Central Bank Communication or the Media's Interpretation: What Moves Markets?

by Scott Hendry

www.bank-banque-canada.ca

Central Bank Communication or the Media's Interpretation: What Moves Markets?

by

Scott Hendry

Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
shendry@bankofcanada.ca

Acknowledgements

Abstract

The goal of this paper is to investigate what type of information from Bank of Canada communication statements or the market commentary based on these statements has a significant effect on the volatility or level of returns in a short-term interest rate market. Two different text mining methods are used to extract interpretable themes from the document set. Bank FAD press release themes emphasizing the balance of risks, effects on GDP, labour, investment, and the CPI, the terrorist attacks of 2001, and the economic effects of SARS, BSE, blackouts, and other shocks all tended to significantly reduce short-term BAX market volatility. In contrast, discussions of oil prices, the Canadian dollar, the inflation projection and whether the economy is at capacity, and certain forward looking statements significantly increased volatility. Market news stories often offset the effects of the Bank's communication statements and were much more likely to increase market volatility while the Bank reduced it. Market stories were also more likely to have significant effects the greater the difference from the Bank news they covered but could still be important when largely replicating the original information.

JEL classification: G14, E58
Bank classification: Financial markets; Asset pricing

Résumé

L'étude a pour but d'examiner quel type d'information véhiculée dans les communications de la Banque du Canada, ou les commentaires ultérieurs des marchés, a une incidence significative sur la volatilité ou le niveau des rendements sur le marché des taux d'intérêt à court terme. L'auteur utilise deux méthodes d'exploration de textes pour extraire les thèmes interprétables de l'échantillon de documents. Les thèmes présents dans les communiqués de la Banque sur le taux directeur qui mettent l'accent sur la résultante des risques, les effets sur le PIB, le travail, l'investissement et l'IPC, les attentats terroristes de 2001 et les conséquences économiques de chocs tels que l'épidémie de SRAS, la découverte d'un cas de maladie de la vache folle et les pannes générales d'électricité ont tous tendance à abaisser substantiellement la volatilité du marché des contrats BAX à court terme. Par opposition, la présence de thèmes concernant les prix du pétrole, le dollar canadien, la projection en matière d'inflation, la capacité de production de l'économie et certaines déclarations prospectives font augmenter la volatilité de façon significative. Les commentaires des marchés font souvent contrepoids aux effets des communications de la Banque et sont beaucoup plus susceptibles d'accroître la volatilité des marchés, alors que les déclarations de la Banque ont un effet modérateur. En outre, plus les commentaires des marchés divergent de l'information véhiculée par la Banque, plus leur effet est important. Néanmoins, ceux qui reproduisent largement l'information d'origine peuvent aussi jouer un rôle notable.

Classification JEL : G14, E58
Classification de la Banque : Marchés financiers; Évaluation des actifs

1.0 Introduction

One question facing central bankers is whether financial markets respond directly to their communication statements only or whether there are (also) subsequent effects as market analysts digest these official statements and then release their own interpretations and opinions of the consequences.

This paper decomposes the effects of Bank of Canada policy statements into the effects of interest rate surprises, the effects of certain themes or topics of the accompanying communication statements, and the effects of subsequent market commentary on the policy statements. Advanced text mining techniques are used to extract the important messages from Bank of Canada communications as well as the subsequent market commentary and investigate whether these messages have significant impact on Canadian interest rates.

This paper is follow up work to Hendry and Madeley (2010) which focused on just communication statements by the Bank of Canada and did not examine whether there was any further effect from the subsequent market commentary. That paper found that Bank of Canada communication statements focusing on major shocks hitting the Canadian economy, the balance of risks to the economic projection, and certain forward looking statements had significant impacts on bond market returns. In addition, the current paper also examines the results using a second text mining technique as a robustness check of the results.

It is found that Bank FAD press release themes related to oil prices, the Canadian dollar, the inflation projection and whether the economy is at capacity, and certain forward looking statements significantly increase BAX market volatility. Emphasis on the balance of risks, effects on GDP, labour, investment, and the CPI, the terrorist attacks of 2001, and the economic effects of SARS, BSE, blackouts, and other shocks all tended to significantly reduce volatility. Many of these themes also had significant effects on BAX returns.

Themes extracted from market news stories could either reinforce or offset the effects of the Bank's themes. For the most part, market commentary tended to increase market volatility. News stories that focused on oil prices, currency effects, bond yields or spreads, terrorist attacks, global demand or imbalances, tight credit conditions, and the economic effects of SARS, BSE and other shocks all had significant positive effects on volatility. Market commentary was more likely to be significant when it focused on FAD press releases than MPR(U)s. There was also greater significance among stories that were more, rather than less, different from the original Bank document they covered.

This paper is close in spirit to Hayo and Neuenkirch (2010a, b) which also examine whether Bank of Canada communication (2010a, b) and media reporting on them (2010b) drive financial markets (bond, stock, and foreign exchange). Bank of Canada communication statements (speeches, monetary policy reports, testimonies, and press releases) are manually classified as either having "monetary policy" or "economic outlook" content. Economic outlook content is subdivided into "positive" and "negative" news while monetary policy is subdivided into "tightening" and "easing" stance. The resulting dummy variables are used in Hayo and Neuenkirch (2010a) to estimate a daily GARCH model of bond, equity, and foreign exchange market returns. They find that economic outlook content across all types of communication significantly affects Canadian markets while monetary policy content is significant only in speeches. In subsequent work, Hayo and Neuenkirch (2010b) estimate a similar daily GARCH model but also investigate the effects of similarly classified media coverage of the Bank of Canada on market returns. Bank of Canada communication on the economic outlook is again found to significantly affect bond market returns. In addition, media coverage of the Bank of Canada also significantly affects returns especially when the coverage offers an interpretation that deviates from the original official wording or when it arrives on days when there is no original Bank communication.

This paper adds to the growing literature on central bank communication and its effects on financial markets. More recently this literature has used automated text mining techniques to extract information from communication statements. See for example recent work using content analysis by Lucca and Trebbi (2009), Bligh and Hess (2009), and Rosa (2007). Much work has also been done through the manual classification of documents into various categories to investigate what affects asset prices. See for example papers on U.S. Federal Reserve communication by Gurkaynak, Sack and Swanson (2005), Kohn and Sack (2004), Ehrmann and Fratzscher (2004), Zovodny and Ginther (2005), Balke and Peterson (2002), Fettig, Rolnick and Runckle (1999), and Cook and Hahn (1988).

Section 2 outlines the methodology used in this paper while Section 3 describes the bond market data and documents summarizing Bank of Canada communication statements and the media coverage of these statements. Section 4 discusses the results of the study and the final section concludes.

2.0 Methodology

2.1 Latent Semantic Analysis

Latent Semantic Analysis (LSA) is used to extract meaning or information from Bank of Canada communications and their associated news coverage. This technique was also used in Hendry and Madeley (2010) as well as Boukus and Rosenberg (2006) to study just central bank documents in Canada and the U.S., respectively. The interested reader is referred to these papers for a more in-depth description of this methodology. LSA is extensively used in the fields of text interpretation, information retrieval, and search engine technology but has not seen many applications to economics or finance.

According to LSA, the meaning of a passage of text is a function of the words it contains. In other words, the meaning of a document is the sum of the meanings of its individual words.

The meaning of an individual word is not explicitly defined but rather is situated with respect to all the other words used in the collection of documents. LSA ignores word order but is still able to extract meaning from a text using the frequency of occurrence of the individual words. Meaning is derived from the circumstances in which a word is and is not used. The meaning of a word is a kind of average of the meaning of all the passages in which it appears and the meaning of a passage is a kind of average of the meaning of all the words it contains (Landauer, Foltz, and Laham, 1998). LSA is used to extract what are referred to as themes from the document set where a theme is a weighted combination of words that summarize a particular concept or idea. Each theme is designed to represent a different idea that is important to the overall set of documents.

LSA is a different text mining technique from the content analysis used in Lucca and Trebbi (2009), Bligh and Hess (2009), and Rosa (2007). Content analysis scores documents along a particular dimension (expansionary, contractionary, optimism, uncertainty, etc.) by counting those words that represent the target idea. The principle benefit of these techniques is the control it gives in the definition of the sentiment being examined. The researcher decides which words to count that best represent the dimension to be measured. In contrast, LSA considers all the words in the documents and extracts different meanings or themes based on how much they contribute to the distribution of meaning across all the documents. What LSA lacks in terms of ability to control or define the themes extracted from the documents, is made up for by its ability to give a good overall representation of all the ideas contained in the documents. The goal of this paper is to examine the broad spectrum of information flowing to markets from the Bank of Canada and associated news stories and allow the data to speak in terms of what is important to interest rate markets.

The first step in implementing LSA is to pre-process the documents.[1] This involves removing formatting and stop words[2] and then stemming[3] the remaining words. Finally, only words that occurred in two or more documents were included.

The next step is to organize the documents into a matrix (X) where each row represents a word (t) and each column is a document (d). Each cell of the matrix is the weighted frequency (x_{ij}) of row t's word in column d's document.

$$X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,n} \end{bmatrix}$$

The X matrix is an mxn matrix where m is the total number of words and n is the total number of documents. Typically, m is much larger than n but this need not be true.

