Universität Zürich[UZH] Faculty of Business, Economics and Informatics

Global market sentiment and exchange rates: evidence from Twitter

Brian Piotrowski (21-740-683)

Advised by Prof. Dr. Mathias Hoffmann & Nicola Benigni

11.08.2023

ABSTRACT

This paper investigates the extent to which market sentiment expressed on Twitter correlates to exchange rate and stock market movements. The main variable of interest is the broad-dollar index–a trade-weighted index of the US dollar against a basket of other major currencies. Global central bank sentiment is found to be a meaningful indicator of dollar depreciations as measured by this index. The paper goes on to compare the effects of the posts of individual central banks on domestic and foreign stock indexes. The paper attempts to use idiosyncratic central bank sentiment as a proxy for the stochastic discount factor. This second goal proves to be inconclusive due to data availability constraints.

1 INTRODUCTION

”Nobody has any idea what will happen with stock prices, inflation, interest rates, or the economy in 2023.”

Marc Andreessen (@pmarca)

Seen by 3.7 million people.

Twitter allows us to look into the thoughts, minds, and feelings of hundreds of millions of financial actors both big and small. Investors like Mr. Andreessen–whose firm has $35 billion under management–share their thoughts in the same public space used by politicians, central bankers, and ordinary citizens. Does Mr. Andreessen's uncertainty about the economy have an effect on markets? It seems like a stretch that a single individual's thoughts, which are confined to 230 characters or less, can have a meaningful effect. But what if the poster is a central bank? Whose announcements carry with them the ability to shape monetary policy.

Using original datasets of tweets, this paper attempts to draw economic insights from the world's ”public square”. The tweets of influential financial posters and central banks are processed by extensively trained machine learning models in order to identify the underlying sentiments of each tweet. This automated process allows for the efficient and objective classification of thousands of tweets. The resulting sentiments are analyzed against exchange rate and stock information. Special attention is given to the effect of central bank posts on both their domestic market and other international markets.

2 LITERATURE REVIEW

As central banks have embraced more transparent policies for communicating with the public, Twitter has become an increasingly useful outlet for information dissemination (Masiandaro, Peia, and Romelli, 2023). In contrast to the 1980s, when central banks were secretive about their discussions and decisions, banks have now realized that their monetary policy decisions can better influence expectations–and thereby drive effective monetary policy–when they are clearly communicated to the public (Masiandaro et al., 2023).

Modern central banks see transparent communication as a way to more effectively drive inflation expectations, strengthen the bank's credibility, and demonstrate that it is fulfilling its mandate (Masiandaro et al., 2023). This transparency helps stabilize the public's expectations about the bank's coming actions. Social media as a communication medium is consistent with the push by many central banks to make their communications more accessible to the general public (Blinder et al., 2022). Although, Blinder himself expresses skepticism about whether central banks have been successful in their attempts to reach a broader audience (Blinder et al., 2022). Among social media platforms, Twitter has been by far the most popular in regards to interaction with central bank announcements. The US Federal Reserve has nearly ten-times as many followers on Twitter as it does on its next most popular social media site, LinkedIn (995,453 Twitter followers versus 99,550 LinkedIn followers as of 9/2022) (Masiandaro et al., 2023). Masiandaro et al. found that central bank tweets about banknote issuances and monetary policy decisions had the most engagement from the public (in the form of retweets) compared to speeches, data releases, exchange rate information, reports, and conferences (Masiandaro et al., 2023).

Natural language processing (NLP) opens the possibilities of analyzing large qualitative data in an objective way. Maylin Sun's innovative use of NLP to draw insights from tens of thousands of news articles is a prime example of how this technology is opening new doors for economic analysis (Sun, 2023). Ms. Sun's work was an inspiration for this project.

The work of Brandt, Cochrane, and Santa-Clara (2006) examines international risk sharing and posits that exchange rates are too smooth relative to equity premia, implying that there must be a large degree of risk sharing and a correlation of marginal utilities across countries. This project will attempt to build on this work by using country-specific banking factors as a proxy for the stochastic discount factor.

3 METHODOLOGY

3.1 Data Selection Criteria

3.1.1 Twitter users Given the sheer size of Twitter and the irrelevance of most of its content to the currency markets, it was necessary to refine the selection to tweets to only financially-relevant actors. The actors selected fall into two categories: central banks, and top financial posters. The central banks have the ability to take actions that directly affect exchange rates, because of this, it is expected that their tweets–which often reflect their outlook on the economy and portend future monetary policy decisions–will have an outsized impact on market sentiment. Unlike central banks, individual posters within the financial world cannot directly affect monetary policy, however, these actors post more frequently and have higher followings than most central banks making then a useful gauge of general market sentiment from the consumer-side.

Masiandaro et al. identified 19 major central banks whose official Twitter accounts have high numbers of followers and engagement on Twitter (Masiandaro et al., 2023). These same 19 central banks were used in this paper. Key financial posters were identified based on a review of the Twitter financial space led by Sentieo (a private equity research firm) and published by Forbes[1] (Shah, 2017). Sentieo identified 120 professional investors who use Twitter as part of their information gathering process. Sentieo then compiled all the accounts that these professional investors follow and published the top 100 accounts that were most followed among this group of investors (Shah, 2017). These accounts also have broad popularity across the general users of Twitter. This sampling ensures that the accounts that this paper pulls data from are useful approximations of general market sentiment.

[1] Forbes: The 100 Best Finance Twitter Accounts You Should Be Following

Dataset	number of tweets	start date	end date
Exploratory set, 2022	1492	1/1/2022	1/1/2023
Analysis set, 2022	4277	1/1/2022	1/1/2023
Supplemental set, 2023	3775	1/7/2022	1/7/2023
News set, 2023	1400	23/7/2023	27/7/2023

Table 1. Breakdown of twitter pulls.

Using the Twitter accounts of major commercial banks as a component of this project was also considered, but was ultimately decided against after examining the content of their tweets. Commercial banks use their Twitter accounts for advertising and customer interaction. They do not post their thoughts on global market sentiment.

Even with this careful selection of users, the financial posters dataset still contained a considerable amount of noise. Because of this, a larger portion of the total dataset was devoted to these users so that the noise would average out better (see Table 2 for the composition of the poster-types used).

3.2 Twitter Scraping

Tweets were pulled from Twitter a total of four different times with each dataset building on the findings of the preceding one (see Table 1 for a breakdown of each pull). The exploratory set was pulled before Twitter required paid API access. It was accessed using SNScrape, a social network datascraping package from Python. All subsequent pulls were done through the Twitter API with paid access using the Python Tweepy package.

Access to the Twitter API allows for up to 10,000 tweets to be accessed per month. The access method through the API allowed for tweets to be searched by search terms (query) or by users, not both simultaneously. This limitation meant that all pulls had to be carefully planned to maximize the amount of usable information acquired. The 10,000 tweet cap eliminated the option of pulling indiscriminately and then sorting for only the useful tweets after the fact.

Tweets were pulled by user and were restricted to the year of 2022. Each of the 18 central banks pulled 100 tweets while each of the 87 financial posters could pull up to 75 tweets (see Table 2: tweet composition of dataset). More tweets were allocated to each central bank than to each financial poster since there were fewer banks than financial posters and because the tweets of the central banks are of more interest to this project.

Another major restriction was that only the most recent 3200 tweets could be accessed for any Twitter user. News outlets were a poster category that would have been interesting to include in this project, but due to their posting frequency, no tweets within the year 2022 could be accessed[2]. In the 2022 data-pull, this meant that the tweets of some of the financial posters were inaccessible. It also caused an uneven distribution of data (as shown in Fig. 1). This uneven distribution was the result of some high-frequency posters either reaching their archive limit (3200 tweets) or exhausting their pull-allowance (100 tweets for central banks and 75 tweets for financial posters) before all dates could be covered. This is why there are many fewer tweets earlier in the year than later. This inconsistency presents a problem for statistical analysis since the later tweets are more stable daily averages than the earlier ones (average of 140 tweets being compared to an average of 5 tweets). This motivated another data pull with a more recent time-frame (supplemental set, 2023), with the hope that this more recent set would not run into the archive limit. While the archive limit was less of a problem here, the inconsistency remained due to the pull allowance (Twitter does not allow it to be any higher than 100 tweets per user). High-frequency posters meant that the data still skewed towards recent observations. For a more detailed discussion of these limitations see Section 4: Notes on Data Quality.

In light of this concern, the original 2022 dataset was used for analysis. Although the quantity inconsistency is less than ideal, since the data itself is being treated as trendless (see Section 3.3.2: De-trending time-series components), it seemed appropriate to continue with the analysis as-is. In their paper "'Whatever it takes' to

[2] A full dataset of tweets from 14 top international news posters was pulled as part of this project. In order to avoid the archive limit, the tweets were set to pull for the days directly preceding when the code was executed. 1400 tweets were pulled, but only three days were covered. This is a striking example of the frequency with which these news outlets post, that every news outlet in the sample had posted more than 100 tweets in the last three days.

