WORKING paper

BANQUE DE FRANCE
EUROSYSTÈME

Blowing against the Wind? A Narrative Approach to Central Bank Foreign Exchange Intervention

Alain Naef [1]

April 2023, WP #911

Abstract

Most countries in the world use foreign exchange interventions, but measuring the success of the policy is difficult. By using a narrative approach, I identify interventions when the central bank manages to reverse the exchange rate based on pure luck. I separate them from interventions when the central bank actually impacted the exchange rate. Because intervention records are daily aggregates, an intervention might appear to have changed the direction of the exchange rate, when it is more likely to have been caused by market news. This analysis allows to have a better understanding of how successful central bank operations really are. I use new daily data on Bank of England interventions in the 1980s and 1990s. Some studies find that interventions work in up to 80% of cases. Yet, by accounting for intraday market moving news, I find in adverse conditions, the Bank of England managed to influence the exchange rate only in 8% of cases. I use natural language processing to confirm the validity of the narrative approach. Using Lasso and a VAR analysis, I investigate what makes the Bank of England intervene during that period. I find that only movement on the Deutschmark and not US dollar exchange rate made the Bank intervene. Also, I find that interest rate hikes were mostly a tool for currency management and accompanied by large reserve sales.

Keywords: Intervention, Foreign Exchange, Natural Language Processing, Central Bank, Bank of England.
JEL classification: F31, E5, N14, N24

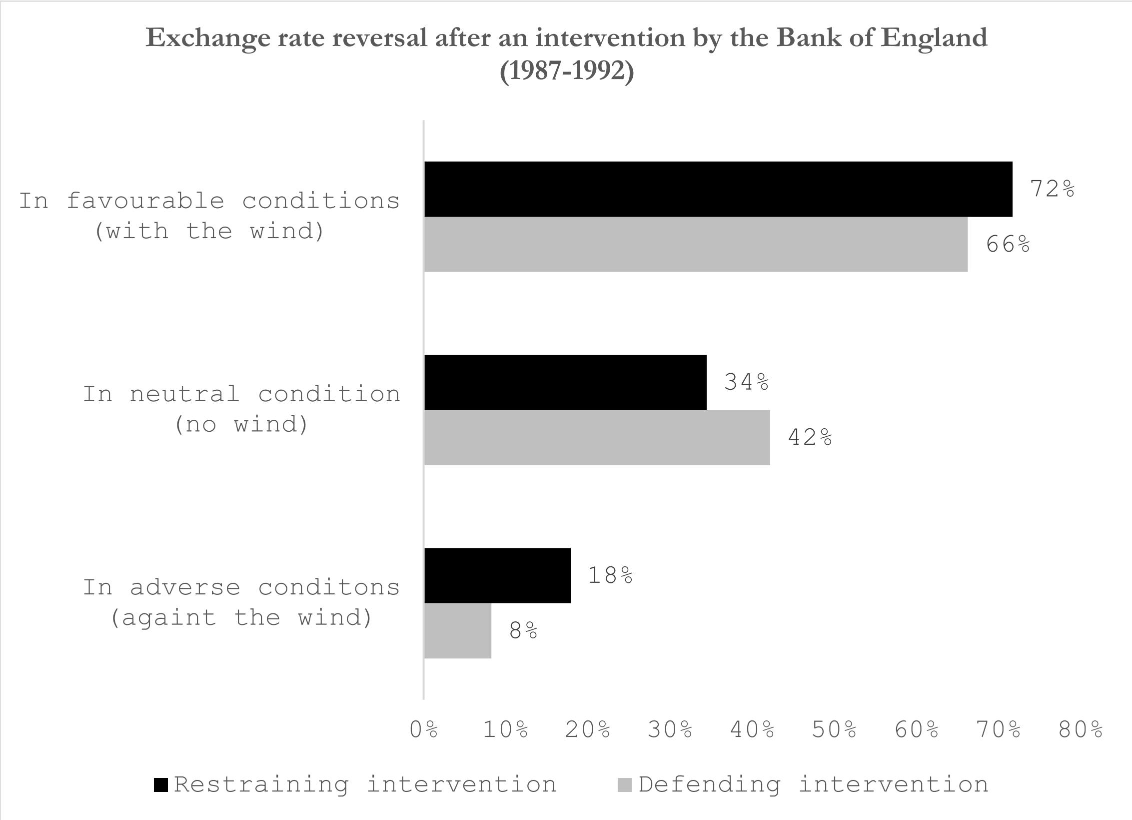
[1] Research Economist, Bank of France, alain-naef@banque-france.fr. This research was undertaken while a postdoctoral fellow at the Department of Economics of the University of California, Berkeley. I thank Jiaxin Lei and Gaurav Gandhi for excellent research assistance. For discussions and comments, I am grateful to Marc Adam, , Victor DeGorce, Brad DeLong, Simon Hinrischsen, Owen Humpage, Walter Jansson, Chris Meissner, Eric Monnet, Emi Nakamura, Kilian Rieder, David Romer, Yasmin Shearmur, Jón Steinsson, Tobias Stöhr, Mauricio Villamizar Villegas and Jacob Weber as well as participants at ASSA 2020, the European Economic History Conference in Paris, all-UC economic history conference, seminars at the Banque of France, University of Neuchatel, Lancaster, UC Berkeley, UC Davis, ASSA 2021, Atlanta Fed monetary history conference, LSE, King's College, ESSEC and Paris XVIII. Part of this research was supported by the Economic and Social Research Council (grant number ESRC KFW/10324121/0) as well as the Swiss National Science Foundation (grant number P2SKP1_181320). The views expressed in this paper do not represent the opinion of the Banque de France or the Eurosystem.

NON-TECHNICAL SUMMARY

Interventions on the foreign exchange market are important. Most central banks still follow exchange rate objectives and over 80% of countries are on fixed exchange regimes. Japan recently renewed with a long tradition of interventions to try to prop up the yen, spending $36bn in a day. Yet, practitioners and academics disagree on the effectiveness of interventions.

Our understanding of central bank interventions is limited, as interventions are often endogenous to market conditions; a central bank intervenes in reaction to a market shock. Here I identify these market shocks at a daily frequency to better measure the impact of interventions on exchange rates. To identify these shocks, I analyze narrative evidence about market conditions written by Bank of England officials. I clearly identify days when the currency is hit by negative news that are not related to the intervention of the central bank. I get this measure by analyzing the text from the daily reports written by Bank of England employees. To test the robustness of my narrative analysis, I rely on both an external assessment and machine learning in the form of Natural Language Processing (NLP).

I find that when discounting for operations not meant to impact the exchange rate, the Bank of England is only successful around 8% of the time (chart below). That is when it is trying to make the exchange rate appreciate against a bearish market (going "against the wind"). These are times when interventions really matter. This complements other studies in the literature that find higher success rates, without accounting for market news. Fratzscher et al. (2019) find success rates of over 80% when central banks aim to manage volatility. Their study is the most comprehensive attempt to understand foreign exchange interventions to date. Presenting evidence from 33 countries, they argue that central bank interventions were effective in attaining the goals set by policymakers from 1995 to 2011. The paper does an excellent job at analyzing new data, but like most papers on the topic does not offer a bulletproof identification strategy.



More than half of the interventions considered successful using the previous standard methodologies no longer count as successful with my approach. When measuring whether good or bad news (independent from central bank interventions) was circulating on a given day, success drastically changes. Intervention is particularly ineffective when attempting to reverse the direction of the exchange rate after negative news affecting the currency. Another finding is that Bank of England intervention was more effective when trying to tame the appreciation of sterling ("restraining intervention") than when trying to avoid a depreciation of sterling ("defending intervention"). This makes intuitive sense. Markets are more likely to believe a central bank with unlimited domestic currency it can print than a central bank with scarce dollar reserves.

The main contribution of this paper is to show that by not accounting for intraday news, most papers on intervention effectiveness provide biased estimates. An intuitive example allows us to understand this shortcoming and how this paper deals with them. Imagine that today, the People's Bank of China (PBOC) was trying to make the renminbi appreciate through foreign exchange intervention. At 10am, they buy renminbi with their dollar reserves, hoping this will bolster the renminbi's price. Now imagine that an hour later at 11am, the American administration announces dropping planned tariffs, sparking a stark appreciation of the renminbi. Most studies on intervention would simply assume that the 10am interventions were successful, completely ignoring the news later which change the direction of the exchange rate. In other terms, they assume that any interventions during that day were going against the wind. That is, that they were going against the market. These studies would count this intervention as successful. My narrative approach accounts for other news during the day to assess whether the intervention was really going against the wind, or if it merely happened to go in the same direction as the market.

Souffler contre le vent ? Une approche narrative des interventions de change des banques centrales

RÉSUMÉ

La plupart des pays du monde ont recours aux interventions sur le marché des changes, mais il est difficile de mesurer le succès de cette politique. En utilisant une approche narrative, j'identifie les interventions lorsque la banque centrale parvient à inverser le taux de change par pure chance. Je les distingue des interventions au cours desquelles la banque centrale a réellement influencé le taux de change. Étant donné que les enregistrements des interventions sont des agrégats quotidiens, une intervention peut sembler avoir changé la direction du taux de change, alors qu'il est plus probable qu'elle ait été causée par des nouvelles de marché. Cette analyse permet de mieux comprendre le succès réel des opérations des banques centrales. J'utilise de nouvelles données quotidiennes sur les interventions de la Banque d'Angleterre dans les années 1980 et 1990. Certaines études montrent que les interventions fonctionnent dans 80 % des cas. Pourtant, en tenant compte des nouvelles qui influencent le marché au cours de la journée, je constate que dans des conditions défavorables, la Banque d'Angleterre n'est parvenue à influencer le taux de change que dans 8 % des cas. J'utilise le traitement du langage naturel pour confirmer la validité de l'approche narrative. À l'aide de régressions LXasso et d'une analyse VAR, j'étudie les raisons qui poussent la Banque d'Angleterre à intervenir. Je constate que seuls les mouvements du taux de change du deutschemark, et non du dollar américain, sont à l'origine de l'intervention de la Banque. Je constate également que les hausses de taux d'intérêt ont été principalement un outil de gestion des devises et qu'elles ont été accompagnées d'importantes ventes de réserves.

Mots-clés : intervention de change, taux de change, traitement du langage naturel, banque centrale, Banque d'Angleterre.

---

1. Introduction

Intervention on the foreign exchange market is important. Most central banks still follow exchange rate objectives and over 80% of countries are on fixed exchange regimes (Taylor 2010; Ilzetzki, Reinhart, and Rogoff 2019). Japan recently renewed with a long tradition of intervention to try to prop up the yen, spending $36bn in a day. Yet, practitioners and academics disagree on the effectiveness of intervention. Central bankers generally believe that intervention has an impact on exchange rates (Neely 2008). Academics, on the other hand, have generated contradictory findings. Most studies trying to assess intervention use daily data and struggle to deal with the intraday changes in market conditions, making it difficult to assess the true effectiveness of intervention.

Our understanding of central bank intervention is limited, as interventions are often endogenous to market conditions; a central bank intervenes in reaction to a market shock. Here I identify these market shocks at a daily frequency to better measure the impact of intervention on exchange rates. To identify these shocks, I use a narrative approach as pioneered by Romer and Romer (1989). I analyze narrative evidence about market conditions written by Bank of England officials. I clearly identify days when the currency is hit by negative news that is not related to the intervention of the central bank. I get this measure by analyzing the text from the daily reports written by Bank of England employees. To test the robustness of my narrative analysis, I rely on both an external assessment and machine learning in the form of Natural Language Processing (NLP).

I find that when discounting for operations not meant to impact the exchange rate, the Bank of England is only successful around 8% of the time. That is when it is trying to make the exchange rate appreciate against a bearish market (going "against the wind"). These are times when interventions really matter. This complements other studies in the literature that find higher success rates, without accounting for market news. Fratzscher et al. (2019) find success rates of over 80% when central banks aim to manage volatility. Their study is the most comprehensive attempt to understand foreign exchange intervention to date. Presenting evidence from 33 countries, they argue that central bank intervention was effective in attaining the goals set by policymakers from 1995 to 2011. The paper does an excellent job at analyzing new data, but like most papers on the topic does not offer a bulletproof identification strategy.