To help distinguish the documents from each other and allow a better representation of the most important words, local and global weights are applied to the elements of the X matrix. This step gives low weights to high-frequency terms (occur in many documents) and high weights to low-frequency terms (occur in some documents). Within-document relative frequency is selected as the local term weighting scheme while Inverse Document Frequency (IDF) is used as the global term weighting function. Finally, to ensure that longer documents do not dominate the results, the documents are normalized using the cosine normalization to adjust for differences in length.

---

[1] We use the TMG program written in Matlab to perform all our LSA. See D. Zeimpekis and E. Gallopoulos (2005, 2006).
[2] Stop words are extremely common words that occur in all or virtually all documents and add little value to the meaning of the texts. These typically include pronouns, articles, prepositions, conjunctions, auxiliary verbs, days, months, etc.
[3] Stemming is the process of removing suffixes (e.g. s, es, er, est, ly, etc.) so that words sharing a common etymological root are mapped into a single "stem."

2.2 Singular Value Decomposition

The final step of LSA is to decompose the resulting *X* matrix into a small number of factors, themes, or concepts that summarize the information contained in each document. As in Hendry and Madeley (2010), singular value decomposition (SVD) is the first technique used to perform this decomposition.

SVD is what describes the relationships between the words across the documents taking account of not only where words are used but where they are not used. SVD is a generalized form of principal component analysis in which any arbitrary rectangular matrix can be exactly decomposed into the product of three characteristic matrices: $X=USV^T$. The columns of *U* (*mxn*) and *V* (*nxn*), referred to as the left and right singular vectors of X, represent the orthonormal eigenvectors of XX^T and X^TX, respectively. *S* (*nxn*) is a diagonal matrix of singular values where the diagonal entries are the non-negative square roots of the eigenvalues of XX^T and X^TX, ordered by decreasing size. SVD yields a unique decomposition of *X* up to the sign of the singular vectors in *U* and *V*.[4]

The *U* matrix describes the derived orthogonal factor values (*n* columns) as vectors of words (*m* rows) while the *V* matrix describes the documents (*n* rows) as vectors of the same factors (*n* columns). These factors may be thought of as underlying concepts or themes that run through the documents. The meaning of each document can then be characterized by a vector of weights indicating the importance of each of these underlying themes. The i,j^{th} element of *U* is the contribution or importance of the ith word to the meaning of the jth theme. Similarly, the i,j^{th} element of *V* represents the importance of the jth theme in the ith document. A theme is thus a combination of words summarizing a particular idea or concept.

The S matrix represents the importance of each theme for explaining the variance of meaning across the documents. With the elements of S ordered by decreasing magnitude, the first

---

[4] The themes were signed using the methodology outlined in Bro, Acar, and Kolda (2007).

theme is thus the most important underlying concept spanning the set of documents. Each

successive theme accounts for less of the subject or meaning variation across the documents. If

the documents are ordered chronologically, then a column of V represents the importance of each

theme in a given document while a row of V defines a time series of the importance of a given

theme. It is these rows of V that define the time series of information measures that will be

investigated for their importance in moving financial markets.

LSA does not use the full decomposition of the X matrix. Instead, a k-dimensional

approximation is found by using only the first k columns of U and V and the kxk upper left matrix

of S. This approximation is found to remove much of the useless "noise" or extraneous

information that is in the document set and focuses on only those factors explaining the important

variation in meaning across documents. Only the 20 most important themes are used in this study.

Two specifications of the model are estimated using SVD. In the first, called Model 1, the

themes are extracted from only the official Bank of Canada documents as in Hendry and Madeley

(2010).[5] Then, through a "folding-in" exercise, the weights for these themes for each news story

document are computed. To derive the theme loadings for a news story document, it is first

filtered in the same manner as the Bank documents. Any words appearing in the news stories that

do not appear in the Bank documents are deleted. The resulting term-frequency matrix Y is

transformed according to $\hat{Y} = Y^T U S^{-1}$. Multiplying the term-frequency matrix by U (which

contains the contribution of each word to each theme) and dividing by S (measuring the relative

importance of each theme) yields a matrix with weights representing the importance of each of

the original themes in the new documents. This new matrix can be appended to the original V

matrix to get the importance of each theme across all the documents. This specification has the

advantage of measuring the market's focus on the themes specified in the official Bank

---

[5] Model 1 in this paper is most similar to Model 2-o in Hendry and Madeley (2010). The difference is that, in their
paper, the data in the X matrix is demeaned as in Boukus and Rosenberg (2006) while in the current paper it is left
in its raw form. This did change and rearrange the themes somewhat but the overall results are more or less the
same.

document. The disadvantage is that new words or themes raised in the market commentaries will be ignored.

The second specification, Model 2, using SVD derives new themes from the complete set of Bank and news story documents. This will allow new themes arising in the market commentaries to play a role in explaining the meaning of the documents.

2.3 Non-Negative Matrix Factorization (NNMF)

Because the SVD weights on words in each theme can be positive or negative, the interpretation of each theme can be difficult. As a result, an alternative set of techniques to decompose the document matrix have been developed. These techniques form a non-negative decomposition of the *X* matrix for which interpretation tends to be easier.

There are several different techniques available to do NNMF. The technique selected, Nonnegative Double Singular Value Decomposition (NNDSVD), was proposed by Boutsidis and Gallopoulos (2008). Using this methodology, it is possible to determine the nonnegative factors *W (m*x*k)*, and *H* (*k*x*n*) such that *X=WH*. It has been shown that there exists a *k≤min(m,n)* such that the decomposition is exact. In most cases, however, smaller values of *k* are used which only yield an approximation of *X*. In particular, *W* and *H* are selected such that

$$X^{(k)} = \operatorname*{argmin}_{W,H} \|X - WH\|_F = \sum_{j=1}^{k} s_j C^{(j)} \tag{1}$$

where $C^{(j)} = u_j v_j^T$ and u_j and v_j are the jth singular factors of $X = \sum_{j=1}^{r} s_j u_j v_j^T$. The operator $\|\bullet\|_F$ represents the Frobenius norm, which is the square root of the sum of squares of all the elements in the matrix. Initial values of *W* and *H* are derived from the non-negative elements of u_j and v_j. In other words, the initialization assumes that the *k* factors in *W* and *H* are the positive elements of the first *k* factors or themes from the SVD of *X*, scaled by the corresponding singular value. From

this initialization, various algorithms can be used to iterate to the final values of W and H satisfying (1). These results are called Model 3.

2.4 Interest Rate Model

The interest rate model follows the same basic format as Hendry and Madeley (2010) with the addition of themes extracted from market commentary as explanatory variables. This borrows from the Weighted Least Squares estimation used in Anderson et al (2007). The first step involves estimating the following return equation using OLS.

$$R_t = \beta_0 + \beta_1 R_{t-1} + \beta_2 F_t + \sum_{k=1}^{K} \sum_{m=-3}^{0} \rho_k S_{k,t+m} + \beta_3 FAD_t + \beta_4 MPR_t$$
$$+ \sum_{i=1}^{I} \gamma_i^{pr} T_{i,t}^{pr} + \sum_{i=1}^{I} \gamma_i^{mpr} T_{i,t}^{mpr} + \sum_{i=1}^{I} \gamma_i^{mkt} T_{i,t}^{mkt} + \varepsilon_t \tag{2}$$

where R_t is the five-minute return of either the 90-day BAX or the CGB, F_t is the five-minute order flow, $S_{k,t}$ is the standardized surprise in US or Canadian macro announcement k, FAD_t is a dummy with a value of one for the five-minute interval after a 9:00am FAD press release, MPR_t is a dummy with a value of one for the five-minute interval after the 10:30am release of an MPR, $T_{i,t}^{pr}$ is the ith theme for press releases, $T_{i,t}^{mpr}$ is the ith theme for MPRs, $T_{i,t}^{ktr}$ is the ith theme for market commentary, and ε_t is the residual. Themes from press releases and MPRs enter the regressions separately so that we can investigate whether these two communication products have differential effects on returns.

While this estimate is consistent, the disturbances are likely to be highly heteroskedastic. To improve the efficiency of the estimates, a weighted least squares (WLS) technique is used. This involves taking the absolute value of the residuals from the OLS estimate of the conditional mean equation above and estimating a time-varying volatility function which is then used to perform a WLS estimation of equation (2). The volatility model fitted to the first stage residuals is

9

$$\left| \hat{\varepsilon}_t \right| = \sum_{j=1}^{J} \alpha_0 C_j + \alpha_1 \left| \hat{\varepsilon}_{t-1} \right| + \sum_{k=1}^{K} \sum_{m=-2}^{1} \lambda_k D_{k,t+m} + \sum_{m=-8}^{2} \alpha_3 FAD_{t+m} + \sum_{m=-2}^{0} \alpha_4 MPR_{t+m}$$

$$+ \alpha_5 NEWS_t + \sum_{i=1}^{I} \phi_i^{pr} T_{i,t}^{pr} + \sum_{i=1}^{I} \phi_i^{mpr} T_{i,t}^{mpr} + \sum_{i=1}^{I} \phi_i^{mkt} T_{i,t}^{mkt} + \omega_t \qquad (3)$$

where C_j is a time-of-day dummy to capture intraday patterns in volatility[6], $D_{k,t}$ is a dummy for the five-minute interval in which macro news announcement k occurs, and ω_t is the residual. $NEWS_t$ is a dummy with a value of one (zero otherwise) for the half hour before and after the release of a news story on the Bank of Canada. This dummy is designed to measure the typical volatility environment into which news stories on the Bank are released. If news stories tend to be released when volatility is high, then the effect will be positive. By accounting for already elevated volatility, the effects of the news story themes on volatility can be better identified and estimated.

