How Long Do Markets Need to Fully React to Monetary Policy Announcements?

Paul L. Tran*

This version: October 25, 2025

[Click here for the latest version]

Abstract

This paper shows that financial markets need more time to fully react to Federal Open Market Committee (FOMC) policy announcements than typically assumed. Using finance literature techniques and neural network methods for text analysis, I systematically estimate that on average, markets fully react to the information content of FOMC statements within an event window *ending at least 30 minutes* after release. This optimal window increases with the underlying maturity of an asset, reaching 50–60 minutes in length for maturities at least two quarters ahead. Additionally, statements with greater complexity, less similarity, and dissents are associated with longer event windows on average. I find that the correlation between monetary policy surprises measured within optimal versus conventional 30-minute windows decreases with asset underlying maturity. These differences alter the forward guidance component of monetary policy shocks and magnify their estimated impact on interest rates, break-even inflation, and equity prices. Furthermore, monetary policy shocks constructed in optimal windows results in the responses of macroeconomic variables to become more precise.

Keywords: Event window studies, FOMC statements, monetary policy shocks, neural networks, natural language processing, text analysis

JEL Codes: C45, E52, E58, G14

*Department of Economics, The University of Texas at Austin. 2225 Speedway, BRB 2.128, C3100, Austin, TX 78712, USA. Email: pltran@utexas.edu. Website: https://paulletran.com/. I am indebted to my committee—Oli Coibion, Chris Boehm, Saroj Bhattarai, and Amy Handlan—for their invaluable guidance and support. For their helpful comments and suggestions, I also wish to thank Andi Mueller, Andres Drenik, Stefano Eusepi, Oliver Pfauti, Andrew Chang, and seminar participants at UT Austin. I am grateful to Chris Boehm and Niklas Kroner for sharing data with me. Finally, I thank Trieu Tran for her research assistance.

1 Introduction

> To establish the right price for a stock the market must have adequate information, but it by no means follows that if the market has this information it will thereupon establish the right price. The market's evaluation of the same data can vary over a wide range, dependent on bullish enthusiasm, concentrated speculative interest and similar influence, or bearish disillusionment. Knowledge is only one ingredient on arriving at a stock's proper price. The other ingredient, fully as important is sound judgement.
>
> Graham (1974)

How much time do financial markets need to fully respond to news? The answer might not be obvious. Consider the following example of the S&P 500 Index price movement shortly after the release of the January 2008 Federal Open Market Committee (FOMC) monetary policy announcement. As shown in Figure 1, measuring the price change up to different times after the announcement leads to different conclusions on how positive was the reaction of the stock market index. In other words, the *time window* chosen to measure financial market price reactions to news can change the answer to the original question.

Because changes in financial prices can serve as proxy for changes in market expectations, applied macroeconomists have used tools such as event window studies to measure the unanticipated financial market response to news and capture the predicted effect of news on macroeconomic outcomes. However, news releases can be dense and difficult to interpret, especially when the information content deals with time further into the future. As a result, whilst markets may need more than a few minutes to react to complex events or policy announcements, longer windows increase the likelihood that the measurement will be contaminated as additional news is revealed, making it difficult to assign market price responses to a particular event. Balancing these two forces is what allows applied macroeconomists to pin down the event window length that should be used to determine the response of financial markets to news releases. Therefore, determining the optimal window length is an empirical question and not a parameter whose value we should simply assume.

A relevant and important application of event window studies is to construct monetary policy surprises, which are unanticipated changes in interest rates around the FOMC policy announcements. The choice of event window size in the literature falls into two categories: Short or long window lengths. Proponents of narrow windows around FOMC announcements (e.g., 30 minutes) argue that this high-frequency identification makes plausible the assumption that observed price changes are caused solely by the policy announcement. This approach and the choice of 30-minute windows have become the popular choice and de facto standard in the literature (e.g., Gürkaynak, Sack, and Swanson, 2005; Gertler and Karadi, 2015; Nakamura and Steinsson, 2018; Gürkaynak, Kisacikoğlu, et al., 2020; Jarociński and Karadi (2020); Handlan, 2022b; Swanson and Jayawickrema, 2023; Bauer and Swanson, 2023; and others). An implicit assumption from this argument is that the initial price reaction of financial markets within small event windows is equal to the full reaction. However, this implicit assumption could be violated due to various issues potentially facing financial markets, such as over-reactions (e.g., Boguth et al., 2023; Bianchi et al., 2024), under-reactions (e.g., Boguth et al., 2023; Bianchi et al., 2024), the presence of noise trading (e.g., Hervé et al., 2019; Ben-David et al., 2022), and algorithmic trading (e.g., Bazzana and Collini, 2020; Ben Ammar and Hellara, 2022). In addition, Ramey (2016) finds that the macroeconomic effects of monetary policy surprises constructed with samples beginning after 1984 declined in precision likely because of how monetary policy was conducted more systematically after this time period. As a result, Ramey (2016) argues that monetary policy shocks now consist mostly of information effects from the FOMC and *noise*. The combination of all these effects could result in monetary policy surprises to lack relevance and precision.

Conversely, those who utilise longer event windows (e.g., 1-hour to 1-day) contend that this approach has a higher probability of capturing the market's full reaction to FOMC announcements, resulting in shock measures constructed from the monetary policy surprises that have higher explanatory power (e.g., Kuttner, 2001; Rigobon and Sack, 2004; Gürkaynak, Sack, and Swanson, 2005; Gürkaynak, Kisacikoğlu, et al., 2020; Bauer, Lakdawala, et al., 2022; Swanson and Jayawickrema, 2023; Boehm and Kroner, 2025b; An et al., 2025; Boehm and Kroner, 2025a; and others). Unfortunately, the trade-off with this alternative size is that longer event window lengths increase the chance that measurements are contaminated with unrelated news and confounding factors, causing a decline in statistical significance and precision of estimates (e.g., Gürkaynak, Sack, and Swanson, 2005; Nakamura and Steinsson, 2018).

Ultimately, the choice of event window length has remained an "ad-hoc" decision. Relatedly, the literature has implicitly assumed that this choice applies to all assets, an assumption that may not hold. For example, markets for assets with short underlying maturities might price in FOMC policy announcements more easily, whilst markets for assets with longer underlying maturities might need more time.[1] Therefore, one might expect that longer event window lengths are required to capture the full reaction of the markets for assets such as longer-term bonds and equities.

The observed trade-offs with different event window sizes bring up a crucial question: how does one choose the appropriate event window length for their study in monetary policy? For a literature that has available a finite sample of policy announcements, the wrong answer to this question could result in monetary policy surprises whose effects are either attenuated with classical measurement error or possess confounding factors.

There are surprisingly very few papers that directly investigate the appropriate length of time over which to measure price reactions in either the empirical macroeconomics and finance literatures. One of the earliest papers looking into this question is Hillmer and Yu (1979), who develop a theory around the assumption that there exist two price reaction processes: One during a period of adjustment to the event of interest and one outside the reaction period. Tests of their model show that as long as markets are still reacting to the new information, the variance of the price reactions within the event window was higher than in periods outside of the window. Using intraday data, the authors find that the event window should last several hours after the initial event of interest. Chang and Chen (1989)

[1] Throughout this paper, I differentiate between "underlying maturity" and "futures maturity". The former typically refers to the maturity of the underlying asset for futures contracts (e.g., 2-year Treasury futures deal with interest rate expectations 2 years into the future). The latter refers to the expiration date of the futures contract themselves, (e.g., front-month 2-year Treasury futures).

apply the methodology of Hillmer and Yu (1979) on daily data, finding that event windows should span several days due to the market continuing to react to the initial event. Krivin et al. (2003) propose various proxy rules for determining the length when looking at a limited number of observations, such as those based on significant abnormal returns. The authors find that rules based on continuing price movements result in event window sizes that correlate with the "size" of the news. Das and King (2021) attempt to provide some systematic structure by measuring the informativeness of earnings announcements using abnormal R^2 and other statistics, discovering that asymmetric, narrower one-day event windows exhibit greater information content. Casini and McCloskey (2025) determine that narrow window sizes aren't necessary sufficient for causal identification, but window size should be selected based on when the surprise component of the monetary policy announcement dominates all other events occurring within that window[2] A recent paper by Boehm and Kroner (2025a) combines the equivalent of a one-step Kalman filter estimation with the heteroscedasticity-based approach of Rigobon and Sack (2004) to identify the event window around FOMC announcements where excess variance is no longer statistically significant. Under this criterion, the authors select an event window length of fourteen hours.

My paper builds upon this literature from an alternative angle. Whilst the aforementioned studies infer the optimal window length using only observed price dynamics, my approach combines these price dynamics with a text-based signal of the market's price response under the minimised average impact of noise to further pin down the optimal window length. In particular, I use neural network methods for text analysis to approximate the underlying relationship between the information content of FOMC statements to the market price response. The text-based signal comes from this approximation and represents the predicted price reaction of the market to the monetary policy announcement, word-for-word. Using this text-based signal, I am then able to estimate the optimal event window length that best reflects the *full* price reaction of the market to the FOMC statement, an interpre-

[2]The authors call this condition relative exogeneity, which specifically states that the variance ratio between the policy shock and that of other variables be infinite.

tation of the signal that is only possible when the impact of noise is *minimised* on average within this window length. Additionally, this interpretation about the optimal event window is related to the findings of Casini and McCloskey (2025) about the relative dominance of the monetary policy surprise. My method provides an empirical procedure for setting this basic step in event window studies for the high-frequency identification literature and a deeper discussion into the effects of event window choice on empirically measured monetary policy surprises and shocks.

With my method, I provide four key findings, two regarding the optimal event window length and two regarding the effects of event window choice on monetary policy impacts: first, in line with the insights from Hillmer and Yu (1979) and Boehm and Kroner (2025a), my results show that financial markets take more time to fully react to monetary policy announcements than the typically assumed 30 minutes of the modern high-frequency identification literature. Specifically, I find that regardless of contract maturity, financial markets for futures and equities fully react to the information content of FOMC statements within an event window starting 10 minutes before and *ending at least 30 minutes* after release.

Second, and unlike those earlier papers, I find the optimal window length varies systematically across asset classes and maturities. Figure 2 offers a stylised summary of the main results. It displays the averaged systematically estimated event window lengths from my method for all considered asset underlying maturities. As mentioned above, asset markets with the shortest underlying maturities are estimated to fully react within a 40-minute event window. As the maturity increases, the average optimal window length rises accordingly, levelling off between 50–60 minutes around the policy communication for underlying maturities at least two quarters ahead. In other words, my paper documents that the optimal event window length varies across asset types and underlying maturities in a systematic way. Relatedly, I document that statements with greater complexity, less similarity, and the presence of dissents have relatively longer event windows on average.

Third, I document that the correlation between monetary policy surprises constructed within optimally-found event window lengths v. the popular 30-minute window decreases for all underlying maturities when increasing the optimal window length. Furthermore, increasing the underlying maturities of the surprises also sees the correlation fall by a larger magnitude under longer optimal window lengths. Whilst interest-rate surprises for current and next FOMC meetings are similar regardless of event window choice, the differences result in a decline in the correlation by 10% for surprises dealing with fifteen years out. I find that these differences in the surprises affect monetary policy shocks about forward guidance in a variety of ways, such as larger magnitude in the peaks and troughs of the shocks or "shifts in importance" of shock compositions.

Fourth, opening up the window from the popular 30 minutes to the optimal length causes the responses of interest rates, break-even inflation, and equity prices on monetary policy shocks about forward guidance to become larger. Furthermore, the estimated responses of macroeconomic variables to monetary policy shocks constructed within the optimal window lengths become more precise. As a result, previously documented effects of monetary policy shocks could be attenuated and imprecise from the usage of suboptimal window lengths.

When it comes to general applications of text analysis methods in economics, most empirical studies have employed the techniques of clustering words into topics or using word counts to fit predictive models (Gentzkow et al., 2019). Many such examples exist in literature that studies the effects of central bank communication through text analysis, such as Acosta (2023) in their measures of expansionary versus contractionary monetary policy sentiment and Husted et al. (2020) in their creation of a monetary policy uncertainty index. These techniques are popular because they allow researchers to choose the interpretation and emphasis of individual words. However, a common disadvantage amongst these methods is that more complex features of text, such as context and word interdependencies, are not captured. In contrast, neural networks for text analysis do not overlook these characteristics, allowing these natural language processing algorithms to quantify and use the "full" information content of text. Thanks to the techniques of transfer learning and pre-training from the computer science literature, adapting text-analysis neural networks for finite sample applications, such as studying the effects of monetary policy communication on economic outcomes and expectations, has become feasible.

The most similar examples to the approach of my paper that take advantage of these neural network methods are Handlan (2022b); Doh et al. (2023); and Piller et al. (2025). Doh et al. (2023) uses an artificial neural network for text analysis called the Universal Sentence Encoder to measure differences in hawkish-doveish sentiment between official and alternative FOMC statements and create a stance metric over time. Handlan (2022b) approaches these same types of monetary policy statements with a different text-analysis neural network, XLNet-Base from Yang et al. (2019), to create monetary policy text shocks. With this combination, she is able to remove the "Fed Information Effect" (Miranda-Agrippino and Ricco, 2021) from the forward guidance component. Impulse responses of macroeconomic variables from her text shock follow the same trends and amplitudes as those predicted by theory. Building from these two studies, Piller et al. (2025) expand the set of FOMC communications considered by XLNet-Base to include Federal Reserve speeches. The result are estimated monetary policy shocks that have stronger relevance and mitigate endogeneity concerns. My paper contributes to this growing literature by utilising the same neural network as in Handlan (2022b) and Piller et al. (2025), but for a first step in the process of obtaining monetary policy shocks: systematic determination of optimal event window lengths around FOMC announcements such that financial markets have fully reacted and are minimally impacted by noise, resulting in more relevant and precise monetary policy surprises (Acosta, 2023).

The remainder of this paper is structured as follows. Section 2 provides a framework of the dynamics of asset market prices in response to news. Despite strong and simplifying assumptions imposed, the framework motivates the necessity of my text-analysis neural network approach. Section 3 describes the input and output data of the text analysis method as well as the data used to investigate the effects of event window choice on monetary policy shock impacts. Section 4 details the process of systematically estimating the optimal event window lengths and the neural network approach behind it all. Section 5 presents the estimated length of time that financial markets need to fully react to monetary policy announcements. Section 6 details the effects of event window choice on monetary policy surprises and shocks. Section 7 details how the complexity, similar, and presence of dissents within the FOMC statements can affect the optimal window lengths. Finally, Section 8 concludes.

2 Motivating Framework

In this section I present a motivating framework that shows how the dynamics of asset market prices in response to news can have conflicting noise components and why it isn't always optimal to choose the shortest event window length. I note that the following framework imposes strong linearity and orthogonality assumptions on the noise components for simply illustrating their effects on the market price of an asset and in turn, their effects on the optimal event window length. The method behind systematic estimating the optimal event window length does not impose these assumptions because the true composition of the market and asset price in response to news events are unknown. Therefore, this framework and related simulations serve as motivation for the natural language processing methods used to extract a signal based on the text of news releases.[3]

Consider the price of one asset market at time t, P_t. Assume that the news is released at the beginning of the time period, $t = 0$. As shown in Equation 1, the asset price is the sum of three components:

$$\{P_t = P_t^f + \varepsilon_t^c + \varepsilon_t^n | t \geq 0\}, \tag{1}$$

where P_t^f is the fundamental price component, ε_t^c is cognitive noise, and ε_t^n is unrelated news.

P_t^f represents what asset price should be set by the market due to *only* the news release,

[3]Details behind the simulation process can be found in Appendix A.

meaning it is assumed that $P_t^f = P^f \in \mathbb{R}$.

Cognitive noise can be thought of as the collective "market noise" due to market participants only being able to process and react to so much information in a given amount of time. For aforementioned reasons such as algorithmic and noise trading, potential liquidity constraints, and over- or under-reactions, it is possible that market participants will have incomplete information or noise as well as different beliefs of what the news event means to the price of the asset they are trading in a short amount of time. As time progresses after the news release, the cost of processing and reacting to information will decline for more investors, who will also be able to see each other's price responses and react accordingly. These factors lead to a collective convergence to the fundamental price component. In other words, cognitive noise and its error "decay" over time and converge to zero. The framework reflects this behaviour by representing cognitive noise as the following modified AR(1) process:

$$\varepsilon_t^c = \rho_c \varepsilon_{t-1}^c + e^{-\mathcal{D}t} \nu_t^c,$$

where coefficient $|\rho_c| < 1$, decay term $\mathcal{D} \in \mathbb{R}^+$, and random noise ε_t^c is normally distributed with mean zero and variance σ_c^2. Two necessary assumptions to ensure the cognitive noise process in the framework exhibits the decaying behaviour over time are that $|\frac{\rho_c}{\mathcal{D}}| < 1$ and that the variance of cognitive noise at time $t = 0$ is equal to σ_c^2.

