). government bond yields at the effective lower bound: International evidence. *Contemporary Economic Policy*, *37*(1), 102–120. https://doi.org/10.1111/coep.12238

Maul, D., & Schiereck, D. (2016). The bond event study methodology since 1974. *Review of Quantitative Finance and Accounting*, *48*(3), 749–787. https://doi.org/10.1007/s11156-016-0562-4

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*(1-2), 125–144. https://doi.org/10.1016/0304-405X(76)90022-2

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C., & Lin, C. (2014). e1071: Misc functions of the Department of Statistics (e1071), TU Wien. *R package version*, *1*(3). https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=3ceff42db79be884fc7b18b22fcc6ba0d8ef0faf

Mohl, P., & Sondermann, D. (2013). Has political communication during the crisis impacted sovereign bond spreads in the euro area? *Applied Economics Letters*, *20*(1), 48–61. https://doi.org/10.1080/13504851.2012.674201

Moraleda, J. M., & Pelsser, A. (2000). Forward versus spot interest rate models of the term structure. *Journal of Derivatives*, *7*(3), 9–21. https://doi.org/10.3905/jod.2000.319122

Nelson, D. B., & Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics*, *10*(2), 229–235. https://doi.org/10.1080/07350015.1992.10509902

Peters, A., Hothorn, T., Ripley, B., Therneau, T., Atkinson, B., & Hothorn, M. (2012). Package 'ipred': Improved predictors. In. R package version 0.8-13. Retrieved from http://cran.uni-muenster.de/web/packages/ipred/ipred.pdf

Peterson, P. P. (1989). Event studies: A review of issues and methodology. *Quarterly Journal of Business and Economics*, *28*(3)36–66. https://www.jstor.org/stable/40472954

Piazzesi, M. (2001). An econometric model of the yield curve with macroeconomic jump effects. (NBER Working Papers 8246). National Bureau of Economic Research, Inc. Handle: RePEc:nbr:nberwo:8246

Roache, S. K., & Rossi, M. (2010). The effects of economic news on commodity prices. *The Quarterly Review of Economics and Finance*, *50*(3), 377–385. https://doi.org/10.1016/j.qref.2010.02.007

Rossi, M. M. (1998). Economic announcements and the timing of public debt auctions. International Monetary Fund. https://doi.org/10.5089/9781451846386.001

Sabet, S. A. H., Cam, M.-A., & Heaney, R. (2012). Share market reaction to the BP oil spill and the US government moratorium on exploration. *Australian Journal of Management*, *37*(1), 61–76. https://doi.org/10.1177/0312896211427321

South African Reserve Bank. (2021). Discussion paper on Prudential Authority Government Bond Curve Review. Retrieved from https://www.resbank.co.za/en/home/publications/publication-detail-pages/prudential-authority/pa-documents-issued-for-consultation/2021/Document-Document-for-Comment-PA-Government-Bond-Curve-Review

Swanson, E. T. (2011). Let's Twist Again: A High-Frequency Event-Study Analysis of Operation Twist and Its Implications for QE2. *Brookings Papers on Economic Activity*, 2011(1), 151–188. https://doi.org/10.1353/eca.2011.0006

Takahashi, S., Takahashi, M., Takahashi, H., & Tsuda, K. (2007). Analysis of the relation between stock price returns and headline news using text categorization. In B. Apolloni, R. J. Howlett, & L. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems*. KES 2007. Lecture Notes in Computer Science(), vol 4693. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-74827-4_167

Urbschat, F., & Watzka, S. (2020). Quantitative easing in the Euro Area – An event study approach. *The Quarterly Review of Economics and Finance*, *77*, 14–36. https://doi.org/10.1016/j.qref.2019.10.008