Alternative specifications of the interest rate model will also be investigated in which the themes for market news stories are separated by a) type of Bank document reviewed (press releases versus MPR(U)s), b) whether it is a verbatim reprinting of a Bank document or not, and c) distance from the appropriate Bank document covered.

While the themes extracted by the methodologies above are orthonormal this does not mean they will not be correlated especially in subsamples when the data is split between press releases and MPR(U)s. As such, the equations above are estimated using orthogonalized themes in which each theme (after the first one) is first regressed against the proceeding orthogonalized themes. The resulting residual is the new orthogonalized theme with new information not contained in the proceeding themes. This is done separately for press releases, MPR(U)s, and news stories. News stories are also orthogonalized with respect to information already contained in press releases and MPR(U)s.

---

[6] Dummies are included for each five-minute interval from 8:25am to 9:00am and for each half-hour interval from 9:00am to 3:00pm. As such, there are J=19 subintervals each day.

3.0 Data

3.1 Bank of Canada Text Documents

The document set for Bank of Canada communications is the same as used in Hendry and Madeley (2010). This comprises 95 Bank of Canada text documents from two different sources, interest rate press releases and Monetary Policy Reports or Updates, spanning the 64 Fixed Announcement Dates (FAD) and 31 MPR(U)s from the start of FADS in December 2000 until September 2008.

The interest rate press releases are posted at 9am on the day of the FAD, usually a Tuesday, and contain the decision, an explanation of the decision and, possibly, a forward looking statement giving a broad flavour for the future path of rates. We deleted the beginning of each press release document which simply states the title, the rate action taken, and the city in which the decision took place.

The Monetary Policy Report (MPR) and Update (MPRU) are both published twice a year thereby providing 31 documents ranging from January 2001 to July 2008. The full reports were published in April and October, while the updates were published in January and July and were shorter in length. The MPR and MPRUs outline a detailed account of the Bank of Canada's policies and strategies. They shed light on the economic climate and its implications for inflation and often discuss press releases, speeches and publications that had been made in the months prior. We removed the bibliography section when processing the documents.

3.2 Market Commentary Documents

The market commentary documents include all Reuters news stories referring to the Bank of Canada interest rate decision or policy report on the 95 FAD and MPR(U) dates in the sample. Market news stories were all released after that day's Bank communication statement. In total, there were 610 news stories ranging from only a few words relaying the interest rate decision, to

in-depth independent analysis, to verbatim restatements of the Bank document.[7] In circumstances where there were multiple news releases within a five-minute window, all of the releases were combined together as if they were one story.

The Bank of Canada and market commentary documents were pre-processed (combining of certain linked words, applying the stop list to exclude words of little value, and stemming of words into their root) in the same manner as in Hendry and Madeley (2010).

3.3 Interest Rate Data

Intraday tick-by-tick quote and transaction data on the three-month Canadian Bankers' Acceptance futures (BAX) contract were obtained from the Montreal Exchange. This is one of the most actively traded interest rate markets in Canada and the best intraday data that is available. This data covers the period from January 2002 to June 2008. There were approximately 1114 trades/day for the BAX contract. Only days with a full intra-day sample were used. As a benchmark for Canadian short-term interest rates, the BAX futures are used for hedging over-the-counter derivatives and were the first interest rate contracts to be listed on the Montreal Exchange.

As in Hendry and Madeley (2010), BAX returns were calculated from the front contract's midpoint of the bid and ask prices for the end of each five minute interval. Order flow was calculated as the sum of the number of buys for a five-minute interval less the number of sells.

3.4 Macro News Announcements

The expected value and actual outcome for the following Canadian macro news announcements were collected from Bloomberg: Bank of Canada target rate, international securities transactions, Ivey purchasing managers index, and month over month growth in Building Permits, GDP, retail sales, and wholesale sales. The Bank of Canada target overnight

---

[7] News releases announcing official noon exchange rates or results of Treasury auctions were excluded.

rate is released at 9:00am, the Ivey purchasing managers index at 10:00am, and the rest at 8:30am. Standardized surprise values are calculated as $(Act_t$-$Exp_t)/Var_t$ where Act_t is the actual value, Exp_t is the expected value, and Var_t is the variance of the difference between them over the full sample of releases.

Announcement surprises for the following U.S. macro news releases were also included: FOMC rate decision, change in Nonfarm Payrolls, Consumer Confidence Index, ISM Manufacturing Index, Initial Jobless Claims, New Homes Sales, Retail Sales less Autos, CPI , annualized GDP, PPI, ISM non-manufacturing, and Durable Goods Orders. The FOMC decision comes out at 2:15 pm, Consumer Confidence, ISM, and New Homes Sales are issued at 10:00 am and the rest are released at 8:30 am.

4.0 Results

Tables 1 to 3 show the top 10 words and their weights in each of the top 20 themes for the three models. For Model 1, the first theme extracted from only the Bank of Canada communication statements emphasizes words like *gdp, real, project(ion), dollar, and United St(ates).* This theme summarizes the discussion of the economic projection and the importance of real factors such as gdp, business investment, and labour factors as well as external factors related to the dollar and oil prices. In Model 2, the first theme extracted from the combined set of Bank and news story documents emphasizes the projection for inflation, the inflation target, movements in bond yields, growth, and the dollar. For Model 3, using the non-negative matrix factorization on the full set of documents, the first theme focuses on the balance of risks and the inflation target. Similar interpretations can be offered for the remaining themes. Interpreting some of the themes can be quite difficult especially those derived from the first two Models using SVD since the weights on the words can be either positive or negative. The interpretation of the themes for Model 3 is substantially easier and will be the main focus of the analysis.

The remaining tables contain the results for themes in the WLS estimation of the model in equations (2) and (3). Tables 4 and 5 contain the results for the volatility and return equations, respectively, for Model 1. It can be immediately seen that certain themes had a significant effect on market volatility when communicated by the Bank of Canada through press releases at FAD decisions but that there was no significant effect when these themes were emphasized in MPR(U) documents or market news stories. The information contained in FAD press releases is already incorporated into financial markets by the time market news stories begin to be released. A similar general result can be seen for the return equation (Table 4) but here we see that there is some significance for a couple of themes when emphasized by MPR(U)s or news stories.

In general, the interpretation of the themes extracted from the documents using SVD (Models 1 and 2) was problematic given the themes had words that were both positively and negatively weighted. As such, the discussion of the interpretation of themes will be kept at a high level until the analysis of Model 3 in which the non-negative decomposition was used. Bank FAD press release discussions significantly affected bond market volatility when they focused on such topics as the forward looking statement that a "reduction in monetary stimulus will be required" (Theme 2), the price of oil and geopolitical risk (Themes 4 and 8), and terrorist attacks, confidence levels, oil prices, and employment (Themes 12 and 13). FAD press releases also significantly affected volatility when they emphasized real factors such as GDP, labour, and business investment (Theme 2), economic recovery, and accommodative policy (Theme 13), as well as energy prices and the forward looking statement that "some further reduction in monetary stimulus will be required" (Theme 8). Note that the two forward looking statements were almost the same but had opposite effects on market volatility.

In the return equation, bond prices were affected by FAD discussions of geopolitical risk, uncertainty, confidence levels, and terrorist attacks (Theme 3), the effect of these factors on business and household spending (Theme 9), the dollar and the exchange rate (Theme 15), the forward looking statement that "some modest further increase in interest rates may be required"

(Themes 5 and 6) and oil prices, the dollar and appreciation (Theme 5). Bond prices also moved significantly when the FAD press releases focused on the economic projection, housing, exports, credit, and the balance of risks (Theme 3), employment, external environment, and exports (Theme 15), and SARS, BSE, and the dollar (Theme 6). Some of the other themes have some minor significance as well but these are the primary effects.

In general, we can see that several of the important themes emphasized by Bank of Canada communication statements do significantly impact markets but the interpretation of these themes is challenging.

The themes contained in Bank of Canada communications do not again significantly impact markets when repeated in subsequent market news stories. These market news stories, however, may emphasize other themes as well and these may be significantly important. Models 2 and 3 are used to investigate this effect. These models extract themes using all the documents allowing us to investigate the importance of ideas raised in market news stories but not necessarily discussed in official communications.

Tables 6 and 7 report the results for Model 2. It is immediately obvious that there is significant information contained in market news stories that carries information over and above that contained in official Bank of Canada communication statements. Over half of the themes emphasized by market news stories are significant in the volatility and return equations. This is more than is significant for the FAD press releases. MPR(U)s continue to have no significant effect on market volatility but do now have some impact on bond returns.

For the volatility results in Table 6, note that in three of the four cases where a theme was significant for both the FAD press releases and the market news stories the coefficients had opposite signs. Since the news story themes have been orthogonalized with respect to the information contained in press releases, this is not the result of any correlation between the themes for the two types of statements. At least some of the volatility effects of FAD press releases are undone by the commentary in subsequent market news stories.