The effect of news unrelated to the event of interest on the asset price is modelled as a random walk:

$$\varepsilon_t^n = \varepsilon_{t-1}^n + \nu_t^n,$$

where random noise ν_t^n is normally distributed with mean zero and variance σ_n^2. Similar to cognitive noise, the variance of the unrelated news process at time $t = 0$ is equal to zero.

2.1 The Effects of Noise on the Variance of Asset Prices

The expression of P_t as Equation 1 allows for the derivation of the variance of the asset price for all time $t \geq 0$ through iterative substitution.

At time $t = 0$ when news is released, one can express the variance of the asset price as $\text{Var}(P_0) = \text{Var}(\varepsilon_0^c) + \text{Var}(\varepsilon_0^n) = \sigma_c^2$. Similarly, the expression at time $t = 1$ is $\text{Var}(P_1) = \text{Var}(\varepsilon_1^c) + \text{Var}(\varepsilon_1^n) = \sigma_c^2(\rho_c^2 + e^{-2\mathcal{D}}) + \sigma_n^2$. Continuing this iterative process yields the following expression of

$$\begin{aligned}
\text{Var}(P_t|t \geq 0) &= \left[\sum_{i=0}^{t} \rho_c^{2(t-i)} e^{-2\mathcal{D}i}\right] \sigma_c^2 + t\sigma_n^2 \\
&= \left[\frac{\rho_c^{2(t+1)} - e^{-2(t+1)\mathcal{D}}}{\rho_c^2 - e^{-2\mathcal{D}}}\right] \sigma_c^2 + t\sigma_n^2,
\end{aligned} \tag{2}$$

where $\lim_{t\to\infty} \left[\frac{\rho_c^{2(t+1)} - e^{-2(t+1)\mathcal{D}}}{\rho_c^2 - e^{-2\mathcal{D}}}\right] = 0$ by the aforementioned framework assumptions.

I define t^{one} as the time where the variance of P_t is minimised. Solving for this time horizon yields the following indirect expression that provides important insight into the factors influencing the appropriate event window length:

$$t^{one} : \mathcal{D}\left[e^{-2(t+1)\mathcal{D}}\right] + ln(\rho_c)\rho_c^{2(t+1)} = \left[\frac{(e^{-2\mathcal{D}} - \rho_c^2)}{2}\right] \frac{\sigma_n^2}{\sigma_c^2}. \tag{3}$$

Two important findings from Equation 3 are that $\frac{\partial t^{one}}{\partial \sigma_n^2} < 0$ and $\frac{\partial t^{one}}{\partial \sigma_c^2} > 0$. In other words, an increased presence of unrelated news (cognitive noise) results in the time where the variance of the asset price is minimised to decrease (increase).[4] The dynamics caused by these two components give formal insight into why strictly using narrow event windows isn't appropriate.

[4] Because Equation 3 is unable to have t isolated on one side, I numerically verify the dynamics of the t^{one} for various values of σ_c^2 and σ_n^2 in the indirect expression whilst holding the other parameters constant.

2.2 Estimator Form

The dynamics described by Equation 3 provide insight into the trade-offs for a *single* news announcement. To apply this framework empirically such that the methodology of this paper is motivated, this concept must be extended to the *average* price reaction across many news releases. Therefore, the goal of this paper is not to find t^{one} for any single event, but to systematically estimate t^*, the time at which the market fully reacts to news announcements *on average*. Formally, consider N news announcements and the price of one asset market. For each news i that is released at time t, the asset price $P_{i,t}$ and its components will respond to the announcement. Therefore, I define t^* as the optimal time horizon that minimises the mean squared error (MSE) between $P_{i,t}$ and $P^f_{i,t}$ for all N news:

$$t^* : \min_t \frac{1}{N} \sum_{i=1}^{N} \left(P_{i,t} - P_i^f\right)^2 = \min_t \frac{1}{N} \sum_{i=1}^{N} \left(\varepsilon^c_{i,t} + \varepsilon^n_{i,t}\right)^2 \tag{4}$$

However, assume that the fundamental component is unobservable to the econometrician. Instead, suppose that the econometrician observes a noisy signal of the fundamental component, $s_i = P_i^f + \xi_i$, where $\xi_i \sim \mathcal{N}(0, \sigma_s^2)$. By making the core assumption that the sum of the asset price noise components, $(\varepsilon^c_t + \varepsilon^n_t)$, and the signal noise are *independent*, I am able to derive the following expression for the minimisation problem of the MSE:

$$\begin{aligned}
t^* : \min_t \frac{1}{N} \sum_{i=1}^{N} (P_{i,t} - s_i)^2 &= \min_t \frac{1}{N} \sum_{i=1}^{N} \left(P_i^f + \varepsilon^c_{i,t} + \varepsilon^n_{i,t} - P_i^f - \xi_i\right)^2 \\
&= \min_t \frac{1}{N} \sum_{i=1}^{N} \left(\varepsilon^c_{i,t} + \varepsilon^n_{i,t} - \xi_i\right)^2 \\
&= \min_t \frac{1}{N} \sum_{i=1}^{N} \left[\left(\varepsilon^c_{i,t} + \varepsilon^n_{i,t}\right)^2 + \xi_i^2 - 2\xi_i \left(\varepsilon^c_{i,t} + \varepsilon^n_{i,t}\right)\right] \\
&= \min_t \left\{\mathbb{E}\left[\left(\varepsilon^c_{i,t} + \varepsilon^n_{i,t}\right)^2\right] + \mathbb{E}\left[\xi_i^2\right] - 2\,\mathbb{E}\left[\xi_i\right]\mathbb{E}\left[\left(\varepsilon^c_{i,t} + \varepsilon^n_{i,t}\right)\right]\right\} \\
\implies t^* : \min_t \frac{1}{N} \sum_{i=1}^{N} (P_{i,t} - s_i)^2 &= \min_t \left[\frac{1}{N} \sum_{i=1}^{N} \left(\varepsilon^c_{i,t} + \varepsilon^n_{i,t}\right)^2 + \sigma_s^2\right]
\end{aligned} \tag{5}$$

Because the addition of a constant does not affect the optimisation problem of the MSE with respect to time t, this derivation implies that the econometrician is still able to solve for the optimal time horizon, t^*, by observing *only* the noisy signal of the fundamental component, s_i. Note that this optimisation problem is an asymptotic result. In other words, when given infinitely many news events, the precision of the signal has zero effect on finding the optimal time horizon, t^*. However, in finite sample applications such as the empirical monetary policy literature,[5] the precision of the signal can determine the feasibility of estimating t^*. An equivalent but important interpretation of the optimal time horizon is that t^* is the time in which the average impact of the noise components on the asset price is minimised, resulting in P_i^f (and subsequently s_i) to have the largest "share" in the asset price.

2.3 Framework Takeaways

I simulate the asset price process over time for multiple news announcements in order to illustrate the effects of cognitive noise and unrelated noise on the optimal time horizon and demonstrates the necessity of a good signal for this MSE minimisation. I discuss the main implications of the simulation in the main text, but Appendix A details the process.

Three scenarios are considered for the asset market price responding to the release of news: one that exhibits cognitive noise but little unrelated news, one with unrelated news but little cognitive noise, and one with both cognitive noise and unrelated news. The simulations show that regardless of scenario, the estimated time horizon that minimises the MSE using only the noisy signal s_i is essentially equal to that estimated when minimising the MSE observing the fundamental price component P_i^f. Therefore, a "good" signal allows for the possibility of estimating the time horizon that reflects when markets have fully reacted to news on average. However, recall that this the motivating framework and related simulations have strong linearity and orthogonality assumptions imposed to simply demonstrate the effects of noise on the estimation for the optimal time horizon. In reality, the composi-

[5] For example, the FOMC only release eight scheduled monetary policy announcements in a year.

tions and processes of the fundamental price and noise components are unknown for equity markets, yet still influence the movement of the overall asset price.

Abstracting away from the simple framework, I argue in the rest of this paper that my text-analysis neural network approach does not need to know what the true noise prices behind asset price movements are at any point of time in order to approximate of the underlying mapping between monetary policy communications and asset price changes. Therefore, my method is able to still extract a "good" signal.

3 Data

The sample period for my analysis runs from May 1999 to October 2019. Because the FOMC is the principal entity of the Federal Reserve System that determines monetary policy for the U.S., the Committee is closely watched by markets, which react to regularly scheduled announcements about the its policy actions and forecasts. Of particular importance are the released FOMC statements, which are the initial announcements of monetary policy. Furthermore, these statements are considered to be the primary source of monetary policy news to market participants.[6] Because financial markets have been known to dissect and react to these statements word-by-word,[7] understanding when financial markets fully react to the information contained in these headline text is important for event study applications in monetary policy. However, the mapping from FOMC statements to financial market asset price reactions is nonparametric and difficult to fully capture with popular text analysis methods in empirical macroeconomics. Approximating this underlying relationship—necessary for obtaining the signal needed for estimating the optimal event window length—motivates the text-analysis neural network approach in my paper. Before detailing the application of my method for systematically estimating optimal event window lengths, I first provide

[6]Appendix Figure E1 shows U.S. interest spikes on Google Trends for phrases such as "FOMC meeting" and "FOMC statement" during scheduled meeting dates. Further support for financial market reactions coming primarily from the FOMC statements is the fact that the 1^{st}–3^{rd} query results on Google Search direct to the Board of Governors of the Federal Reserve System website.

[7]E.g., CNBC coverage of the January 2025 FOMC statement text.

descriptive information on the FOMC statements (i.e., the inputs) and the financial market asset prices considered in this paper (i.e., the outputs). Appendix Table F2 in the appendix provides an overview of the employed inputs and outputs data.

3.1 Inputs: FOMC Statements

The decisions of the FOMC are typically announced in a press release shortly after its meeting concludes. Scheduled meetings occur eight times per years and are typically 6–8 weeks apart. Occasionally, the FOMC will hold unscheduled meetings whenever they deem policy action is needed before the next scheduled meeting occurs. I drop the statements of these unscheduled meetings from my sample because I want to ensure the change in asset prices within event window lengths around the publication time is driven by the content of the statements themselves and not confounded with the surprise that there was an FOMC meeting. I source the scheduled statements from the Board of Governors of the Federal Reserve System website.[8] This selection process results in the current version of this paper to have a sample size of 165 statements.

Text Pre-processing To prepare my sample of FOMC statements for the neural network approach, I pre-process all FOMC statements by converting them into plain text format, changing the text coding into standardised UTF-8 format (e.g., change length of "-"), and ensuring the spacing between words is one space. In addition, I remove all URLs, the FOMC member voting record found at the end of each statement, the list of regional bank request approvals, and mentions of release timestamps. Although papers such as C Madeira and J Madeira (2019) have shown that voting records affect stock market reactions, no effect was found for interest-based assets such as Treasuries. I prioritise removing the voting records such that the main economic discussions within the FOMC statements are entirely considered by the neural network with its hard constraint on the number of words when approximating the underlying relationship. More information about this limitation is discussed in Subsec-

[8] https://www.federalreserve.gov/monetarypolicy/fomc.htm

tion 4.3.[9]

Content and Evolution of Statements A typical FOMC statement begins with a general discussion of current macroeconomic conditions and then communicates the Committee's expectations for the future of the macroeconomy. The statement then mentions the Committee's goal of achieving maximum employment and price stability concludes with the new Federal Funds and discount rates. Following the Great Recession of 2008–2009, these statements included discussion about unconventional monetary policy, such as quantitative easing programmes. Because this period of time saw an increase in complexity in the language of FOMC communications and the inability of the Committee to use rate changes to influence market expectations about monetary policy, the statements grew rapidly in length up to 2014, as seen in Figure 3. This trend gradually reverses after 2014, but the statements are still longer on average than those before the Great Financial Crisis.

The following example is an excerpt of the FOMC statement released on 01 August, 2018:

> Information received since the Federal Open Market Committee met in June indicates that the labor market has continued to strengthen and that economic activity has been rising at a strong rate. [. . .]
>
> The Committee expects that further gradual increases in the target range for the federal funds rate will be consistent with sustained expansion of economic activity, strong labor market conditions, and inflation near the Committee's symmetric 2 percent objective over the medium term. Risks to the economic outlook appear roughly balanced.
>
> In view of realized and expected labor market conditions and inflation, the Committee decided to maintain the target range for the federal funds rate at 1-3/4 to 2 percent. The stance of monetary policy remains accommodative, thereby supporting strong labor market conditions and a sustained return to 2 percent inflation.

3.2 Outputs: Interest-rate and Equity Futures

The output data for the neural network approach of approximating the underlying relationship are interest-rate and equity futures prices. All asset prices come from the Thomson

[9]Another reason why I choose to remove the voting record excerpt from the statement text is because the vast majority of my approach outputs are interest-rate futures price changes, meaning the removal should not have much impact on the systematic estimation of the optimal event window length for these markets. Nonetheless, futures versions of this paper will have the approach inputs contain the voting records.

Reuters Tick History database and are obtained from LSEG. I use these asset prices at the intraday frequency for all analyses because it allows for the consideration of event window lengths frequently seen in the high-frequency identification literature, such as the most popular choice of 30 minutes. More specifically, the change in expected interest rate paths can be measured via the variation in prices of interest rate futures in event windows surrounding FOMC announcements.

Federal funds futures and Eurodollar futures contracts are considered because many high-frequency monetary policy surprises are constructed based on the change in expected interest rates at the underlying maturities measured in these asset types (Kuttner, 2001; Gürkaynak, Sack, and Swanson, 2005; Gertler and Karadi, 2015; Nakamura and Steinsson, 2018; Jarociński and Karadi, 2020; Handlan, 2022b; and others). The market for federal funds futures is liquid enough to measure expectations of monetary policy roughly three months out, specifically capturing expectations about the federal funds rate after the current and next FOMC meetings (i.e., $FF1, FF2, FF3, FF4$). Eurodollar futures can be used as instruments to measure expectations beyond three months. As an example in Acosta, Brennan, et al. (2024), a market participant trading on date t of quarter q based on their expectations on interest rates two quarters from now would trade the second-outstanding or 2-quarter Eurodollar future, with the settlement towards the end of quarter $q + 1$. This paper considers Eurodollar futures capturing market expectations from about four months to one year ahead (i.e., $EDcm2, EDcm3, EDcm4$). For longer underlying maturities, systematic estimation of optimal event window lengths is performed on financial markets for Treasury futures with horizons two ($TUc1, TUc2$), five ($FVc1, FVc2$), ten ($TYc1, TYc2$), and thirty years ($USc1, USc2$), whose underlying maturities were approximated by Gürkaynak, Kisacikoğlu, et al. (2020) to be two, four, seven, and fifteen years ahead, respectively. I also focus on equity markets, particularly the S&P 500 Index and its E-mini futures (i.e., $SPX, ESc1, ESc2$), for two reasons. First, aside from interest rates, equities are one of the most studied asset class in the empirical monetary policy literature. Second, stock index futures are traded outside of regular trading hours, providing sufficient data quality for me to consider a wide variety of event window lengths.

Dependent Variable Construction Price levels for these futures contracts are collected at 10-minute intervals, starting from 10 minutes before to 18 hours after an FOMC statement release. The output of interest for the neural network approach is the log-price differences of each futures contract constructed within event window lengths starting 10 minutes before and $t+n$ minutes after the FOMC announcement,

$$DP_{t+n} = ln\left(\frac{P_{t+n}}{P_{t-10}}\right), \tag{6}$$

I have the neural network approach of the systematic estimation only consider price log-differences within event window up to 70 minutes in length.[10]

3.3 Data to Investigate Impacts of Monetary Policy Shocks

My investigations on the effects of event window choice on the impacts of monetary policy surprises and shocks use data on Treasury yields, industrial production (IP), the consumer price index (CPI), and the excess bond premium (EBP). Daily data for Treasury yields come from Gürkaynak, Sack, and Swanson (2005). Daily data for Treasury-inflation-protected-security (TIPS) yields and break-even inflation (i.e., the difference between nominal and real rates from TIPS) come from Gürkaynak, Sack, and Wright (2010).[11] Monthly data on IP and CPI are collected from the Federal Reserve Economic Data. The monthly measure for the EBP come from Gilchrist and Zakrajšek (2012), which is available on the Board of Governors of the Federal Reserve System website.[12]

[10]Price log-differences calculated within longer event windows are not considered due to computational and financial constraints, but are still outputs of interest for future versions of this paper.

[11]Treasuries: https://www.federalreserve.gov/data/nominal-yield-curve.htm; TIPS and break-even inflation: https://www.federalreserve.gov/data/tips-yield-curve-and-inflation-compensation.htm

[12]https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/files/ebp_csv.csv

4 Systematically Estimating Optimal Event Windows

At the core of systematically estimating optimal event window lengths is the underlying, nonparametric relationship between the monetary policy communications and the financial market price reactions. Obtaining the text-based signal needed for the estimation is only possible if this mapping can be approximated with precision. A general requirement for this precision is that the approximation method is able to consider and quantify the full information content contained within the FOMC statement text. Otherwise, the method would fail to account for all of the possible dimensions of monetary policy communication from the statement driving financial market reactions.