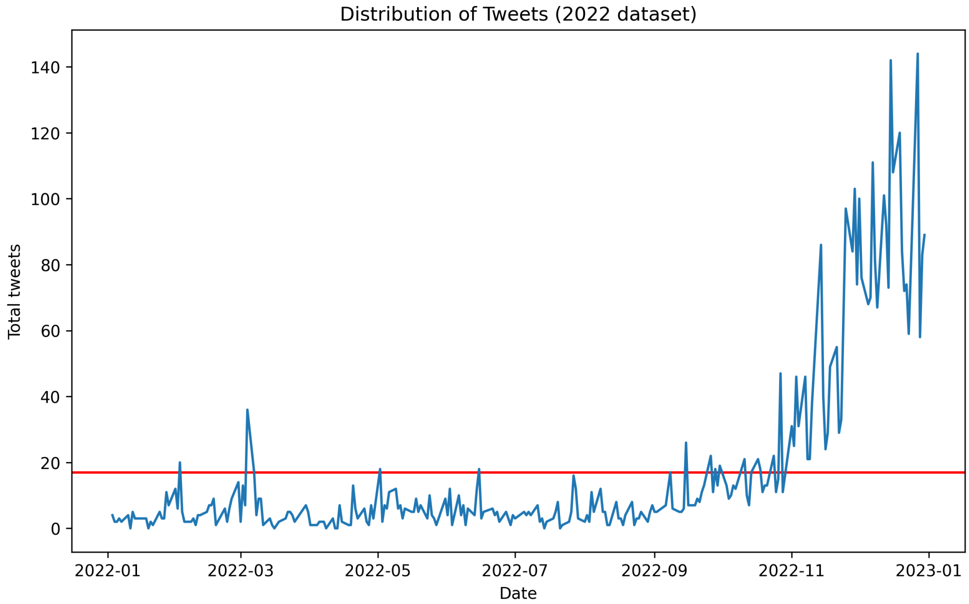


Fig. 1: Distribution of daily tweets pulled in 2022

User type	Count of users	Max tweets per user	Total tweets	% of dataset
Central Banks	18	100	1800	20%
Financial Posters	87	75	2402	80%
Totals	105		4202	

Table 2. Tweet composition of dataset.

change belief: evidence from Twitter" the authors also deal with Twitter data that where the number of posts varies over time (Stiefel and Vives, 2022). Although not a perfect comparison since the variance in post counts for their data was more evenly spaced throughout the sample, it does provide some reassurance that an analysis can be done on averages of varying depth.

For complete lists of the central banks and financial posters used, please see Sections 9.1-9.3 within the appendix.

3.3 Data Consolidation

3.3.1 Outcome variables Both the nominal broad-dollar index and the S&P 500 index were used as dependant variables for this project. These time-series were gathered from the Federal Reserve Bank of St. Louis (FRED) database (FRED, 2023). The nominal broad-dollar index is a trade-weighted comparison of the US dollar against a basket of 26 other major currencies[3] (Von Beschwitz et al., 2019). It is important to note that the broad-dollar index increasing represents an appreciation of the US dollar against the comparison currencies while a decrease in the broad-dollar index represents a depreciation.

For the international component of this project, the main stock indexes of each country of interest were used. These stock indexes were denominated in the local currency of each country. The indexes included can be seen in Table 3 and were all retrieved from Yahoo Finance (Yahoo Finance, 2023). Monthly average exchange rates for all currencies of interest were also used and were retrieved from the Organization for Economic Cooperation and Development (OECD, 2023).

[3] The countries included in this index–in order of weight from highest to lowest–are: Euro Area, China, Canada, Mexico, Japan, United Kingdom, Korea, India, Switzerland, Brazil, Taiwan, Singapore, Hong-Kong, Australia, Vietnam, Malaysia, Thailand, Israel, Indonesia, Philippines, Chile, Columbia, Saudi Arabia, Argentina, Russia, Sweden (Von Beschwitz et al., 2019).

Country	Country Abbr.	Stock Index	Ticker	p-value of ADF
United States	US	S&P 500	GSPC	1.91e-28
Euro Area	EU	STXE 600	STOXX	4.12e-27
United Kingdom	UK	FTSE 100	FTSE	4.53e-23
Australia	AU	S&P/ASX 200	AXJO	2.89e-29
Canada	CA	S&P/TSX Composite	GSPTSE	2.16e-25

Table 3. Stock indexes used and the p-values from Augmented Dickey-Fuller stationarity tests of their first-differences.

3.3.2 De-trending time-series components All outcome variables had to be de-trended so that regressions could be run without the interference of a time-dependant trend. To do this, the first difference of each time-series was used. An Augmented Dickey-Fuller test was uses to test whether the first-differencing had successfully resolved the time-series component. For all outcome variables, the result of this test was a rejection of the null hypothesis that there exists a unit root (Table 3). Therefore, it can be concluded that the resulting series are stationary and that the regressions used in the coming sections are appropriate.

3.3.3 Handling non-trading days Neither the US broad-dollar index nor the S&P 500 is recorded on weekends or on US bank holidays. The other stock indexes of interest also do not trade on weekend or on holidays. However, relevant tweets from global central banks and financial actors still occur on these days. The effects of these tweets on an index can be seen by observing the difference in price between the last day before the non-trading day (or days) and the first day after trading resumes. To match this effect to the tweets that may have caused it, the tweets on non-trading days are treated as belonging to the next upcoming trading day (ie. the tweets sent on Saturday and Sunday of a normal week will be treated as belonging to Monday since their effect can only be seen in the daily change recorded on Monday).

3.4 Sentiment Analysis

3.4.1 Natural langauge processing model Sentiment analysis was performed using a pre-trained cross-lingual language model (XLM model) from Barbieri et al. (2022). This is a multilingual model that was trained on 198M tweets and then fine-tuned for sentiment analysis on eight major languages (Arabic, English, French, German, Hindi, Italian, Spanish, and Portuguese) (Barbieri et al., 2022). Although the model is most optimally tuned to these eight languages, it was trained on over thirty languages and operates well in all of them (Barbieri et al., 2022). This multilingual capability is important to the analysis of central bank tweets since some banks post exclusively in their local language [4]

The XLM model was used to assign sentiment probabilities to every tweet in the dataset. For each tweet, the model returns three categories: negative, neutral, and positive. Each category is assigned a value between 0 and 1 with the three categories always summing to 1 (examples of this can be seen in Table 4).

3.4.2 Defining dummy variables sec:3.4.2 Dummy variables were used to denote the dominate sentiment as identified by the NLP model. Since the posts by the financial posters generally had strong emotional valences, it was sufficient to classify them based on their largest sentiment value. The classification method had to be adjusted for the central bank tweets since the banks tended to have more neutral phrasings and were rarely as strongly emotioned as the financial posters. 84% of the tweets from the central banks had "neutral" as the dominate emotion while only 47% of the tweets from financial posters were predominately neutral. To compounded the problem, when one central bank made an non-neutral statement there were almost always other banks making neutral statements that day and weighing the average bank sentiment back to neutral. Only about 1% of the trading days in the 2022 dataset had non-neutral average central bank when using the dominate emotion criteria defined above. While this plurality emotion classification was sufficient for use with the financial posters' tweets, a more sensitive metric was needed to capture the subtlety of the central banks' statements.

[4] Banks that post partially or exclusively in their local language: Central Bank of Argentina, Saudi Central Bank, Deutsches Bundesbank, Banque de France, Central Bank of Indonesia, Banca d'Italia, South Africa Reserve Bank, Central Bank of Brazil, Central Bank of Mexico

Example tweet	Negative	Neutral	Positive
”Rough day in the markets today.”	**.91**	.07	.02
”Central bank considers changing currency denominations”	.10	**.71**	.19
”Investors cautiously optimistic as COVID concerns recede.”	.10	.32	**.58**

Table 4. Example phrases and XLM classifications.

A threshold method was used to classify central bank sentiments. If the positive or negative categories for a given tweet exceeded 0.2, the tweet was classified as belonging to that category. If neither positive nor negative exceeded 0.2, it was classified as neutral. If both positive and negative exceeded 0.2, it was classified based on the larger of those two categories. The threshold was set at 0.2 after testing with multiple threshold levels and finding that 0.2 yielded the most useful results in regressions.

4 NOTES ON DATA QUALITY

4.1 XLM classification concerns

The XLM model runs a sentiment analysis to create the positive, neutral, and negative scores used for this analysis and makes these classifications based on the tone, word-use, and phrasing of each tweet as opposed to the financial content. An example of how this can be problematic is shown below using a real tweet from the dataset:

”The Monetary Policy Committee voted by a majority of 7-2 to increase #BankRate to 5%.”

- Bank of England (22/06/23)

The model is not trained to recognize that an increase in rates is generally seen as a negative indicator for the economy. The model gave this tweet a majority neutral rating, but was fairly close to calling it positive[5]. This occurrence, and others like it, are a reminder of why a sufficiently high threshold is needed to set the associated dummy variable to 1. These occurrences must also be kept in mind when interpreting the regression results; the takeaway being that the XLM model rates only sentiment–the feelings expressed in the tweet–and not the economic implications of the tweet. There is usually a fair amount of alignment between the sentiments expressed in the tweet and the message for the economy[6], but there are exceptions to this rule, especially among central banks where contractionary events are phrased diplomatically.

Complex sentence structure can also be a point of confusion for the XLM model. The use of many negative words in the following tweet led to a faulty negative classification:

”The 2023 stress test shows that large banks are well positioned to weather a severe recession and continue to lend to households and businesses even during a severe recession, and the largest banks’ trading books were resilient to the rising rate environment tested. (1/4)”

- US Federal Reserve Bank (28/06/23)

The model struggled to classify this tweet, showing a fairly close split between all the sentiment categories[7]. The fact that ”severe recession” was mentioned twice along with a complicated message (the message of potential resilience amidst a difficult economic environment) confused the XLM model.

Although the systematic problems described above do result in misclassifications by the XLM sentiment model, the vast majority of tweets are classified in a useful manner by this model and the use of the model allows for both an objectivity in scoring and a scalability that would not have been achievable if done manually.

[5] This Bank of England tweet received sentiment scores of positive: 0.17, neutral: 0.78, negative: 0.06. The threshold to be considered positive is 0.2 for a central bank poster.

[6] posters usually write optimistically about positive/expansionary economic events and write pessimistically about negative/contractionary events.

[7] This Federal Reserve tweet received sentiment scores of positive: 0.23, neutral: 0.42, negative: 0.34

4.2 Relevance concerns

As a result of the new Twitter API rules (in-place April 2023 onward), tweets could either be pulled by username or by query and not by both. The nature of this project meant that having the right users (the central banks) was more important than having the right keywords. This was a tradeoff that meant that some irrelevant tweets were included in the datasets. Consider the following central bank tweets:

"God Save The King!"