BAX market volatility responded significantly to press release themes discussing the projection for inflation, demand conditions, bond yields, the dollar, oil prices, and the real effects on labour, GDP, business investment, and exports. Themes from market news stories had significant effects when discussing the projection for inflation, the inflation target, bond yields, growth, the dollar, geopolitical factors, uncertainty, confidence levels, the balance of risks, credit conditions, housing, and the real effects on labour, GDP, business investment, and exports. Some of the major factors overlapped between the two sources of information but not all.

Table 7 shows that more of the themes were important for returns than volatility. In fact, only two of the top 20 themes were not significant for any of the information sources. All of the other themes showed some significant effect on bond returns based on discussions of different real factors, inflation, forward looking statements, risks, and shocks.

Turning to Tables 8 and 9 with the results for the extraction of themes based on the NNDSVD technique, there is a similar set of results with broad significance of most of the top 20 themes for either bond return levels or volatility. Markets are responding to information in both Bank of Canada documents (primarily press releases) as well as market news stories discussing and interpreting these documents.

The themes extracted using NNDSVD are much easier to interpret (see Tables 1 to 3) given that only positive weights are permitted. The themes are also more obviously distinct from each other with each generally emphasizing a different set of words. The themes extracted using SVD tended to exhibit overlap in their most important words even though the themes were orthogonal to one another by construction. Examining plots of the themes (not shown) reveals that some of the Model 3 themes are actually not representative of ideas expressed in FAD press releases despite being significant for such documents in either the volatility or return equation. In these cases, the theme has a very small positive weight for one or two of the FAD press releases and a

zero value everywhere else.[8] Therefore, these particular themes are acting more as a dummy variable for the five-minute interval after 9:00am on a certain FAD day than as a measure of the importance of that particular theme for that FAD. This is true for themes 2, 3, 7, 8, 9, 13, 19, and 20. These themes represent ideas obviously derived from market news stories (market reactions to FAD decisions, ex-post economist polls and forecasts, market coverage of press conferences, and dealer commentary) and not from Bank communication. These themes will be ignored for the Bank of Canada documents despite any significance they may show.

Looking closer at the results in Table 8 for BAX market volatility, it can be seen that volatility is significantly increased by discussions in FAD press releases of: oil prices and the forecast for core inflation (Theme 6); the forward looking statement that a reduction of monetary stimulus will be required (Theme 10); the Canadian dollar (Theme 11); reducing rates combined with the inflation target, the target operating band for the overnight rate, and demand conditions (Theme 17); and the projection for inflation and production when combined with comments on whether the economy is operating near capacity (Theme 18).

Some of these positive volatility effects are reinforced when picked up by market commentary. Themes 6 (oil prices), 11 (Canadian, U.S., Japanese, and Australian currencies), and 17 (reducing rates and the inflation target) also increase volatility in BAX markets when emphasized in market news stories. In addition, market news stories can also increase volatility when they focus on: bond yields and spreads from the U.S. (Theme 2); instant commentary by economists on Bank of Canada releases (Theme 3); the effects of terrorist attacks, confidence levels, and geopolitical risk on businesses and households (Theme 5); comments by the governor or a deputy governor at a press conference (Theme 7); aggressive cuts or slashed interest rates in the U.S. by the Federal Reserve and the effects on Canada (Theme 9); global aggregate demand, imbalances, and China (Theme 12); tight credit conditions, excess supply and the projection

---

[8] A stricter convergence criteria in the NNDSVD estimation procedure did ensure some of these themes were zero for all FADs and MPR(U)s. But other themes remained and continued to act as dummy variables for particular observations.

(Theme 14); and the economic effects of SARS, BSE, insurance premiums, blackouts, and forest fires (Theme 16).

The Bank of Canada was also able to reduce BAX market volatility when it emphasized the following topics: the balance of risks, the inflation target, and household or housing effects (Theme 1); effects on real GDP, labour, business investment, and the CPI (Theme 4); the terrorist attacks of 2001, confidence levels, geopolitical risk, and the effects on businesses and consumers (Theme 5); the forward looking statement that some modest further increase in rates may be required and the balance of risks (Theme 15); and the economic effects of SARS, BSE, insurance premiums, blackouts, and forest fires (Theme 16). Note that this last theme, as well as the one on the terrorist attack, reduces volatility when emphasized by the Bank but increase it when discussed in market commentary. The Bank can calm volatility when discussing the economic effects of major negative shocks but this is often offset by subsequent market commentary.

Market commentary did little to reduce market volatility. Only market commentary focused on the balance of risks, the inflation target, and household or housing effects (Theme 1) or the global outlook and the forward looking statement that stimulus will be reduced (Theme 10) had negative effects on BAX volatility. Focus on the balance of risks (Theme 1) reinforced the effects of Bank discussions while any focus on the global outlook in Theme 10 had the opposite effect as the Bank.

Bank of Canada commentary on major macroeconomic developments and the most important risks facing the economy generally succeeded in reducing market volatility. In contrast, market news stories were much more likely to increase market volatility. It is particularly interesting to note that Bank of Canada focus on most macro factors affecting the economy reduce volatility while discussions of oil prices and the Canadian dollar tend to increase interest rate volatility.

Turning now to the results for the BAX return equation in Table 9, we see again that all but two of the themes are significant for at least one of the three categories of documents. Ignoring

again themes that are significant for Bank documents but are virtually always zero and actually do not represent Bank content, there are seven significant themes for FAD press releases, three for MPR(U)s, and 13 for market news stories. BAX returns significantly increase in response to FAD press releases that emphasize the Canadian dollar (Theme 11) or the projection, inflation, production, and whether the economy is operating near capacity (Theme 18). Comments in the MPR(U)s increased returns when focused on SARS, BSE, insurance premiums, blackouts, and forest fires (Theme 16). Market news stories had positive effects on returns when concentrating on: bond yields and spreads to the U.S. (Theme 2); commentary by market economists on Bank decisions (Theme 3); real GDP, labour, business investment, and the CPI (Theme 4); comments from the Governor or a Deputy Governor at a news conference (Theme 7); and the projection, inflation, production, and whether the economy is operating near capacity (Theme 18). Only this last theme reinforced the effects of the Bank documents in increasing returns.

Short-term market returns were reduced by Bank press release discussions of: the balance of risks, the inflation target, and effects of or on households and housing (Theme 1); global aggregate demand, imbalances, and China (Theme 12); tight credit conditions, excess supply and the projection (Theme 14); and the economic effects of SARS, BSE, insurance premiums, blackouts, and forest fires (Theme 16). MPR(U)s focused on the currency (Theme 11) significantly reduced market returns as did discussions of the balance of risks and the forward looking statement that some modest further increase in rates would be required (Theme 15).

The negative effects of Themes 1 and 14 in press releases were significantly reinforced when emphasized in market commentary. Interestingly, when market news focused on terrorist attacks, confidence, geopolitical risk, and effects on consumers and businesses (Theme 5) it reduced markets returns which is contrary to the effects of the press releases and what one would have expected. Similarly, market commentary focused on different currencies (Theme 11) reduced BAX returns in contrast to the effects of press releases. Returns were also reduced by market news stories on: rate cuts in the U.S. (Theme 9); the global outlook and the forward

looking statement that stimulus will be reduced (Theme 10); and the results of dealer polls and forecasts (Theme 13).

It is also interesting to look at what themes were insignificant across the three types of documents. Only Theme 8 was insignificant for FAD press releases, MPR(U)s, and market news stories for both the volatility and return equations. The words emphasized by this theme are more related to the document style or possibly author residence than to any economic concept. The most important words are *nation*, *factbox*, *London*, *official*, and *group*. It is understandable why a theme emphasizing such words would not significantly more financial markets. Theme 19 was another that simply summarized the names of the authors of articles rather than any important economic ideas. And yet the theme was significant for some of the document types in either the volatility or return equation. As mentioned above, the significance for Bank documents is more because the variable is acting like a dummy for a particular FAD than a measure of real information. The significance for market news stories may imply that some authors are actually more influential than others no matter what they are writing about.

Table 10 shows estimates of the basis point impact of the themes from the NNDSVD model in Tables 8 and 9. In this experiment, each theme is increased, ceteris paribus, from its median value (typically zero or a very small positive) to its 95th percentile (a typical value for a document that puts a reasonably strong emphasis on that theme). Only results for themes significant in Table 8 or 9 are shown. Bank of Canada press releases have increased the standard deviation of BAX returns by up to 52 basis points for emphasizing the forward looking statement that a reduction of monetary stimulus will be required (Theme 10) or by 24 basis points when discussing reduced rates, the inflation target, the target operating bands, and demand conditions (Theme 17). In contrast, BAX market volatility was reduced by about 42 basis points by discussions of the economic effects of SARS, BSE, hikes in insurance premiums, blackouts, and forest fires (Theme 16) and by 27 basis points when emphasizing the forward looking statement that some modest further increase in rates may be required (Theme 15). The biggest effect from

market news stories was a 23 basis point increase in BAX standard deviation in response to stories about the effects of terrorist attacks, confidence levels, and geopolitical risk (Theme 5). In the return equation, the effects were much more muted with significant themes moving BAX returns by about 0.5 to just above one basis point.[9]

Further analysis of the importance of the news stories can be done by grouping the documents into sub-categories. For instance, by dividing the news stories into those covering FAD decisions and those covering MPR(U)s it is revealed that most of the significance of these stories comes from those covering press releases.[10] There is some significance to those stories covering MPR(U)s but it is much less prevalent (only Themes 5, 6, 7, 16, and 17).