When considering the text analysis methods that are popular in the monetary policy literature, most can be described as "fitting predictive models on simple counts of text features" (Gentzkow et al., 2019). However, these methods often miss how words within and across sentences relate to each other. For example, the phrases "*employment* went up, but *inflation* did not" and "*inflation* went up, but *employment* did not" would produce the same measures. Alternatively, methods that measure the frequency of neighbouring words, called "n-grams", often miss information spread out across a sentence. For the following phrase, "economic growth slowed, but is expected to pick up pace later this year", a trigram would count "economic growth slowed", but would miss the point of the sentence that economic growth will reverse direction and expand later in the year. In other words, these popular methods cannot realistically capture the "full" information content of text.[13]

In contrast, text-analysis neural networks can approximate the complex relationships between words such as context and interdependencies. Using these methods, the words "slowed" and "pick up pace" could both be associated with "economic growth" for prediction, even though these words are not all adjacent within the sentence. The promise of these neural networks is also mathematically justified by the Universal Approximation Theorem

[13] If it was possible to account for every possible combination of words and phrases of interest, then these frequency measures would be able to quantify the "full" information content of text like FOMC statements.

(Hornik et al., 1989; and others), which states that any continuous function—including the underlying relationship between monetary policy communications and financial market price reactions—can be approximated by a network with at least one hidden layer to any desired accuracy.[14] Indeed, with other social sciences observing success in applying natural language processing, economics could witness similar benefits when using these methods to study monetary policy communication (Gentzkow et al., 2019).

4.1 Approach: XLNet-Base, a Neural Network for Text Analysis

The approximation approach employed behind the systematic estimation is the XLNet-Base language model from Yang et al. (2019), which builds upon the transformer architecture introduced by the seminal paper of Vaswani et al. (2017).[15] The neural network abandons the literature technique of "corrupting" input text through masking the words of prediction-interest and assuming that these masked words are independent of each other. The former practice can lead to discrepancies when subsequently training the network on a different set of text without masking. The latter assumption is often violated.[16] Instead, XLNet-Base considers all possible information from all permutations of words around the words of interest in its pre-training phase. As a result, the neural network can learn context and dependencies bi-directionally without the usage and weaknesses of masking. Furthermore, XLNet-Base is better suited for tasks prioritising context and semantics compared to recently popular generative pre-trained transformer models (e.g., ChatGPT), which can only process and learn

[14]The Universal Approximation Theorem is an existence theorem. In other words, the theorem states that for any continuous function, there exists a neural network, of a *certain structure with at least one hidden layer*, capable of approximating to any desired accuracy. However, the theorem does *not* specify this certain structure. In practice, the machine learning literature has argued that adding more layers into the network architecture reducing the number of parameters per node function, therefore reducing computational, training, and data requirements for function approximation.

[15]The text-analysis neural network is open-source, pre-trained, and has 12 hidden layers (each with 768 dimensions), 12 self-attention heads (each with 64 dimensions), a vocabulary size of 32,000 word tokens, and two feed-forward layers with 768 and 3072 dimensions, respectively, ultimately resulting in approximately 117 million parameters. More information about the neural network can be found on the open-source platform https://huggingface.co/xlnet/xlnet-base-cased.

[16]Consider the example phrase, "The Federal Open Market Committee decided to increase its target rate". The words "target" and "rate" are masked. However, these words still share implicit relations to each other.

text in a unidirectional manner.[17] The permutation language modelling objective allows for XLNet-Base to be highly versatile and suitable for being fine-tuned for tasks beyond simple word prediction. Relevant to this paper is the regression task of approximating the underlying relationship between FOMC statement text and financial market price reactions. I start with the off-the-shelf architecture, parameters, weights, and general English language skills of the XLNet-Base model. I then fine-tune the neural network to "transfer" its knowledge into the task of approximating the underlying relationship and obtaining the text-based signal.[18]

4.2 The Inputs, Outputs, Approach, and Optimal Window Length

Recall from the motivating framework in Section 2 that within any given event window, the asset price changes will be mixtures of the fundamental price components and noise factors. It is *only* within the optimal event window length that on average, the impact from the noise factors will be minimised, simultaneously resulting in the fundamental price components to have the largest "shares" of the price changes. Within a given event window, the text-based signal can be interpreted as the *change in futures prices predicted by the neural network's approximation of the underlying relationship between the FOMC statement text and the observed futures price change.* Therefore, only within the optimal event window length will the average impact of noise on the price changes be minimised, allowing the neural network's approximation to have the best performance at obtaining the most precise text-base signal. This simultaneous relationship between the optimal window length and precise text-based signal is what drives my systematic estimation method, which I provide an overview below:

1. For a given event window length, I have XLNet-Base regress the price changes of a futures contract on only the text content of FOMC statements and assess its predictive

[17] All large language models are defined by predicting a word using only those that came before it. This "left-to-right" operation is why such models are commonly labelled as "generative" models.

[18] This application is only possible thanks to Yang et al. (2019) and others in the machine learning literature demonstrating that the transfer learning approach reduces training data requirements whilst achieves similar accuracy in new tasks.

performance and signal precision.

2. I have the neural network perform this regression for event window lengths increasing up to 70 minutes.

3. The event window length that resulted in XLNet-Base to have the best approximation of the underlying relationship through its predictive performance is what I define as the optimal event window length for that futures contract.

4.3 Stratified-sampling Cross Validation

To prepare the neural network for systematic estimation, I split my sample into training and testing subsamples five times, where 20 per cent of FOMC statements belong in each testing subsample. Splitting the statement observations is conditioned by the type of decision the FOMC made to the target federal funds rate, who the FOMC Chair was, if the date was pre- or post-2007, and by imposing both subsamples to have equal distribution of statements according to word count. This process yields five distinct training-testing subsample splits such that no testing subsample shares FOMC statements with one another.[19]

These splits can be thought of as consisting of input-output pairs, where the inputs are the FOMC statements and the outputs are the calculated futures price log-differences. For each 10-minute interval and each futures contract maturity, I train XLNet-Base by refitting its parameters to approximate the relationship between the text of FOMC statements and the price log-differences on the training subsample and then tested on the held-out subsample. This process is repeated five times, once for each split. The final accuracy for each event window length is the average performance across all five test splits. This methodology ensures a robust measure of out-of-sample performance and accounts for any variation in the network's predictions coming from the splits themselves.[20]

[19] This process is called "stratified sampling 5-fold cross validation in the machine learning literature.

[20] Stratified sampling 5-fold cross validation minimises the differences between the population distribution of FOMC announcement characteristics and the subsample distributions of these characteristics, which is necessary for training neural networks on finite samples through transfer learning.

I restrict the context window of XLNet-Base (i.e., the number of word tokens the neural network considers as a single input and retains in its memory) to be a sequence length of 512 word tokens for each FOMC statement.[21] Because the average number of words found in FOMC statements is roughly 327 (with a peak of 800 words), I assume that XLNet-Base should still have adequate headroom to understand the full meaning of most statements. Relatedly, a list of the neural network main hyperparameters for the fine-tuning process can be found in Appendix Table F3.

4.4 Accuracy Metrics

For each split, the primary metric considered by the neural network to judge its accuracy during fine-tuning is a *generalised* R^2 statistic defined in Hawinkel et al. (2024), which I denote as R^2_{OOS} for the rest of the paper. This statistic is chosen primarily because of two reasons. First, recall that the conventional definition of R^2 is the proportion of variance in the data that is explained by a given model. This definition breaks down when applied to non-linear methods like neural networks because the variance of the observations is no longer comprised of only the model and residual variances. In other words, squaring the Pearson correlation coefficient does not equal R^2. Second, the formula for R^2_{OOS} is adjusted to more appropriately assess out-of-sample—not in-sample—model performance.[22] Mathematical details of the generalised R^2 statistic can be found in Appendix B.

As is standard in the machine learning literature, the predictive quality of models is compared to the common baseline that states there is *no* relationship between the FOMC statement text and futures price changes and naively predicts with the in-sample average. When using the neural network to instead approximate this underlying relationship and predict asset price changes for statements in the testing subsample, the R^2_{OOS} therefore

[21] Tokens are the converted input text fed into the neural network. The practice is done to reduce the input vocabulary size. For example, {"decreasing", "increasing"} can be broken into the following tokens: {"de", "in", "creas", "ing"}.

[22] When evaluating the predictive performance of the neural network under the conventional R^2 statistic, the systematic estimation of optimal event window lengths and predictive quality of the network are the same and similar, respectively.

states how much predictive error from naively using the in-sample average has been reduced.

Other related metrics that are tracked by XLNet-Base during its fine-tuning process are the out-of-sample Pearson correlation coefficient ρ between the predicted and actual observations, the out-of-sample mean absolute error, and the in-sample mean squared error. The last criterion is important for determining if the network is learning the relationship between the FOMC statement text and the price log-differences or not.

The decisive criterion for systematic estimation is the out-of-sample R^2 *averaged across sample splits* for each 10-minute interval of each futures contract maturity:

$$\overline{R^2_{OOS}} = \frac{\sum_{i=1}^{K} R^2_{OOS}}{K}, \tag{7}$$

where $K = 5$ is the number of sample splits. The statistic measures the average improvement in performance of the neural network approximating the relationship between FOMC statements and the price reactions of markets within a given event window length, relative to naively predicting with the in-sample average. In other words, the optimal window is the length that yields the largest average improvement in generalisability by the neural network, a property that is only possible when the observed asset price changes are minimally impacted by cognitive noise and unrelated news.

4.5 Fine-tuning the Neural Network for Approximation

It is important to note that the individual parameters of neural networks do not have the same interpretability as estimators from parametric models (Athey and Imbens, 2019). Therefore, my paper cannot describe the causal effect of one word over another within FOMC communications on asset market price reactions. Instead, the training process behind the neural network allows me to approximate this underlying, complex relationship and use its predictions to pin down what event window size best reflects the markets' full reaction to monetary policy communications.

When fine-tuning the network, overfitting becomes an inevitable issue due to XLNet-Base having 110 million parameters. On a per-split basis, this issue contributes to the difference between R^2_{OOS} and the equivalent calculated for the training sample. I minimise this difference by limiting the neural network training through the number of training iterations used and the learning rate. Deploying too many training iterations causes the neural network and its weights to perfectly match the training data whilst deteriorating its out-of-sample prediction accuracy. In other words, too many training iterations causes XLNet-Base to become less generalisable. In contrast, allowing the neural network with too few training iterations can prevent the neural network from fully learning the relationship between the FOMC statements and futures price log-differences. The learning rate is essentially the rate at which a neural network updates its parameters to the training data. Too high of a learning rate will result in the network forgetting information from prior training iterations during its current iteration (i.e., the weights of the network are updated too dramatically). This issue results in the out-of-sample predictions of the network to either degrade rapidly over fewer training iterations or remain constant because the network can no longer learn. Conversely, too low of a learning rate results in XLNet-Base to never fully learn the mapping between FOMC communications and asset price log-differences. Because I am applying the transfer learning approach, the initial values of XLNet-Base's parameters are optimised to understanding the general English language. Therefore, I can ensure the weighting of the network's parameters remain generalisable to new data by limiting how much the parameters can update both within and across training iterations.

Ultimately, the process of training a neural network is a balancing act: I train XLNet-Base to meaningfully approximate the underlying relationship between FOMC statements and financial market price reactions, but do so without overfitting to the point that the predictions are non-generalisable and a bad signal. I perform this balancing act through the following two-step process and imposed restrictions. First, I set the maximum number of considered training iterations to be 2040 steps (120 epochs). Second, I perform what the literature calls a "hyperparameter sweep" on the learning rate. Specifically, Bayesian optimisation is performed on the set of learning rates XLNet-Base uses when fine-tuning up to 2040 training steps.[23] During the sweep, the accuracy metrics mentioned in Subsection 4.4 are tracked throughout the training process for a given learning rate. Training XLNet-Base is stopped at the iteration where in-sample accuracy increases, but out-of-sample accuracy decreases. At the end of the sweep duration of 24 hours for each split, the learning rate and training iteration that yields the highest R^2_{OOS} is selected. Because out-of-sample predictions can be volatile during applications of transfer learning on finite sample sizes, I select the version of XLNet-Base from the training iteration that exhibits this simultaneity in accuracy only if future iterations yield a permanently worse out-of-sample prediction accuracy. The platform behind this fine-tuning process is Simple Transformers from Rajapakse et al. (2024), which simplifies the process of training, experimenting, and performing inference with Transformer model architectures for natural language processing tasks (e.g., Simple Transformers automatically selects the appropriate tokenizer when selecting XLNet-Base).

4.6 Addressing Look-ahead Bias

A valid concern when using off-the-shelf neural networks for text analysis is look-ahead bias. As explained in Sarkar and Vafa (2024), look-ahead bias occurs whenever these networks predict values in the past using information in the future. This issue stems from how most text-analysis neural networks are initially trained with a large amount of data from many sources, which increases the probability that information from the future was used in the pre-training of the network weights.[24] I argue that the effects of look-ahead bias on my neural network method are mitigated for two reasons. The first reason is from the pre-training corpora of XLNet-Base itself and the second pertains to the language composition of the

[23]The set of considered learning rates are $[1e-5, 9e-5]$ and Bayesian optimisation assumes a uniform distribution.

[24]The issue of look-ahead bias is especially prevalent for modern large language models (e.g., ChatGPT) because their architectures are not open-source and their pre-training data sources are neither fully known nor publicly accessible.

FOMC statements.

First, the authors of XLNet-Base address the concern of look-ahead bias by restricting the pre-training data of the network to *only* BookCorpus and the entire English Wikipedia. This restriction makes it highly unlikely for the neural network to learn any information about the intersection of FOMC communications and futures markets price reactions at any point of time during its pre-training phase. I also note that Yang et al. (2019) showed that this limited pre-training data does not negatively impact the generalisability of the neural network when compared to other models.

Look-ahead bias is also mitigated in my application of XLNet-Base for systematic estimation because of the temporal anonymisation of the information content in FOMC statements. Similar to Glasserman and Lin (2023), the language composition of the pre-processed statements has no references to relative times t and $t + 1$. As a result, even if XLNet-Base was trained with FOMC statements that were in the future relative to the statement it is being asked to predict the futures price change, the network does not have any necessary information to chronologically order the FOMC statements and subsequently "figure out" that its weights have been trained on future data.

5 How Long Until Markets Fully React to MP News?

In short, systematic estimation of the optimal event window length can be summarised as the following exercise: the neural network is searching for the event window length where the average impact of noise on the variation of asset prices is minimised, allowing the network to approximate the underlying mapping from FOMC statement text to asset price changes such that its out-of-sample predictions are the most generalisable and represent the most precise signal on average.

Figure 4 presents the results of this exercise for selected federal funds, Eurodollar, and Treasury futures of various underlying and futures maturities and the S&P 500 Index. Appendix Figures E2–E8 show the window estimation results for all interest-rate and S&P 500 E-mini futures. The horizontal axis of each figure depicts the end time of the event window lengths (e.g., $t+20$ represents a 30-minute event window starting at 10 minutes before and 20 minutes after FOMC statement release) and the vertical axis is the generalised R^2 statistic discussed in Subsection 4.4. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window length with the largest averaged out-of-sample R^2_{OOS}. For each event window size, box-and-whisker plots are shown surrounding the corresponding averages in order to better present the distribution of predictions from the neural network across sample splits.[25]

The cross points show that regardless of futures maturity and asset type, the event window length that yields the largest $\overline{R^2_{OOS}}$ from the neural network is always at least 40 minutes in length, beginning 10 minutes before and ending 30 minutes after FOMC statement release. Additionally, this implication is robust to the act of averaging the out-of-sample R^2_{OOS} of each split, as similar conclusions are reached when examining other presented metrics of the box-and-whisker plots, such as either the 25^{th} and 75^{th} percentiles: financial markets fully react to the information content of FOMC statements within a 40-minute event window *at minimum*, implying that the popular choice of 30-minute event windows is too short.[26]

An interesting observation is that the event window length where XLNet-Base achieves the highest $\overline{R^2_{OOS}}$ increases with the underlying maturity of the considered futures. This trend can be seen when referring back to stylised summary visual of Figure 2. Specifically, the average systematically estimated event window for the front-month and one-month-ahead Federal Funds futures—assets used to measure market expectations about future monetary policy from the upcoming FOMC meeting—is 40 minutes in length, whereas the average window length for the two- and three-month-ahead Federal Fund futures is 45 minutes. For

[25]Typical standard errors and confidence bands are not calculated because there are too few data points, i.e., the number of sample splits is $k = 5$. Whilst this could be changed such that the neural network performs systematic estimation using leave-one-out cross validation, the computation and financial costs are too great to make this alternative exercise feasible for the current version of this paper.

[26]Although not shown to avoid visual clutter, this conclusion holds true when using the median R^2_{OOS}. The optimal event window is still never 30 minutes in length.