- Bank of England (08/05/23)

"Learn more about Governor Bowman: (link)"

- US Federal Reserve Bank (26/06/23)

"INR 2000 Denomination Banknotes – Withdrawal from Circulation; Will continue as Legal Tender"

- Reserve Bank of India (22/05/23)

As is clear from these examples, central bank tweets are not always financially relevant, and without the ability to query tweets before pulling, all tweets from a user are considered in the dataset. Part of this problem is mitigated by the fact that the Federal Reserve tweet and the RBI tweet above were classified as neutral, as were many irrelevant tweets. However, the Bank of England tweet received a majority positive sentiment score. The XLM model is performing correctly here, however the irrelevant content of the given tweet means that the positive classification is misleading.

This problem is also present for the financial posters dataset. These user's posts are less focused than the central bank tweets, but are higher in quantity so average out better.

4.3 Exclusion of high-frequency posters

As discussed in Section 3.2: Twitter Scraping, the Twitter API only allows for 100 tweets to be pulled per user and pulls these tweets from the end-date backwards. If a poster regularly posts dozens of tweets per-day (as most news organizations do), this will only result in a few days worth of data. Additionally, Twitter only archives the most recent 3200 tweets from each user for access via the API. This means that some high-frequency posters did not have any tweets available in the main year of interest, 2022, because they had already posted more than 3200 times since the end of 2022.

These restrictions are the reason why major news outlets are not included in this dataset. It is also why some financial posters appear in the 2023 dataset but not in the 2022 dataset. The asymmetric distribution of tweets is also a result of these restrictions. As seen in Fig. 1, many more tweets are collected towards the end of the year. This also caused the central bank data to be clustered towards the last quarter of the year, which proved problematic in the analysis of country-specific sentiment as compared to exchange rates.

4.4 Data size relative to noise

As explored in this section, there is a fair amount of noise in the Twitter datasets. This noise becomes less of a problem as the number of observations increases. However, given the API restriction of pulling no more than 10,000 tweets per month, along with the archiving restrictions described in sections 3.2 and 4.3, it was not possible to access the dataset sizes that previous papers had used before the API changes. The size of the 2022 tweet dataset used for this analysis (N= 4277) is not negligible–resulting in an average of 17 tweets per business-day–but given the amount of noise present in many of these tweets, a larger dataset would be needed to reach stronger conclusions.

4.5 Regression set-up: broad-dollar regressions

The goal of these first regressions is to isolate moments from the sentiment data that are statistically significant in predicting movements in dollar exchange rates as measured by the nominal broad-dollar index. There were dozens of possible explanatory variables collected and computed from the Twitter data and sentiment analysis. For these OLS regressions, the focus was on examining causality rather than maximizing predictive power. For this reason, many of the explanatory variables were not used here.

In order to match the tweets to broad-dollar changes, the data is evaluated per-day. This means that multiple tweets for a given day had to be aggregated together. Each day had between 0-140 tweets. The tweets were denoted by poster category, either ”central bank” or ”financial poster”, and these categories were kept separate in the daily aggregation since the foreign exchange markets are likely affected differently by central bank tweets as opposed to the tweets of ordinary financial posters. Within each category, both the average and variance of the day’s tweets are taken within each of the emotion categories (negative, neutral, positive). Dummy variables were then assigned for each emotion-category pair using the procedure described in Section 3.4.2: Defining dummy variables. This results in six dummy variables, three for each poster category.

The daily emotion variances were originally recorded for both the negative and positive categories for each poster-type (four variance variables total). However, the variances in negative and positive were not significant on their own, so were consolidated into a single variance observation for each poster-type which is the sum of the variance in the negative and positive categories. This way, the variance variable can be interpreted as the overall divergence in opinions posted within a category on a given day, accounting for differences in both positive and negative sentiment. For example, if $bank\_variance$ is high, it means that the central banks posting that day differed in their messages; some gave positive messages while others gave negative messages.

5 RESULTS

5.1 OLS Regressions

Three versions of the OLS regressions were run. The first regression includes the positive and negative dummy variables for banks and for financial posters. The ”neutral” category is represented when both ”positive” and ”negative” dummies for a poster-type are marked as zero. The constant picks up the value for both the ”neutral” days and the days where no tweets were posted. The second regression includes only the variance metrics. One variance metric for banks and one for financial posters. Each metric is measuring the total variance in sentiments for the poster-type during a given day (sum of the variances in positive and negative sentiment within poster-type). The third regression combines the variables from the first two. It includes both the sentiment dummies for positive and negative for each poster-type, as well as the type-specific variances.

Regressions 2 and 3 had smaller sample sizes due to the inclusion of the variances. The use of dummy variables before had allowed for the use of days where one or both poster-types did not post (this would be represented as a zero in both positive and negative dummies). However, since the variances have no way to represent missing data, the analysis can only be run on days where data is present for both poster-types. This drastically reduces the dataset.

Regression 1 results in one significant coefficient: the dummy variable for the central banks’ sentiment being negative ($central\_bank\_negative$) is significant at the 10% level (see Table 5 for regression results). This indicates that the dollar weakens against the index of other major currencies when global central banks post negative news.

Regression 2 shows the most significant coefficient of these three regressions; the coefficient representing the overall variance in tweet sentiment among central banks ($bank\_variance$) is significant at the 1% level. This indicates that the more central banks disagree with each other in terms of the sentiments of their posts, the more the dollar depreciates that day. Global central bank uncertainty is leading to a weaker dollar.

The final regression, Regression 3, maintains the results of Regression 2 while controlling for the emotion dummies of Regression 1. The significance level of the $bank\_variance$ coefficient is weaker (now only at the 10% level). This reduction of significance is expected since there is a correlation between $bank\_variance$ and $central\_bank\_negative$, which has caused colinearity here. These two variables have a correlation coefficient of 0.55 (see Fig. 2: Correlation Matrix).

Regressions on 2022 dataset

	1	2	3
$central\_bank\_positive$	-0.0199		0.0602
	(0.814)		(0.669)
$central\_bank\_negative$	-0.3051^{*}		-0.2298
	(0.081)		(0.432)
$fin\_poster\_positive$	0.0982		0.3103
	(0.470)		(0.182)
$fin\_poster\_negative$	-0.1016		0.0059
	(0.222)		(0.972)
$bank\_variance$		-4.5434^{***}	-3.4874^{*}
		(0.008)	(0.075)
$fin\_poster\_variances$		0.8123	0.6865
		(0.368)	(0.500)
constant	0.0534	0.0503	-0.0239
	(0.222)	(0.766)	(0.894)
Prob. (F-stat)	0.277	0.0228**	0.0829*
AIC	275.5	129.0	132.9
N	199	76	69
MSE (out of sample)	0.2601	0.2249	0.2148
Variance of dollar_diff	0.2327	0.2903	0.2903

Note: $^{*}\, p < 0.1$; $^{**}\, p < 0.05$; $^{***}\, p < 0.01$

Table 5. Regressions on 2022 dataset.

While the coefficients related to the financial posters were not individually significant, their inclusion in this analysis was important because it controlled for general market sentiment, allowing the central bank coefficients to better identify the variance in exchange rates related to banking tweets as opposed to the banking coefficients also picking up exchange rate variation that could be mapped back to general market sentiment.