The market news stories can also be sub-divided into categories depending on their difference or distance from the original Bank document that they cover. A cosine distance measure was used to divide the news stories into the closest 25% of documents, the furthest 25%, and the group in between. Performing the same WLS estimation with themes separated as such reveals that even market news commentary that is close in content to the original Bank documents can have significant effects on BAX volatility and returns. Themes extracted from news stories that are in the middle distance category are the most consistently significant followed by those that are the most different from the original documents followed by those stories that are most similar. While stories that are more different carry more significance, it is surprising just how much information is still carried in market news stories that are similar to the original documents they cover. If the category of stories that are close to the original document is narrowed to only the closest 10% or 5%, then fewer themes are found to be significant but there is still some that are. Market news stories can significantly move BAX returns even when relating content very similar to the original Bank document. It is possible that the significance of these

---

[9] It should also be noted that for some themes there was a substantial difference between the effects of an increase to the 95th percentile and an increase to the maximum value for the theme. The maximum impact observed in the sample for some of the themes was therefore quite a bit larger than reported in Table 10.

[10] These results are not shown but are available on request.

stories that are close in content to Bank documents comes from those documents that are released soon after the Bank document basically implying that lags on the Bank documents would be important at least some of the time. However, initial tests of this possibility revealed that some themes were still significant for market news stories that were close in content to the original Bank document but followed well after the release of the original.

In one final robustness check of the stability of the identified themes, a version of Model 3 was estimated after randomly choosing 30 (about 5%) of the market news stories to be excluded. This is a small enough reduction that the sample can still be considered large and yet big enough that the results could change significantly if the methodology is prone to instability. Of the 20 themes extracted, 18 were essentially the same as the full sample result. The final two themes were quite different but these were relatively unimportant in the regression results above and tended to combine a number of separate ideas into one theme rather than represent one subject. While changes in the sample of documents could affect some of the resulting themes, the results were quite robust to the addition or subtraction of new information.

Overall, the results show that Bank of Canada communication and the related media coverage can significantly move markets. This is similar to the findings in Hayo and Neuenkirch (2010a, b). They found that content focused on the economic outlook was significant but monetary policy content was only significant for speeches. Here, it was found that most of the significant themes were related to the economic outlook but there was also some significance to themes summarizing forward looking statements i.e. monetary policy content. Hayo and Neuenkirch (2010b) also found more significance in media coverage that deviated most from the original Bank documents. This result is generally replicated here in that more market news story themes are significant when classified as being far from the Bank statements they cover. There were, however, still significant themes even among those news stories that were largely replicating the original communication.

5.0 Concluding Remarks

The goal of this paper is to investigate what type of information from Bank of Canada communication statements or the market commentary based on these statements has a significant effect on the volatility or level of returns in a short-term interest rate market. Latent Semantic Analysis is used to extract themes from the document set using different types of factorization techniques. A non-negative factorization methodology is found to be very effective at extracting interpretable themes that show significant effects on both the volatility and level of returns for both the original Bank documents and the associated market commentary. For example, Bank FAD press release themes related to oil prices, the Canadian dollar, the inflation projection and whether the economy is at capacity, and certain forward looking statements are found to significantly increase BAX market volatility. Emphasis on the balance of risks, effects on GDP, labour, investment, and the CPI, the terrorist attacks of 2001, and the economic effects of SARS, BSE, blackouts, and other shocks all tended to significantly reduce volatility. Many of these themes also had significant effects on BAX returns.

Themes extracted from market news stories could either reinforce or offset the effects of the Bank's themes. For the most part, market commentary tended to increase market volatility. News stories that focused on oil prices, currency effects, bond yields or spreads, terrorist attacks, global demand or imbalances, tight credit conditions, and the economic effects of SARS, BSE and other shocks all had significant positive effects on volatility. Overall, Bank communications were much more likely to reduce market volatility while market news stories tended to increase it.

Market commentary was more likely to be significant when it focused on FAD press releases than MPR(U)s. There was also greater significance among stories that were more, rather than less, different from the original Bank document they covered.

References

Anderson, Torben, Tim Bollerslev, Francis Diebold, Clara Vega. (2007) "Real-time price discovery in global stock, bond and foreign exchange markets." *Journal of International Economics*. 73. pp. 251-277.

Balke, Nathan S., and D'Ann Petersen. (2002) "How Well Does the Beige Book Reflect Economic Activity: Evaluating Qualitative Information Quantitatively." *Journal of Money Credit and Banking*. 34. 114-136.

Bligh, Michelle and Gregory Hess. (2009) "A Quantitative Assessment of the Qualitative Aspects of Chairman Greenspan's Communications." Unpublished manuscript.

Boukus, Ellyn, and Joshua V. Rosenberg. (2006) "The Information Content of FOMC Minutes." Unpublished manuscript.

Boutsidis, C. and E. Gallopoulos. (2008) "SVD Based Initialization: A Head Start for Nonnegative Matrix Factorization." Pattern Recognition. 41(4) pp. 1350-1362.

Cook, Timothy, and Thomas Hahn. (1988) "The Information Content of Discount Rate Announcements and Their Effect on Market Interest Rates." *Journal of Money, Credit and Banking*. 20. No. 2. pp. 167-180.

Ehrmann, Michael, and Marcel Fratzscher. (2005) "Central Bank Communication: Different Strategies, Same Effectiveness?" Working Paper Series 488, European Central Bank.

Fay, C. and T. Gravelle. (2008). "Has the Inclusion of Forward Looking Statements in Monetary Policy Communications Made the Bank of Canada more Transparent?" Bank of Canada unpublished manuscript.

Fettig, David, Arthur J. Rolnick, and David E. Runkle. (1999) "The Federal Reserve's Beige Book: A Better Mirror Than Crystal Ball." *Federal Reserve Bank of Minneapolis The Region*. 10-13. pp. 28-32.

Gürkaynak, Refet S., Brian Sack, and Eric T. Swanson. (2005) "Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking*. 1. pp. 55-93.

Hayo, B. and M. Neuenkirch. (2010a) "Domestic or U.S. News: What Drives Canadian Financial Markets?" *Economic Enquiry*, forthcoming.

_____. (2010b) "Bank of Canada Communication, Media Coverage, and Financial Market Reactions." MAGKS Joint Discussion Paper Series in Economics No. 20-2010.

Hendry, Scott and Alison Madeley. (2010). "Text Mining and the Information Content of Bank of Canada Communications." Bank of Canada Working Paper 2010-31.

Kohn, Donald and Brian Sack. (2004) "Central Bank Talk: Does It Matter and Why?" *Macroeconomics, Monetary Policy, and Financial Stability*. Ottawa: Bank of Canada. pp. 175-206.

Landauer, Thomas, Peter Foltz, and Darrell Laham. (1998) "An Introduction to Latent Semantic Analysis." *Discourse Processes*, 25, pp. 259-284.

Lucca, David and Francesco Trebbi. (2009) "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements". Unpublished manuscript.

Rosa, Carlo. (2007) "Providing Content to ECB Announcements". *Rivista Internazionale di Scienze Sociali*, CXV (4), pp. 525-546.

Zeimpekis, D. and E. Gallopoulos, (2006) TMG: A MATLAB toolbox for generating

    term document matrices from text collections, In "Grouping Multidimensional Data:

    Recent Advances in Clustering" (J. Kogan, C. Nicholas, and M. Teboulle, eds.),

    Springer, Berlin, pp. 187-210.

Zeimpekis, D. and E. Gallopoulos, (2005) "Design of a MATLAB toolbox for term-

    document matrix generation". In Proc. Workshop on Clustering High Dimensional

    Data and its Applications, (held in conjunction with 5th SIAM Int'l Conf. Data

    Mining), (I.S. Dhillon, J. Kogan and J. Ghosh eds.), pp. 38-48, Newport Beach,

    California.

Zavodny, Madeline, and Donna K. Ginther. (2005) "Does the Beige Book Move

    Financial Markets?" *Southern Economic Journal.* 72, No. 1, 138-151.