Eurodollar futures, which are used to measure expectations about the path of interest rates between two and four quarters into the future, the neural networks predicts markets fully react to the information content of FOMC statements within a 50-minute event window. 50 minutes is also the systematically estimated window length for both 2- and 5-year Treasury futures (when averaging across the first- and second-month contract maturities), but the trend increases for 10- and 30-year Treasury futures, which both have optimal event window lengths of 60 minutes. One possible explanation behind this increasing trend overall is that market participants trading futures with interest-rate expectations further into the future are more exposed to the "soft" information of monetary policy announcements (e.g., forward guidance) which are harder to process and interpret (Indriawan et al., 2021). Therefore, these longer-horizon assets could be more influenced by belief and information uncertainty, which is ultimately reflected in significant and larger risk premia (Piazzesi and Swanson, 2008). With regards to asset liquidity, observed differences in this dimension could interestingly be a symptom of these optimal event window lengths. For example, Fleming and Piazzesi (2005) document that Treasury markets with longer underlying maturities experience abnormal trading volumes for longer durations of time after an FOMC announcement. Additionally, Okada (2025) document that longer-duration Treasury bonds are more effected by the FOMC announcement premium in contrast to short-horizon because of interest rate risk.

With regards to equities, the S&P 500 Index and its front-month E-mini futures contract are estimated to fully react to FOMC statements within 50 minutes, whilst the second-month E-mini futures market fully reacts within a 60 minute window on average. The average optimal event window length estimated by the neural network for equities is 53.3 minutes. One possible explanation for the optimal window length being longer than 30 minutes is because equities and their options are less directly impacted by monetary policy actions, resulting in a larger degree of investor disagreement and misinterpretation of interest rate decisions (Zhang et al., n.d.). Although both the E-mini futures share the same underlying maturity of infinity with the S&P 500 Index, one explanation behind the differences in estimated optimal window lengths are that both futures fully react to monetary policy announcements based on both the underlying and futures maturities. For example, market participants for the second-month E-mini futures contract ($ESc2$) have the underlying maturity of infinity due to the S&P 500 Index reflecting the net present value of all expected future cash flows. But because the contract has an expiration date farther into the future than the front-month futures contract, they also react to the FOMC statement based on changes to expectations up until the expiry date. When combined with less liquidity compared to $ESc1$ and popularly traded exchange-traded funds of the S&P 500 Index, this weighted expectations can result in the market for $ESc2$ to fully react to FOMC announcements in a longer amount of time because participants put higher weight on the "soft" information in monetary policy statements.

Overall, the systematic estimation of optimal event window lengths by the neural network shows that $\overline{R^2_{OOS}}$ broadly exhibits a "hump" shape. The measure of generalisability increases with event window length, peaks at 30–50 minutes after statement publication, then declines. The motivating framework factors of cognitive noise and unrelated news could be at play here. As time increases, the effects of cognitive noise decline as financial markets start to fully react to the monetary policy announcements. Eventually, the effects of unrelated news dominate the price reaction of futures markets, causing $\overline{R^2_{OOS}}$ of the neural network to shrink. The optimal event window length is the length where the average impact from both of these components are minimised, allowing the neural network's approximation of the underlying relationship to be the most generalisable on average relative to naively predicting with the in-sample average.

It is important to observe that for all considered event window lengths of all interest-rate and equity futures, the predictive quality of neural network achieves a positive $\overline{R^2_{OOS}}$, indicating that there is indeed an underlying relationship between the FOMC statement text and asset price changes. Relative to naively predicting with the in-sample average (i.e., asserting there is no such relationship), the generalisability of the neural network increases by an additional 2–17 p.p. (7.4 p.p. on average) within the optimal event window compared to that in conventional 30-minute windows.[27] This relative improvement in predictive performance is only possible within the optimal event window lengths because of the minimised average impact of cognitive noise and unrelated news on the asset price changes, thereby allowing the neural network to approximate the underlying relationship.

5.1 Different Event Windows, Different Market Responses

With the optimal event window length systematically estimated for every interest-rate and equity futures contract, I compare the market price responses measured within these windows to those calculated within the popular 30-minute window of the literature. Figure 5 depicts this exercise for the same selected interest-rate futures of various underlying and futures maturities and the S&P 500 Index. Appendix Figures E9–E15 present this exercise for the entire set of interest-rate and equity futures. The horizontal (vertical) axis of each sub-figure represents the price log-difference within the optimal (30-minute) event window. The blue dots are the market reactions on scheduled FOMC meetings in my sample. The 45-degree line is depicted as grey and dashed.

For a simple ex post characterisation of whether markets for certain asset types and maturities under- or over-react to FOMC statements on release, I regress the price log-differences within the optimal event window lengths on those calculated within a 30-minute window through OLS. If the slope coefficient is less than one and its difference with one is statistically significant for at most $\alpha = 0.05$, then I state that the futures market under-reacts to the information content of FOMC statements on release. Otherwise, the price reactions of the market between these two event windows are not different in general. I colour the regression lines as red (grey) for the former (latter) scenario in each sub-figure.

Financial markets for all interest-rate futures appear to under-react to the information

[27] Appendix Table F4 summarises these results for each interest-rate and equity futures contracts.

content of FOMC statements on release, ex post. Furthermore, the slope of the regressions decreases as the underlying maturity of these assets increases. The movement of these markets could potentially be related to the post-FOMC announcement drift described in Indriawan et al. (2021) and Brooks et al. (2023), with both papers pointing to the limitations in processing information for bond market under-reactions. As described in the former paper, interest-rate markets are more exposed to "soft" information in monetary policy announcements like forward guidance, which is costlier to process and interpret. The authors of the latter paper find that mutual fund investors slowly adjusting their interest rate expectations play a role in the price dynamics after the policy announcements. A notable example of this under-reaction can be seen in the 16 December, 2008 FOMC meeting, where the federal funds rate range was cut down to between 0% and 0.25% in response to the Great Recession. In addition to being the first time that the federal funds rate would be below 1%, the decision was also a surprise to financial markets, who were expecting a reduction to only between 0.25% and 0.5%. When combined with how the FOMC statement itself stated that the committee would "employ all available tools to promote the resumption of sustainable economic growth and to preserve price stability", these characteristics possibly contributed to the markets needing more time than typically assumed to fully react to the information describing the weakening U.S. macroeconomy. The top sub-figure of Figure 6 presents the market price reactions for all considered Eurodollar futures maturities on this FOMC meeting date. One can immediately see that as the event window length expands, the change in price becomes increasingly positive compared to the initial reaction within a 30-minute window.

Financial markets for the S&P 500 Index were also found to systematically under-react to the FOMC statement releases, ex-post, corroborating findings from Neuhierl and Weber (2024) about stock return movements in the Center for Research in Security Prices value-adjusted index and from Golez et al. (2025) about belief under-reaction of individual equities, possibly due to individual equity market participants having difficulty reacting to the full information contained in FOMC announcements. However, it should be noted that the S&P 500 Index was also the market that saw the most price reversals, particularly from negative to positive log-differences as time progressed after the publication of some FOMC statements. One good example can be seen in the S&P 500 Index market response to the 16 September, 2008 FOMC statement, depicted in bottom sub-figure of Figure 6. Within 20 minutes after statement release, equity markets responded negatively to the release. This trend reversed within a window length at least 40 minutes long and continued to increase until market close. One possible explanation behind the reversal could be that the beliefs of equity markets themselves reversed when reading the FOMC's decision to maintain interest rates. In other words, the equity market could only fully react to the relatively positive information behind the interest rate decision with more time, allowing cognitive noise to die out and the associated negative "knee-jerk" price reactions to reverse. Indeed, Bordalo et al. (2024) documents that over-reactions are a common occurrence in stock markets and will reverse as market beliefs "correct" themselves with time.

Overall, these figures from the systematic estimation and ex post characterisation support the idea that the optimal event window length is not a parameter whose value we should simply assume. Instead, I document how the event window length that best reflect market full reactions to FOMC statements varies across asset types and underlying maturities systematically.

6 What Happens to Monetary Surprises and Shocks?

The systematic estimation of this paper has shown that regardless of asset type and underlying maturity, the optimal event window length is always longer than the standard literature assumption of 30 minutes. However, how are monetary policy surprises and shocks affected when changing to the optimal window lengths?

Proper investigation into this question requires that I use a common event window length when constructing the surprises because the “invertibility”, or “spanning” (Duffee, 2013), of information about macroeconomic variables from the yield curve is well-defined for a temporal cross-section of yields. In contrast, it is not currently known if these macro-finance term structure models are well-defined for different points of time on the yield curve. As a result, I construct monetary policy surprises for each of the window lengths found to be optimal by the neural network: *40-, 50-, and 60-minute windows.* The rest of this section uses the same data and considered FOMC meeting dates as described in Section 3.

6.1 Effects of Window Choice on Monetary Policy Surprises

For each FOMC meeting date, I construct interest-rate surprises within 30 minutes and the three found optimal event window lengths. Following Kuttner (2001), Gürkaynak, Sack, and Swanson (2005), Nakamura and Steinsson (2018), and others, I use federal funds rate to cover interest rate expectations up to three months out, Eurodollar futures to capture expectations from about four months to one year ahead, and Treasury futures for interest rate expectations out to fifteen years. Appendix C provides details behind the conversion from changes in futures contract prices to interest-rate surprises.

Figure 7 plots the correlation between the monetary policy surprises calculated within the optimal event window lengths v. 30 minutes. The horizontal axis depicts the underlying maturity of the surprises (e.g., “0–1M” represents the interest-rate surprise for the FOMC meeting happening in the current or next month). The red-dotted, yellow-dashed, and green-mixed lines represent the correlations calculated for surprises measured within the optimally-found 40, 50, and 60 minutes, respectively. The blue-solid line represents the correlations averaged across the three optimal window lengths. For all underlying maturities of all surprises, we see that the correlations decline when increasing the optimal event window length. Increasing the underlying maturity of the surprises also sees the correlation fall by a larger magnitude under longer optimal window lengths. For example, whilst monetary policy expectations about current and next FOMC meetings are more or less identical regardless of optimal window choice, the differences between 30- and 50-minute windows result in a 10% decline in the coefficient when looking at monetary policy surprises out to fifteen years ahead. In other words, as expectations about monetary policy are further into the future, calculating interest-rate surprises become more sensitive to the choice of event window lengths imposed around FOMC statement releases.

6.2 Effects of Window Choice on Monetary Policy Shock Impacts

In this subsection, I use the interest-rate surprises calculated within 30 minutes and the optimal event window length as financial instruments to identify exogenous variation in monetary policy through several popular methods in the literature. I choose to include interest-rate surprises from Treasury futures in my instrument set for two reasons. First, recent studies (e.g., Brennan et al., 2024; An et al., 2025) have shown that using instruments with longer underlying maturities can help prevent the understatement of monetary policy stimulus during the effective lower bound period.[28] Second, my findings from Figure 7 imply that only changing the event window length causes larger effects to Treasury futures markets than shorter maturity assets.[29] For conciseness, I choose to use the *median* of the window lengths found to be optimal by the neural network: *10 minutes before and 40 minutes after statement release.*

Shock Construction I construct monetary policy shocks following the construction methods of Gürkaynak, Sack, and Swanson (2005), Nakamura and Steinsson (2018), and Jarociński and Karadi (2020) (henceforth GSS, NS, and JK, respectively).[30] Rather than keep track of the entire expectations path of interest rates implied by futures prices, the authors of these papers reduce this dimensionality using principal component analysis on multiple

[28]For my sample period, the effective lower bound of the federal funds rate is defined as defined as 16 December, 2008 through 16 December, 2015.

[29]As a robustness check, all exercises performed in this subsection are re-done using the original instrument set of the authors behind the monetary policy shock construction methods. The corresponding figures and tables can be found in Appendix D and show that the results from the main text are overall robust to the choices of interest rate-surprises as instruments.

[30]I summarise the names, notation, and construction of the monetary policy shock series in Appendix Table F5.

monetary policy surprises to extract the common variation in one or two dimensions.[31,32] GSS use the first and second principal components to identify two shocks (factors) due to the multi-dimensional nature of monetary policy. The authors rotate both components such that the first component drives changes in interest-rate surprises for the current federal funds rate, whilst the second component has no effect on interest-rate surprises for the current federal funds rate. This rotation yields the "target" and "path" factors of monetary policy, which I call GSS_T and GSS_P, respectively. Because principal components have no interpretable units, GSS re-scale both factors such that GSS_T is one-for-one with the interest-rate surprise for the current federal funds rate whilst GSS_P is re-scaled to have equal impact on the implied rate surprise from the four-quarter Eurodollar futures contract. NS and JK use only the first principal component because, as explained in Bauer and Swanson (2023), the first principal component is essentially a weighted average of the target and path factors. NS re-scale their first principal component to be one-for-one with the daily change in the zero-coupon, nominal one-year Treasury yield, which I denote as NS_{MP}. JK first differentiate from prior studies by re-scaling the first principal component to have equal standard deviation with that of the implied rate surprise in the four-quarter Eurodollar futures. To account for possible information effects from the central bank, JK then impose sign restrictions on the first principal component based on its co-movement with stock market changes.[33] The series that has negative (positive) co-movement with stock prices is considered to be the monetary policy (central bank information) shock, which I call JK_{MP} (JK_{CBI}).

For easy interpretation and comparison between all monetary policy shock series, I re-scale all shock series to be one-for-one with the daily change in the zero-coupon, nominal

[31]Principal component analysis uses eigenvalue decomposition of a given dataset's covariance matrix to project the data onto new dimensions based on its variance. The first principal component, representing the first coordinates on the new dimensions, captures the largest common variation of the original variables. The second principal components captures the second largest common variation and so on.

[32]The main results of JK use the implied interest-rate surprises from $FF4$. However, the first principal component of multiple surprises are used in the modern update of the paper's shock: https://github.com/marekjarocinski/jkshocks_update_fed

[33]I calculate stock market changes within the same event window length used to measure the monetary surprises.

one-year Treasury yield.[34] Appendix Table F6 provides summary statistics of each of the shock series. Because these methods together account for the various characteristics of monetary policy, such as its impact beyond the immediate horizon, its multi-dimensional impact, and the possible existence of information effects, any differences found in the shocks and their effects from this exercise would be the impact of event window choice.

Visual Shock Differences To begin my investigation, I plot the shock series constructed in both event window lengths, displayed in Figure 8. For each sub-figure, the horizontal axis depicts the FOMC statement release dates. The vertical axis is in units of percentage points after re-scaling each shock to be one-for-one with the daily change in the nominal one-year Treasury yield. The black-solid and red-dotted lines represent the shock series derived from 30-minute and optimal window lengths, respectively. Regardless of construction method, we see that much of the peaks and troughs of shocks derived within 30 minutes see larger values in magnitude at many of those FOMC statement release dates for shocks constructed using optimal window lengths.[35] For example, both the December 2008 and March 2009 FOMC meetings see larger negative shocks by 3–6 basis points in magnitude when using the optimal event window length, possibly suggesting that the full impacts of the information content within the FOMC statements aren't yet captured within 30 minutes.

Another episode worth pointing out is the August 2011 FOMC meeting date, where the statement maintained the federal funds rate range between zero and 25 basis points, but was the first communication that included explicit dates on the expected path of monetary policy in its forward guidance language (Crump et al., 2013). We can see that the optimal event window length is associated with a more negative monetary policy shock in both GSS_P and NS_{MP} on this release date. When imposing the sign restrictions of Jarociński and Karadi (2020), the negative central bank information shock increased in magnitude by

[34] The robustness checks performed in Appendix D show that the choice of scaling has no overall effect on the findings of the main text.

[35] Plots for GSS_T are not shown because regardless of window choice, the shocks are essentially identical. This similarity makes sense since markets of all underlying maturities need little time to fully react to the immediate federal funds rate change in the FOMC statements.

3.6 basis points whilst the monetary policy shock stayed roughly the same. One possible interpretation of these differences is that because the federal funds rate was already at the zero lower bound, language indicating that the future path of the federal funds rate would remain at these low levels could be interpreted as an information surprise about weaker economic conditions than initially anticipated. The full extent of this negative information shock might only materialise when financial markets are given enough time to react to the statement.

On the other hand, changing event window lengths can result in "shifts in importance" in the composition of the monetary policy shock. When comparing JK_{MP} and JK_{CBI} derived with a 30-minute window on the September 2008 FOMC meeting, the monetary policy shock is the main effect from the statement release, whilst the central bank communication shock is close to zero. When using a 50-minute window however, this interpretation flips. Looking at the statement from this FOMC meeting, we can see that the forward guidance language from the central bank suggested moderate expected growth due to accommodative monetary policy and its efforts to promote liquidity in financial markets. Combined with its decision to hold the federal funds rate range steady, market participants would likely update upwards their views about economic prospects. In line with a positive central bank information shock, the S&P 500 increased in value during trading hours.

When comparing the identified shocks from JK overall, JK_{CBI} sees a higher frequency of increased amplitude when increasing the event window length. Although many FOMC meeting dates see increased amplitude for the JK_{MP} shock, the actual size of these differences due to event window choice is relatively smaller. One likely explanation for the larger differences in the central bank information shock is a combination of the changes to the S&P 500 Index from event window choice and the imposed sign restrictions of the construction method.