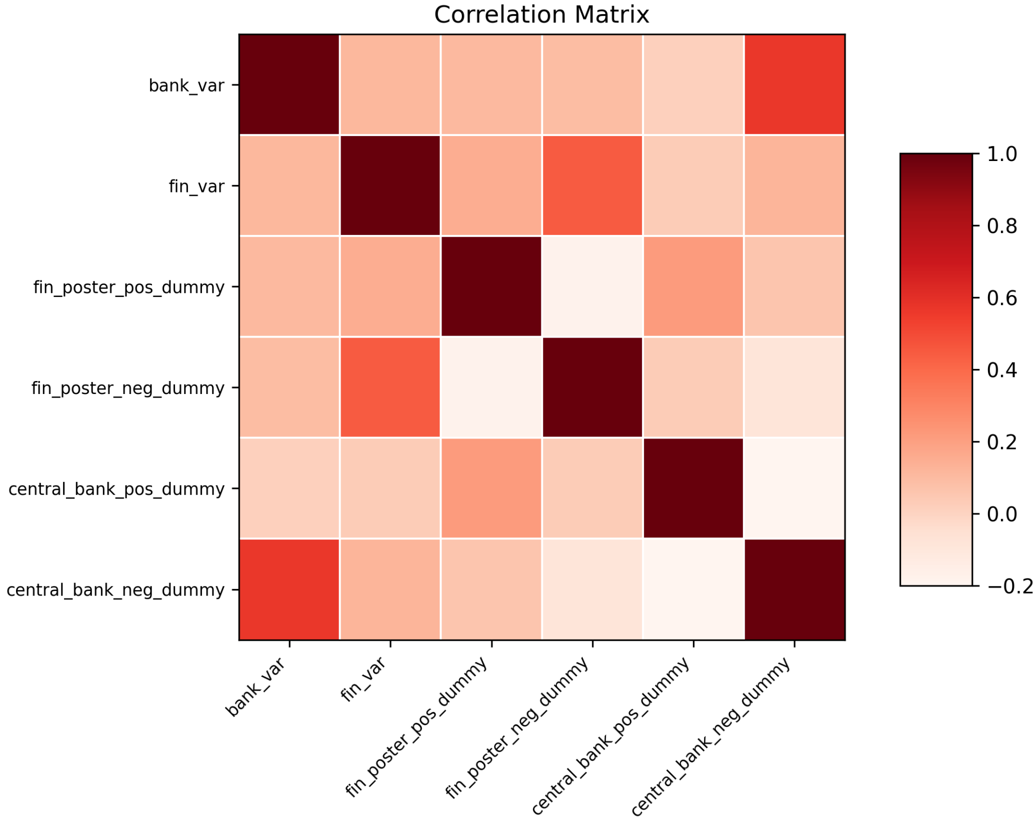


Fig. 2: Correlation matrix for 2022 dataset

5.2 Country-specific analysis

Regressions on 2022 dataset, country dummies			
	4	5	6
$central\_bank\_positive$	0.0459	0.0435	0.0558
	(0.627)	(0.650)	(0.553)
$central\_bank\_negative$	−0.2830	−0.2677	−0.2951
	(0.162)	(0.183)	(0.139)
$fin\_poster\_positive$	0.0999	0.0816	0.0838
	(0.458)	(0.549)	(0.531)
$fin\_poster\_negative$	−0.0632	−0.0720	−0.0795
	(0.462)	(0.403)	(0.342)
$US\_positive$	$1.153e{-16}$		$-4.7e{-16}^{*}$
	(0.148)		(0.058)
$US\_negative$	-0.3570^{*}	-0.2878^{*}	-0.2961^{*}
	(0.060)	(0.097)	(0.086)
$EU\_positive$	0.1585		
	(0.326)		
$EU\_negative$	−0.2417	−0.1663	
	(0.449)	(0.562)	
$UK\_positive$	−0.0948		
	(0.420)		
$UK\_negative$	0.0460	0.0567	
	(0.812)	(0.766)	
$AU\_positive$	0.4351^{**}		0.3483^{*}
	(0.048)		(0.075)
$AU\_negative$	0.8996^{*}	0.5829	0.8570^{*}
	(0.085)	(0.255)	(0.094)
$CA\_positive$	-0.2714^{**}		-0.2741^{**}
	(0.027)		(0.022)
$CA\_negative$	−0.3397	−0.0977	−0.2055
	(0.259)	(0.737)	(0.472)
constant	0.0614	0.0464	0.0623.
	(0.176)	(0.297)	(0.155)
Prob. (F-stat)	0.160	0.9315	0.0871
AIC	280.0	282.1	275.1
N	199	199	199
MSE (out of sample)	0.2855	0.2813	0.2846
Variance of dollar_diff	0.2327	0.2327	0.2327

Note: $^{*}\, p < 0.1$; $^{**}\, p < 0.05$; $^{***}\, p < 0.01$

Table 6. Regressions on 2022 dataset, including dummies for central bank of major economies.

5.2.1 Regressions on broad-dollar index The international analysis (Table 6) breaks apart the central bank dummy variables from the previous regressions. In these models, the central banks of the United States, Euro Area, United Kingdom, Australia, and Canada are given their own dummy variables. This is done to isolate the outsized effects that major central banks can have on exchange rates. These regressions are also used to identify a model that will later be used to predict the returns of each country's primary stock index.

These country-specific dummy variables are marked as 1 if and only if the average sentiment within that specific central bank for that day exceeds the threshold value of 0.2 and is greater than the opposite sentiment value[8]. The banks that have their own individual dummy variables have been removed from the general $central\_bank\_positive$ and $central\_bank\_negative$ variables. In these regressions, $central\_bank\_positive$ and $central\_bank\_negative$ represent the average sentiment among all other central banks.

Regression 4 includes all the major-country and minor-country banking dummies, as well as the financial posters dummies as a control for non-banking sentiment. Regression 5 takes only the negative dummies, with the logic being that negative central banking announcements are more likely to lead to monetary policy changes and therefore changes in exchange rates. The final specification, Regression 6, keeps only the major countries whose central bank sentiment had significance of some level in Regression 4. This means that the ECB and the Bank of England were dropped from Regression 6.

Negative US central banking sentiment, $US\_negative$, is marginally significant in all three regressions. This is expected since the US dollar is the currency being examined and, as noted before, negative sentiment has the most convincing channel through which it can have an effect on markets. As in the prior regressions, the coefficient on this variable is negative, indicating dollar depreciations. The Canadian and Australian central banks also show significant coefficients in these regressions, indicating that they may be either affecting changes in the broad-dollar index, or reacting to changes in this exchange rate.

Regressions 5 and 6 provide no meaningful increase in out-of-sample predictive power over Regression 4. Because of this, the specification of Regression 4 is used in the next set of models which use the country-specific central banking variables to predict stock market returns within each country.

5.2.2 Regressions on country-specific stock indexes The regressions in Table 7 take the explanatory variables used in Regression 4 and set the dependant variable to now be the first difference of the key domestic-based stock index of each country. (The stock indexes used and the Augmented Dickey-Fuller tests for the stationarity of their first differences can be seen in Table 3 along with the country abbreviations that are used in the following tables.)

The column labels in Table 7 correspond to the different dependant variables which are the major stock indexes of each country (denominated in the country's own currency). Highlighted coefficients in Table 7 indicate the effect of a given central bank's tweets on its own market. The significance levels of these coefficients is of interest to this project because they can help identify the extend to which a country's own central bank sentiment affects its domestically-based stock index. The results in Table 7 show that the US Federal Reserve and the Bank of England are the only two central banks whose posts can be correlated with movements in their domestic stock markets. The small magnitude of the US_pos_dummy in relation to the S&P 500 index means that this effect is not economically significant (a standard deviation for the daily change in the S&P 500 is 62).

The Bank of England's sentiments correlate with its domestic stock index but also with all the other indexes in this panel. This indicates that its tweets may be picking up a global pricing factor. Upon closer examination of the data, the content of the Bank of England's tweets is not remarkably different than that of any other banks. However, the bank's tweeting activity is concentrated in Q4 of 2022. Given this concentration of data, and the fact that many of the bank's late-year tweets were about non-monetary topics (the replacement of banknotes was a recurring topic in the bank's posts), it is most likely that this correlation with a global pricing factor is spurious. For a deeper look at the Bank of England's tweeting activity, please see Fig. 4 in the appendix.

5.2.3 Bilateral loading In this part of the analysis, the country-specific central bank sentiment is used as an approximation of the stochastic discount factor. This is based on the work of Brandt et al. (2006), where the authors argue that bilateral exchange rates depreciate in accordance with the difference between the discount factors in each country. The central bank sentiments indicate what an informed actor thinks about the state of the domestic economy and for this reason may be a usable proxy for the domestic discount factor.

[8] Ie. A bank will only show as positive if its positive value for that day is greater than 0.2 and greater than its negative value. The second criterion is important for ensuring that the positive and negative dummy variables are mutually exclusive as it is possible for both the positive and negative values to exceed the threshold simultaneously.

Regressions on major country stock indexes

	US	EU	UK	AU	CA
fin_poster_pos_dummy	-57.52	-0.42	-15.94	-3.64	30.39
fin_poster_neg_dummy	-46.55	-0.53	-4.34	6.55	-0.46
central_bank_pos_dummy	-139.04^{***}	-0.43	7.10	14.80	-41.88
central_bank_neg_dummy	32.03	1.92	25.05	41.27	42.44
US_pos_dummy	$\mathbf{1.6e-13^{***}}$	-2.4e-16	-4.0e-15	1.0e-14	1.2e-14
US_neg_dummy	**-97.84**	0.76	-9.67	-12.31	-19.97
EU_pos_dummy	-141.08	**-1.22**	-6.77	-34.30	-72.79
EU_neg_dummy	-228.37	**-0.72**	16.44	-19.42	-84.03
UK_pos_dummy	-394.15^{***}	0.80	**-3.48**	17.21	17.67
UK_neg_dummy	-201.09^{**}	5.91^{**}	$\mathbf{90.04^{**}}$	126.73^{***}	207.22^{**}
AU_pos_dummy	113.75	-2.79	-8.84	**2.38**	45.26
AU_neg_dummy	-90.99	-0.08	-6.45	**-9.68**	-124.26
CA_pos_dummy	-153.97^{**}	0.75	12.53	-9.45	**65.63**
CA_neg_dummy	-149.99	156.44	40.33	31.63	**125.55**
const	4209.35^{***}	-0.62	-1.27	-6.43	-9.21

Table 7. Regressions on major country stock indexes using country-specific central bank dummies. Column labels correspond to the different dependant variables which are the major stock indexes of each country (denominated in the country's own currency). Highlighted are the effects of a country's own central bank posts on its domesticly-based stock index.

Sentiment ratios vs. EX rates

Country pairing	Sent. ratio (pos.)	Sent. ratio (neg.)	Currency/USD
EU - US (Nov.)	0.3828	0.4245	0.980
EU - US (Dec.)	0.2933	0.6505	0.945
UK - US (Nov.)	0.5408	0.8174	0.852
UK - US (Dec.)	0.4977	0.7288	0.822
AU - US (Nov.)	0.6832	0.2813	1.517
AU - US (Dec.)	0.7531	0.2162	1.482
CA - US (Nov.)	0.3041	0.5360	1.344
CA - US (Dec.)	0.3868	0.3656	1.359

Table 8. Comparison of the ratio of negative sentiments between domestic central banks and the related exchange rate. Sentiment ratios are calculated as: domestic bank sent. / Federal Reserve sent.

Unfortunately, due to data limitations, this analysis was severely under-powered. Quarterly average sentiment observations could not be used because of the sparsity of the central bank data in the earlier parts of the year. Monthly averages were used instead but proper data for all banks was only available during the last two months of the year. In order to compare in the same units, the analysis seen below uses ratios instead of differences. This is a departure from Brandt et al. (2006), but was necessary in order to normalize the sentiment ratings and make them comparable to exchange rates.

Table 8 outlines these findings while the tables in the appendix show the same data graphically (Fig. 5 - Fig. 9). As can be seen in the accompanying graphs, the positive sentiment ratio and the exchange rate move in the same direction month-over-month for three of the four country pairings (the exception being AUD - USD). While interesting, the data here is far too limited to draw any conclusions.