Table 1: U matrix – Weights on words in top 20 themes for Model 1 Folding-in

Theme1		Theme2		Theme3		Theme4		Theme5		Theme6		Theme7	
gdp	0.1704297	reduce1	0.2423446	geopolit	0.2230011	willb	0.3239639	rais	0.247466	bear	0.1788066	nearterm	0.1681456
real	0.1574417	willb	0.2359291	uncertain	0.1877746	reduce1	0.323558	somemodest	0.2350933	sar	0.1720117	further1	0.1641648
project	0.154749	requir	0.1813409	confid	0.1833658	oil	0.1752054	increase1	0.2350933	dollar	0.1558396	credit	0.1552227
dollar	0.1368022	risk	0.1696654	recoveri	0.1635226	requir	0.1543366	mayb	0.1987688	prospect	0.1465835	condit	0.1428746
unitedst	0.1288716	rais	0.1669774	reduc	0.1576455	stimul	0.1326765	further1	0.1741173	bse	0.1458103	likelytob	0.1339508
rise	0.123883	stimul	0.1560253					oil	0.1543001			dollar	0.1321154
high	0.1214689	bankrat	0.1438265	risk	-0.1407498	employ	-0.1272267			rais	-0.1387173	sar	0.1317992
labour	0.1188356	downsid	0.1378012	consistent1	-0.1447744	unitedst	-0.1334735	credit	-0.1597979	further1	-0.178664		
busi	0.1187779			upsid	-0.1519241	consistent1	-0.1541798	upsid	-0.1699872	mayb	-0.1956213	reduc	-0.1314637
oil	0.118713	real	-0.1379535	downsid	-0.1519681	low	-0.1563151	reduce1	-0.1733175	somemodest	-0.2196321	action	-0.1478835
		gdp	-0.1631453	project	-0.2152188	reduc	-0.1706578	willb	-0.1892744	increase1	-0.2196321	consistent1	-0.2392077

Theme8		Theme9		Theme10		Theme11		Theme12		Theme13		Theme14	
geopolit	0.2469126	confid	0.2317341	employ	0.2690684	somefurth	0.192161	terrorist	0.1556249	background	0.2507326	extern	0.2190043
east	0.1567448	terrorist	0.1794119	recoveri	0.2555125	appropriate1	0.180124	confid	0.1547282	shortterm	0.187189	present	0.1633731
appropriate1	0.1431518	attack	0.1491081	nearterm	0.1731472	energi	0.1698438	employ	0.1447807	rais	0.1615725	context	0.1593052
oil	0.139526	consistent1	0.1387335	furthermeasur	0.1650899	converg	0.1648864	nearterm	0.1373264	oil	0.1547109	inform	0.155605
rais	0.13223	busi	0.1287891	likelytob	0.1280217	accommod	0.1565401	oil	0.1346658	attack	0.1154831	accumul	0.1535406
converg	0.1294751					major	0.1404527	counteract	0.1328749	except	0.113058	nearterm	0.1535101
		extern	-0.1326929	attack	-0.1280823			net	0.127029			nation	0.1365607
energi	-0.1255594	rise	-0.1327607	confid	-0.1351749	consistent1	-0.1234883			confirm	-0.1132315		
reduc	-0.1365768	futur	-0.1332971	energi	-0.1364188	further1	-0.1319779	gross	-0.1202423	recoveri	-0.1398707	somefurth	-0.1502644
slowdown	-0.1376722	oil	-0.1724112	terrorist	-0.1643438	timelyremov	-0.1428581	sar	-0.1284105	real	-0.1539573	resumpt	-0.1525108
reduce1	-0.1495896	control	-0.1726046	geopolit	-0.2222893	geopolit	-0.1471179	bear	-0.1428905	gdp	-0.1674982	employ	-0.1844633

Theme15		Theme16		Theme17		Theme18		Theme19		Theme20	
dollar	0.1961214	sar	0.2122251	gst	0.1689226	nearterm	0.1919645	analysi	0.211484	inventori	0.1585511
prior	0.1650323	bse	0.1826009	east	0.148496	thought	0.1631791	presenc	0.1435694	final	0.1344009
background	0.1571249	somefurth	0.1652235	accumul	0.1428443	east	0.1205535	factor	0.1378911	due	0.1303896
part	0.1541013	shortterm	0.1509758	recoveri	0.1315031	realign	0.1160647	realign	0.126752	deceler	0.1193453
exchang	0.147398	tone	0.1089794	credit	0.125854			reduc	0.107328	reduce1	0.1135858
pronounc	0.1400575			confirm	0.1254806	ahead	-0.1283644			outlin	0.1125359
data	0.1301288	include	-0.1118193			credit	-0.1322643	timelyremov	-0.1240898	firm1	0.1123647
		prospect	-0.1151916	action	-0.1237592	furthermeasur	-0.1500695	except	-0.1450212		
environ	-0.1445259	mayb	-0.1213128	slowdown	-0.1264885	context	-0.1629009	somefurth	-0.1455809	plai	-0.110761
extern	-0.1620059	forward	-0.1214624	employ	-0.1713339	market	-0.1630721	watch	-0.17265	market	-0.1241778
employ	-0.3655646	relat	-0.123424	converg	-0.1899114	extern	-0.1675837	bear	-0.2279441	analysi	-0.1418738

Table 2: U matrix – Weights on words in top 20 themes for Model 2 All documents

Theme1		Theme2		Theme3		Theme4		Theme5		Theme6		Theme7	
project	0.1860746	bond	0.4240184	cut	0.2159014	gdp	0.1584212	cut	0.1981129	oil	0.3365243	governor	0.228936
inflat	0.1725104	yield	0.3592264	think	0.1865732	real	0.1470327	confid	0.1931498	hike	0.1686129	told	0.2091321
bond	0.1638331	cent	0.1896207	economist	0.1764806	busi	0.135636	uncertain	0.1757336	stimul	0.140232	boc	0.201524
yield	0.1388671	dollar	0.1252905	sai	0.1362509			bankrat	0.1541415	rais	0.1329724	confer	0.1612213
target	0.1307336	cut	0.1188656	dealer	0.1270736	overnight	-0.1360989	reduc	0.1503457			dollar	0.1475427
grow	0.1188789	treasuri	0.1087641	hike	0.1177231	risk	-0.141036	activ	0.1471777	downsid	-0.1276477	thursdai	0.1436704
dollar	0.1179189			forecast	0.1146745	yield	-0.1456061	geopolit	0.1428137	upsid	-0.1379805	new	0.1287092
demand	0.1171418	cpi	-0.1068713			bond	-0.1598765			risk	-0.1467087		
price	0.1112158	target	-0.1133642	project	-0.1101505	oper	-0.1702324	balanc	-0.1335592	project	-0.1621158	instant	-0.1209312
global	0.1098187	inflat	-0.1185073	bond	-0.3010463	band	-0.1719588	risk	-0.1610771	credit	-0.1659479	think	-0.1520244
		project	-0.1881487	yield	-0.3080486	target	-0.192413	project	-0.2748387	cut	-0.260396	economist	-0.1826008

Theme8		Theme9		Theme10		Theme11		Theme12		Theme13		Theme14	
nation	0.2804773	cut	0.32604	dealer	0.1968224	dollar	0.2927388	cut	0.2275282	confid	0.1998219	credit	0.2452717
factbox	0.27738	oil	0.2715319	reduce1	0.181798	currenc	0.1991915	oil	0.1995859	attack	0.1513743	condit	0.1572835
london	0.2513357	forecast	0.1780643	willb	0.1690268	york	0.1547644	global	0.159749	somemodest	0.1470617	tight	0.1571882
offici	0.2455705	wti	0.165036	global	0.1660372	sar	0.1538178	outlook	0.1265923	increase1	0.1442834	geopolit	0.1571599
group	0.2069582	project	0.161906	relat	0.148641	demand	0.1182884	relat	0.1244321	requir	0.1299364	willb	0.1419223
borrow	0.1763144	dealer	0.1256461	cut	0.1334934	versu	0.1138393	economist	0.1217664	further1	0.1239337	stimul	0.1415486
aim	0.1678053	barrel	0.1236903	requir	0.1281333							oil	0.1227436
move	0.1501001					confid	-0.1208504	reduc	-0.1199681	secur	-0.1202931		
actual	0.1494999	offici	-0.1257648	total	-0.1312919	yield	-0.1217528	action	-0.124506	poll	-0.1356033	govern	-0.1170175
		london	-0.1395977	oil	-0.1468166	governor	-0.1505675	dealer	-0.1299133	dealer	-0.183279	increase1	-0.184458
governor	-0.1425058	factbox	-0.1526523	inflat	-0.1814204	bond	-0.1825316	hike	-0.2672097	sar	-0.2103893	somemodest	-0.1927554

Theme15		Theme16		Theme17		Theme18		Theme19		Theme20	
sar	0.2223905	sar	0.1928393	reduc	0.2909634	project	0.2454168	hous	0.1752266	york	0.2906337
further1	0.1791039	confid	0.1566807	slow	0.1591121	condit	0.1664533	inflat	0.1066582	cad	0.2457719
requir	0.1647103	hike	0.1538372	pace	0.1457453	open	0.1375104	insur	0.0992975	balanc	0.137189
somemodest	0.157837	geopolit	0.1147551	stimul	0.1278387	credit	0.1114887			risk	0.1228271
increase1	0.1538908							support	-0.1013061	dollar	0.109953
suppli	0.1464088	louis	-0.1122758	confid	-0.1276312	held	-0.1123484	cumul	-0.10435	boc	0.1064579
demand	0.1462664	egan	-0.1122758	geopolit	-0.1300847	forecast	-0.1145493	oil	-0.1067646	insur	0.097625
cut	0.1406207	inform	-0.1245783	busi	-0.1308668	risk	-0.1183068	aggreg	-0.1191711		
		boc	-0.1318013	somemodest	-0.1319983	outlook	-0.1246721	terrorist	-0.1213042	schedul	-0.1014582
consistent1	-0.1216475	edit	-0.1412546	increase1	-0.1350069	repeat	-0.1246894	confid	-0.1341145	strong	-0.1063151
dollar	-0.1328198	oil	-0.1690164	oil	-0.1384257	gdp	-0.1405645	busi	-0.1590786	appreci	-0.1119133