Yet, more than half of the interventions considered successful using the previous standard methodologies no longer count as successful with my approach. When measuring whether good or bad news (independent from central bank interventions) was circulating on a given day, success drastically changes. Intervention is particularly ineffective when attempting to reverse the direction of the exchange rate after negative news affecting the currency. Another finding is that Bank of England intervention was more effective when trying to tame the appreciation of sterling ("restraining intervention") than when trying to avoid a depreciation of sterling ("defending intervention"). This makes intuitive sense. Markets are more likely to believe a central bank with unlimited domestic currency it can print than a central bank with scarce dollar reserves.

The effectiveness of sterilized intervention has long been questioned, and the debate is still ongoing.[1] Another strategy in the literature to better identify the effectiveness of intervention is the use of high frequency data (Chang and Taylor 1998; Echavarría, Melo-Velandia, and Villamizar-Villegas 2018; Menkhoff 2010; Hofmann, Shin, and Villamizar-Villegas 2019; Aslam et al. 2020) [2]. These studies, however, rarely have access to confidential intervention data like what is done here.[3] Equally, they can also suffer from the same bias suggested here; that is, that news instead of intervention can be what moves the exchange rate from one period to the other. The literature also uses IV approaches, but these often fail to completely deal with the endogeneity issues (Menkhoff, Rieth, and Stöhr 2021; Naef and Weber 2021; Adler, Lisack, and Mano 2019). Yet these studies rarely convincingly address causality.[4]

[1] For an overview of the literature on central bank intervention, see Sarno and Taylor (2001) and Neely (2005), more recent papers by Adler, Lisack, and Mano 2019; Echavarría, Melo-Velandia, and Villamizar-Villegas 2018; Hu et al. 2016; Blanchard, Adler, and Filho 2015; Adler and Mano 2021.

[2] Note that there are papers taking into account intraday foreign exchange rate news such as Andersen et al. (2003), Dominguez and Panthaki (2006), Evans and Lyons (2008) and Ehrmann and Fratzscher (2005) but these papers do not directly deal with intervention effectiveness.

[3] And other scholars have looked into secret interventions before, see Klein (1993), Fischer (2006) and Dominguez and Panthaki (2007).

[4] Anecdotal proof of this inability to demonstrate causality is that despite being an important macroeconomic policy, no empirical paper on foreign exchange intervention has been published in any top-5 journal in recent years.

The main contribution of this paper is to show that by not accounting for intraday news, most papers on intervention effectiveness provide biased estimates. An intuitive example allows us to understand this shortcoming and how this paper deals with them. Imagine that today, the People's Bank of China (PBOC) was trying to make the renminbi appreciate through foreign exchange intervention. At 10am, they buy renminbi with their dollar reserves, hoping this will bolster the renminbi's price. Now imagine that an hour later at 11am, the American administration announces dropping planned tariffs, sparking a stark appreciation of the renminbi. Most studies on intervention would simply assume that the 10am interventions were successful, completely ignoring the news later which change the direction of the exchange rate. In other terms, they assume that any interventions during that day were going against the wind. That is, that they were going against the market. These studies would count this intervention as successful. My narrative approach accounts for other news during the day to assess whether the intervention was really going against the wind, or if it merely happened to go in the same direction as the market.

Another limitation to understand central bank intervention is the lack of data, as central banks keep their intervention records secret. Here, I unveil hand-collected intervention data from the UK, spanning from 1987 to 1992. The data are original, as it is mainly composed of secret interventions, which were not communicated to the public (this is still how many central banks operate today). All interventions in the data set are sterilized, a British institutional feature. As operations were run with the funds of the Exchange Equalisation Account and not the Bank of England directly, all operations had a counterparty transaction in government bonds.

In brief, the main contribution of this paper is to demonstrate how intervention success is much lower when intraday market news is accounted for. I especially show that interventions trying to prop up a falling currency are rarely successful. I use a narrative approach to better identify intervention success. The paper also offers a previously secret intervention database, available for replication and further study. And finally, this is one of the first papers to offer a methodology relying on machine learning to check the validity of a Romer and Romer narrative approach.

2. New confidential data

Various empirical studies on foreign exchange intervention use the same datasets and focus only on countries with public intervention records such as Turkey, Colombia or Japan. Looking at secret interventions, my findings have implications for all central banks intervening in secret, which is under-researched (Mohanty and Berger 2013; Chamon et al. 2019). Only around 4% of the operations in my sample were publicized (66 out of 1,533), all others are secret. Note that secret intervention does not always mean that the market is unaware of the intervention; it means that the central bank does not officially announce it. Dominguez (2003) suggests that traders in the 1990s usually knew that the Fed was intervening at least one hour before any news outlets would report it.

This paper presents a daily database of intervention by an advanced economy. The data spans 22 February 1987 (when the dealers start reporting narrative evidence and the UK starts shadowing the Deutschmark) to Black Wednesday when the bank stopped intervening (16 September 1992). Figure 1 presents the data in 1992-US dollars. The data offers an aggregate amount of all Bank of England intervention operations during a given day. This includes intervention in any currency. In practice, interventions were mainly in dollars until 1987 and in Deutschmark thereafter (see the historical section in the appendix for more details).

Central bank intervention can either be sterilized (with simultaneous bill purchases that leave the monetary conditions unaffected) or unsterilized (with no asset purchases, thus affecting the monetary conditions). Unsterilized intervention affects the exchange rate through changes in interest rates, making the currency more or less attractive to investors. There is more debate on the working of sterilized intervention.[5]

[5] The literature identifies three channels through which sterilized intervention works: portfolio-balance, signaling and coordination. The portfolio-balance channel has received the most attention in the literature (Cavallino 2019; Gabaix and Maggiori 2015; Fatum 2015; Dominguez and Frankel 1993). The channel works through investors' portfolio shocks, which affect the amount of bonds in circulation and their risk premia, and by doing so affect the exchange rate. Signaling works through central banks giving hints of future monetary policy stances to which the market reacts (Fatum and Hutchison 1999). The coordination channel works when the market is thin and traders have lost confidence in the ability of macroeconomic fundamentals to inform the price, the central bank can step in and provide direction (Reitz and Taylor 2008). In this paper, I do not choose a channel through which intervention could work, but focus on methodological issues in current papers on the topic.

The data are negative for sales and positive for purchases of foreign currency. Intervention data come from the Bank of England dealers' reports, which offer daily records.[6] The reports were written by the dealers of the Bank of England, foreign exchange operators who managed sterling on behalf of the government. The archive of the Bank of England kept printed copies of the reports, which I copied individually to put my database together.

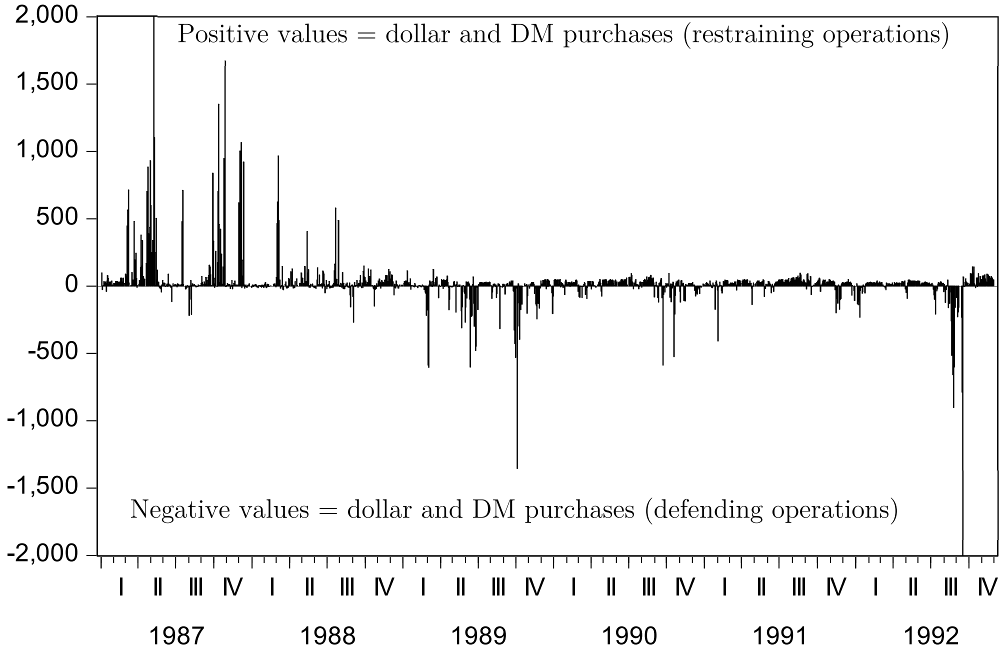


Figure 1 – Intervention by the Bank of England in million of 1992-dollars. Source: The data have been copied from reports written with typewriters kept at the archives of the Bank of England (archive reference C8).
NOTE: The data are cropped at $1bn to improve readability but the figures go up to -22$bn for Black Wednesday on 16 September 1992.

Another benefit of this data is that they come directly from policymakers, without any filter or control. Published data by central banks are likely to be processed before publication and might not include all operations. It is well known to foreign exchange traders today that, for example, South Africa and Brazil, publish some of their operations while keeping a large unpublished derivatives' book. The dataset presented here is much more precise than those of other studies in the field, which rely on proxies, such as changes in reserve levels or press reports.[7]

[6] Bank of England Archives, Cashier's Department: Foreign Exchange and Gold Markets – Dealers' Reports, C8.
[7] Adler et al. (2021) recently released a dataset offersing a comprehensive intervention data for several countries, but unlike the data presented here, it is based on proxies and not real data. Also, it offers monthly, not daily, data.

3. Intervention success rates

When assessing intervention, most papers ignore the intention of the central bank. To better tackle intervention success, I use a narrative approach, detailed below.[8] The advantage of looking into history is that the reasoning behind the intervention decisions of policymakers is available. The Bank of England recently changed its information access policy and now opens most of its archival documents to researchers after a 20-year period. As the last intervention occurred in 1992, we have recently gained access to the reasoning of central bankers as they were intervening. As a narrative approach can contain some subjective assessment, the robustness section uses both human- and machine-based techniques to control for potential subjectivity in my assessment.

Narrative approach – Reading policymakers' minds

Intervention occurs within a specific context where policymakers react to adverse market conditions (what central bankers call "leaning against the wind"). If the exchange rate is depreciating because of poor trade figures, for example, it is likely that intervention will be less effective than if the central bank intervenes on a day with more positive news associated with the currency. Similarly, if traders are bullish about the currency (because of a positive GDP forecast for example), it will be more difficult for the central bank to tame an increase in the currency.

Starting in April 1986, the foreign exchange dealers of the Bank of England changed the way they reported market activity. They started to provide a small paragraph assessing the situation of the pound for every trading day. These memos were sent to the Treasury (remember that at the time the Treasury was in charge of monetary and exchange rate policy in the UK, not the Bank of England). They concisely list whether any exogenous factors were putting pressure on the sterling exchange rate during the day. Table 1 below provides some examples and Table 6 in the appendix presents a broader sample.

[8] Narrative approaches have been used for other questions but this paper is the first to use the methodology in the context of foreign exchange intervention. For more on narrative approaches, see Romer and Romer (1989, 1994, 2014) or Monnet (2014).

These data are invaluable as they not only list exogenous factors influencing the exchange rate (say, the publication of a large trade deficit) but also how the market perceived this in comparison to expectations. This is essential information as bad market news for a currency, such as a large trade deficit, could actually lead to the currency's appreciation if the market was expecting worse figures. Being at the center of the foreign exchange market and in daily contact with all the main foreign exchange dealers, Bank of England employees had a good overview of what the market was expecting. They not only noted any market-moving news but also detailed how it compared to market expectations. The data are accurate as they were directly recorded at the end of the trading day. The information is also superior to any information that can be found in newspapers, as the dealers spoke to investment banks daily and had access to insider information. They also knew before other dealers if there would be changes in the Bank Rate. The reports are consistent and constant, which makes them ideal for our purposes.