Responses of Interest Rates I now look at how the event window choice affects the impact of monetary policy shocks on nominal and real interest rates for different maturities. I use the daily change in Treasury yields and daily change in TIPS yields to represent nominal and real interest rates, respectively. For both sets of dependent variables, I calculate the daily changes using the end-of-day yields for the day of and before the FOMC announcement. Note that both the 30-minute and optimal event window lengths used to construct the monetary policy shocks are nested within the daily window for Treasury and TIPS yield changes. This timing restriction prevents the yield changes to affect the shocks. The regression specification is as follows:

$$y^{i,j} = \beta_0^{i,j,k,l} + \beta_1^{i,j,k,l}(Shock)^{k,l} + \varepsilon^{i,j,k,l} \tag{8}$$

where i indicates daily changes in Treasury or TIPS yields as the dependent variable, j indexes the horizon of the yield, k denotes the considered monetary policy shock used as the regressor, and l denotes whether the shock was constructed within the optimal event window length or 30 minutes.

Regression results for the responses of nominal and real interest rates to the monetary policy shocks within different event windows are presented in Table 1 and Table 2, respectively. In both tables, each estimate comes from a separate OLS regression of the form described in Equation 8. Starting from the left, columns 2–5 (2–4) of Table 1 (Table 2) contain estimates coming from regressions where the dependent variable is the daily change in nominal (real) yields and the independent variable is a monetary policy shock series constructed within 30 minutes around scheduled FOMC announcements. Estimates in columns 6–9 (5–7) have equivalent meaning, but the shocks are constructed within the median optimal event window of 50 minutes instead. Bolded columns 10–12 (8–10) are the differences between the corresponding coefficient estimates, representing the effect from event window choice.

When opening up the event window around scheduled FOMC announcements from 30 minutes to the optimal length, all monetary policy shocks that capture the effects of forward guidance have larger effects on both nominal and real interest rates. For example, increasing the event window length results in JK_{MP} to have an effect that is 26, 24, and 23 basis points larger on the 2-, 5-, and 10-year real yields measured using TIPS, respectively. The statistical significance of the coefficients at the 1% level remain when using the optimal event window length. Interestingly, the target factor under the methodology of Gürkaynak, Sack, and Swanson (2005), GSS_T, has smaller coefficient estimates for nominal interest rates at all horizons when using the optimal event window length, but an unclear direction for real yields. Only estimates for the 2-year nominal yield are statistically significant when the target factor is constructed under both window lengths. One likely explanation for these results is that the full sample period includes the effective lower bound, a time period where the stimulus of monetary policy through changes in the current federal funds rate was limited and greater emphasis was put into forward guidance.

The main takeaway from the regression results of both tables is that measuring the reaction of financial markets to monetary policy announcements within too small of an event window prevents market participants from fully processing the "soft" information content of the announcements, especially those pertaining to the future path of interest rates. As a result, the effects of monetary policy shocks about forward guidance on financial market variables can become hampered. Indeed, these regression differences support an interpretation suggested by Gürkaynak, Sack, and Swanson (2005), which is that financial markets can fully react to announced changes in the federal funds rate target within a short amount of time. However, information within the FOMC statements about the outlook of policy and economic conditions are more complex and require additional time to fully process and price in by market participants.

Responses of Break-even Inflation I also investigate the effect of window size on the impact of monetary policy shocks on break-even inflation as measured by the difference between nominal Treasury rates and TIPS rates. The regressions follow the same form as Equation 8, but the dependent variable is now the break-even inflation measure. Starting from the left in Table 3, the columns 2–4 contain estimates coming from regressions where the dependent variable is the daily change in break-even inflation at various maturities and the independent variable is the monetary policy shock, both constructed within 30 minutes. Columns 5–7 have equivalent meaning, but both variables are constructed within the median optimal event window length. Columns 8–10 contain the percentage change between the corresponding coefficient estimates.

Similar to the findings of Nakamura and Steinsson (2018), a majority of the estimates are not statistically significant. However, there are two interesting insights from the exercise. First, expanding the event window size to the median optimal length of 50 minutes results in an increase in the statistical significance of for certain shock estimates. For example, the decline in break-even information at the furthest horizon from NS_{MP} is statistically significant at the 5% level when using the optimal window length. A similar conclusion is reached when looking at the effects to break-even inflation from GSS_T. Second, there are several instances where monetary policy shocks containing forward guidance negatively impact break-even inflation by several basis points. Similar to tha takeaway from the regression results of interest rates, expanding the event window length to the optimal size results in the impact of monetary policy shocks to become stronger and more precise.

Responses of Equities I finish my analysis on the effects of event window size on financial variables by studying equity prices. Table 4 presents analogous results for the S&P 500 Index and its E-mini futures. The form of the regressions is the same as Equation 8, except the dependent variable is now 100 times the price log-difference of equities. From the left, the columns 2–4 contain estimates coming from regressions where the dependent variable is the price change in the S&P 500 Index and its E-mini futures and the independent variable is the monetary policy shock, both constructed within 30 minutes. Columns 5–7 have equivalent meaning, but both variables are constructed within the median optimal event window length. Columns 8–10 contain the percentage change between the corresponding coefficient estimates.

Increasing the event window around scheduled FOMC announcements from 30 minutes to the optimal length causes all monetary policy shocks containing forward guidance to impact stock prices more negatively. Opening up the event windows results in the decline of the S&P 500 Index from JK_{MP} to grow by 18.25%, whilst GSS_P and NS_{MP} negatively affect stock prices by a more severe 11.51% and 1.23% when using the optimal event window length, respectively. The differences in negative impacts to equity prices become larger in magnitude because of event window choice when examining the responses of the second-month S&P 500 E-mini futures contract. Interestingly, the larger effects of event window choice on the stock price changes have much larger variance across the monetary policy shocks when compared to the effects seen on interest rate changes.

Similar to the earlier investigation on interest rates, the response of stock prices to the target factor from GSS becomes weaker by 11.99% when using the median optimal window length. Although the estimates are still statistically significant at the 1% level and go in the correct direction, the weaker effects are possibly due to the event window lengthening causing a "shift in importance" towards the path factor. Using the optimal window length allows more time for the stock market to fully react to the FOMC statements. The longer window captures market participants quickly reacting to a negative shock to the current federal funds rate, then fully reacting to process the effects of tightening forward guidance and other "soft" information. With regards to the central bank information shock, one possible explanation for the difference in effects could have to do with the aforementioned combination of the larger idiosyncratic volatility of stock prices (relative to Treasury yields) and the imposed sign restrictions of JK. As discussed in Subsection 5.1, the S&P 500 Index under-reacts to monetary policy announcement ex post, but also experiences frequent instances of over-reaction through price reversals. By expanding the event window to the median optimal length of 50 minutes, these two factors can result in different decompositions of the central bank information shock, amplifying the series in both directions and weakening its overall effect on stock prices.

I note that the effects of event window choice on the Jarociński and Karadi (2020) shocks shouldn't be interpreted in isolation because of the sign restrictions method behind extracting JK_{MP} and JK_{CBI}. Because the two shocks are defined based on the co-movement with changes in equity prices, this restriction should naturally cause both shocks to have large impacts on the S&P 500 Index and its E-mini futures. Nonetheless, it is still seen that other monetary policy shocks containing forward guidance, like GSS_P, negatively impact equity prices and futures by similar magnitudes to those from JK_{MP}.

Responses of the Macroeconomy via Local Projections To study the effects of event window choice on the impact of monetary policy transmission into macroeconomic variables, I estimate impulse responses using a local projection approach. Following Gertler and Karadi (2015) and Bauer and Swanson (2023), I include the log of CPI, the log of IP, the nominal two-year Treasury yield, and the EBP. Measured by Gilchrist and Zakrajšek (2012), the EBP represents the effective risk-bearing capacity of the financial sector and is essentially the difference between private and public bonds. Including the variable in the model specification is what allows for the transmission of monetary policy shocks from financial markets into economic variables. In order to have all variables be the same frequency, I convert all shock series into monthly frequency such that months with FOMC meetings have a monetary policy shock of zero. Summary statistics for the monthly specification variables and shock series can be found in Appendix Tables F7 and F8.

I use the lag-augmented local projection method from Olea and Plagborg-Møller (2021) to estimate and graph the impulse response functions of each outcome variable from each monetary policy shock series. The primary reason why I choose this approach is because adding an additional lag to the local projection specification as controls results in correct inference with using only Eicker-Huber-White heteroscedasticity-robust standard errors, avoiding the need to correct for serial correlation in the local projections. Therefore, I run a separate regression of each outcome variable on each shock:

$$y_{t+h}^{i,l} = \theta^{i,k,l}(Shock)_t^{k,l} + controls + \eta^{i,k,l} \tag{9}$$

where i indexes one of the four outcome variables, k denotes the considered monetary policy shock, l denotes whether the shock was constructed within the optimal event window length or 30 minutes, and h is the number of months in the future. The impulse response functions are responses of macroeconomic variables to a 100 basis point increase in the monetary policy shock from the common re-scaling to be one-for-one with the zero-coupon, nominal one-year Treasury yield. For all the shock series, this represents a contractionary policy shock.

Figure 9 visually compares the impulse responses of CPI to the different monetary policy shock series constructed in either the optimal event window length or 30 minutes.[36] Initial impressions are that the point estimates of the impulse responses are qualitatively similar regardless of event window choice for all outcome variables. Whilst there exists some differences (e.g., the decline in CPI is more severe in response to NS_{MP} at farther horizons when the shock is constructed within the optimal window length), the lengthening of the event window doesn't result in dramatic changes in the impulse responses. However, what is visually noticeable is that the confidence intervals from the optimal window length are smaller than those from the 30-minute window for most responses. To formally measure this difference, I first calculate the width of the confidence intervals for each time horizon of each macroeconomic variable response to each monetary policy shock series constructed within the optimal window length or 30 minutes. I then take the ratio of the confidence interval width using the optimal window length to the confidence interval width using the 30-minute window. This ratio being less (greater) than one implies that the impulse responses to monetary policy shocks constructed using the optimal window length are more (less) precise compared to those when using 30 minutes. Calculating the average and median

[36]Because the discussed results in the main text are robust to the response comparisons of the other macroeconomic variables, they are visually presented in Appendix Figures E16–E18.

confidence interval width ratio yields 0.9173 and 0.9410, respectively. In other words, constructing monetary policy shocks within the optimal median window length of 50 minutes results in *more precise* responses of macroeconomic variables.

Overall, the heterogeneity in the differences amongst the responses of interest rates, stock prices, and the broader macroeconomy support the findings of Brennan et al. (2024): the data on long-term rates and construction method used for monetary policy shocks are important contributors. My paper provides evidence that the dimension of event window length must also be considered. By choosing the popular 30-minute window in the literature, the effects of monetary policy shocks about forward guidance on financial variables can more or less be dampened because markets have not yet fully reacted to the policy announcements. Furthermore, constructing monetary policy shocks within the event window length where the average impact of noise is minimised results in the estimated responses of macroeconomic variables to be more precise.[37]

7 Statement Features and Their Effects on Windows

Recall that systematic estimation of the event window lengths that best reflect the full reactions of markets is the neural network regressing the changes in asset prices within different windows on the text of FOMC statements. In other words, systematic estimation could be thought of as "jointly" estimating the and optimal event window length and text-based signal with greatest precision. Unfortunately, the financial and mandatory hardware constraints required by the neural network to perform this "joint" estimation are why the current version of this paper is only able to consider event window lengths ending up to 60 minutes after FOMC statement release.

An alternative method that potentially circumvents these constraints and provides some initial insights is what I call the "one signal" approach. First, assume that the predicted price changes within the "jointly" estimated optimal window (i.e., the most precise signal)

[37]These confidence interval width ratios are robust to the choice of monetary policy surprise instruments.

is *constant* for all time. As a result, event window lengths longer than an hour after FOMC statement release can estimated by following the motivating framework more closely. Before detailing the analyses, I note that the core assumption of the "one signal" approach is a strong one. Recall that systematic estimation is the neural network searching for the simultaneous combination of optimal event window length and signal. Because the weights and embeddings of models like XLNet-Base are sensitive to changes in both its input and output variables, it is entirely possible that the "jointly" estimated event window length is different from that which has the largest $\overline{R^2_{OOS}}$ under the "one signal" approach.[38]

The "one signal" approach uses the text-based signal from the "jointly" estimated optimal event window and calculates $\overline{R^2_{OOS}}$ for all event window lengths longer than the optimal window length. This shortcut provides two advantages. First, I can observe how the different dimensions of monetary policy communication potentially affect the optimal event window lengths. Second, I can check if there exists an event window beyond the optimal window length that has a larger $\overline{R^2_{OOS}}$, which I perform and discuss in the Appendix D.

7.1 How FOMC Statement Characteristics Affect Window Lengths

An interesting question to consider is if different aspects within the text of FOMC statements affect the optimal event window length. I use the "one signal" approach to provide some insight into whether the complexity of the information content in FOMC statements, the degree of similarity between sequential statements, and the mention of dissents within the statements change the optimal event window lengths. These heterogeneity exercises rely on the core assumptions that the text-based signal associated with the "jointly" estimated optimal event window is *constant* for all time *and* that this signal is still relevant when conditioning the FOMC statements into subsamples.[39]

Text Complexity I first observe how the optimal event window length changes for fu-

[38] A rough analogy is how iterative generalised method of moments does not always not converge.

[39] The ideal solution would be to acquire large subsamples of FOMC statements satisfying the these characteristics, then have XLNet-Base systematically estimate the optimal event window lengths for all characteristic subsamples.

tures maturities when conditioned on the complexity of the FOMC statements. I assess the *semantic* complexity using the standard readability formula of the Flesch-Kincaid Grade Level index. These readability formulas determine the grade level required for a reader to properly read and comprehend a provided text based on sentence structure (i.e., average sentence length), word structure (i.e., percentage of difficult words in a sentence), and word phonology (i.e., average number of syllables per word).[40]

I first split my FOMC statements depending on whether their grade levels are below or above the median grade level (i.e., 16.5, which is the comprehensive level of some degree of a master's education). I then calculate the MSEs under the "one signal" approach for each subsample of FOMC statements for every futures maturity. A summary of this exercise is found in the first two columns of Table 5, which displays the minimised MSEs and associated event window lengths averaged across all asset types and maturities. Immediately, one can see that FOMC statements with a required readability level of at least some of a master's degree have a longer average optimal event window of 71 minutes. In contrast, "simpler" statements with a grade level up to some of a master's degree is associated with a shorter average event window length of 60 minutes. Interestingly, both subsamples of FOMC statements possess similar averaged MSEs. One potential explanation for this combination is that the price responses of financial markets to "complicated" statements are relatively affected more by cognitive noise. This association is consistent with the idea that more complex information results in differing interpretations and reactions to said information, causing increased market volatility and trading volume. Indeed, papers such as Smales and Apergis (2017) have shown that the increasing complexity of FOMC statement language over time is associated with increasing volatility across equity, bond, and currency markets. In other words, financial markets might need more time to fully react to FOMC statements with more

[40]Appendix Figure E19 displays the evolution of the Flesch-Kincaid Grade Level for my sample period of FOMC statements. Descriptive statistics for the text complexity measure can be found in right column of Appendix Table F9.

complicated language.[41]

Text Similarity Prior studies (e.g., Acosta and Meade, 2015; Handlan, 2022a) have documented that whilst FOMC statements are relatively concise, the semantics and language have varied more often than typically assumed over time. Using the "one signal" approach, I also analyse how the optimal event window length changes when conditioned on the degree of similarity between the FOMC statements. The method I use for this exercise is a bag-of-words model, which represents text as an unordered, weighted-frequency of words. Importantly, this method assumes that documents with similarly weighted frequencies for words are discussing similar topics. By transforming the FOMC statements into vectors of word frequencies, I can perform vector analysis and create a measure of similarity between any two document vectors.

The weighted frequency of words frequently used in text analysis is called Term Frequency-Inverse Document Frequency (TFIDF). Conversion of my sample of FOMC statements into the TFIDF matrix is done through the Python text analysis library of sklearn. Specifically, each word in a document is multiplied by the number of times it occurs within a particular document over the frequency of documents where the word appears. The purpose behind this ratio is to properly weigh down the importance of frequently occurring words that don't actually provide much meaning with regard to distinguishing FOMC statements from one another. For example, words such as "a" and "the" appear in all documents with high frequency, but don't signal anything about any particular statement. Therefore, the value for these words are weighed down by the high number of occurrences across documents, yielding a *weighted frequency*. Additionally, words like "unemployment", "inflation", and "federal funds rate" will also be weighed down because monetary policy discussed in every FOMC statements has these components. However, it should be noted that these words are weighed

[41] Further research is needed into understanding how the overall complexity of monetary policy communications affects the ability of people, even technically trained professionals like financial market participants, to fully comprehend and respond to these announcements. As shown in McMahon (2023) for example, efforts by the Bank of England to simplify their communications have resulted in decreases to semantic complexity over time, but a simultaneous increase in motivating complexity over time.

down by a smaller degree relatively speaking because they could still convey some importance when discerning the degree of similarity between two statements (e.g., if one statement discusses about inflation relatively more than another statement that focuses more on labour market conditions). Lastly, words that do not appear in every FOMC statement, such as "persists" or "tight", are given a higher weighted frequency because they potentially signal a unique economic environment when the FOMC meeting took place compared to other meeting dates. The differences in these unique and informative words are what allow me to calculate the degree of similarity between FOMC statements.