5.3 Non-parametric regressions and out-of-sample validation

5.3.1 Ridge Regressions Ridge regressions were applied to this dataset as a way to reduce overfitting. The ridge regression discourages single large coefficients through the use of a penalty term, α. The loss equation that

Ridge Regressions on Analysis dataset

	1	2	3
$central\_bank\_positive$	-0.0165		0.0472
$central\_bank\_negative$	-0.2624		-0.0013
$fin\_poster\_positive$	0.0886		0.3079
$fin\_poster\_negative$	-0.0983		-0.3614
$bank\_variance$		-0.3791	-0.2542
$fin\_poster\_variances$		0.1613	0.1516
constant	0.0509	-0.0636	-0.0899
N	199	76	69
Ridge MSE (out of sample)	0.25069	0.1480	0.1768
OLS MSE (out of sample)	0.2601	0.2249	0.2148
Variance of dollar_diff	0.2327	0.2903	0.2903

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 9. Regressions on Analysis dataset.

a ridge regression minimizes is as follows:

$$SSE = \Sigma_{i=1}^{n}(y_i - \hat{y_i})^2 + \alpha\Sigma_{j=1}^{p}(\beta_j)^2$$

Where there are p coefficients.

Each coefficient carries with it a penalty, α, proportional to the square of its size. This regularization helps prevent overfitting and attempts to make the regression more generalizable to out-of-sample observations. The ridge regressions shown here all use an alpha parameter of 1.2. Results can be seen in Table 9.

The significance values are not included in these regressions since the regularization biases the coefficients, making significance values uninterperable. What can be seen in Table 9 is that the ridge regression has forced the coefficients to be smaller. Fig. 3 tests different alpha parameters against out-of-sample prediction and highlights the optimal alpha for the Regression 1 parameters. The ideal alpha parameter is about 31, forcing the coefficients to be vanishing small and for the estimate to be based solely on the constant. Table 10 confirms this finding by showing that the out-of-sample MSE of all the regressions tested thus far exceeds the variance of the dependant variable within the test sample. This finding shows that using the parameters outlined in the regression leads to worse predictions than simply predicting using the mean of $dollar\_diff$ for all observations.

5.3.2 Out of sample validation Unfortunately, neither the OLS nor Ridge regressions were able to predict better than the mean when being tested for out-of-sample validation on the 2023 dataset (N=3775). Regression trees and forests were also tested to see if machine learning approaches could improve predictability here. The most successful machine learning model was a regression tree with a depth of 6. This tree was also unable to predict the 2023 dataset. The inability to predict in this large out-of-sample test speaks to how dynamic and multifaceted exchange rate variations are. The variables that predicted within 2022 may not be relevant in the same way in 2023. At the same time, the variables identified in these regressions are only some of the many factors that move exchange rates. They are insufficient on their own for predictive purposes.

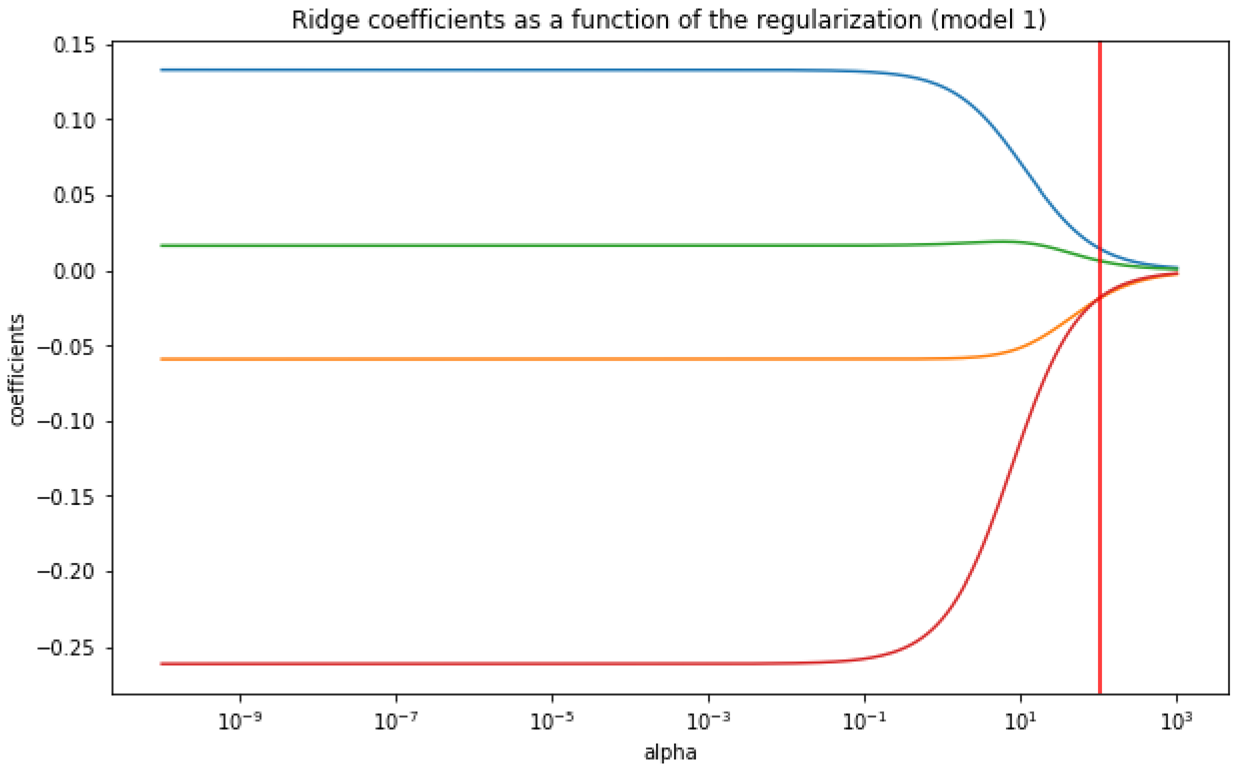


Fig. 3: Convergence of coefficients as alpha increases for Ridge Regression 1. The alpha that creates the best out-of-sample predictions is represented with the vertical line.

Model	Test MSE	Variance of $dollar\_diff$
1: No variance terms, OLS	0.1514	0.1446
1: No variance terms, Ridge	0.1508	0.1446
2: Only variance terms, OLS	0.1490	0.1110
2: Only variance terms, Ridge	0.1154	0.1110
3: All terms, OLS	0.1375	0.1110
3: All terms, Ridge	0.1143	0.1110
RT6: Regression tree (6)	0.1700	0.1110

Table 10. Validation of models using 2023 sample set (out of sample)

6 CONCLUSIONS

Although there are some important limitations to the data used in this analysis (laid out in detail in Section 4: Notes on Data Quality), it is still worthwhile to examine the conclusions of the three OLS regressions and to discuss the implications of the country-specific analysis.

6.1 Conclusion 1: Negative global central bank sentiment correlates to dollar depreciations

6.1.1 Empirical finding Regression 1 shows the $central\_bank\_negative$ dummy to be negative and significant at the 10% level. Regression 1 is the most powerful of the three regressions (N=199), so although the coefficient's p-value of 0.08 is only marginally significant, the correlation still provides useful information. The significance of $central\_bank\_negative$ disappears once the two variance terms are added in regression 3. This is a result of the correlation between $central\_bank\_negative$ and $bank\_variance$. When $bank\_variance$ is excluded, there is omitted variable bias causing $central\_bank\_negative$ to be larger than it otherwise would be and distorting the p-value.

6.1.2 Theoretical framework The fact that negative tweets from central banks move FOREX markets while positive ones do not makes sense when put into the context of which actions typically follow each type of bank tweet. Negative announcements from central banks portend monetary policy changes, which prompt investors to react. Positive announcements indicate less involvement in the markets by the central bank, a laissez-faire stance that does not require investors to change their positions. Hence, the negative central bank announcements–which predict policy action–are the ones that affect changes in the broad-dollar index.

The Regression 1 result indicates that the global central banks having a negative average sentiment as opposed to a neutral sentiment or no tweets that day correlates to a 0.63 SD decrease in the change in broad-dollar index. This is an economically meaningful change, one that can often determine whether the dollar has appreciated or depreciated on a given day. The direction of this change can be explained when put into the context of a global financial cycle where the monetary policy of large central banks drives the policy of smaller central banks. Miranda-Agrippino and Rey (2021) outline such a model where policy changes by a large central bank such as the US Federal Reserve or the European Central Bank cause the central banks of other countries to follow suit (Miranda-Agrippino and Rey, 2021). Degasperi, Hong, and Ricco (2021) go on to demonstrate empirically that global monetary policy tightens in response to contractionary policy changes by the US Federal Reserve (Degasperi et al., 2021). The depreciation of the dollar in response to negative central bank tweets makes sense within this model because a negative statement from one major central bank increases the perception that it, and others, will lower interest rates. The lowering of interest rates has been associated with US dollar depreciation through the channel of increased global liquidity (Degasperi et al., 2021).

6.2 Conclusion 2: Uncertainty among central banks correlates to dollar depreciations

6.2.1 Empirical finding Both Regression 2 and Regression 3 support the claim that increased variance among bank sentiments negatively correlates with changes in the broad-dollar index. This correlation is especially significant in Regression 2; however, the significance level presented in regression 3 is more accurate since it eliminates the omitted variable bias of excluding the $central\_bank\_negative$ dummy. Since $bank\_variance$ is positively correlated with $central\_bank\_negative$, the explanatory power of $bank\_variance$ alone was being inflated when $central\_bank\_negative$ was not also present in the regression. The positive correlation between these explanatory variables makes sense since a high variance in central bank sentiments suggests that at least one central bank posted something strongly negative that day (it may be easier to see this correlation in the inverse where a low amount of bank variance usually means neutral tweets and the negative dummy being equal to zero).

6.2.2 Theoretical framework Building off of the conclusions from Section 6.2.1, the increase in central bank variance is affecting dollar exchange rates in two ways: (1) through its correlation with $central\_bank\_negative$ as described above and (2) by the changing expectations among investors that result from the implied uncertainty of conflicting central bank attitudes. To explore channel (2) further; an increase in uncertainty across different banks leads investors to expect that some banks may soon be intervening in the markets to a larger extend than others. Even without knowing which banks will act, this uncertainty increases the currency risk that investors take on when purchasing foreign assets (which are often dollar-denominated) and encourages them to invest locally, decreasing the demand for dollars. By investing in their domestic currency, investors are insulated from uncertainty in the FOREX markets. Their investments and their obligations will be in the same currency, so a foreign bank changing interest rates will not put their investments at risk the way buying dollar-denominated foreign assets would.