I classify the dealers' assessment of market conditions into three categories depending on the news regarding the value of sterling. Each day either displays good news for the currency (for example better trade figures than the market expected), neutral news (no significant news or change in conditions), or bad news (for example worse than expected unemployment figures).

Table 1 shows examples of the three types of news as expressed by dealers. The Bank of England dealers are also aware of aspects that technical traders observe, for example, a psychological threshold of 3DM per sterling. Other technical traders known as chartists would also look at momentum and sell after a certain number of days of currency increase, or other such rules familiar to Bank of England dealers. These subtleties were also noted by the Bank of England dealers in their records and might not be found in news reports.

These reports are valuable as they show how better than expected news does not always influence the exchange rate as expected in statements such as "Sterling ignored better than expected Q2 GDP figures" (Dealers' Report, July 22, 1994). The dealers report not only general market expectations, which they gather from their daily market interactions, but also how the different news items are reflected in intraday price changes.

To see exactly how I classified these statements into good, neutral and bad, Table 6 in the appendix shows the choices I made on a random sample from the reports. I use content analysis to assess the dealers' reports on market conditions. Content analysis includes a wide series of tools to extract meaning from text (Krippendorff 2018; Neuendorf 2016). I read each paragraph on market conditions and I assessed whether the general conditions indicated good conditions for the currency of intervention. Data were then coded into a dummy variable: value 1 for positive news; 0 for days with unclear trends or little market activity; and -1 for days with adverse news.

	Examples of key sentences
GOOD NEWS for sterling	"Sterling benefited from the weekend opinion polls and press comment" "The dollar and sterling both gained on German interest rate rumours" "After an uncertain start, sterling came into strong demand from Europe during the morning, helped by the trade figures." "Sterling was pulled higher by the strong dollar" "[…] moved steadily higher after the better than expected trade figures" "Sterling was in good demand, helped by the reassuring PPI data and a perception that the recovery is 'on track'."
NEUTRAL or NO NEWS	"The markets were again quiet" "Sterling was on the sidelines for most of the day"
BAD NEWS for sterling	"New York continues to take a more bearish view of sterling, where more weight is given to devaluation rumours." "There was also some short covering in front of tomorrow's Mansion House speech by the Chancellor" "Dealers were unimpressed by the CBI survey and sterling tended to move lower with the dollar" "{sterling} tended to soften along with the dollar, and failed to benefit from better than expected output data (industrial production +0.8%, manufacturing +0.6%)" "Sterling ignored better than expected Q2 GDP figures and struggled {…}"

Table 1 – Examples of good, neutral and bad news for the pound. The assessment is done by the author and the robustness section shows assessment by different methods. Table 6 in the appendix shows twenty randomly selected full quotes from reports with their coding. Source: Bank of England archives, Dealers' reports, reference C8.

As content analysis entails a portion of subjective judgment, the robustness section later in the paper offers several ways of controlling for potential subjectivity. The first was by replacing my personal judgement with a machine learning algorithm; the second, using Amazon Mechanical Turk (MTurk) to make third parties assess the same paragraphs I assessed. But first let us look at how to assess intervention success.

A naïve event study approach

A large proportion of papers in the intervention literature rely on success counts methodologies to understand central bank effectiveness. These measures are not causal and here I use my narrative approach to show how these estimates can be biased. The success methodology presented here relies on three intervention success criteria (SC) and is inspired by a methodology by Bordo, Humpage, and Schwartz (2015) and similar to Fratzscher et al. (2019). SC_1 measures whether intervention leads to an appreciation/depreciation of sterling against the Deutschmark between the previous day's market close and the current day's market close. SC_2 measures whether the exchange rate depreciates/appreciates less after intervention between the day's market close and the previous day's market close than it did over the immediately preceding period (also called smoothing). The final criterion, SC_3, combines the first two. The three criteria take the form of a binary variable and are formalized in the equations below.

$$SC_1 = \begin{cases} 1 & \begin{cases} if\ I_t > 0, and\ \Delta S_t < 0, \\ \qquad or \\ if\ I_t < 0, and\ \Delta S_t > 0 \end{cases} \\ 0 & otherwise \end{cases}$$

$$SC_2 = \begin{cases} 1 & \begin{cases} if\ I_t > 0, and\ \Delta S_{t-1} > 0\ and\ \Delta S_t \ \geq 0, and\ \Delta S_t < \Delta S_{t-1} \\ \qquad or \\ if\ I_t < 0, and\ \Delta S_{t-1} < 0\ and\ \Delta S_t \ \leq 0, and\ \Delta S_t > \Delta S_{t-1} \end{cases} \\ 0 & otherwise \end{cases}$$

$$SC_3 = \begin{cases} 1 & SC_1 = 1\ or\ SC_2 = 1 \\ 0 & otherwise \end{cases}$$

where I_t designates foreign exchange intervention on day t. Positive values are purchases of foreign exchange (called restraining interventions) and negative values are sales of foreign exchange (defending interventions). A purchase is expressed as $I_t > 0$ and a sale as $I_t < 0$. $\triangle S_t$ is the difference between the spot closing rate on the day of the intervention and the spot closing rate the day before the intervention. It shows the effect of the intervention on the exchange rate during the day.

The focus is on the daily effect. This is justified by the type of intervention by the Bank of England, which was on the market most days, as shown in the data section. Kearns and Rigobon (2005) show that most of the impact of an intervention occurs during the day it is conducted.

Table 2 presents the naïve results. It is separated into the three success criteria presented above, as well as into defending (to make the exchange rate appreciate) and restraining interventions (to make the exchange rate depreciate).

		Reversing exchange rate (SC1)		Smoothing appreciation or depreciation (SC2)		Total success (SC3, sum of SC1 and SC2)	
	Day count	Count	Percentage of intervention being successful	Count	Percentage of intervention being successful	Count	Percentage of intervention being successful
Defending interventions	357	102	29%	61	17%	163	46%
Restraining interventions	957	327	34%	179	19%	506	53%
Total	1314	429	33%	240	18%	669	51%

Table 2 – Intervention success according to the three criteria presented above.

Note that these "success" rates in this preliminary approach do not imply causality; it could be that some of these "successes" are due to other factors (as we will see in the narrative approach). Naïve results in Table 2 show the Bank was more "successful" in restraining than defending interventions, that is, it was more successful when it tried to tame sterling than bolster it. This holds true for all three success criteria. The intuition is that markets take a central bank more seriously when it is intervening with its own currency, which is available in unlimited amounts, than when intervening with scarce dollar or Deutschmark reserves.

These results compare with other findings in the literature. For example, using the same methodology, Bordo, Humpage, and Schwartz (2015) found that the Federal Reserve was successful in reversing the Deutschmark/dollar exchange rate 29% of the time, which happens to be the same number as the results of my study for Deutschmark/sterling. However, they found higher success rates on the yen/dollar exchange rate, going up to 70%. Fratzscher et al.

(2019), using a different but somewhat similar methodology, found that from 1995 to 2011, countries in free floating regimes could manage to reverse exchange rates (the equivalent of SC1) in 61% of the episodes.[9] This is higher than the 24% reversal rates in this paper, which is likely due to the fact that the Bank of England intervened more frequently, leading to less success.

Combining the naïve event study with the narrative approach

Figure 2 shows the reversal success rate (SC1) according to whether interventions were going with the market, without any significant market direction, or against the market as explained above. The figure shows how interventions are almost 10 times less successful when they go against intraday market conditions. This is expected. However, most studies on intervention effectiveness miss this distinction and measure the effect of intervention ignoring intraday market information. This sample shows that this assumption is wrong: almost half of the Bank of England's interventions during 1987-92 were not going against market forces. The appendix gives more detailed result than presented here and also shows SC2 and SC3 results. SC3 follows a similar pattern. SC2 offers a different pattern, but these cases are much less frequent and therefore bear less on the total sum of results. SC1 is the criteria most comparable with other studies.

The success rate of 8% for defending interventions going against market trends can be compared to the benchmark rate of 19%, obtained in the previous section when not discriminating for market conditions (see Table 2). Put differently, around 50% of the interventions that were judged successful using the previous standard methodology now no longer count as successful when accounting for market conditions.

What are concrete examples of the Bank intervening with the wind that we do not want to count? It can be two things, either the Bank is trying to reinforce a market trend. Or it can be that the "wind" turned during the day. Below is an example for each case.

[9] The sample in Fratzscher et al. (2019) is over 76% of restraining interventions, where my sample contains 69% of restraining interventions. Regarding methodological differences, Fratzscher et al. (2019) use several days events, which has certain advantages but can lead to endogeneity problems as it becomes hard to see if intervention was successful because of central bank operations or because of normal changes in the exchange rate. The next section tackles this endogeneity issue in more detail.

On July 16, 1992 the Bank of England sold Deutschemark for 19 million dollars' worth to strengthen the pound. This intervention was counted as "successful" by the traditional criteria as the exchange rate reversed from the day before. Dealers reported: "Sterling opened softer on precautionary selling ahead of today's trade figures. The announcement of a smaller-than-expected deficit caused sterling to rally quickly to its highs." In this case, the dealers probably intervened in the morning to support the pound and the intervention was only "successful" because of the announcement of the positive trade figures in the afternoon.

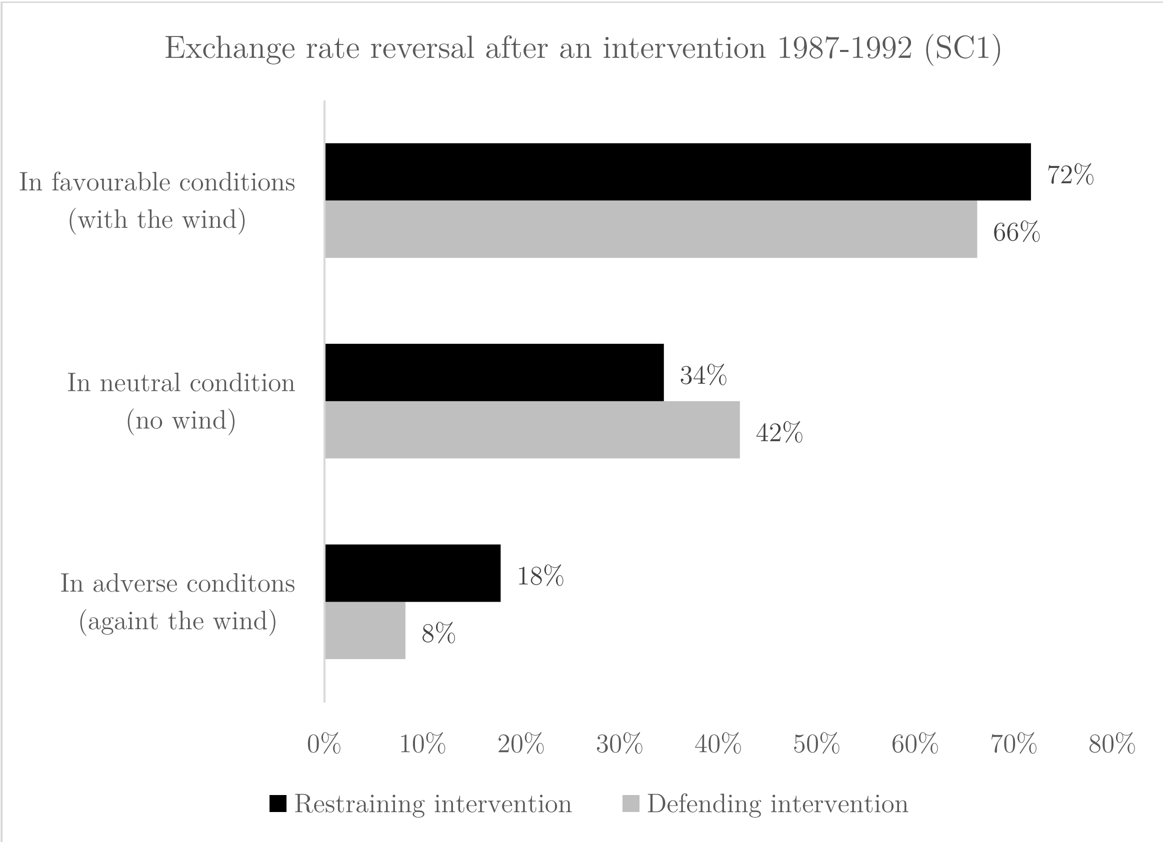


Figure 2 – Success Criteria 1 (SC1) or reversal success count taking into account market conditions.