The calculation of the TFIDF matrix is as follows: Let D be the set of FOMC statements and T be the set of all terms that appear within and across all statements in D. Let a single document and term be indexed by $d \in D, t \in T$, respectively. Therefore, the term frequency of a particular term t in document d can be defined as:

$$tf_{d,t} = \ln\left(\frac{tc_{d,t}}{nt_d}\right) + 1, \tag{10}$$

where $tc_{t,d}$ is the number of times term t appears in document d, nt_d is the total number of words in d, and both the natural logarithm and addition term serve to smooth out the frequency measure. We can then calculate the inverse document frequency as follows:

$$idf_{d,t} = \ln\left(\frac{nd}{df_{d,t} + 1}\right) + 1, \tag{11}$$

where nd is the number of documents in set D and $df_{t,d}$ is the number of documents that term t appears. Multiplying the two terms gives us the weighted frequency of each term t in document d:

$$TFIDF_{d,t} = tf_{d,t} * idf_{d,t}. \tag{12}$$

Calculating $TFIDF_{d,t} \forall d \in D, t \in T$ and combining them yields a $D \times T$ matrix, where each element $TFIDF_{d,t}$ represents the number of times a term appears in a particular document, divided by the number of words in said document, then multiplied with the document frequency (ratio between the number of documents where the term appears and the total number of documents). Essentially, the higher the TFIDF value is for a term, the more informative that word is with regard to distinguishing the information content of those FOMC statements from others.

Although the TFIDF matrix can be produced on the current pre-processed FOMC statements, the results might not be accurate in representing the similarity of the intended or comprehended information content of the statements. For example, the method could potentially view the terms "Federal Funds Rate" and "federal funds rate" as different. In other words, this bag-of-words model is not as "sophisticated" at discerning context from the FOMC statements in their original grammatical structure compared to the neural network, causing the $TFIDF_{d,t}$ values to be inaccurate. I mitigate this issue by deploying additional preprocessing steps on my sample of FOMC statements. First, I make every word lowercase. Second, I remove common words that convey little semantic meaning, such as articles, pronouns, and conjunctions.[42] The last step involves converting all words into their "base" form (e.g., the words "increases", "increasing", and "increase" are all combined into the base form of "increas"). Calculating the TFIDF matrix on the additionally cleaned FOMC statements from May 1999 to October 2019 yields a matrix with 165 rows (statements) and 966 columns (terms).[43]

With the TFIDF matrix calculated, I am able to produce a corresponding matrix whose values represent a measure of similarity between any two pairs of FOMC statements. Specifically, this document similarity matrix is created by multiplying the TFIDF matrix with its transpose:

$$\text{Document Similarity Matrix} = TFIDF \cdot TFIDF^T. \tag{13}$$

[42] I choose to retain numbers because the FOMC's targetted federal funds rate or discussions about forward guidance convey important semantic content.

[43] Appendix Table F10 presents a list of 30 base terms that have the highest TFIDF scores for my sample of FOMC statements. Recall that terms with higher TFIDF values are more informative terms with regard to differentiating the FOMC statements with them from statements without.

Consider a document d row vector of the TFIDF matrix. For each term $t \in T$, the $TFIDF_{d,t}$ values are only positive if the terms appear in document d, otherwise the value is zero. Additionally, the vectors for every document all have the same magnitude due to normalisation by the number of documents and terms in the matrix. Because the product between the TFIDF matrix and its transpose is essentially the dot product between every pair of document row vectors in the sample of FOMC statements, I can define the degree of similarity between any two FOMC statements as the cosine of the angle between these documents, also called the *cosine similarity*:

$$S^{A,B} := \cos\theta = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|}, \tag{14}$$

where $\mathbf{A}, \mathbf{B}$ are the vectors for FOMC statements A and B, respectively. The more similar two FOMC statements are, the angle between the two corresponding vectors approaches zero, resulting in the cosine of the angle to go to one. A cosine similarity value of one between two FOMC statements means they are identical, whilst a value of zero means the two documents have no common words. In other words, the more base terms shared between two FOMC statements, the higher cosine similarity value is assigned to the pair.

The Document Similarity Matrix for FOMC statements from May 1999 to October 2019 allows for me to compare the degree of similarity between any two documents for any points of time d and d'. The series of document-pairs $(d, d-1)$ represents the degree of similarity between sequential FOMC statements, which I call as S^1 for the rest of the paper and is plotted in Figure 10.[44] As time progresses, the cosine similarity between sequential FOMC statements increases. The increased persistence and standardisation of FOMC statement wording begins roughly between early-2009 and early-2010, which coincides with the period when the FOMC was in the process of reducing the federal funds rate towards the zero lower bound. As discussed in Handlan (2022a), one explanation behind this occurrence is that the FOMC wanted to avoid market surprises through reducing the variation in statement

[44] Appendix Figure E20 represents the Document Similarity Matrix for my sample of FOMC statements as a heat map. Descriptive statistics for S^1, the cosine similarity between sequential FOMC statements, can be found in Appendix Table F9.

vocabulary after the financial crisis. Even when the zero lower bound ended with the Board raising the federal funds rate at the end of 2015, the cosine similarity measures remained high through 2019 due to the efforts by the FOMC to increase the transparency of its monetary policy under Chairs Yellen and Powell.

I split my FOMC statements depending on whether S^1 value is below or above the median (i.e., 0.88), labelling the subsample of statements as being "different" or "similar" relative to their immediately previous document, respectively. Calculating the MSEs for each subsample of FOMC statements for every futures maturity, I present the results of this exercise in the third and fourth columns from the left of Table 5. "Different" FOMC statements observe the smallest MSEs within an event window length of 62 minutes when averaged across all asset types and maturities, whilst "similar" statements see minimised MSEs within a 51-minute window on average. Similar to the condition of text complexity, both subsets of statements have similar minimised MSEs on average. Intuitively, if an FOMC statement contains less similar base terms and associated ideas with its predecessor, financial markets could require more time in order to discern and react to the new information of the current FOMC statement. Conversely, markets would presumably react to statements that share a lot of terms and concepts with past policy announcements in a short amount of time because most of said information has already been incorporated into the price.

Presence of Dissents The final characteristic of the FOMC statements I consider under the "one signal" approach is whether the presence of dissents in the monetary policy announcements affects the optimal event window lengths of asset types. Similar to prior studies, I find that roughly 40% of FOMC statements in my sample have recorded dissents. Whilst this suggests dissents are not rare, they still provide important signals. According to Federal Reserve tradition, dissents are usually recorded only when Board members find the majority's opinion unacceptable (C Madeira and J Madeira, 2019). In other words, the very existence of a recorded dissent provides additional information for markets to process, which in turn could affect their reaction time. To investigate this characteristic, I split my sample of FOMC statements depending on whether the policy votes were unanimous or not.[45] The fifth and sixth columns of Table 5 present the calculated MSEs for each subset of statements. On average, statements with(out) dissents observe the smallest MSE under the "one signal" approach within an event window length of 83 (53) minutes. The longer reaction time for statements with dissents may be because they introduce additional layers of information regarding policy transparency and internal disagreement, requiring more time for markets to process compared to the unanimous announcements. It should be noted that prior studies (e.g., C Madeira and J Madeira, 2019; Tsang and Yang, 2024) have shown that the presence of dissents have a nuanced impact on financial markets. Whilst equity markets have a strong reaction to disclosed information contained in dissents, interest-rate markets exhibited muted impulse responses. The findings of my paper show that when averaged across all asset types (i.e., interest-rate based and equities), financial markets are impacted through needing more time to fully react to FOMC statements with the presence of dissents.

Overall, I show that monetary policy communication affects the response of financial markets along multiple dimensions beyond the "hard" information within the FOMC statements (i.e., the change to the federal funds rate). Because monetary policy communication is expressed with varying degrees of semantic complexity, similarity to prior policy announcements, and heterogeneous voting records, these characteristics overall can influence the time financial markets need to fully respond to the "soft" information within FOMC statements, such as forward guidance. Indeed, future research towards understanding the effects of these dimensions could provide deeper insight into how monetary policy communication exactly affects financial markets and in turn the broader macroeconomy.

8 Conclusion

The choice of event window length for measuring financial market responses to news has largely remained an "ad-hoc" decision in the empirical monetary policy literature. This

[45] I do not record and distinguish between dissent votes for tighter policy, easier, policy, or other reasons.

paper challenges the conventions that have emerged from this practice by asking a crucial question: how does one choose the appropriate event window length for their study in monetary policy? By combining observed price dynamics with a text-based signal derived using neural networks methods for text analysis, I develop and implement a methodology to systematically estimate the event window lengths that best reflect the full market reaction to the information content of FOMC statements.

This paper's systematic estimation yields two key findings that directly confront the common assumptions in the literature. First, the common 30-minute window is insufficient time for markets to fully react to FOMC announcements. My results show that, on average, markets fully react within an event window *ending at least 30 minutes after release*, corresponding to a total window length of at least 40 minutes. A 30-minute window is never found to be optimal for any asset type considered. Second, the optimal window length is not a universal parameter. In line with this paper's title, one window does not fit all; the optimal duration increases systematically with asset underlying maturity. This estimated length rises from 40 minutes for assets with the shortest underlying maturities to 50-60 minutes for futures with underlying maturities of two quarters or more. Relatedly, statements with greater complexity, less similarity, and the presence of dissents are associated with longer event windows on average.

The implications of these findings are not merely methodological, as the choice of event window has a tangible impact on the construction and interpretation of monetary policy surprises and shocks. I document that the correlation between surprises measured within the optimal windows and those from a conventional 30-minute window decreases with asset underlying maturity. These discrepancies are economically meaningful, resulting in observed differences in derived monetary policy shocks, particularly the forward guidance component. Furthermore, the responses of interest rates and break-even inflation to monetary policy shocks about forward guidance become larger when opening up the event window to the optimal length. This larger estimated impact suggests that shorter, conventional windows attenuate the measured effects of policy. Similar results are found in the response of stock prices to these shock series. With regards to the broader macroeconomy, I document that monetary policy shocks constructed within the optimal event window length result in the estimated responses of macroeconomic variables to become more precise.

Ultimately, this paper argues that determining the optimal event window length is an empirical question, not a parameter whose value should be simply assumed. The provided methodology offers a systematic approach for this basic but critical step in event window studies. The results indicate that by allowing for longer, asset-specific reaction times, we can construct monetary policy surprises that more accurately reflect the market's full response to central bank communication, leading to a more precise understanding of the effects of monetary policy. Indeed, as the literature considers broader measures of central bank communication in order to construct more holistic monetary policy shocks (e.g., Neuhierl and Weber, 2019; Swanson and Jayawickrema, 2023; Bauer and Swanson, 2023; and others), I argue that when the primary source of reaction is communicated through text, the methodology of this paper can be used to measure when markets fully react to these policy communications, resulting in more relevant and precise measurements of unanticipated monetary policy.

References

Acosta, Miguel (2023). "The Perceived Causes of Monetary Policy Surprises". Published manuscript.

Acosta, Miguel, Connor M. Brennan, and Margaret M. Jacobson (2024). *Constructing high-frequency monetary policy surprises from SOFR futures*. Tech. rep. Finance and Economics Discussion Series 2024-034. Board of Governors of the Federal Reserve System.

Acosta, Miguel and Ellen E. Meade (2015). *Hanging on Every Word: Semantic Analysis of the FOMC's Postmeeting Statement*. Tech. rep. FEDS Notes. Board of Governors of the Federal Reserve System.

An, Phillip, Karlye Dilts Stedman, and Amaze Lusompa (2025). *How High Does High Frequency Need to Be? A Comparison of Daily and Intradaily Monetary Policy Surprises*. Tech. rep. Research Working Paper no. 25-03. Federal Reserve Bank of Kansas City.

Athey, Susan and Guido Imbens (2019). *Machine Learning Methods Economists Should Know About*.

Bauer, Michael D., Aeimit Lakdawala, and Philippe Mueller (2022). "Market-Based Monetary Policy Uncertainty". In: *The Economic Journal* 132.644, pp. 1290–1308.

Bauer, Michael D. and Eric T. Swanson (2023). “A Reassessment of Monetary Policy Surprises and High-Frequency Identification”. In: *NBER Macroeconomics Annual* 37.1, pp. 87–155.

Bazzana, Flavio and Andrea Collini (2020). “How does HFT activity impact market volatility and the bid-ask spread after an exogenous shock? An empirical analysis on S&P 500 ETF”. In: *The North American Journal of Economics and Finance* 54, p. 101240.

Ben Ammar, Imen and Slaheddine Hellara (2022). “High-frequency trading, stock volatility, and intraday crashes”. In: *The Quarterly Review of Economics and Finance* 84, pp. 337–344.

Ben-David, Itzhak, Francesco Franzoni, Byungwook Kim, and Rabih Moussawi (2022). “Competition for Attention in the ETF Space”. In: *The Review of Financial Studies* 36.3, pp. 987–1042.

Bianchi, Francesco, Sydney C Ludvigson, and Sai Ma (2024). *What Hundreds of Economic News Events Say About Belief Overreaction in the Stock Market*. Working Paper 32301. National Bureau of Economic Research.

Boehm, Christoph E and T Niklas Kroner (2025a). “Monetary Policy without Moving Interest Rates: The Fed Non-Yield Shock”. SSRN Working Paper No 3812524.

Boehm, Christoph E and T Niklas Kroner (2025b). “The U.S., Economic News, and the Global Financial Cycle”. In: *The Review of Economic Studies*, rdaf020.

Boguth, Oliver, Adlai J. Fisher, Vincent Gregoire, and Charles Martineau (2023). “Noisy FOMC Returns? Information, Price Pressure, and Post-Announcement Reversals”. SSRN Working Paper No 878091.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2024). “Belief Overreaction and Stock Market Puzzles”. In: *Journal of Political Economy* 132.5, pp. 1450–1484.

Brennan, Connor M., Margaret M. Jacobson, Christian Matthes, and Todd B. Walker (2024). *Monetary Policy Shocks: Data or Methods?* Tech. rep. Finance and Economics Discussion Series 2024-011r1. Board of Governors of the Federal Reserve System.

Brooks, Jordan, Michael Katz, and Hanno Lustig (2023). *Post-FOMC Announcement Drifts in U.S. Bond Markets*. Working Paper 25127. National Bureau of Economic Research.

Casini, Alessandro and Adam McCloskey (2025). *Identification and Estimation of Causal Effects in High-Frequency Event Studies*. arXiv: 2406.15667 [econ.EM].

Chang, Son J and Son-Nan Chen (1989). “Stock-price adjustment to earnings and dividend surprises”. In: *Quarterly Review of Economics and Business* 29.1, pp. 68–81.

Crump, Richard K., Stefano Eusepi, and Emanuel Moench (2013). *Making a Statement: How Did Professional Forecasters React to the August 2011 FOMC Statement?* Liberty Street Economics 20130107. Federal Reserve Bank of New York.

Das, Somnath and Alexander Z. King (2021). “Measuring the informativeness of earnings announcements: The role of event windows”. In: *The Quarterly Review of Economics and Finance* 82, pp. 350–367.

Doh, Taeyoung, Dongho Song, and Shu-Kuei Yang (2023). “Deciphering Federal Reserve Communication via Text Analysis of Alternative FOMC Statements”. Federal Reserve Bank of Kansa City Working Paper.

Duffee, Gregory (2013). "Forecasting Interest Rates". In: *Handbook of Economic Forecasting*. Ed. by Graham Elliott and Allan Timmermann. Vol. 2. Handbook of Economic Forecasting. Elsevier. Chap. 7, pp. 385–426.

Fleming, Michael J. and Monika Piazzesi (2005). *Monetary Policy Tick-by-Tick*. Tech. rep. Working Paper. Federal Reserve Bank of New York.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy (2019). "Text as Data". In: *Journal of Economic Literature* 57.3, pp. 535–74.

Gertler, Mark and Peter Karadi (2015). "Monetary Policy Surprises, Credit Costs, and Economic Activity". In: *American Economic Journal: Macroeconomics* 1, pp. 44–76.

Gilchrist, Simon and Egon Zakrajšek (2012). "Credit Spreads and Business Cycle Fluctuations". In: *American Economic Review* 102.4, pp. 1692–1720.

Glasserman, Paul and Caden Lin (2023). *Assessing Look-Ahead Bias in Stock Return Predictions Generated By GPT Sentiment Analysis*.

Golez, Benjamin, Peter Kelly, and Ben Matthies (2025). "FOMC News and Segmented Markets". In: *Journal of Accounting and Economics* 79.2, p. 101767.