6.3 Conclusion 3: Central bank tweets do not have a discernible effect on domestic markets

6.3.1 Empirical finding The stock index regressions using country-specific dummies for each country's central bank (Table 7) show that bank sentiments as captured through Twitter largely do not have an effect on their domestic stock markets. This can be seen from the lack of significance among the highlighted coefficients in Table 7.

6.3.2 Theoretical framework Before making conclusions about central bank statements not having meaningful effects in and of themselves, the communication channel in question must first be considered. Although popular for communication with the general population, Twitter is not the main communication channel that central banks use to communicate important and timely decisions to informed investors. While the aggregate of global bank tweet sentiment proved useful in Conclusions 1 and 2, the tweets of any individual bank are too noisy to be reliably evaluated on their own.

7 REFERENCES

Barbieri, F., Espinosa Anke, L., Camacho-Collados, J. (2022). *XLM-T: Multilingual Language Models in Twitter for Sentiment Analysis and Beyond.* Proceedings of the Thirteenth Language Resources and Evaluation Conference. Marseille, France. European Language Resources Association. link

Blinder, A.S., Ehrmann, M., de Haan, J., Jansen, D. (2022). *Central Bank communication with the general public: promise or false hope?* European Central Bank.

Brandt, M. W., Cochrane, J. H., Santa-Clara, P. (2006). *International risk sharing is better than you think, or exchange rates are too smooth.* Journal of Monetary Economics.

Degasperi, R., Hong, S., Ricco, G. (2021). *The Global Transmission of US Monetary Policy.* Technical report, Working Paper, University of Warwick.

Federal Reserve Bank of St. Louis (2023). *Nominal Broad U.S. Dollar Index (DTWEXBGS).* https://fred.stlouisfed.org/series/DTWEXBGS

Masciandaro, D., Peia, O., Romelli, D. (2023). *Central bank communication and social media: From silence to Twitter.* Journal of Economic Surveys.

Miranda-Agrippino, S., Rey, H. (2021). *The Global Financial Cycle.* Handbook of International Economics.

Organization for Economic Cooperation and Development (OECD). (2023). *Monthly Monetary and Financial Statistics (MEI): Exchange rates (USD monthly averages).* link

Shah, A. (2017). *The 100 Best Finance Twitter Accounts You Should Be Following.* Forbes. link

Stiefel, M., Vivès, R. (2022). *'Whatever it takes' to change belief: evidence from Twitter.* Review of World Economics.

Sun, H. (2023). *Dollar funding shortage - a new database and new evidence.* Working Paper.

Von Beschwitz, B., Collins, C. G., Datta, D. D. (2019). *Revisions to the Federal Reserve Dollar Indexes.* FED Notes. Federal Reserver Bank of the United States. link

Yahoo Finance (2023a). *FTSE 100 (FTSE).* link

Yahoo Finance (2023b). *S&P/TSX Composite index (GSPTSE).* link

Yahoo Finance (2023c). *S&P/ASX 200 (AXJO).* link

Yahoo Finance (2023d). *STXE 600 PR.EUR (STOXX).* link

8 STATUTORY DECLARATION / AFFIDAVIT

I hereby declare that the thesis with title

Global market sentiment and exchange rates: evidence from Twitter

has been composed by myself autonomously and that no means other than those declared were used. In every single case, I have marked parts that were taken out of published or unpublished work, either verbatim or in a paraphrased manner, as such through a quotation. This thesis has not been handed in or published before in the same or similar form.

Place: Needham, USA

Date: 11.08.2023

Brian Andrew Piotrowski

9 APPENDIX: LARGER TABLES AND DIAGRAMS

9.1 List of central banks used

Country	Name	Twitter handle
Argentina	Banco Central Argentina	@BancoCentral_AR
Australia	Reserve Bank of Australia	@RBAInfo
Brazil	Banco Central Brazil	@BancoCentralBR
Canada	Bank of Canada	@bankofcanada
France	Banque de France	@banquedefrance
Germany	Deutsche Bundesbank	@bundesbank
India	Reserve Bank of India	@RBI
Indonesia	Bank Indonesia	@bank_indonesia
Italy	Banca d'Italia	@bancaditalia
Japan	Bank of Japan	@Bank_of_Japan_e
Korea	Ministry of Economy and Finance - Korea	@moefkorea_eng
Mexico	Banco de México	@Banxico
Russia	Bank of Russia	@bank_of_russia
Saudi Arabia	Saudi Arabian Monetary Authority	@SAMA_GOV
South Africa	South African Reserve Bank	@SAReserveBank
Turkey	Central Bank of Türkiye	@CentralBank_TR
United Kingdom	Bank of England	@bankofengland
United State of America	Federal Reserve	@federalreserve
European Union	European Central Bank	@ecb

9.2 List of financial posters used

	Name	Twitter handle
0	John_Hempton	@John_Hempton
1	Barbarian Capital	@BarbarianCap
2	MuddyWatersResearch	@muddywatersre
3	Marc Cohodes	@AlderLaneEggs
4	Citron Research	@CitronResearch
5	brattlestcap	@brattlestcap
6	Kerrisdale Capital	@KerrisdaleCap
7	modest proposal	@modestproposal1
8	Market Folly	@marketfolly
9	Event Driven	@EventDrivenMgr
10	Activist Insight Shorts	@ActivistShorts
11	Carl Icahn	@Carl_C_Icahn
12	LST	@LongShortTrader
13	Donut Shorts	@DonutShorts
14	Spruce Point Capital	@sprucepointcap
15	Bluegrass Capital	@BluegrassCap
16	George	@NoonSixCap
17	Diogenes	@WallStCynic
18	Gotham City Research	@GothamResearch
19	Herb Greenberg	@herbgreenberg
20	ValueWalk - Exclusive hedge fund info (below)	@valuewalk
21	Union Square Research Group	@UnionSquareGrp
22	Plan Maestro	@PlanMaestro
23	The Reformed Broker blog	@reformedbroker
24	Skeletor	@SkeleCap
25	Fat Tail Capital	@FatTailCapital
26	Shortsighted Capital	@ShortSightedCap
27	footnoted	@footnoted
28	jacob Wolinsky	@JacobWolinsky
29	zerohedge	@zerohedge
30	FundyLongShort	@FundyLongShort
31	˜	@MugatuCapital
32	DumbLuckCapital	@DumbLuckCapital
33	Berna Barshay/HedgeFundGirl	@Hedge_FundGirl
34	Prescience Point Capital Management	@PresciencePoint
35	Marc Andreessen – e/acc	@pmarca
36	Fundamental Investor	@fundiescapital
37	PAA Research	@ActAccordingly
38	EquityNYC	@EquityNYC
39	No Sunk Costs	@nosunkcosts
40	Aron Pinson	@MicroFundy
41	Mike Bergen	@BergenCapital
42	marginal idea	@marginalidea
43	Jesse Livermore	@Jesse_Livermore
44	PainCapital	@PainCapital
45	Ed Borgato	@EdBorgato
46	Alex Rubalcava	@AlexRubalcava
47	FOHF	@LadyFOHF
48	tae kim	@firstadopter
49	Warren Buffett	@WarrenBuffett
50	The Wall Street Journal	@WSJ
51	Donald J. Trump	@realDonaldTrump

9.3 List of financial posters used (cont.)

	Name	Twitter handle
52	SoL	@xuexishenghuo
53	Jacob Ma-Weaver	@cablecarcapital
54	Probes Reporter, Now Disclosure Insight @DIrep...	@probesreporter
55	GRANT'S	@GrantsPub
56	Bloomberg	@business
57	DennyCrane	@DennyCrane550
58	StaleyRdCap	@StaleyRdCap
59	AV	@AureliusValue
60	FMV	@Find_Me_Value
61	David Einhorn	@davidein
62	Soren (CRE Appraiser)	@valuedude
63	Michael Fritzell (Asian Century Stocks)	@Fritz844
64	Post McTroll	@plainview_
65	Matt Wolfson	@TMTanalyst
66	MOI Global	@manualofideas
67	Quoth the Raven	@QTRResearch
68	Matt Levine	@matt_levine
69	Liberty	@LibertyRPF
70	AZ Value	@AZ_Value
71	FCFYield	@FCFYield
72	GlaucusResearch	@GlaucusResearch
73	hardcorevalueNA	@HardcoreValue
74	Voltaire	@PhilipEtienne
75	Hedgeye Energy	@HedgeyeENERGY
76	Tigre Capital	@TigreCapital
77	Copperfield Research	@CopperfieldRscr
78	The Ox	@adoxen
79	Michael Mauboussin	@mjmauboussin
80	Jeffrey Gundlach	@TruthGundlach
81	Bespoke	@bespokeinvest
82	Underwater Capital	@UnderwaterCap
83	Sterling Capital	@jay_21_
84	schaudenfraud	@schaudenfraud
85	John Huber	@JohnHuber72
86	Dow	@mark_dow