On June 27, 1989, the Bank of England explicitly went with the wind after some positive market move. The bank followed the direction of the exchange rate and this intervention should not be counted as successful. The dealers report explain that the "unexpectedly modest May deficit" figures "sparked of a strong rally in sterling" which they followed with "aggressive support" in the same direction as the market. In this case, the Bank explicitly followed the market trend. With my methodology this would not be counted as a successful against the wind intervention when in other studies it would. The cause of the reversal of exchange rate from the previous day is clearly due to the trade figures, and not the Bank which simply followed an existing market trend.

4. Testing the robustness of the narrative approach – Humans and machines

A frequent criticism to narrative approaches is their subjectivity. What one researcher might classify a certain way, another might do differently. To mitigate the issue, I use two forms of robustness check. First I use Amazon Mechanical Turk (MTurk) to have third party subjects replicate my assessment. Second, I use a machine learning algorithm to see whether my results are consistent. Neither of these methodologies offer the same richness of data analysis as the narrative approach, but they do enable confirmation that the results are unbiased. These two checks do not measure temporality. When I assessed the direction of the wind by reading the dealers' reports, I specifically made sure that news affecting the currency occurred at the end of the day. For example, if the day started with positive news but ended with negative news, I would record it as a negative day, as this had the most impact on the closing exchange rate. The machine learning algorithm on the other hand only looks at the overall sentiment in the extracts. Equally, as I did brief assessors on Amazon Mechanical Turk to look for news at the end of the day, it is unclear whether all assessors understood this instruction well. However, despite the shortcomings of these tests, they both confirm that the choices made in my assessment are not arbitrary.

Amazon Mechanical Turk

MTurk is frequently used in research in psychology, marketing and experimental economics. For example, Ambuehl, Niederle, and Roth (2015) use MTurk to question participants' willingness to take part in a medical trial depending on the size of compensation. The quality of the results obtained is variable, but the advantage is that the workers are unbiased as they are only presented with the text from the dealers' reports to analyze and have no stake in the study.

I randomly selected 100 excerpts from the dealers' reports out of the 1,679 trading days I coded as good, neutral or bad. I then copied the text of these 100 dealers' reports into a document so that they are available in digital form. The respondents on Amazon Mechanical Turk are asked to perform what can be referred to as sentiment analysis. They are asked to assess whether the Bank of England dealer perceived market conditions as good, bad, or neutral for sterling. Each statement is reviewed by 10 different workers on Amazon Mechanical Turk. The answers take the form of a dummy variable taking value 1 for good news, 0 for no news and -1 for bad news, just as for my assessment. I then take the mode (most frequent answer) of these 10 observations and compare it with my answer. Using the mode controls for the variability in the answers of different respondents and weeds out lower quality responses while using the consensus.[10] On average, each respondent spent 50 seconds per abstract and was given up to 2 minutes to respond. Extracts in Table 6 compares my assessment with that of the 10 reviewers for the 20 first statements.

Table 3 below measures the agreement on the randomly selected sample. Just by chance, agreeing with one of the three choices (1, 0 or -1) should be 33%. Agreement rates of 77% for positive assessments and 57% for negative assessment are unlikely to be random, whereas the agreement rates for neutral situation are not clearly better than random. While these results do not categorically attest to the objectivity of the analysis, they still show significant overlap for both my positive and negative assessments and those of MTurk.

	Positive assessment	Neutral assessment	Negative assessment	Total
My assessment	31	41	28	100
Most common answer by 10 MTurk reviewers (mode)	51	26	23	100
Agreement rate	77%	37%	57%	
N = 100 text samples				

Table 3 – comparing answers by MTurk and the author.

[10] Taking the average and rounding it up leads to similar answers but is less precise, as it includes responses from respondents who might not have read the question.

Natural Language Processing algorithm

As a second form of robustness check, I use sentiment analysis done by an algorithm. Natural Language Processing (NLP) is a set of techniques that use computational power to analyze large datasets of natural language. The field recently blossomed with advances in machine learning, allowing for a much better understanding of human language. I use a Python script named natural language toolkit (or nltk in short) set up by Bird, Klein, and Loper (2009). This algorithm is widely used. Each of the 100 digitized statements presented above are analyzed with the algorithm. Unlike my assessment or the one done by MTurk assessors, the algorithm does not provide a dummy, but a continuous score from -1 (negative sentiment) to 1 (positive sentiment).

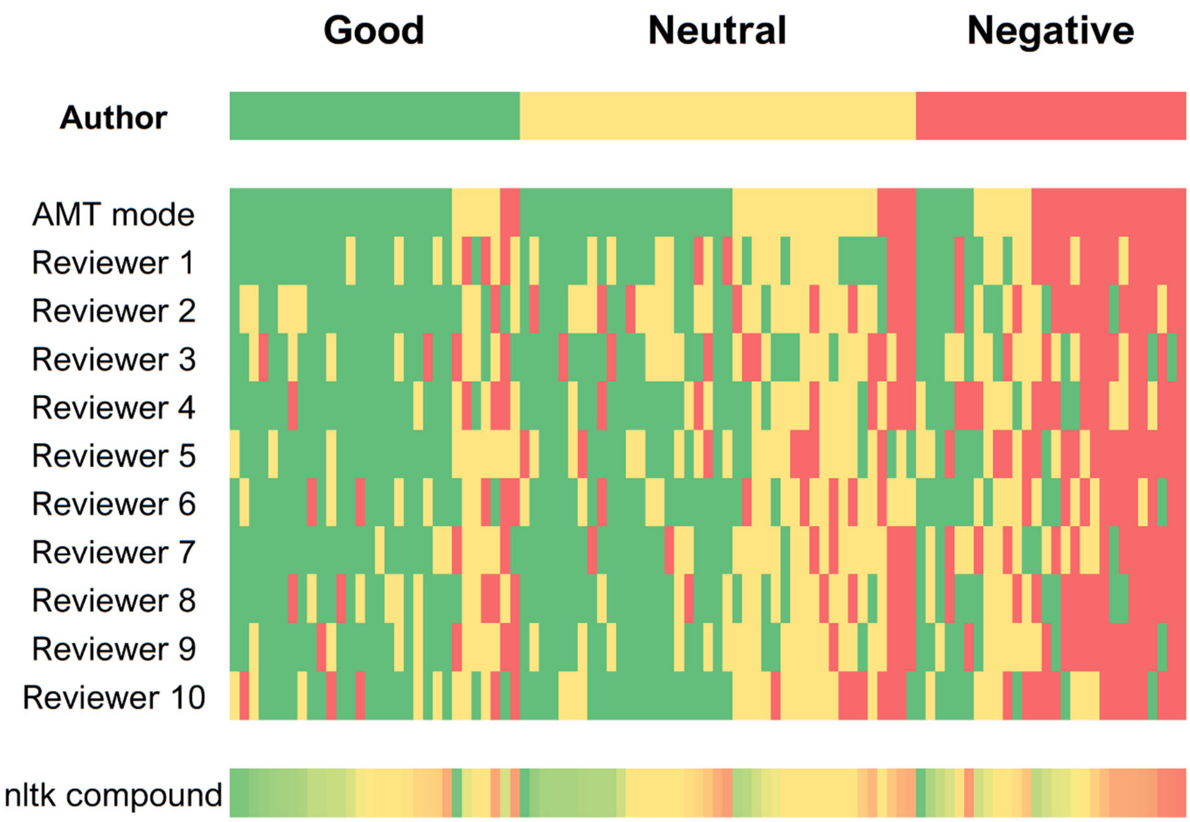


Figure 3 – Heat map of answers by the author, MTurk reviewers and the nltk algorithm. The scale has three colors using green for positive, yellow for neutral and red for negative. See text for how the answers were collected.

Figure 3 above shows a heatmap of the different answers and Table 5 below compares my assessment with the one made by both nltk and MTurk. All the reports I assessed as negative are also assessed as negative by both nltk and MTurk on average (as the negative coefficient shows). The positive assessments by both other techniques also yield qualitatively similar results. When it comes to neutral assessment, it seems that my assessment was more negative than both other assessment methods (as all my neutral assessments where more often assessed as positive by the two other methodologies). Note that the fact that there is disagreement between manual analysis and NLP measures is not a surprise, and has been documented in the literature (Jongeling et al. 2017).

	Average score – nltk algorithm	Average score – Amazon Mechanical Turk mode	Average author score
Assessed as bad by the author	-0.08	-0.36	-1
Assessed as neutral by the author	0.12	0.44	0
Assessed as good by the author	0.23	0.70	1

Table 4 – Correlation matrix of answers of the author, MTurk and the nltk algorithm. The color coding is from green (good news) to red (bad news).

What accounts for differences in assessment? In a few cases, it was a clear mistake on my side, where I either misjudged or mistyped the assessment. But the large disagreement over neutral days is justified as the following examples show. “Sterling was quiet and sluggish after some light, technical selling at the opening”. I rated this statement neutrally; the modal MTurk response was -1; and the nltk algorithm scored it -0.40. I gave this statement a neutral score because there seemed to have been little market activity - as suggested by the word “quiet” (a word often used by the dealers). The nltk algorithm, however, saw the statement as negative, potentially picking up on negative keywords like “sluggish”. MTurk responses were surprisingly homogenous, with 8 out of 10 saying that the statement was negative, and only 2 labelling it as neutral.

A statement I deemed to be neutral but the other systems deemed positive reads: “Sterling remained quietly on the sidelines and gained ground in effective terms despite a further erosion in oil prices." Here the MTurk consensus was 1 and the nltk algorithm granted a 0.69 score. Here again my justification for the 0 rating was that the market was mainly quiet, meaning that any news or action by the central bank would be likely to move the exchange rate, unlike if there was clear market activity due to specific news moving the price. These examples show that the assessment retains a certain amount of subjective judgement. However, unlike other narrative approaches that rely on the reader trusting the assessor, here I have endeavored to benchmark and cross-check my own judgement against assessments gleaned from two very different approaches. Table 5 shows that, on average my assessment was confirmed by both the algorithm and the external assessors.

5. What makes the Bank of England intervene?

After understanding how successful the Bank was at intervening, I ask what made the Bank intervene. To do that I rely on both lasso and a VAR analysis using a vast dataset collected from the archives of the Bank of England and other European central banks.[11]

Because there are many explanatory variables and to avoid overfitting, I rely on a lasso model. Lasso is stands for least absolute shrinkage and selection operator. It allows shrinking the number of explanatory variables to avoid overfitting. In a recent horse race among different machine learning algorithms, Pellet and Sciacovelli (2022) find that lasso performs best at predicting historical data, better than random forests for example.

Lasso shrinks the size of the coefficients of the independent variables depending on their predictive power. It also allows to deal with a large number of explanatory variables and pick the most relevant. As our model here has 55 explanatory variables, using lasso allows us to only pick the ones that are relevant and avoid overfitting of the ones selected. To make the variables comparable, I run a Z score normalization on most variables.[12] This allows for all variables to be comparable (but comes at the price of somewhat complicating interpretation of the coefficients).

The model has two parts. First a linear regression followed by a penalty term. The penalty term is run through several iterations, shrinking the size of some coefficients depending on their

[11] For a similar exercice but on the 1950s and 1960s, see Naef (2020).

[12] This is done by subtracting the mean and dividing by the standard deviation.

$$\overbrace{\hphantom{xxxxxxxxxxxxxxxxxxxx}}^{\text{linear model}} \quad \overbrace{\hphantom{xxxxxxxxx}}^{\text{penalty term}}$$

predictive power. If the shrinking of the coefficient reaches 0, the independent variables are removed from the model. The lasso model used is as follows:

$$J = \frac{1}{2m}\sum_{i=1}^{m}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right) + \lambda\sum_{j-1}^{p}|\beta_j|$$

where y_i is the dependent variable, x_{ij} are the independent variables, β_j are the coefficients, p the number of independent variables and m is the number of data points. The coefficient β_j is then obtained by minimizing J on the left-hand side of the model. The penalty term adds a λ adjustable parameter (set to 0.0001 in the model here).