Table 3: W matrix – Weights on words in top 20 themes for Model 3 - NNDSVD

Theme1		Theme2		Theme3		Theme4		Theme5		Theme6		Theme7	
risk	0.4666961	bond	0.5977583	think	0.35895	gdp	0.2460732	confid	0.4694178	oil	0.6410387	governor	0.3893562
upsid	0.2654672	yield	0.5245306	economist	0.3492923	real	0.2424189	attack	0.2663033	barrel	0.2512438	told	0.3761752
consistent1	0.2592805	cent	0.221732	instant	0.2121722	labour	0.2154302	terrorist	0.257866	wti	0.2327099	boc	0.3178725
balanc	0.2373632	treasuri	0.1564638	reaction	0.2080395	busi	0.168181	busi	0.2543366	price	0.2110674	confer	0.2878353
downsid	0.1928829	close	0.1175197	bit	0.180923	cpi	0.1497999	geopolit	0.2019801	crude	0.2035116	thursdai	0.2291351
target	0.1865442	previou	0.1066846	sai	0.1656708	increas	0.1489349	consum	0.1756344	core	0.1759241	new	0.2192387
hous	0.1821018	unitedst	0.1055668	senior	0.1625787	invest	0.1476078	normal	0.1716784	inflat	0.1703176	sai	0.1665994
inflat	0.1651455	spread	0.1031657	fact	0.1473962	sector	0.1362709	uncertain	0.1673612	forecast	0.1647889	deputi	0.1498591
household	0.1620971	bill	0.0985614	pretti	0.1454871	project	0.1351564	underpin	0.1602244	total	0.1257351	ottawa	0.1433527
current	0.1471824	gain	0.091483	don	0.1447936	household	0.1345998	improv	0.1575391	texa	0.1053494	report	0.1180561

Theme8		Theme9		Theme10		Theme11		Theme12		Theme13		Theme14	
factbox	0.3730754	cut	0.7652418	willb	0.3419713	dollar	0.4425729	global	0.2443568	dealer	0.4644074	credit	0.4588195
london	0.3385865	point	0.1759926	reduce1	0.3313867	york	0.2675595	aggreg	0.2274201	move	0.3373083	condit	0.3395001
nation	0.3375388	federalreserv	0.1663807	global	0.2695679	currenc	0.2547444	realign	0.2192622	poll	0.3085121	project	0.2510587
offici	0.3352961	aggress	0.1294831	requir	0.2446211	versu	0.2322105	demand	0.2111561	action	0.2955588	tight	0.2025153
group	0.2811362	slash	0.1145121	outlook	0.2225289	yen	0.1907005	includ	0.1988073	secur	0.2830186	suppli	0.1806656
aim	0.240827	centr	0.1040101	stimul	0.213044	trade	0.1849309	imbal	0.1639865	merril	0.1514272	excess	0.1545469
borrow	0.2224311	forecast	0.0892364	develop	0.1818249	unitedst	0.1584175	relat	0.1601687	lynch	0.1514272	demand	0.1497774
actual	0.2151285	bank	0.0887355	relat	0.181273	cad	0.1511939	solid	0.1579804	cibc	0.1490167	likelytob	0.1296944
cycl	0.1798057	eas	0.0882361	oper	0.1759681	australian	0.1459442	china	0.1545167	casgrain	0.1394666	signific	0.1240563
major	0.1692318	martin	0.0881777	target	0.1721018	cent	0.1330733	presenc	0.1458611	deutsch	0.1390032	balanc	0.117481

Theme15		Theme16		Theme17		Theme18		Theme19		Theme20	
somemodest	0.4184688	sar	0.5139439	reduc	0.274036	project	0.6195882	write	0.3526691	hike	0.5359114
increase1	0.4099095	bse	0.2840685	inflat	0.2169995	product	0.2180695	edit	0.3516001	rais	0.2025263
further1	0.3316952	slack	0.2539606	target	0.2048927	period	0.1930826	louis	0.3399678	tight	0.1707007
wouldb	0.256474	insur	0.1419731	inform	0.2035372	inflat	0.1920663	egan	0.3399678	forecast	0.1674062
mayb	0.2273302	forest	0.1418572	bankrat	0.1993066	capac	0.1744695	gallowai	0.2325446	bank	0.1546251
requir	0.21122	blackout	0.1410463	band	0.1946108	slight	0.1602445	peter	0.2078545	need	0.1443482
balanc	0.1956821	inflat	0.1384046	pace	0.1788616	total	0.1434687	wilson	0.1671599	dollar	0.1216756
suppli	0.1715529	ontario	0.1341497	total	0.1605206	oper	0.1390503	mccrank	0.1610089	paus	0.1205936
aggreg	0.1307873	fire	0.1318202	consist	0.1600893	upsid	0.1357964	rob	0.1367426	centr	0.1147152
tilt	0.1265704	rebound	0.1186674	activ	0.1542348	target	0.1236885	hous	0.1331815	statement	0.108271

Table 4: BAX Market Return Volatility Equation Estimate: Folding In Exercise

Variable	Press Releases	MPR(U)s	Market News Stories
THEME1	43.20	15.46	-0.37
THEME2	88.25***	-25.25	-7.58
THEME3	11.21	6.70	6.60
THEME4	19.50**	-0.75	-3.64
THEME5	8.22*	11.58	-2.00
THEME6	9.57	7.64	3.70
THEME7	6.72	-3.83	-1.47
THEME8	28.48***	-11.08	1.85
THEME9	13.12*	8.73	-7.26
THEME10	4.92	-21.17	5.29
THEME11	1.46	3.73	3.43
THEME12	31.57***	-5.94	-9.36
THEME13	46.31***	16.72	-2.61
THEME14	10.16	-8.65	12.99
THEME15	7.04	10.93	0.13
THEME16	-4.01	8.54	-5.31
THEME17	24.90***	-16.39	8.59
THEME18	-0.66	11.43	2.98
THEME19	23.68***	29.02	-1.28
THEME20	8.62	27.29	-6.42
FAD_{t+2}	0.17		
FAD_{t+1}	1.51**		
FAD_t	12.69***		
FAD_{t-1}	40.85***		
FAD_{t-2}	-2.53***		
FAD_{t-3}	2.39***		
FAD_{t-4}	3.93***		
FAD_{t-5}	1.79***		
FAD_{t-6}	1.05*		
FAD_{t-7}	2.04***		
FAD_{t-8}	1.90***		
MPR		-2.15***	
MPR_{t-1}		8.87***	
MPR_{t-2}		2.65***	
$NEWS_t$			0.50***
R-Bar**2	0.17		

Notes: This table presents estimates of equation (3) from the text. The dependent variable is the absolute value of the residual from an OLS estimate of equation (2) for the five-minute return in the BAX market. FAD and MPR variables are dummies with a value of 1 for the five-minute interval after the event and zero otherwise. NEWS is a dummy variable equal to one for the period a news story was released as well as for the half hour before and after, and zero otherwise. Results for dummies representing the time of day and whether there was a particular Canadian or U.S. macro news announcement are omitted to save space. *** is significant at the 1% level, ** is significant at the 5% level, and * is significant at the 10% level. All coefficients have been multiplied by 10^5.

Table 5: BAX Return WLS Estimates: Folding In Exercise

Variable	Press Releases	MPR(U)s	Market News Stories	Other
THEME1	177.96*	-72.76	-1.40	
THEME2	83.10**	8.46	0.40	
THEME3	116.50***	-9.59	-1.95	
THEME4	16.66	-5.27	-4.70	
THEME5	36.05***	13.24	-5.27	
THEME6	-124.26***	-17.24	-6.24	
THEME7	-60.24***	-18.41*	15.90**	
THEME8	17.16	-23.62	-2.69	
THEME9	84.09***	-46.52	3.22	
THEME10	-3.79	-54.88***	6.34	
THEME11	10.06	-110.71***	2.73	
THEME12	-37.65*	-0.66	-5.09	
THEME13	-17.04	-6.17	0.10	
THEME14	-38.29**	51.51	8.07	
THEME15	114.43***	-72.59	-0.36	
THEME16	26.58*	-8.15	0.32	
THEME17	22.73	75.57	31.20***	
THEME18	56.79***	-19.03	7.07	
THEME19	104.52***	102.89	10.90	
THEME20	-21.13*	-19.92	-9.24	
OFLOW				0.99***
R_{surp}				-22.93***
$R_{surp(t-1)}$				-34.39***
$R_{surp(t-2)}$				-5.39***
$R_{surp(t-3)}$				-6.61***
R-Bar**2	0.24			

Notes: This table presents the Weighted Least Squares estimates of equation (2) from the text. The dependent variable is the five-minute return in the BAX market. $THEME_i$ represents the ith theme extracted from the document set. OFLOW is the five-minute order flow. R_{surp} is the announcement surprise for Bank of Canada FAD interest rate decisions. Results for Canadian or U.S. macro news announcement surprises are omitted to save space. *** is significant at the 1% level, ** is significant at the 5% level, and * is significant at the 10% level. All coefficients have been multiplied by 10^5.