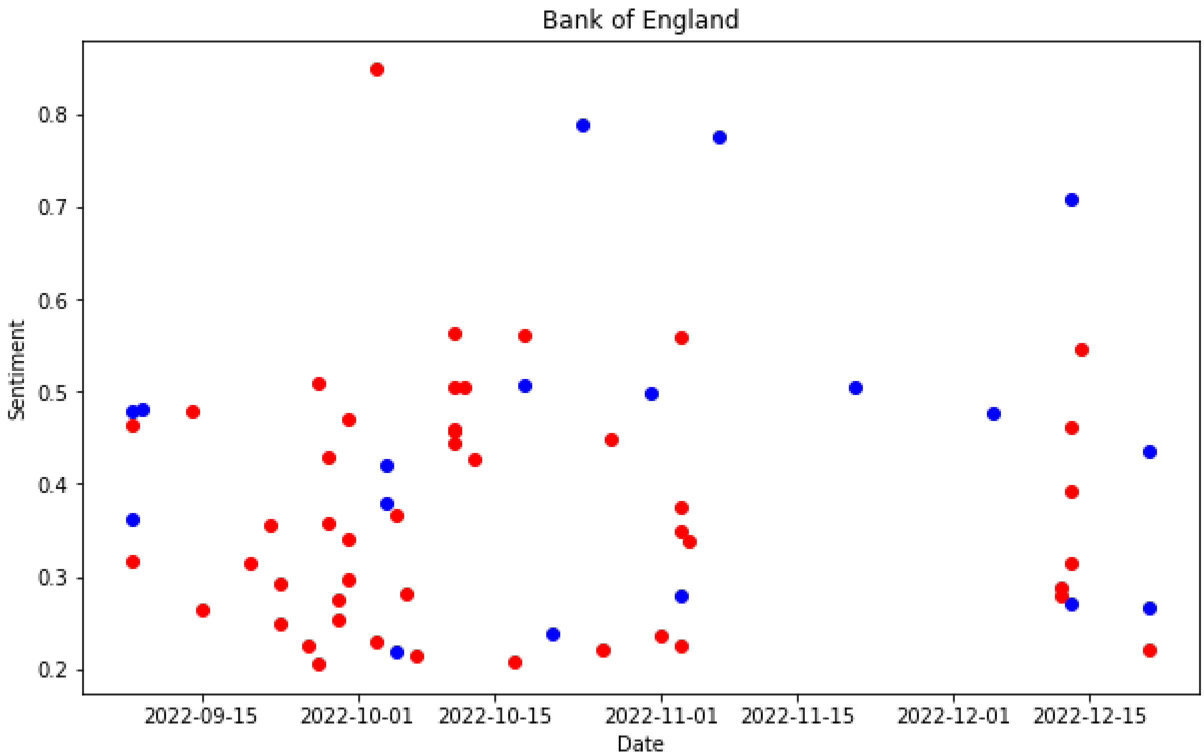
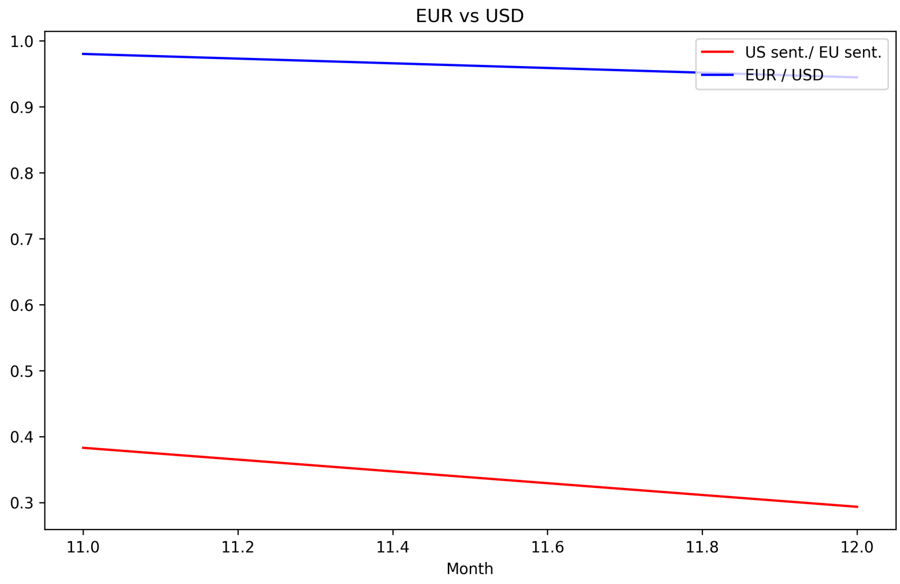
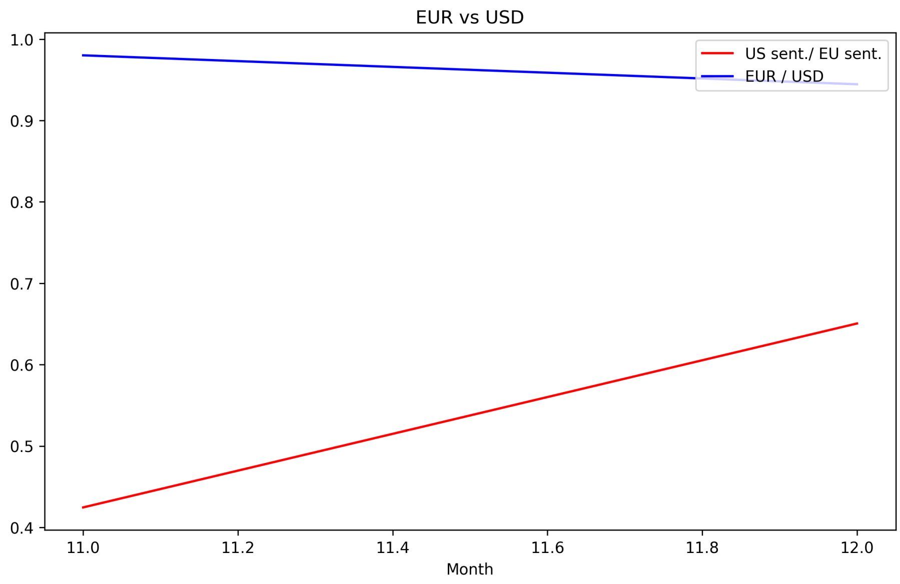


Fig. 4: Bank of England tweets which triggered a positive or negative classification (sentiment ¿ 0.2)