The independent variables x_{ij} reflect many potential economic indicators that could affect the central bank decisions to intervene. To control for the movement of other market players, I also use new data on the market intervention of other European countries. The data has recently been digitized by Eichengreen and Naef (2022) and is available at a daily frequency. The idea is that if another country is also heavily intervening on currency markets, it could affect the exchange rate of the pound. I use the aggregate of all the interventions in all currencies by other European countries as a variable.

Additionally, I use a broad set of controls. As lasso selects the relevant variables, adding controls will unlikely bias the results, as irrelevant controls will not be used in the chosen specification. Table 7 in the appendix lists all the variables used.

I also decide to include in the list of variables lasso cannot exclude the following: the GBP-USD exchange rate, the USD-DEM rate and the EMP exchange market pressure index for all Exchange Rate Mechanism (ERM) countries.[13] These variables are all important for Bank of England intervention and should be in the model.[14]

Headline results are shown in Table 5 below, and Table 7 in the appendix give the full picture with all 55 variables, ranked by significance. Positive coefficients in the figure mean reserve accumulation. Negative coefficients mean reserves sales, to defend the pound. They are expressed

[13] The ERM was a European-wide system of exchange rate pegs in place at the time.
[14] On why the DEM-USD rate matters, see Eichengreen and Naef (2022)

in million US dollars, though the right-hand side variables are normalized and bear no meaning that can be expressed in units.

Table 5 – lasso regression, selected variables

Dependent variable: interventions by the Bank of England	
Normalised difference of the GBP-DEM exchange rate	26.13 (5.15)***
Normalised difference of the GBP-USD exchange rate	-0.63 (6.57)
Normalised Exchange Market Pressure Index for European countries	-495.93 (28.7)***
UK Bank Rate	-47.28 (6.87)***
Adjusted R-squared	0.33
N	1277
Number of variables inputted	55
Number of variables kept	44

Note: detail of all variables presented in Table 7 in the appendix. As variables are normalized, units are not directly relevant but comparable

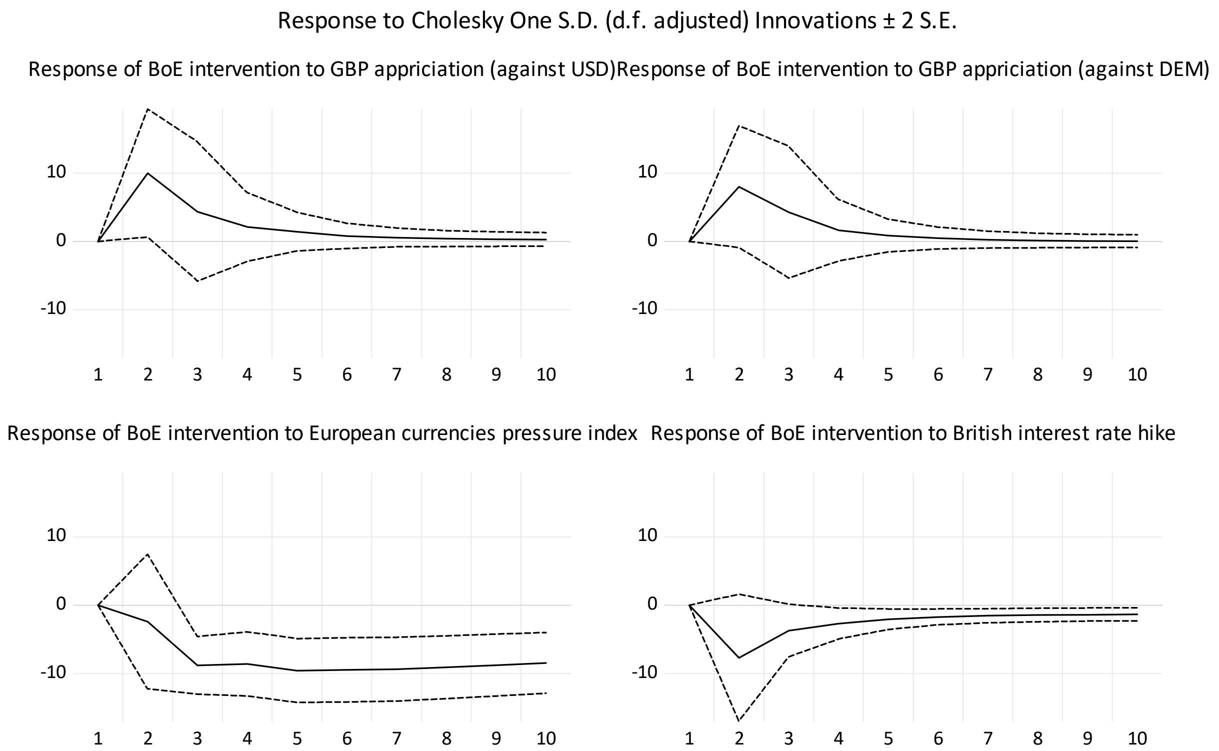
So, what makes the Bank of England intervene? The figures in table 5 tell us that when the GBP-DEM rate depreciates, the Bank of England intervenes to defend the pound (or accumulates reserves when the rate appreciates). The GBP-USD rate on the other hand is not significant at all. The Bank was exclusively targeting the Deutschmark rate, not the dollar. If we believe our reaction function, exchange rate pressure on all European countries also affects the pound. But pressure on single countries, for example Italy, does not make the Bank defend the pound (see positive coefficient in Table 7). It could be that when international investors have other fish to fry, the pound is left alone. Interest rate hikes by central banks in Europe also lead to more intervention (Table 7 in the appendix). This makes intuitive sense as higher interest rates abroad, all things equal, will lead to capital outflow, as capital chases higher returns.

What might appear more puzzling is that higher interest rates at home also lead to more intervention to defend the pound. Here a reverse causality issue is likely at play; the Bank in this period often raised rates on the same day as large intervention to defend the pound. For example, Black Wednesday is the day with the highest rate hikes but also highest interventions. More on this last element in the VAR section below.

VAR approach

Lasso provides a good overview of the drivers of intervention. Taking it a step further, I run a vector autoregression with the same question in mind: what makes the Bank of England intervene? The result of a VAR approach, using 2 lags, is presented below. I use the following variables: the GBP-USD exchange rate, the USD-DEM rate and the EMP exchange market pressure index for all Exchange Rate Mechanism (ERM) countries. These four variables are the same as used in the lasso approach above, minus all the control variables.

Note that when reading the graphs, positive shocks mean the Bank of England was able to accumulate reserves. Negative shocks mean the Bank of England had to sell reserves to defend the pound.



In a nutshell (and being generous when it comes to confidence intervals), an increase in the pound against the dollar or the deutschmark allows the Bank of England to accumulate reserves. So far this is hardly groundbreaking and was expected.

Here is the more interesting part (and this was also reflected in the lasso analysis). A hike in interest rates by the Bank of England often causes ("is accompanied by" is maybe more elegant here) the Bank having to defend the pound and sell its reserves. A better reading of this is that heavy currency crises often lead the Bank to using all its tools to defend the pound: foreign exchange interventions along with interest rate hikes making the pound more attractive internationally. The other interesting factor is looking at the average Exchange Market Pressure index (EMP) from most other European currencies.[15] Think of this as a measure that highlights a currency crisis in most European currencies. In such a situation, the pound is not unaffected. But, as we saw, when only one single currency has a crisis, the pound is unaffected (for this look at Table 7 in the appendix). A crisis in Italy or Norway is not a bad thing for the pound as it probably keeps speculators occupied elsewhere, but a European-wide crisis is bad news for the pound and forces the Bank of England to intervene.

6. Conclusion

This study offers a new way of assessing central bank intervention accounting for intraday shocks. This avoids overstating the impact of central bank intervention and mistaking intervention for markets simply picking up on news. Not controlling for the intraday conditions of the currency is problematic when assessing intervention success. Good market news (or even bad news that is better than expected) can lead a test to show intervention success when it is only changes in market conditions.

Presenting a novel dataset, this paper uses a narrative approach to tackle this issue. By reading the daily reports of policymakers at the time, I show how news affecting the exchange rate during the day can influence intervention outcomes. Far from the intervention success rates of 80% in certain studies, I show that when controlling for market conditions, success rates drop as low as 8%. In particular, I show that the Bank of England performs poorly when trying to

[15] Again, for construction see Eichengreen and Naef (2022).

make sterling appreciate in negative market conditions. And in a fixed exchange rate system, this is when it matters most, as speculators know the central bank has limited dollar reserves. Despite the low success rates, the Bank still intervened as it was acting on behalf of the Treasury and did not have independence in setting the exchange rate.

The paper also provides some interesting insight on why the Bank of England intervened. It focused exclusively on the Deutschmark and not the dollar. Interest rate hikes by the Bank were mostly accompanied by large reserves sales; that means that monetary policy was at the service of exchange rate management. And finally, crises elsewhere in Europe were relatively good news for the pound, unless they became Europe-wide, in which case the pound also suffered.

These findings are specific to their historical settings and are not directly comparable to intervention today. Yet, the approach presented here offers a word of caution for existing studies on foreign exchange intervention.

7. References

Adler, Gustavo, Kyun Suk Chang, Rui Mano, and Yuting Shao. 2021. 'Foreign Exchange Intervention: A Dataset of Public Data and Proxies'. 2021/047. *IMF Working Papers*. IMF Working Papers. International Monetary Fund. https://ideas.repec.org/p/imf/imfwpa/2021-047.html.

Adler, Gustavo, Noëmie Lisack, and Rui C. Mano. 2019. 'Unveiling the Effects of Foreign Exchange Intervention: A Panel Approach'. *Emerging Markets Review*, June, 100620. https://doi.org/10.1016/j.ememar.2019.100620.

Adler, Gustavo, and Rui C. Mano. 2021. 'The Cost of Foreign Exchange Intervention: Concepts and Measurement'. *Journal of Macroeconomics* 67 (March): 103045. https://doi.org/10.1016/j.jmacro.2018.07.001.

Ambuehl, Sandro, Muriel Niederle, and Alvin E. Roth. 2015. 'More Money, More Problems? Can High Pay Be Coercive and Repugnant?' *The American Economic Review* 105 (5): 357–60.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Clara Vega. 2003. 'Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange'. *American Economic Review* 93 (1): 38–62. https://doi.org/10.1257/000282803321455151.

Aslam, Faheem, Saqib Aziz, Duc Khuong Nguyen, Khurrum S. Mughal, and Maaz Khan. 2020. 'On the Efficiency of Foreign Exchange Markets in Times of the COVID-19 Pandemic'. *Technological Forecasting and Social Change* 161 (December): 120261. https://doi.org/10.1016/j.techfore.2020.120261.

Bird, Steven, Ewan Klein, and Edward Loper. 2009. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. 1 edition. Beijing ; Cambridge Mass.: O'Reilly Media.

Blanchard, Olivier, Gustavo Adler, and Irineu de Carvalho Filho. 2015. 'Can Foreign Exchange Intervention Stem Exchange Rate Pressures from Global Capital Flow Shocks?' *NBER Working Paper*, July. http://www.nber.org/papers/w21427.

Bordo, Michael D., Owen F. Humpage, and Anna J. Schwartz. 2015. *Strained Relations: US Foreign-Exchange Operations and Monetary Policy in the Twentieth Century*. Chicago: University of Chicago Press.

Burkhard, Lukas, and Andreas M. Fischer. 2009. 'Communicating Policy Options at the Zero Bound'. *Journal of International Money and Finance* 28 (5): 742–54. https://doi.org/10.1016/j.jimonfin.2008.06.006.

Cavallino, Paolo. 2019. 'Capital Flows and Foreign Exchange Intervention'. *American Economic Journal: Macroeconomics* 11 (2): 127–70. https://doi.org/10.1257/mac.20160065.

Chamon, Marcos d, David J. Hofman, Nicolas E. Magud, and Alejandro M. Werner. 2019. *Foreign Exchange Intervention in Inflation Targeters in Latin America*. *Foreign Exchange Intervention in Inflation Targeters in Latin America*. International Monetary

Fund. https://www.elibrary.imf.org/view/books/071/25455-9781484375686-en/25455-9781484375686-en-book.xml.