Gorodnichenko, Yuriy and Walker Ray (2017). *The Effects of Quantitative Easing: Taking a Cue from Treasury Auctions*. Working Paper 24122. National Bureau of Economic Research.

Graham, Benjamin (1974). *The Renaissance of Value: The Proceedings of a Seminar on the Economy, Interest Rates, Portfolio Management, and Bonds Vs Common Stocks, September 18, 1974*. Financial Analysts Research Foundation.

Gürkaynak, Refet S., Burçin Kisacikoğlu, and Jonathan H. Wright (2020). "Missing Events in Event Studies: Identifying the Effects of Partially Measured News Surprises". In: *American Economic Review* 110.12, pp. 3871–3912.

Gürkaynak, Refet S., Brian Sack, and Eric T. Swanson (2005). "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements". In: *International Journal of Central Banking* 1.1.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright (2010). "The TIPS Yield Curve and Inflation Compensation". In: *American Economic Journal: Macroeconomics* 2.1, pp. 70–92.

Handlan, Amy (2022a). "FedSpeak Matters: Statement Similarity and Monetary Policy Expectations". Published manuscript.

Handlan, Amy (2022b). "Text Shocks and Monetary Surprises: Text Analysis of FOMC Statements with Machine Learning". Published manuscript.

Hawinkel, Stijn, Willem Waegeman, and Steven Maere (2024). "Out-of-Sample R2: Estimation and Inference". In: *The American Statistician* 78.1, pp. 15–25.

Hervé, Fabrice, Mohamed Zouaoui, and Bertrand Belvaux (2019). "Noise traders and smart money: Evidence from online searches". In: *Economic Modelling* 83, pp. 141–149.

Hillmer, S.C. and P.L. Yu (1979). "The market speed of adjustment to new information". In: *Journal of Financial Economics* 7.4, pp. 321–345.

Hornik, Kurt, Maxwell Stinchcombe, and Halbert White (1989). "Multilayer feedforward networks are universal approximators". In: *Neural Networks* 2.5, pp. 359–366.

Husted, Lucas, John Rogers, and Bo Sun (2020). "Monetary policy uncertainty". In: *Journal of Monetary Economics* 115, pp. 20–36.

Indriawan, Ivan, Feng Jiao, and Yiuman Tse (2021). "The FOMC announcement returns on long-term US and German bond futures". In: *Journal of Banking & Finance* 123, p. 106027.

Jarociński, Marek and Peter Karadi (2020). "Deconstructing Monetary Policy Surprises—The Role of Information Shocks". In: *American Economic Journal: Macroeconomics* 12.2, pp. 1–43.

Krivin, Dmitry, Robert Patton, Erica Rose, and David Tabak (2003). "Determination of the Appropriate Event Window Length in Individual Stock Event Studies". SSRN Working Paper No 466161.

Kuttner, Kenneth N. (2001). "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market". In: *Journal of Monetary Economics* 47.3, pp. 523–544.

Madeira, Carlos and João Madeira (2019). "The Effect of FOMC Votes on Financial Markets". In: *The Review of Economics and Statistics* 101.5, pp. 921–932.

McMahon Naylor, Matthew (2023). *Getting Through: Communicating Complex Information*. Tech. rep. Staff Working Paper No. 1047. Bank of England.

Miranda-Agrippino, Silvia and Giovanni Ricco (2021). "The Transmission of Monetary Policy Shocks". In: *American Economic Journal: Macroeconomics* 13.3, pp. 74–107.

Nakamura, Emi and Jón Steinsson (2018). "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect". In: *The Quarterly Journal of Economics* 133.3, pp. 1283–1330.

Neuhierl, Andreas and Michael Weber (2019). "Monetary policy communication, policy slope, and the stock market". In: *Journal of Monetary Economics* 108. "Central Bank Communications:From Mystery to Transparency"May 23-24, 2019Annual Research Conference ofthe National Bank of UkraineOrganized in cooperation withNarodowy Bank Polski, pp. 140–155.

Neuhierl, Andreas and Michael Weber (2024). *Monetary Momentum*. Chicago Booth Research Paper, Fama-Miller Working Paper, University of Chicago, Becker Friedman Institute for Economics Working Paper, LawFin Working Paper 20-05, 2020-39, 37. Chicago Booth Research Paper Series, Chicago Booth: Fama-Miller Working Paper Series, Becker Friedman Institute for Economics Working Paper Series, DFG Center for Advanced Studies - Foundations of Law and Finance.

Newey, Whitney K. and Kenneth D. West (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". In: *Econometrica* 55.3, pp. 703–708.

Okada, Masayuki (2025). "Interest Rate Risk around Monetary Policy Announcements and Asset Duration". Working paper.

Olea, José Luis Montiel and Mikkel Plagborg-Møller (2021). "Local Projection Inference is Simpler and More Robust Than You Think". In: *Econometrica* 89.4, pp. 1789–1823. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA18756.

Piazzesi, Monika and Eric T. Swanson (2008). "Futures prices as risk-adjusted forecasts of monetary policy". In: *Journal of Monetary Economics* 55.4, pp. 677–691.

Piller, Alexander, Marc Schranz, and Larissa Schwaller (2025). "Using Natural Language Processing to Identify Monetary Policy Shocks". Working Paper.

Rajapakse, Thilina C., Andrew Yates, and Maarten de Rijke (2024). "Simple Transformers: Open-source for All". In: *Proceedings of the 2024 Annual International ACM SIGIR*

Conference on Research and Development in Information Retrieval in the Asia Pacific Region. SIGIR-AP 2024. Tokyo, Japan, pp. 209–215.

Ramey, V.A. (2016). "Chapter 2 - Macroeconomic Shocks and Their Propagation". In: *Handbook of Macroeconomics*. Ed. by John B. Taylor and Harald Uhlig. Vol. 2. Elsevier, pp. 71–162.

Rigobon, Roberto and Brian Sack (2004). "The Impact of Monetary Policy on Asset Prices". In: *Journal of Monetary Economics* 51.8, pp. 1553–1575.

Sarkar, Suproteem and Keyon Vafa (2024). "Lookahead Bias in Pretrained Language Models". SSRN Working Paper No 4754678.

Smales, L.A. and N. Apergis (2017). "Does more complex language in FOMC decisions impact financial markets?" In: *Journal of International Financial Markets, Institutions and Money* 51, pp. 171–189.

Swanson, Eric T. and Vishuddhi Jayawickrema (2023). "Speeches by the Fed Chair Are More Important than FOMC Announcements: An Improved High-Frequency Measure of U.S. Monetary Policy Shocks". Unpublished manuscript.

Tsang, Kwok Ping and Zichao Yang (2024). *Agree to Disagree: Measuring Hidden Dissent in FOMC Meetings*. arXiv: 2308.10131 [econ.GN].

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin (2017). "Attention Is All You Need". In: *CoRR* abs/1706.03762.

Yang, Zhilin, Zihang Dai, Yiming Yang, Jaime G. Carbonell, Ruslan Salakhutdinov, and Quoc V. Le (2019). "XLNet: Generalized Autoregressive Pretraining for Language Understanding". In: *CoRR* abs/1906.08237.

Zhang, Yu, Konstantina Kappou, and Andrew Urquhart (n.d.). "Post-FOMC Drift in the Equity Options Market". SSRN Working Paper No 5464630.

Figures and Tables

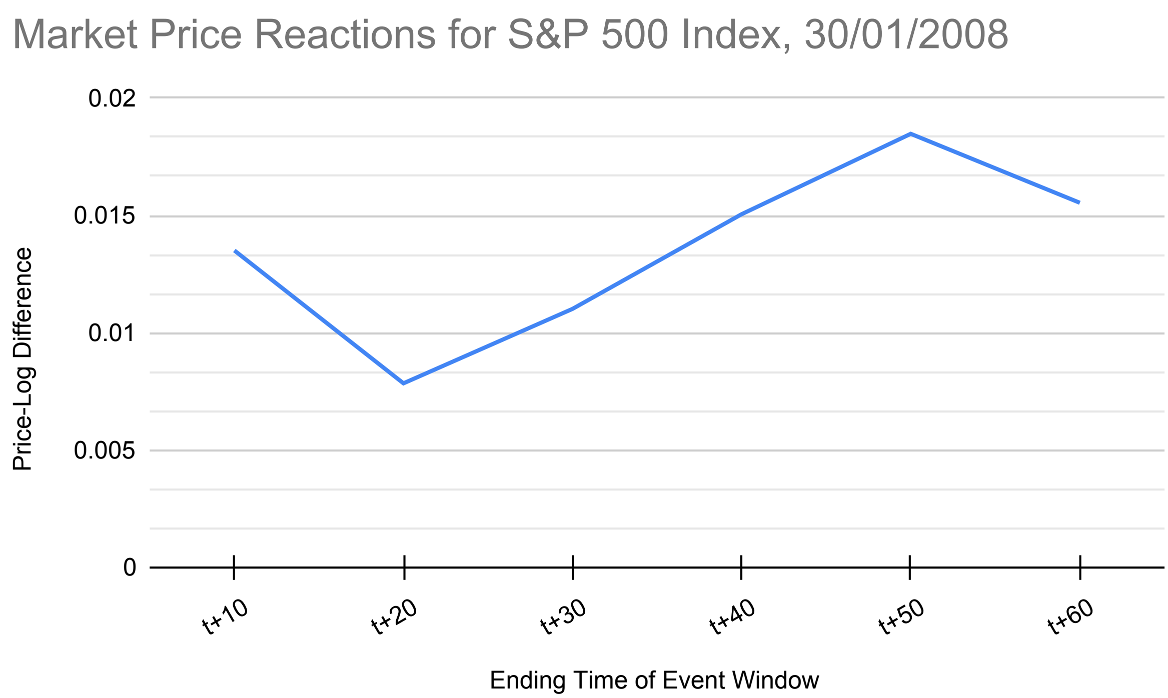


Figure 1: Responses of S&P 500 Index Market to FOMC Statements

Notes: The price log-differences of the S&P 500 Index market for different event window lengths after the release of the 30 January, 2008 FOMC statement is depicted. The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the price-log differences.

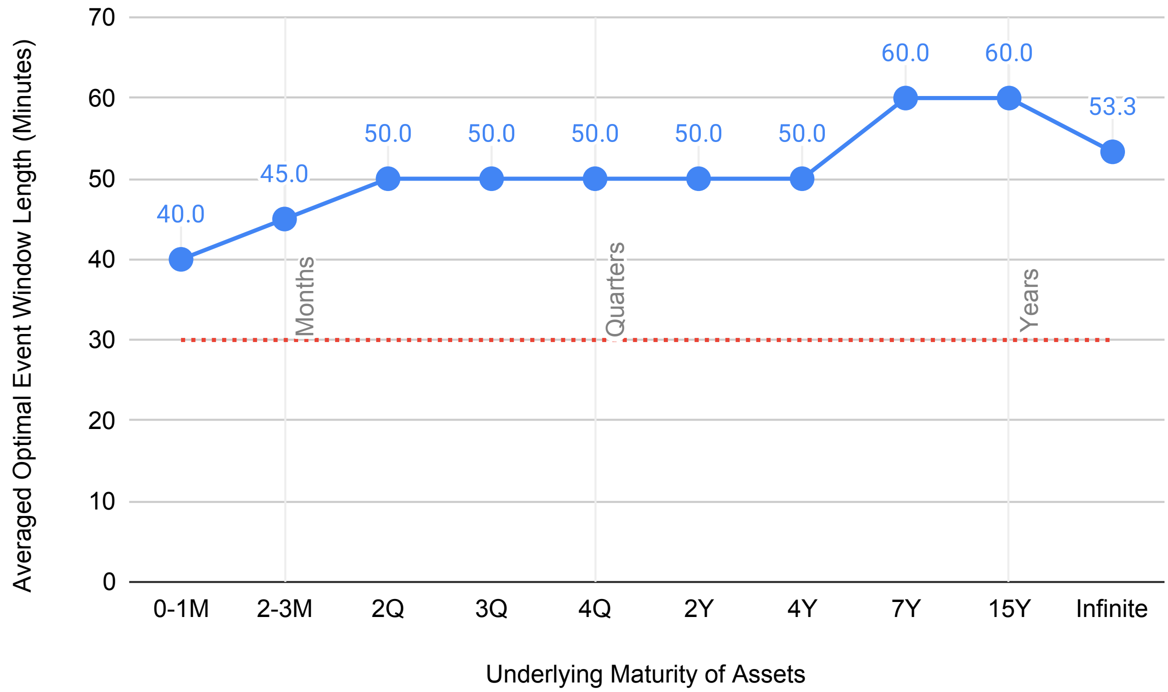


Figure 2: Averaged Optimal Event Window Lengths for Various Asset Underlying Maturities

Notes: The horizontal axis depicts the horizons of all considered futures contracts and equities. The vertical axis depicts the event window lengths in minutes, with all windows starting at 10 minutes before FOMC statement release. The red dotted line depicts the 30-minute window length, common in the literature. The systematically estimated event window lengths are averaged across futures maturities for each asset type. "0–1M" is the average of the event window lengths for front-month and 1-month-ahead Federal Funds futures. "2–3M" is the average of the event window lengths for 2-month and 3-month-ahead Federal Funds futures. "Infinite" is the underlying maturity of the S&P 500 Index and considered E-mini futures contracts. The underlying maturities for the Treasury futures contracts are approximated by Gürkaynak, Kisacikoğlu, et al. (2020).

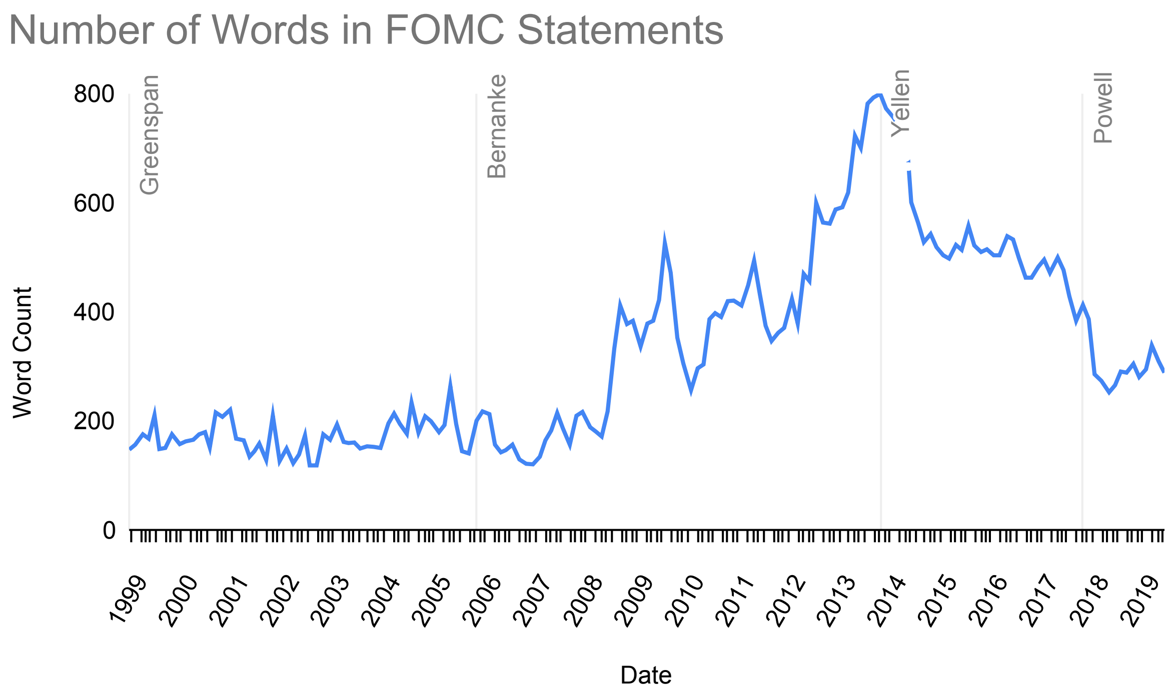
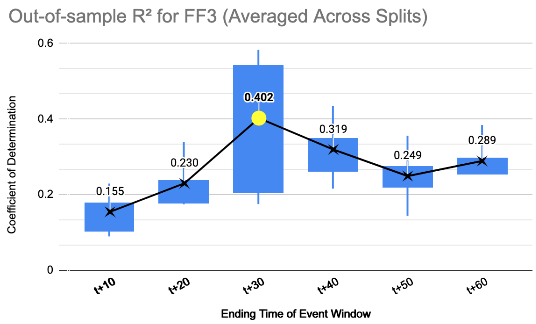
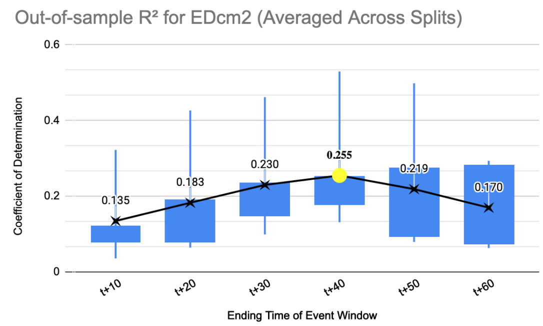


Figure 3: Number of Words in FOMC Statements, May 1999–October 2019

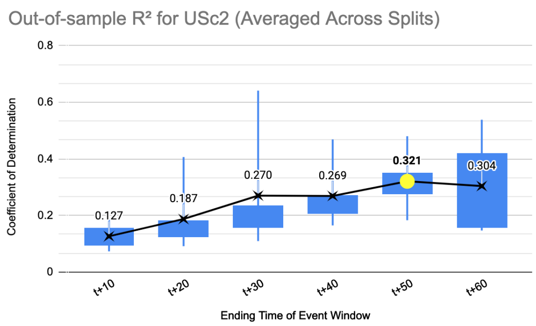
Notes: The above counts are for FOMC statements that have underwent pre-processing, which is explained in Subsection 3.1. From left to right, the vertical grey lines indicate the first FOMC meeting with Greenspan, Bernanke, Yellen, and Powell as Fed Chair.