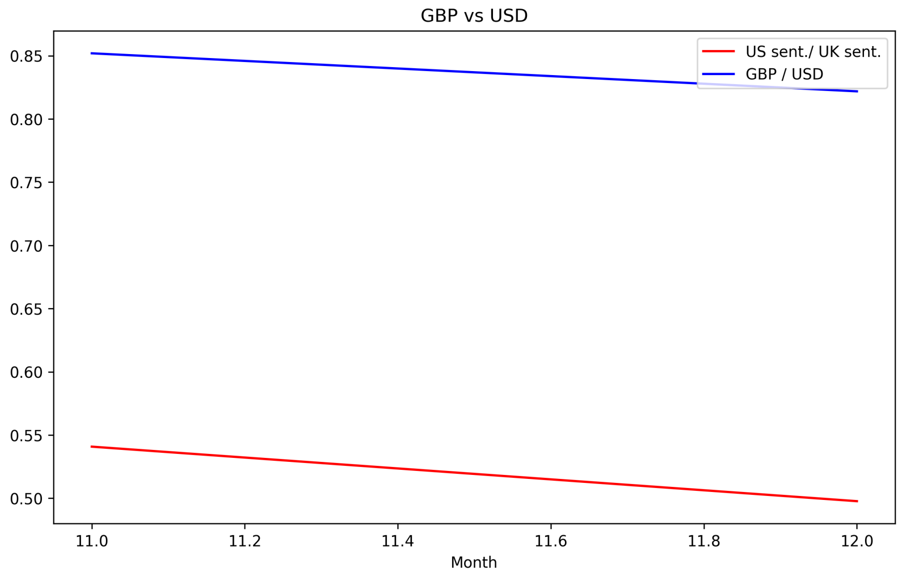


(a) Positive sentiment

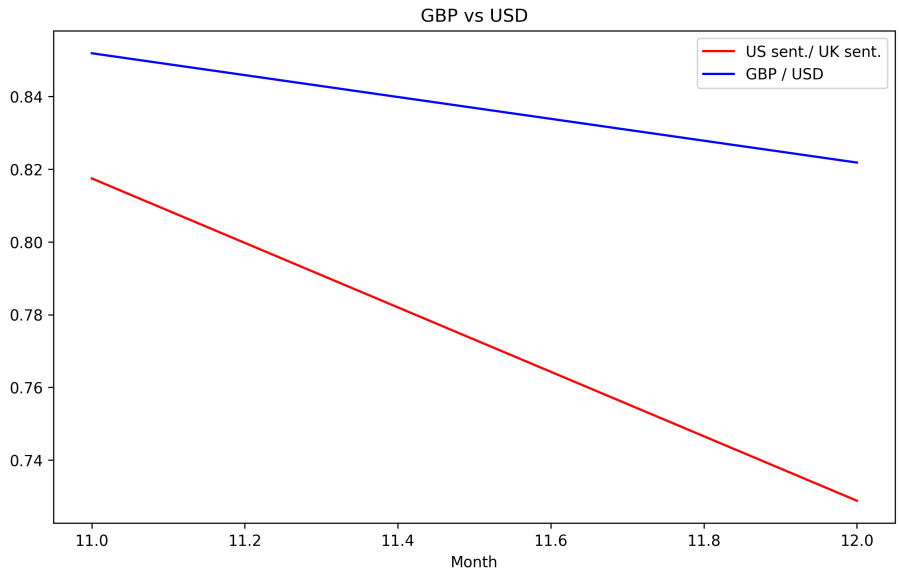


(b) Negative sentiment

Fig. 5: European Union

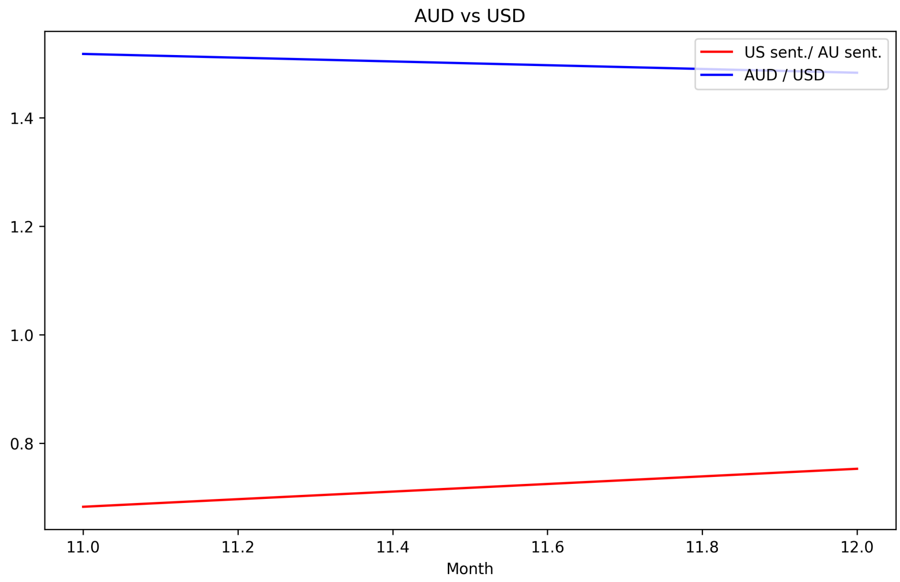


(a) Positive sentiment

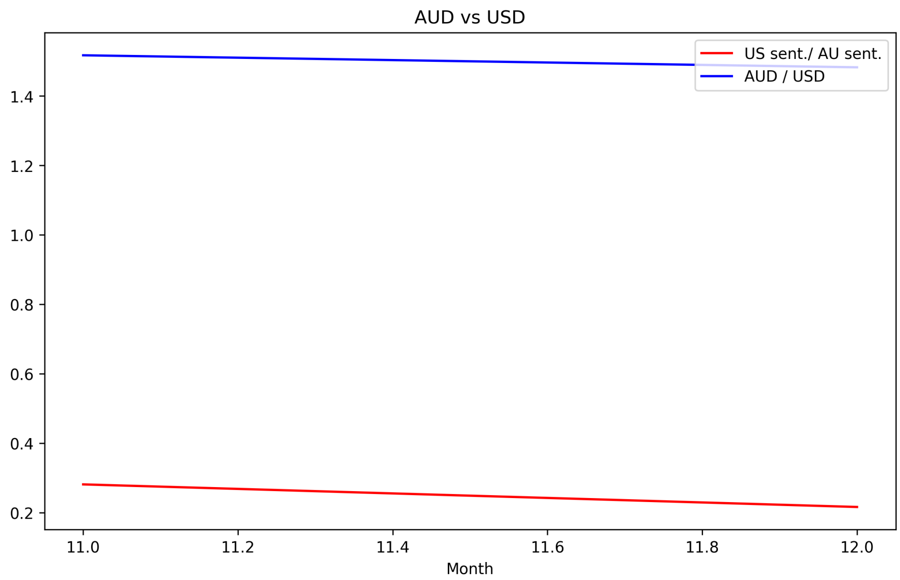


(b) Negative sentiment

Fig. 6: United Kingdom

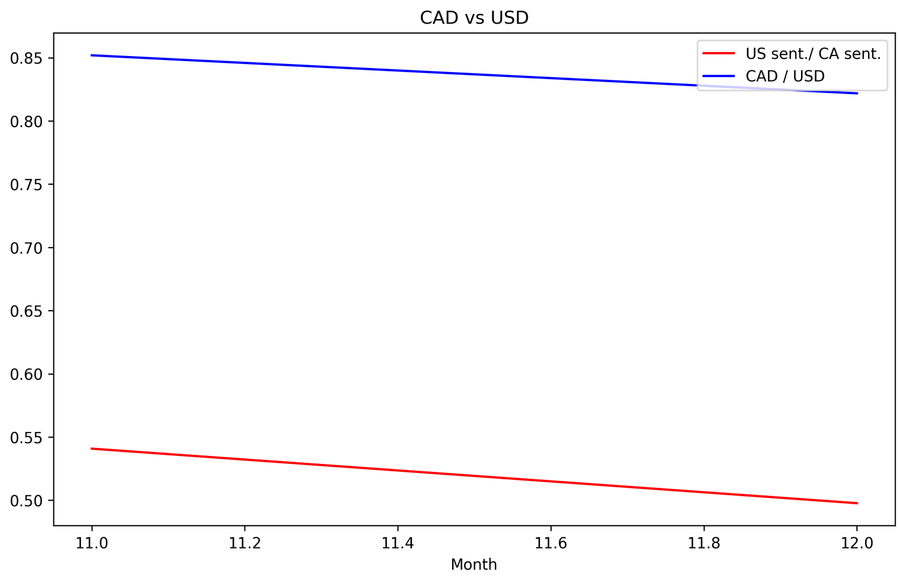


(a) Positive sentiment

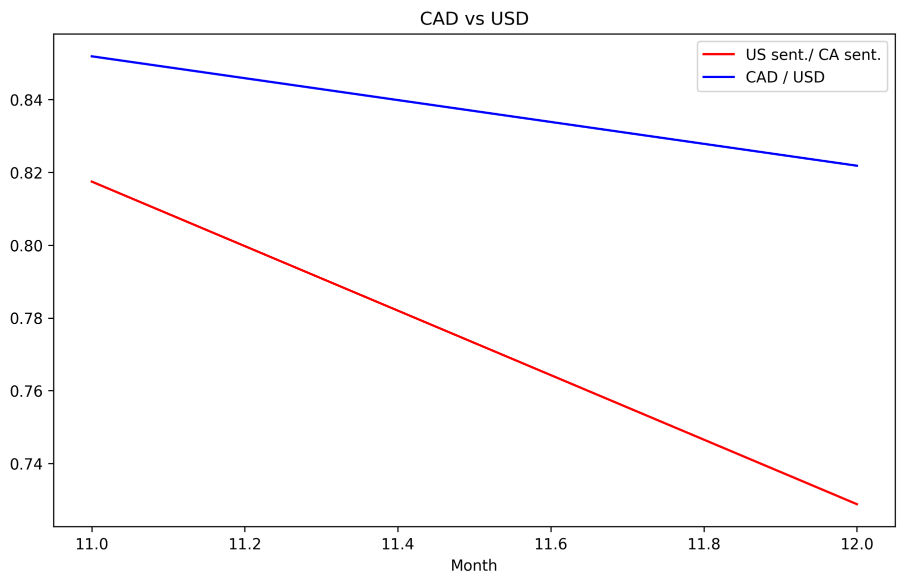


(b) Negative sentiment

Fig. 7: Australia

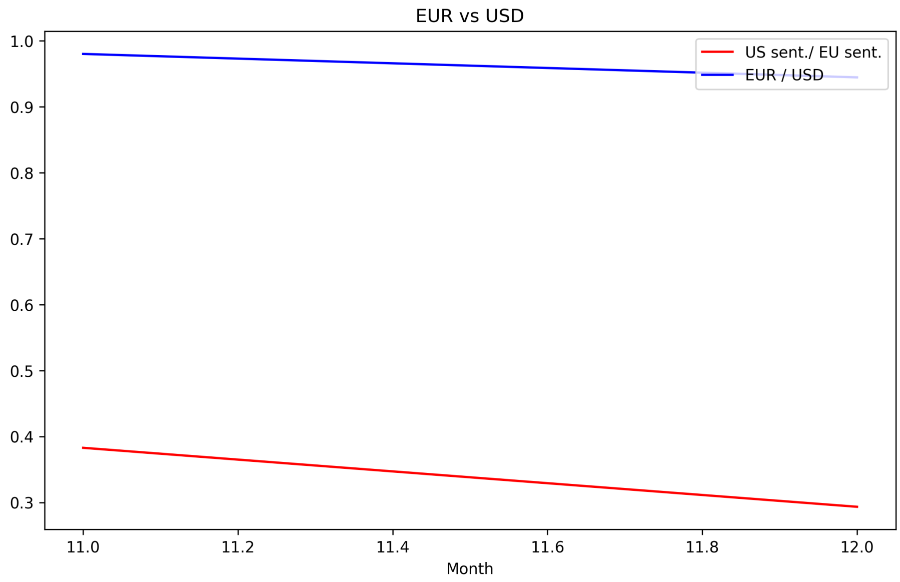


(a) Positive sentiment

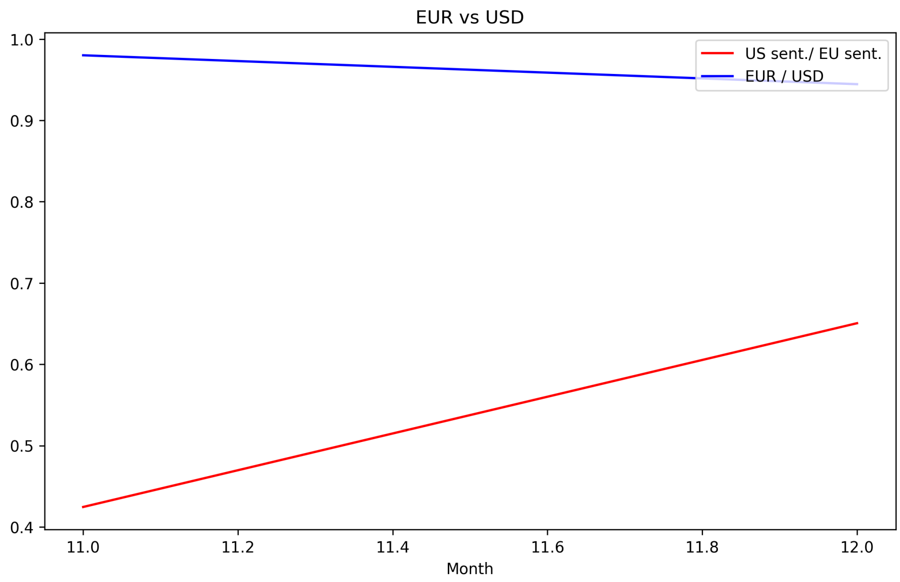


(b) Negative sentiment

Fig. 8: Canada



(a) Positive sentiment



(b) Negative sentiment

Fig. 9: Euro Area