Chang, Yuanchen, and Stephen J. Taylor. 1998. 'Intraday Effects of Foreign Exchange Intervention by the Bank of Japan1We Thank Richard Baillie, JIMF-LIFE and the Sixth SFM Conference Participants for Their Helpful Comments. We Would Also like to Thank Olsen and Associates for Providing the Intraday Exchange Rates and Reuters' News Data.1'. *Journal of International Money and Finance* 17 (1): 191–210. https://doi.org/10.1016/S0261-5606(97)00056-9.

Dominguez, Kathryn, and Jeffrey A. Frankel. 1993. 'Does Foreign-Exchange Intervention Matter? The Portfolio Effect'. *The American Economic Review* 83 (5): 1356–69.

Dominguez, Kathryn M. 2003. 'The Market Microstructure of Central Bank Intervention'. *Journal of International Economics* 59 (1): 25–45. https://doi.org/10.1016/S0022-1996(02)00091-0.

Dominguez, Kathryn M. E., and Freyan Panthaki. 2006. 'What Defines "News" in Foreign Exchange Markets?' *Journal of International Money and Finance*, Foreign Exchange Markets, 25 (1): 168–98. https://doi.org/10.1016/j.jimonfin.2005.10.009.

———. 2007. 'The Influence of Actual and Unrequited Interventions'. *International Journal of Finance & Economics* 12 (2): 171–200. https://doi.org/10.1002/ijfe.326.

Echavarría, Juan J., Luis F. Melo-Velandia, and Mauricio Villamizar-Villegas. 2018. 'The Impact of Pre-Announced Day-to-Day Interventions on the Colombian Exchange Rate'. *Empirical Economics* 55 (3): 1319–36. https://doi.org/10.1007/s00181-017-1299-1.

Ehrmann, Michael, and Marcel Fratzscher. 2005. 'Exchange Rates and Fundamentals: New Evidence from Real-Time Data'. *Journal of International Money and Finance*, Exchange Rate Economics, 24 (2): 317–41. https://doi.org/10.1016/j.jimonfin.2004.12.010.

Eichengreen, Barry, and Alain Naef. 2022. 'Imported or Home Grown? The 1992–3 EMS Crisis'. *Journal of International Economics* 138 (September): 103654. https://doi.org/10.1016/j.jinteco.2022.103654.

Evans, Martin, and Richard Lyons. 2008. 'How Is Macro News Transmitted to Exchange Rates?' *Journal of Financial Economics* 88 (1): 26–50.

Fatum, Rasmus. 2015. 'Foreign Exchange Intervention When Interest Rates Are Zero: Does the Portfolio Balance Channel Matter after All?' *Journal of International Money and Finance* 57 (October): 185–99. https://doi.org/10.1016/j.jimonfin.2015.07.015.

Fatum, Rasmus, and Michael Hutchison. 1999. 'Is Intervention a Signal of Future Monetary Policy? Evidence from the Federal Funds Futures Market'. *Journal of Money, Credit and Banking* 31 (1): 54–69. https://doi.org/10.2307/2601139.

Fischer, Andreas M. 2006. 'On the Inadequacy of Newswire Reports for Empirical Research on Foreign Exchange Interventions'. *Journal of International Money and Finance* 25 (8): 1226–40. https://doi.org/10.1016/j.jimonfin.2006.09.006.

Fratzscher, Marcel, Oliver Gloede, Lukas Menkhoff, Lucio Sarno, and Tobias Stöhr. 2019. 'When Is Foreign Exchange Intervention Effective? Evidence from 33 Countries'. *American Economic Journal: Macroeconomics*. https://doi.org/10.1257/mac.20150317.

Gabaix, Xavier, and Matteo Maggiori. 2015. 'International Liquidity and Exchange Rate Dynamics'. *The Quarterly Journal of Economics* 130 (3): 1369–1420. https://doi.org/10.1093/qje/qjv016.

Hofmann, Boris, Hyun Song Shin, and Mauricio Villamizar-Villegas. 2019. 'FX Intervention and Domestic Credit: Evidence from High-Frequency Micro Data', March. https://papers.ssrn.com/abstract=3355009.

Howson, Susan. 1980. *Sterling's Managed Float: The Operations of the Exchange Equalisation Account, 1932-39*. International Finance Section, Department of Economics, Princeton University. https://www.princeton.edu/~ies/IES_Studies/S46.pdf.

Hu, May, Yunfeng Li, Jingjing Yang, and Chi-Chur Chao. 2016. 'Actual Intervention and Verbal Intervention in the Chinese RMB Exchange Rate'. *International Review of Economics & Finance* 43 (May): 499–508. https://doi.org/10.1016/j.iref.2016.01.011.

Ilzetzki, Ethan, Carmen M. Reinhart, and Kenneth S. Rogoff. 2019. 'Exchange Arrangements Entering the Twenty-First Century: Which Anchor Will Hold?' *The Quarterly Journal of Economics* 134 (2): 599–646.

Jongeling, Robbert, Proshanta Sarkar, Subhajit Datta, and Alexander Serebrenik. 2017. 'On Negative Results When Using Sentiment Analysis Tools for Software Engineering Research'. *Empirical Software Engineering* 22 (5): 2543–84. https://doi.org/10.1007/s10664-016-9493-x.

Kearns, Jonathan, and Roberto Rigobon. 2005. 'Identifying the Efficacy of Central Bank Interventions: Evidence from Australia and Japan'. *Journal of International Economics* 66 (1): 31–48. https://doi.org/10.1016/j.jinteco.2004.05.001.

Klein, Michael W. 1993. 'The Accuracy of Reports of Foreign Exchange Intervention'. *Journal of International Money and Finance* 12 (6): 644–53. https://doi.org/10.1016/0261-5606(93)90030-F.

Krippendorff, Klaus. 2018. *Content Analysis: An Introduction to Its Methodology*. Fourth edition. Los Angeles: SAGE Publications, Inc.

Menkhoff, Lukas. 2010. 'High-Frequency Analysis of Foreign Exchange Interventions: What Do We Learn?' *Journal of Economic Surveys* 24 (1): 85–112. https://doi.org/10.1111/j.1467-6419.2009.00582.x.

Menkhoff, Lukas, Malte Rieth, and Tobias Stöhr. 2021. 'The Dynamic Impact of FX Interventions on Financial Markets'. *The Review of Economics and Statistics* 103 (5): 939–53. https://doi.org/10.1162/rest_a_00928.

Mohanty, Madhusudan S., and Bat-el Berger. 2013. 'Central Bank Views on Foreign Exchange Intervention' No. 73 (October).

Monnet, Eric. 2014. 'Monetary Policy without Interest Rates: Evidence from France's Golden Age (1948 to 1973) Using a Narrative Approach'. *American Economic Journal: Macroeconomics* 6 (4): 137–69. https://doi.org/10.1257/mac.6.4.137.

Naef, Alain. 2019. *Sterling and the Stability of the International Monetary System, 1944-1971*. Cambridge: University of Cambridge. https://doi.org/10.17863/CAM.32540.

Naef, Alain. 2022. *An Exchange Rate History of the United Kingdom: 1945–1992*. Studies in Macroeconomic History. Cambridge: Cambridge University Press. https://www.cambridge.org/core/books/an-exchange-rate-history-of-the-united-kingdom/68B7E57D9884394B815C76D48ACD3FB6.

Naef, Alain. 'Dirty float or clean intervention? The Bank of England in the foreign exchange market' *European Review of Economic History*, Volume 25, Issue 1, February 2021, Pages 180–201, https://doi.org/10.1093/ereh/heaa011

Naef, Alain, and Jacob Weber. 2021. 'How Powerful Is Unannounced, Sterilized Foreign Exchange Intervention?' *SocArXiv*. https://doi.org/10.31235/osf.io/bfehz.

Neely, Christopher. 2005. 'An Analysis of Recent Studies of the Effect of Foreign Exchange Intervention'. Federal Reserve Bank of St. Louis Working Paper No. 2005-030B.

Neely, Christopher J. 2008. 'Central Bank Authorities' Beliefs about Foreign Exchange Intervention'. *Journal of International Money and Finance* 27 (1): 1–25. https://doi.org/10.1016/j.jimonfin.2007.04.012.

Neuendorf, Kimberly A. 2016. *The Content Analysis Guidebook*. Second edition. Los Angeles: SAGE Publications, Inc.

Pellet, Thomas, and Giovanni Sciacovelli. 2022. 'Machine Learning about the Latin Monetary Union'. *BNB Conference Proceedings*.

Reitz, Stefan, and Mark P. Taylor. 2008. 'The Coordination Channel of Foreign Exchange Intervention: A Nonlinear Microstructural Analysis'. *European Economic Review* 52 (1): 55–76. https://doi.org/10.1016/j.euroecorev.2007.06.023.

Romer, Christina D., and David H. Romer. 1989. 'Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz'. In *NBER Macroeconomics Annual*, by Olivier Blanchard and Stanley Fischer, 4:121–70. Cambridge: MIT Press. https://www.journals.uchicago.edu/doi/abs/10.1086/654103.

———. 1994. 'Monetary Policy Matters'. *Journal of Monetary Economics* 34 (1): 75–88. https://doi.org/10.1016/0304-3932(94)01150-8.

———. 2014. 'The Incentive Effects of Marginal Tax Rates: Evidence from the Interwar Era'. *American Economic Journal: Economic Policy* 6 (3): 242–81. https://doi.org/10.1257/pol.6.3.242.

Sarno, Lucio, and Mark P. Taylor. 2001. 'Official Intervention in the Foreign Exchange Market: Is It Effective and, If so, How Does It Work?' *Journal of Economic Literature* 39 (3): 839–68. https://doi.org/10.1257/jel.39.3.839.

Taylor, Alan M. 2010. ‘Global Finance after the Crisis’. *Bank of England Quarterly Bulletin* 50 (4): 366–77.

8. Appendix (for online publication)

Historical background

The postwar history of the pound can be separated into three clear phases when it comes to Bank of England operations on the foreign exchange market. From 1985 to 1987, British policymakers mainly managed the pound against the dollar; from 1987 to 1992, the deutschmark was the reference currency and from 1992 to today, the pound was left to float freely.

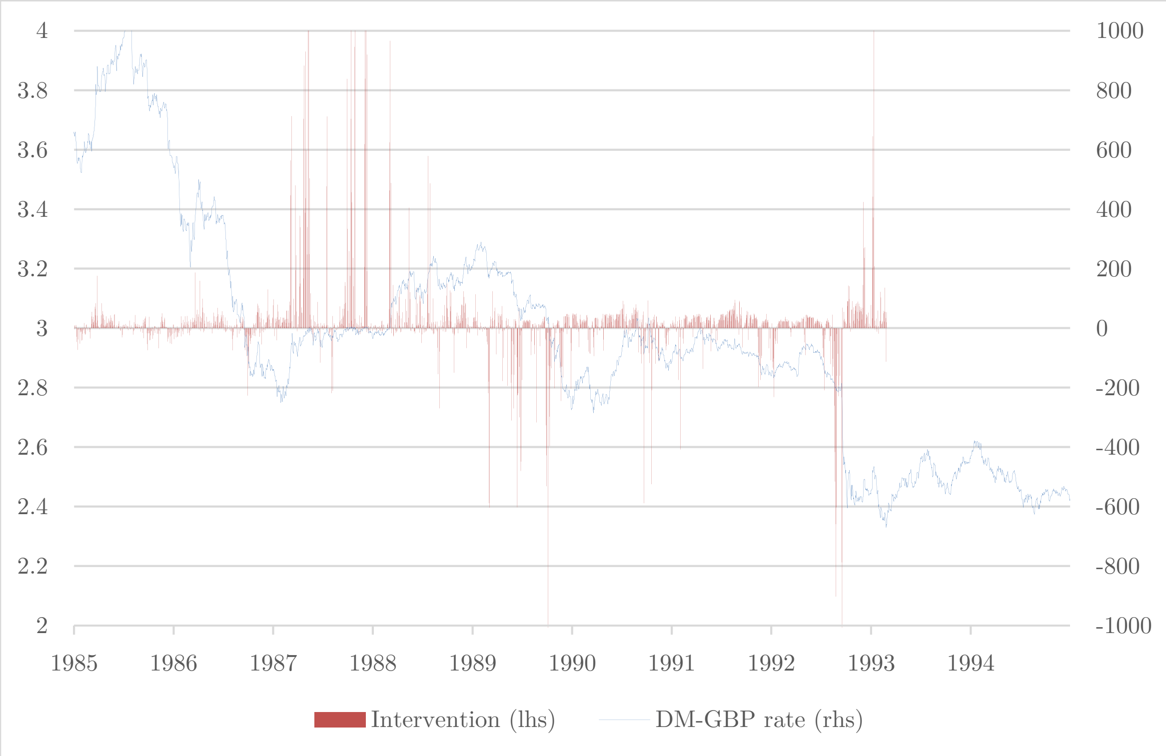
Main events

22 September 1985	Plaza Accord: coordinated interventions to depreciate the dollar.
22 February 1987	Louvre Accord depreciates the dollar; the UK starts shadowing the deutschmark.
Early 1988	End of official deutschmark shadowing, but the deutschmark remains the main focus.
1 October 1990	Britain joins the European Exchange Rate Mechanism (ERM) with a band of +/- 6% with the ECU (and *de facto* with the deutschmark).
16 September 1992	Black Wednesday: the UK leaves the ERM, floating the pound.