Table 6: BAX Market Return Volatility Equation Estimate: SVD

Variable	Press Releases	MPR(U)s	Market News Stories
THEME1	25.85	-42.56	38.63***
THEME2	453.90***	35.38	9.57
THEME3	327.21***	8.17	-9.49
THEME4	-176.60***	31.30	65.99***
THEME5	29.92	77.06	58.10***
THEME6	-10.17	-9.77	4.42
THEME7	-204.94***	36.47	20.93***
THEME8	-12.09	-14.59	-24.67***
THEME9	1.73	-9.39	0.71
THEME10	236.29***	-11.93	-33.49***
THEME11	-69.60***	-21.98	1.81
THEME12	-1.35	-51.79	-18.62***
THEME13	-18.32	27.10	12.24*
THEME14	-34.67*	-4.42	11.03*
THEME15	-15.84	46.64	2.60
THEME16	16.50	-26.59	14.67**
THEME17	-6.14	-8.52	-34.81***
THEME18	-47.38	-53.47	26.96***
THEME19	-84.41***	-63.65	-43.48***
THEME20	-143.30***	49.56	2.48
FAD_{t+2}	-0.43		
FAD_{t+1}	0.90		
FAD_t	16.55***		
FAD_{t-1}	37.69***		
FAD_{t-2}	-2.30***		
FAD_{t-3}	2.18***		
FAD_{t-4}	3.75***		
FAD_{t-5}	1.50**		
FAD_{t-6}	0.39		
FAD_{t-7}	1.39**		
FAD_{t-8}	1.50**		
MPR		-3.33***	
MPR_{t-1}		6.99***	
MPR_{t-2}		2.30***	
$NEWS_t$			0.62***
R-Bar**2	0.17		

Notes: This table presents estimates of equation (3) from the text. The dependent variable is the absolute value of the residual from an OLS estimate of equation (2) for the five-minute return in the BAX market. FAD and MPR variables are dummies with a value of 1 for the five-minute interval after the event and zero otherwise. NEWS is a dummy variable equal to one for the period a news story was released as well as for the half hour before and after, and zero otherwise. Results for dummies representing the time of day and whether there was a particular Canadian or U.S. macro news announcement are omitted to save space. *** is significant at the 1% level, ** is significant at the 5% level, and * is significant at the 10% level. All coefficients have been multiplied by 10^5.

Table 7: BAX Return WLS Estimates: SVD

Variable	Press Releases	MPR(U)s	Market News Stories	Other
THEME1	118.84	-311.18**	44.59***	
THEME2	891.84**	-192.32***	8.95	
THEME3	79.84	-13.55	25.95***	
THEME4	-301.98**	41.49	69.52***	
THEME5	383.60***	186.70***	15.46	
THEME6	-6.85	-5.83	30.78***	
THEME7	-328.88**	-193.12**	-7.49	
THEME8	442.31***	-135.95**	-30.16***	
THEME9	-245.67***	35.51	3.91	
THEME10	-243.04**	6.04	-12.23*	
THEME11	-294.08***	203.70***	85.96***	
THEME12	-80.71	2.10	90.56***	
THEME13	68.29	-75.84	3.81	
THEME14	-374.81***	-59.35	-26.45***	
THEME15	-201.18***	267.39***	-0.68	
THEME16	327.65***	-18.26	-45.31***	
THEME17	-83.02	-33.39	-59.33***	
THEME18	-98.93	282.11***	27.62***	
THEME19	-35.89	227.95	-60.31***	
THEME20	128.20	-23.14	5.29	
OFLOW				0.98***
R_{surp}				-19.51***
$R_{surp(t-1)}$				-31.99***
$R_{surp(t-2)}$				-4.22***
$R_{surp(t-3)}$				-6.55***
R-Bar**2	0.24			

Notes: This table presents the Weighted Least Squares estimates of equation (2) from the text. The dependent variable is the five-minute return in the BAX market. $THEME_i$ represents the ith theme extracted from the document set. OFLOW is the five-minute order flow. R_{surp} is the announcement surprise for Bank of Canada FAD interest rate decisions. Results for Canadian or U.S. macro news announcement surprises are omitted to save space. *** is significant at the 1% level, ** is significant at the 5% level, and * is significant at the 10% level. All coefficients have been multiplied by 10^5.

Table 8: BAX Market Return Volatility Equation Estimate: NNDSVD

Variable	Press Releases	MPR(U)s	Market News Stories
THEME1	-17.77***	13.44	-14.63***
THEME2	-498.42***	-3.15	4.91***
THEME3	-30360.00***	-989.94	4.10**
THEME4	-39.05***	3.08	1.92
THEME5	-30.75***	6.55	61.16***
THEME6	16.14**	5.54	8.25***
THEME7	504.05*	-7.36	9.61***
THEME8	-4.15	-2.14	-1.08
THEME9	-380.23***	-3.30	10.97***
THEME10	19.32***	-5.57	-23.68***
THEME11	298.64***	28.78	10.75***
THEME12	5.98	-4.34	19.61***
THEME13	-48.32	-327.10	-3.32
THEME14	6.88	-13.03	23.44***
THEME15	-20.17***	-27.76	0.61
THEME16	-54.44***	14.02	19.64***
THEME17	35.06***	-21.38	22.53***
THEME18	19.67**	16.91	-4.61*
THEME19	-175.83**	278.09	6.95**
THEME20	-444.49***	478.48	-3.33
FAD_{t+2}	-0.43		
FAD_{t+1}	0.90		
FAD_t	13.03***		
FAD_{t-1}	37.43***		
FAD_{t-2}	-2.09***		
FAD_{t-3}	2.37***		
FAD_{t-4}	3.80***		
FAD_{t-5}	1.55**		
FAD_{t-6}	0.05		
FAD_{t-7}	1.35**		
FAD_{t-8}	1.66***		
MPR		-2.94***	
MPR_{t-1}		6.39***	
MPR_{t-2}		2.20***	
$NEWS_t$			0.62***
R-Bar**2	0.17		

Notes: This table presents estimates of equation (3) from the text. The dependent variable is the absolute value of the residual from an OLS estimate of equation (2) for the five-minute return in the BAX market. FAD and MPR variables are dummies with a value of 1 for the five-minute interval after the event and zero otherwise. NEWS is a dummy variable equal to one for the period a news story was released as well as for the half hour before and after, and zero otherwise. Results for dummies representing the time of day and whether there was a particular Canadian or U.S. macro news announcement are omitted to save space. *** is significant at the 1% level, ** is significant at the 5% level, and * is significant at the 10% level. All coefficients have been multiplied by 10^5.

Table 9: BAX Return WLS Estimates: NNDSVD

Variable	Press Releases	MPR(U)s	Market News Stories	Other
THEME1	-22.17**	-25.36	-8.95***	
THEME2	2.90	-49.83**	12.37***	
THEME3	9080.00***	736.09***	2.56**	
THEME4	-49.99	-5.14	15.43***	
THEME5	132.14***	0.27	-17.00**	
THEME6	8.24	-3.47	0.81	
THEME7	1950.00**	270.97**	5.68***	
THEME8	124.81	-61.24	-1.04	
THEME9	-1370.00***	-61.42	-27.86***	
THEME10	2.47	-29.38	-7.95***	
THEME11	708.99***	-307.25***	-35.34***	
THEME12	-21.00*	10.97	-0.41	
THEME13	295.59***	-507.95*	-19.90***	
THEME14	-33.39***	-12.62	-22.60***	
THEME15	12.46	-96.90***	3.63	
THEME16	-49.45***	41.76**	-9.16	
THEME17	-2.61	28.03	-5.50	
THEME18	101.43***	28.72	2.32***	
THEME19	306.61***	3120.00***	-20.44***	
THEME20	410.39***	-93.83	-2.29	
OFLOW				0.99***
R_{surp}				-13.35***
$R_{surp(t-1)}$				-31.76***
$R_{surp(t-2)}$				-4.78***
$R_{surp(t-3)}$				-6.53***
R-Bar**2	0.24			

Notes: This table presents the Weighted Least Squares estimates of equation (2) from the text. The dependent variable is the five-minute return in the BAX market. $THEME_i$ represents the ith theme extracted from the document set. OFLOW is the five-minute order flow. R_{surp} is the announcement surprise for Bank of Canada FAD interest rate decisions. Results for Canadian or U.S. macro news announcement surprises are omitted to save space. *** is significant at the 1% level, ** is significant at the 5% level, and * is significant at the 10% level. All coefficients have been multiplied by 10^5.

Table 10: Basis Point Effect of a Shock to a Theme: NNDSVD

	Volatility Equation			Return Equation		
	Press Release Effect	MPR Effect	Market News Story Effect	Press Release Effect	MPR Effect	Market News Story Effect
THEME1	-39.53		-17.75	-0.77		-0.15
THEME2	-4.24		17.41		-0.19	0.59
THEME3	-5.54		11.35	0.02	0.03	0.09
THEME4	-23.55					0.19
THEME5	-18.89		23.40	1.10		-0.09
THEME6	25.79		9.83			
THEME7	16.05		19.26	0.82	0.18	0.15
THEME8						
THEME9	-9.85		13.15	-0.45		-0.43
THEME10	52.31		-19.74			-0.09
THEME11	34.78		12.21	1.25	-0.53	-0.52
THEME12			13.09	-0.38		
THEME13				0.25	-0.13	-0.13
THEME14			11.85	-1.03		-0.15
THEME15	-26.70				-0.35	
THEME16	-41.65		12.23	-0.60	0.23	
THEME17	24.16		13.14			
THEME18	14.98		-3.99	1.02		0.02
THEME19	-4.20		2.93	0.09	0.35	-0.10
THEME20	-15.36			0.19		

Notes: This table contains the estimate, in basis points, of how much BAX returns and return volatility (standard deviation) would change in response to the value for one theme increasing, ceteris paribus, from its median value to its 95th percentile. The estimates are given for the NNDSVD model. Estimates are only shown for themes which were significant at the 90% significance level.