The short timeline above outlines the history of these exchange rate systems from 1952 to the present day. Most of the time, the pound was either in a fixed exchange rate system or in a managed float. Only in 1992 was the currency left to float freely.

Characteristics of Bank of England intervention

How did the Bank of England intervene? In the sample presented, interventions by the Bank of England were frequent and secret. Only 66 of the 8,429 interventions were publicly communicated, less than 1% of the sample. Secret intervention makes study of the Bank of England extremely relevant as most central banks today still intervene in secret as well. Surveying central bankers, Mohanty and Berger (2013) found that only 18% of central banks frequently communicated their intervention practices. This means that most central banks intervene in secret, despite the literature arguing that communicating intervention is more effective (Burkhard and Fischer 2009). Overall, the Bank was in the market 89.96% of the trading days from 1987 to 1992.



Sterilized intervention by design

An institutional feature of the Bank of England ensured that all interventions were automatically sterilized. The Bank of England operation were all done through the EEA (Exchange Equalisation Account), which was independent from the Bank of England and belonged to the Treasury. Howson (1980) showed that the institutional structure of the EEA meant that all intervention operations had a counterparty in Treasury bills. When the Bank of England was selling dollars against pounds, it would reinvest the newly acquired pounds into British Treasury bills and doing so, sterilizing the operation. Conversely, when the Bank wanted to buy US dollars, it first had to sell Treasuries at the EEA to obtain sterling to purchase dollars. This meant that any operation was automatically sterilized as a feature of the EEA.

Goals of Bank of England interventions

When assessing intervention, it is essential to understand what the central bank was trying to achieve. Today, as during the 2009 crisis, central banks mainly want to reduce exchange rate volatility (Fratzscher et al. 2019; Mohanty and Berger 2013; Blanchard, Adler, and Filho 2015). However, during the period observed, the goal of the Bank of England was different. Interviews with policymakers of the time and archival records show that policymakers wanted to influence the exchange rate in one direction or the other.

Although objectives change and are not set in stone, historical analysis shows clear patterns in the goals of the Bank of England. The Bank intervened either to make the exchange rate appreciate or depreciate. Below, I present several reports written by the very people intervening: Bank of England dealers. By analyzing their own assessment of interventions, the underlying goals of intervention become clear.

On April 7, 1988, as sterling was appreciating against the deutschmark, the dealers' reports read: "Sterling was mostly steady, but dipped this afternoon following a well-publicised round of co-ordinated sales by ourselves and the Bundesbank". Here the goal was to make the appreciating currency depreciate. The operation, according to the Bank, seemed successful.

On June 23, 1989, Bank of England dealers commented that "Sterling's early weakness was met by a round of well-publicised official intervention, after which the pound drifted quietly into the weekend". The goal of the intervention was to counter sterling's weakness and the impact (as assessed by the Bank itself) was visible over the weekend.

On September 16, 1992 (a day before Black Wednesday), the reports read (emphasis added): "Several rounds of overt intervention only had momentary **success**: selling pressure at the margin increased as the Bank's early morning money market round passed without a move on interest rates." Here success is defined as increasing the sterling-deutschmark exchange rate, which was falling against the backdrop of a growing crisis in the ERM.

These examples show the Bank trying to move the exchange rate, up or down. The goal was not simply, as it is today, to reduce volatility, but to *push* the rate in a given direction. This is an important point to bear in mind when assessing what counts as success.

Extract of Amazon Mechanical Turk assessment

Randomly selected date	Text	Most common answer by 10 reviewers (mode)	My assessment
22/01/1987	Sterling remained on the sidelines	0	0
12/2/1987	Sterling remained quietly on the sidelines and gained ground in effective terms despite a further erosion in oil prices.	1	0
26/02/1987	Sterling steadied as the oil price climbed back above $16 per barrel.	1	1
18/03/1987	Sterling encountered steady demand throughout the day reflecting the favourable response to the budget and the hope that the 1/2% cut in Base rates might leave scope for another reduction soon.	1	1
16/04/1987	Sterling was helped by the stronger dollar and opened firmer in effective and cross rate terms, but was little changed during the day.	0	0
13/05/1987	Sterling was on the sidelines, but was pulled up against third currencies by the stronger dollar.	1	0
29/05/1987	Sterling was on the sidelines but encountered some data commercial demand and touched DM2.97 1/8 at 5 o'clock.	0	0
16/06/1987	Sterling rallied on the better than expected PSBR data (negative borrowing of £374 mn against an expected requirement of £800 mn), but eased against the firmer dollar this afternoon.	1	1
9/7/1987	Sterling was steady in quiet conditions.	1	0
28/07/1987	Sterling was also quiet, but benefited from the encouraging CBI survey.	1	1
25/08/1987	Sterling weakened generally today as the market focused on the recent falls in oil prices and concerns grew about next week's trade figures. Outward investment flows also contributed to the fall but the market was throughout very orderly.	-1	-1
23/09/1987	Sterling firmed against the easier dollar but met no significant upward pressure despite the publication of a bullish CBI survey.	0	0
25/09/1987	Sterling was mostly on the sidelines but enjoyed underlying support as a result of the wider interest differentials against European currencies.	1	1
24/02/1988	Sterling remained quiet but with a firm undertone.	0	0
17/03/1988	Sterling opened on a firm note in the absence of expected official sales at DM3.10 but fell following the Bank's signal of a 1/2% cut in Base Rates. However, good underlying demand led to a partial recovery, and this afternoon the pound regained further ground helped by the firmer dollar.	1	1

26/04/1988	Sterling opened softer after easing in New York last night, but recovered on Middle East demand this morning.	1	0
10/5/1988	Sterling was actively traded in a good two-way market with profit-taking after Yesterday's rise balanced by renewed demand above DM3.15 3/4.	1	0

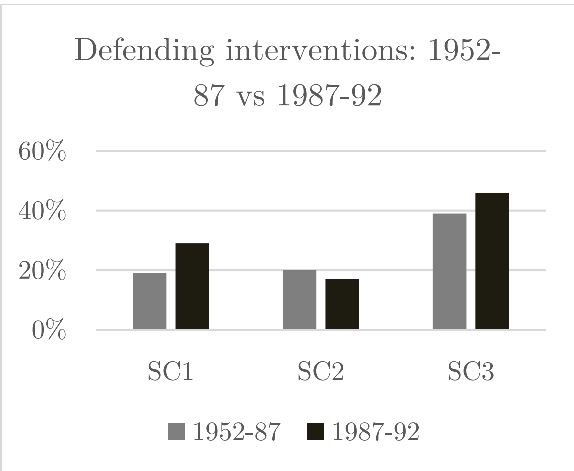
Table 6 – First 20 extract randomly selected from the dealers' reports. The two columns on the left show first the mode of the Amazon Mechanical Turk assessment (-1 being negative, 0 neutral and 1 positive) and then my assessment of the extract.

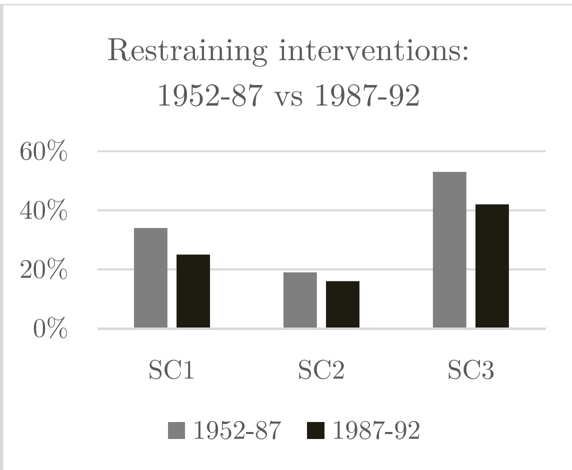
Expanding success count methodology to 1952-1987

To put the results in perspective, the table below shows the success rates for all interventions from 1952 to 1987.

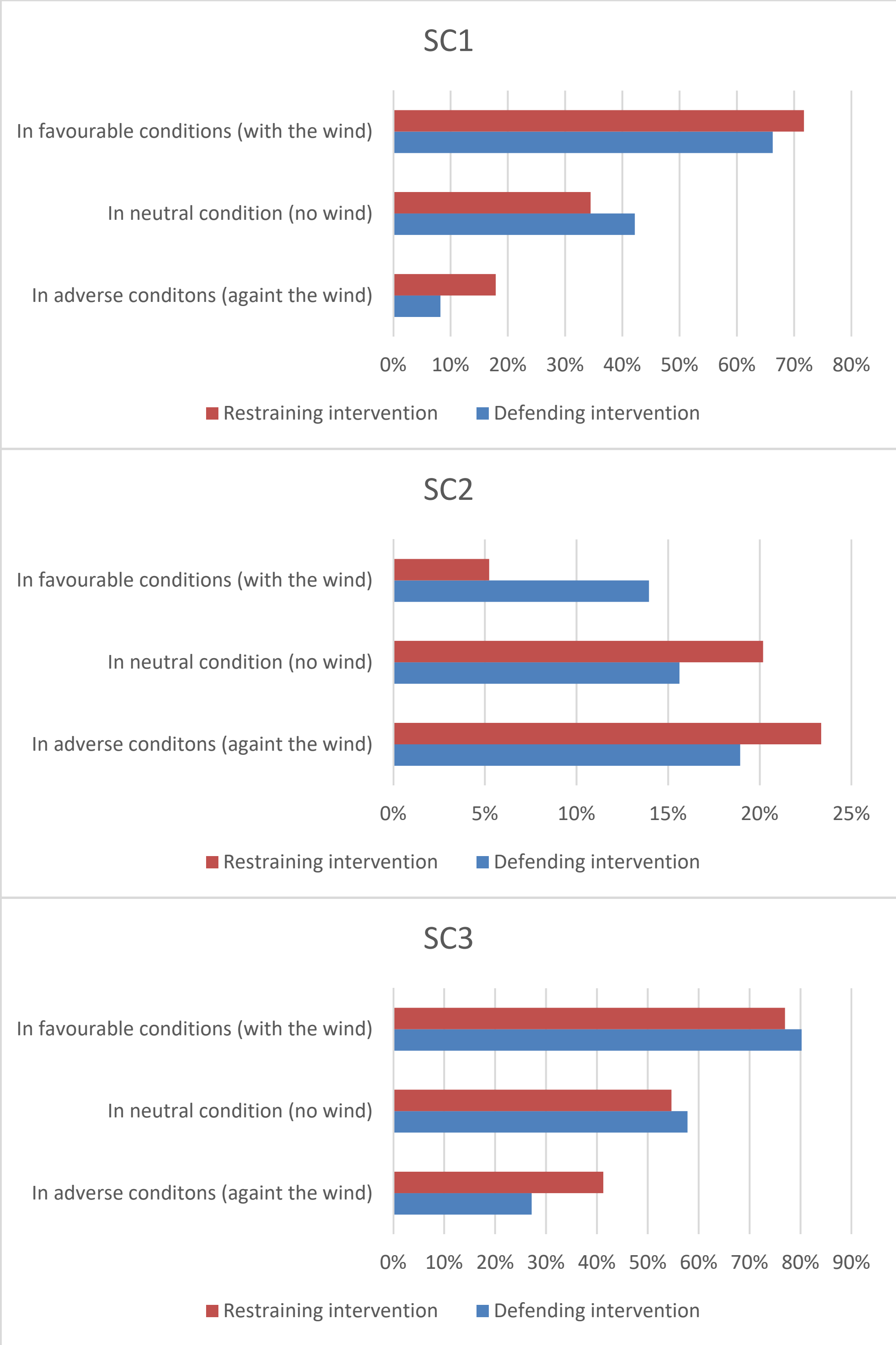
	Day count	Success count	Percentage of successful intervention	Success count	Percentage of successful intervention	Success count	Percentage of successful intervention
Dollar intervention (1952-1987)							
Defending interventions	2298	434	19%	465	20%	899	39%
Restraining interventions	4817	1211	25%	794	16%	2005	42%

When comparing these results with the ones in Table 2, we can see that overall, interventions were more successful in the sample 1952-87 than in the later sample (1987-92). It is likely that a larger size of the foreign exchange market in the UK meant that the Bank of England had more firing power. For more on market size, see Naef (2019). Overall the results are in the same range however.





Breaking down success criteria



Lasso variable selection

Table 7 – full lasso regression including all accepted independent variable

Dependent variable: interventions by the Bank of England	
A_DELTA_FX2_GBP_DEM_NORM	26.13 (5.15)***
EMP_ALL_NORM	-495.93 (28.7)***
EMP_DENMARK_NORM	57.85 (9.59)***
EMP_FINLAND_NORM	107.75 (10.35)***
EMP_NORWAY_NORM	32.85 (6.95)***
EMP_PORTUGAL_NORM	320.28 (53.71)***
INT_PORTUGAL_Z_TOTAL	3.26 (0.72)***
RATE_SWEDEN_DAILY	-9.92 (1.64)***
RATE_UK_DAILY	-47.28 (6.87)***
EMP_BELGIUM_NORM	106.39 (28.85)***
EMP_SPAIN_NORM	148.09 (40.97)***
EMP_IRELAND_NORM	509.37 (170.09)***
RATE_IRELAND_DAILY	189.88 (64.98)***
INT_IRELAND_Z_TOTAL	6.72 (2.62)**
RATE_GERMANY_DAILY	-139.71 (54.45)**
RATE_NETHERLANDS_DAILY	-75.61 (33)**
M_JAN	-39.97 (17.58)**
INT_FRANCE_Z_TOTAL	-0.09 (0.05)*
RATE_NORWAY_DAILY	20.17 (10.99)*
M_FEB	-32.3 (18.2)*
RATE_SPAIN_DAILY	24.9 (14.05)*
EMP_ITALY_NORM	77.81 (45.84)*
RATE_DENMARK_DAILY	-34.71 (21.34)
INT_BELGIUM_Z_TOTAL	0.29 (0.18)
M_AUG	-24.38 (15.72)
M_NOV	-24.45 (17.17)
M_JUN	-22.59 (16.01)
D_FRIDAY	17.99 (12.77)
C	-764.58 (560.45)
RATE_PORTUGAL_DAILY	32.78 (25.15)
M_OCT	-16.54 (17.05)
INT_SPAIN_Z_TOTAL	0.14 (0.17)
DELTA_FX_DM_USD_NORM	-5.61 (6.59)
M_APR	12.71 (15.99)
INT_GERMANY_Z_TOTAL	0.02 (0.03)
D_WEDNESDAY	-9.13 (12.71)
EMP_FRANCE_NORM	16.61 (23.6)
D_TUESDAY	-7.82 (12.73)
M_MAY	5.15 (16.61)
NLTK_NEGATIVE	12.12 (49.53)
INT_ITALY_Z_TOTAL	-0.04 (0.17)
D_THURSDAY	-1.94 (12.86)
A_DELTA_FX2_GBP_USD_NORM	-0.63 (6.57)
INT_FINLAND_Z_TOTAL	0 (0.05)

Table 7 (continued)

Adjusted R-squared	0.33
N	1277
Number of variables inputted	55
Number of variables kept	44

Note: independent variables in italics were included as mandatory variables and not selected by the algorithm. All other variables have been selected by the lasso algorithm. The algorithm compares 100 models to optimize fitting and excluding overfitted variables.

As many of the coefficients in the lasso regression are normalized, most units do not have meaning in this table. Interest rate and intervention variables were not normalized to avoid changing the zero values and as they were already within a restricted range.

Lasso excluded the following 11 variables in the regression: Central bank intervention for Denmark, the Netherlands, Norway. Some time fixed effects (months of July, March and September). The interest rates for Belgium, Finland, France and Italy. As well as the nltk (natural langue toolkit) measure for positive news in the dealers reports (negative news on the other hand remained).

Robustness – Comparison with placebo

To test if the interventions of the Bank of England are better than random, I compared them with placebo interventions. Because over the sample analyzed the Bank was on the market 79.5% of the trading days, days with no intervention do not offer a good comparison as they would present specific characteristics. I therefore compare interventions from 1986 to 1992 to a placebo from 1992-1999. These two periods are similar. In the sample period from 1986-92 the Bank was targeting the Deutschmark exchange rate and from 1992-99 the Deutschmark was also the reference currency for the pound (before the introduction of the euro in 1999).

I run two placebo tests. The first is comparing the actual intervention period (1986-99) with the period from 1992-99 where there are no more interventions (remember the Bank of England stopped intervening after Black Wednesday in September 1992). The second placebo does a similar comparison but takes into account the intraday direction or wind of the market. The first test finds that overall the Bank of England was not better than random at moving the direction of the exchange rate.

Both test use the same success criteria presented in the methodology. Since there are no actual interventions occurring during this period, the test mimics what the Bank would have done. For example, if the exchange rate is dropping from day t-2 to day t-1, the placebo assumes that the Bank of England would have intervened at this time to make the exchange rate appreciate again. In this sense, this methodology counts how often reversal of the exchange rate occurred absent Bank of England operations. This is then compared to the success of the actual operations of the Bank of England. To analyze if the placebo is different from the actual interventions, I calculate if the placebo lies two standard deviations above or below the observed value. The standard deviation is calculated using a hypergeometric distribution presented below (this is common in the literature, see for example Bordo, Humpage, and Schwartz 2015). The three equations below give the formalization of success criteria and match the success criteria presented on page 17.

$$SC_{1\,placebo} = \begin{cases} 1 & \begin{cases} if\ S_t > S_{t-1}\ and\ S_{t-1} < S_{t-2}, \\ \quad or \\ if\ S_t < S_{t-1}\ and\ S_{t-1} > S_{t-2} \end{cases} \\ 0 & otherwise \end{cases}$$

$$SC_{2\,placebo} = \begin{cases} 1 & \begin{cases} if\ S_{t-1} > S_{t-2}\ and\ \Delta S_t < \Delta S_{t-1}\ and\ \Delta S_t > 0 \\ \quad or \\ if\ S_{t-1} < S_{t-2}\ and\ \Delta S_t > \Delta S_{t-1}\ and\ \Delta S_t < 0 \end{cases} \\ 0 & otherwise \end{cases}$$

$$SC_{3\,placebo} = \begin{cases} 1 & SC_1 = 1 \; or \; SC_2 = 1 \\ 0 & otherwise \end{cases}$$

$SC_{1\ placebo}$ covers cases when the exchange rate was going down (from day t-2 to day t-1) and where it reversed on day t. $SC_{2\ placebo}$ covers cases when the exchange rate depreciated from day t-2 to day d-1 and then depreciated less from day t-1 to day t. These values are then compared to the total number of days the exchange depreciated from day t-2 to day d-1, giving a success rate in percentage (see 6th column Table 8 and Table 9). $SC_{3\ placebo}$ is a combination of the first two criteria. And the same logic applies when the exchange rate was appreciating from day t-2 to day d-1.

The assumption behind the placebo is that the central bank will only try to make the currency appreciate if it was depreciating (and depreciating if it was appreciating). This is in line with the evidence presented in the section "Characteristics of Bank of England intervention". Table 8 below presents the result for the whole sample and Table 9 adds the wind dummy created using the narrative approach.

Table 8

Sample (1986-1992) vs Placebo (1992-1999) success count

	Intervention episodes	Intervention success		Placebo episodes	Placebo success		Expected success	Standard deviation	Random range	Is the central bank better than random?
		#	%		#	%				
DEFENDING OPERATIONS										
Success Criteria 1	427	102	24%	848	455	54%	229	3	222 - 236	NO
Success Criteria 2	427	61	14%	848	162	19%	82	3	76 - 87	NO
Success Criteria 3	427	163	38%	848	617	73%	311	4	302 - 319	NO
RESTRAINING OPERATIONS										
Success Criteria 1	1196	327	27%	1000	451	45%	539	8	523 - 556	NO
Success Criteria 2	1196	179	15%	1000	239	24%	286	6	273 - 298	NO
Success Criteria 3	1196	506	42%	1000	690	69%	825	10	805 - 845	NO
Observations sample:	1453									
Observations placebo:	1902									
Exchange rate days in both samples	3355									

The first test shows that overall, the interventions of the Bank of England did not influence the exchange rate differently from the placebo (Table 8). When distinguishing interventions with and against the wind, I find that defending interventions against the wind (the main mission of the Bank of England during this period), did not affect the market more than the placebo. Restraining interventions and intervention going with the wind or with the absence of wind show an effect significantly different from the placebo. This mirrors the overall findings of the paper which shows that the Bank of England performed particularly poorly when trying to avoid the exchange rate from falling when intervening against the wind (this is briefly understood by looking at Table 2 and Table 9).

Table 9

Sample (1986-1992) vs Placebo (1992-1999) success count including direction of the wind

	Intervention episodes	Intervention success		Placebo episodes	Placebo success		Expected success	Standard deviation	Random range	Is the central bank better than random?
		#	%		#	%				
DEFENDING OPERATIONS										
Against the wind										
Success Criteria 1	206	17	8%	101	10	10%	20	1	18 - 23	NO
Success Criteria 2	206	39	19%	101	21	21%	43	2	40 - 46	NO
Success Criteria 3	206	56	27%	101	31	31%	63	2	59 - 67	NO
No wind										
Success Criteria 1	64	27	42%	91	25	27%	18	1	16 - 19	YES
Success Criteria 2	64	10	16%	91	3	3%	2	0	2 - 3	YES
Success Criteria 3	64	37	58%	91	28	31%	20	1	18 - 21	YES
With the wind										
Success Criteria 1	86	57	66%	124	37	30%	26	1	24 - 27	YES
Success Criteria 2	86	12	14%	124	0	0%	0	0	0 - 0	YES
Success Criteria 3	86	69	80%	124	37	30%	26	1	24 - 27	YES
RESTRAINING OPERATIONS										
Against the wind										
Success Criteria 1	441	79	18%	124	16	13%	57	3	52 - 62	YES
Success Criteria 2	441	103	23%	124	28	23%	100	3	93 - 106	Random
Success Criteria 3	441	182	41%	124	44	35%	156	4	148 - 165	YES
No wind										
Success Criteria 1	322	111	34%	91	23	25%	81	3	76 - 87	YES
Success Criteria 2	322	65	20%	91	6	7%	21	1	19 - 24	YES
Success Criteria 3	322	176	55%	91	29	32%	103	3	97 - 108	YES
With the wind										
Success Criteria 1	191	137	72%	101	26	26%	49	2	46 - 52	YES
Success Criteria 2	191	10	5%	101	0	0%	0	0	0 - 0	YES
Success Criteria 3	191	147	77%	101	26	26%	49	2	46 - 52	YES
Observations sample:	1453									
Observations placebo:	1902									
Exchange rate days in both samples	3355									