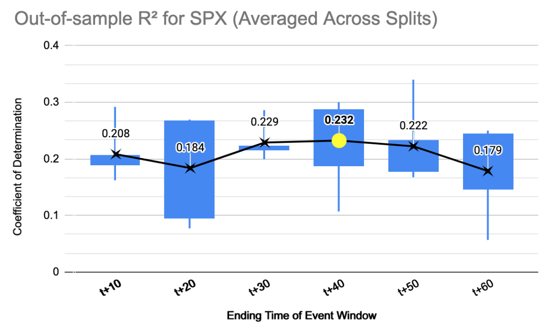
(a) $FF3$



(b) $EDcm2$



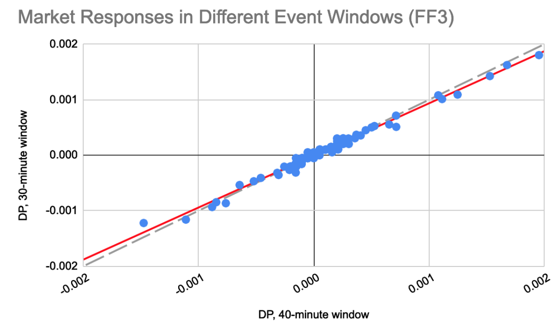
(c) $USc2$



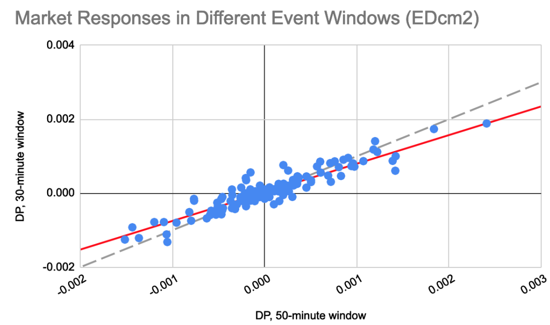
(d) SPX

Figure 4: Systematic Estimation of Optimal Window Sizes for Interest-Rate Futures and S&P 500 Index

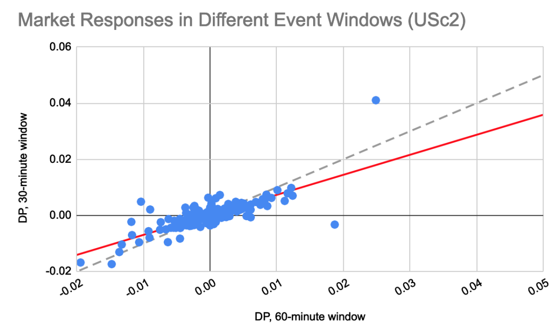
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the percentage points of price changes by the market directly and only due to the FOMC statement text. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $0 \leq \overline{R^2_{OOS}} \leq 1$ represents the proportion of the mean squared error from predicting with the in-sample average explained by the superior out-of-sample performance of XLNet-Base.



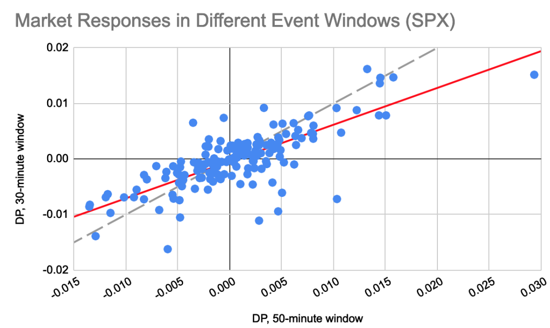
(a) $FF3^{**}$



(b) $EDcm2^{***}$

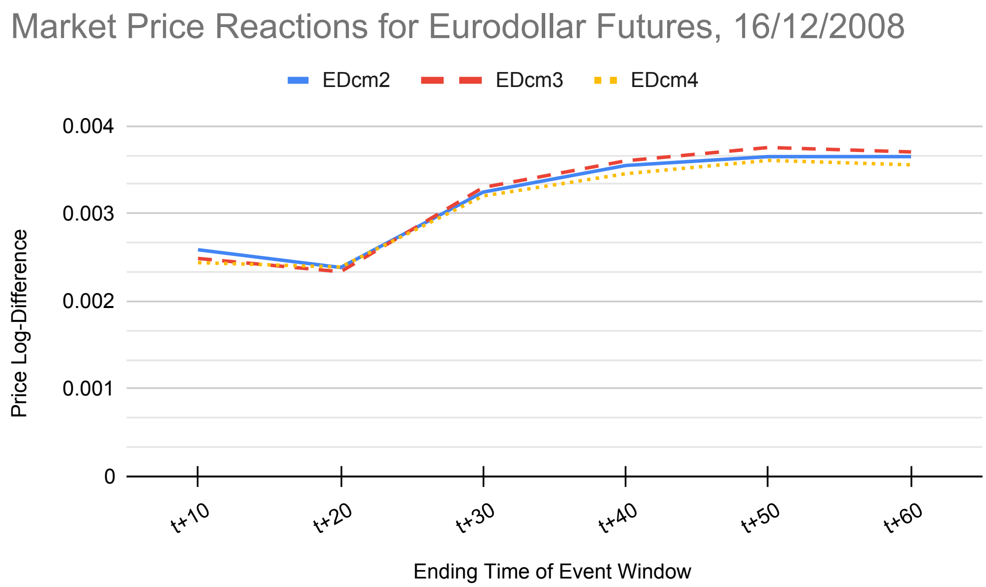


(c) $USc2^{***}$

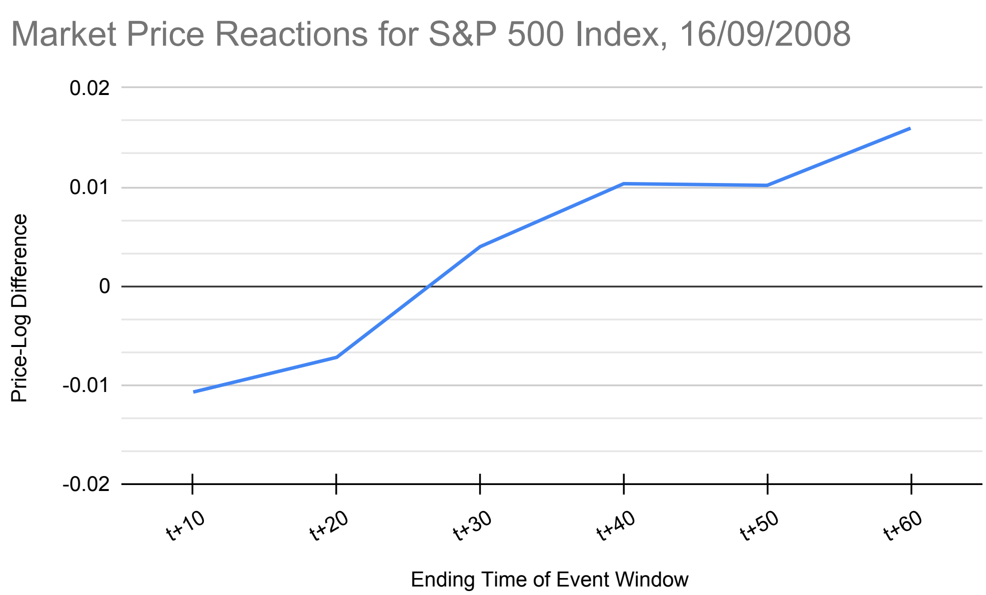


(d) SPX^{***}

Figure 5: Comparing Market Responses in Different Event Windows for Interest-Rate Futures and S&P 500 Index
Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient is greater (less) than one and its difference with one is statistically significant for at most $\alpha = 0.10$, then the financial market under-reacts to the information content of FOMC statements on release, ex post, and is depicted in red. ** sig. at the 5% level, *** sig. at the 1% level.



(a) Eurodollar Futures



(b) SPX

Figure 6: Responses of Eurodollar Futures and S&P 500 Markets to FOMC Statements
Notes: The top (bottom) sub-figure depicts the price log-differences of Eurodollar futures markets (the S&P 500 Index market) for different event window lengths after the release of the 16 December (September), 2008 FOMC statement. The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the price-log differences.

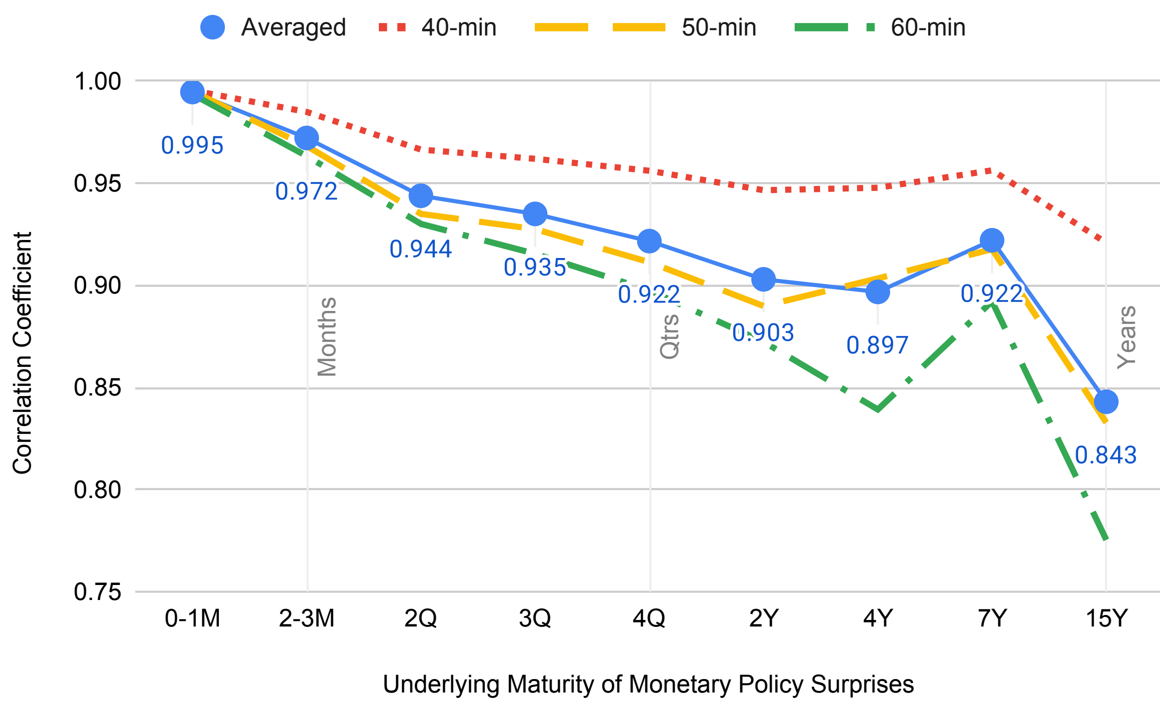
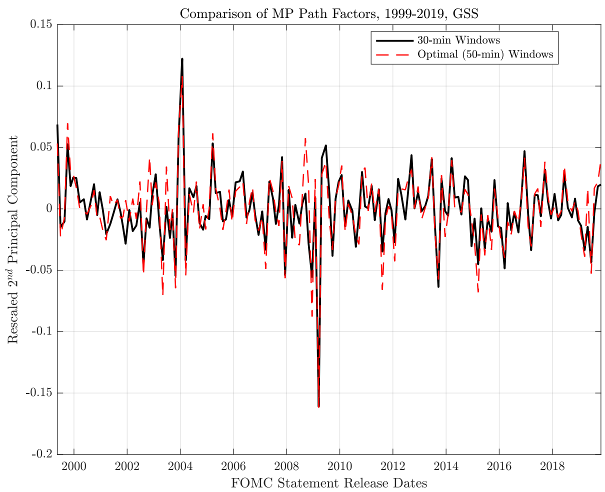
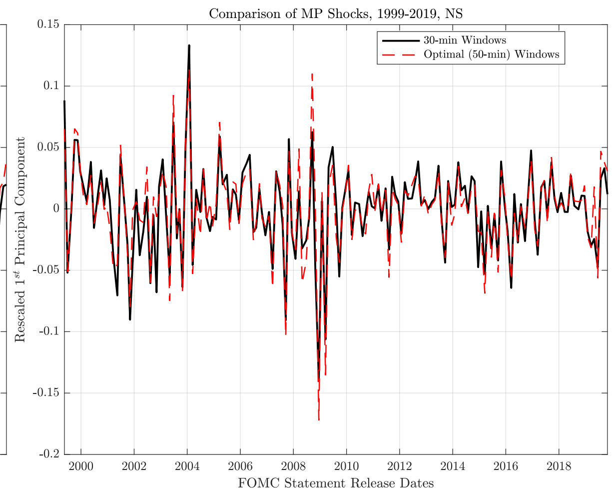


Figure 7: Correlation Between Monetary Policy Surprises Calculated in Optimal v. 30-minute Window Lengths

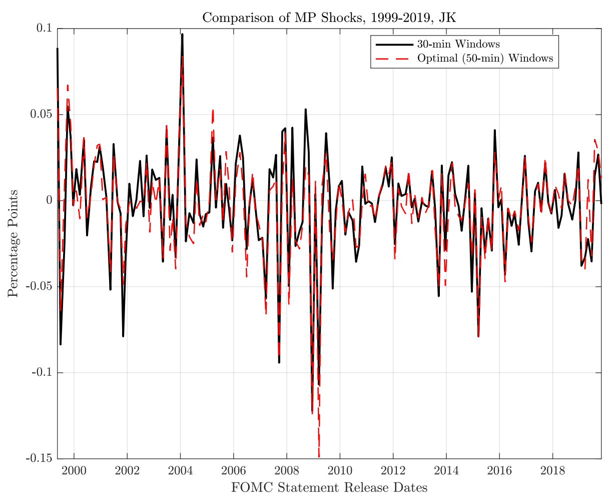
Notes: The horizontal axis depicts the underlying maturities of interest-rate surprises. The vertical axis represents the Pearson correlation coefficient between the surprises calculated within optimal event window lengths v. 30 minutes. The red-dotted, yellow-dashed, and green-mixed lines represent the correlations calculated for surprises measured within the optimally-found 40, 50, and 60 minutes, respectively. The blue-solid line represents the correlations averaged across the three optimal window lengths. "0–1M" is interest rate surprise calculated using information from front-month and 1-month-ahead Federal Funds futures. "2–3M" represent expectations calculated from 2-month and 3-month-ahead Federal Funds futures. The underlying maturities for the Treasury futures contracts and resulting surprises are approximated by Gürkaynak, Kisacikoğlu, et al. (2020).



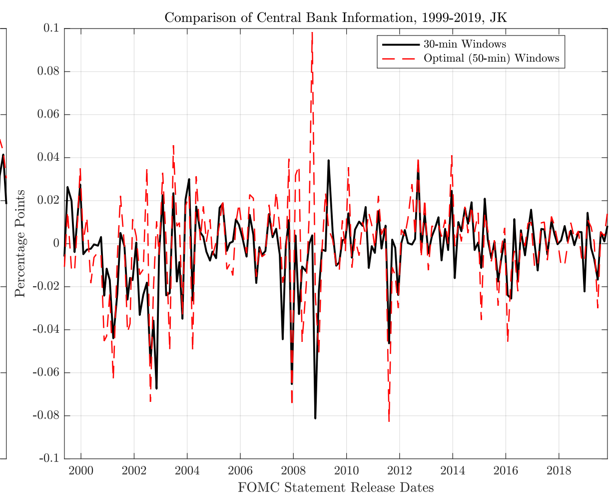
(a) GSS_P



(b) NS_{MP}



(c) JK_{MP}



(d) JK_{CBI}

Figure 8: Comparing Monetary Policy Shock Series Derived from Optimal Window Length v. 30 Minutes

Notes: For each sub-figure, the horizontal axis represents FOMC statement release dates. The vertical axis depicts percentage points after rescaling each shock to be one-for-one with the daily change in the zero-coupon, nominal one-year Treasury yield. For all construction methods, the black-solid and red-dotted lines represent the shocks derived from surprises measured within 30 minutes and the median optimal event window length of 50 minutes, respectively.

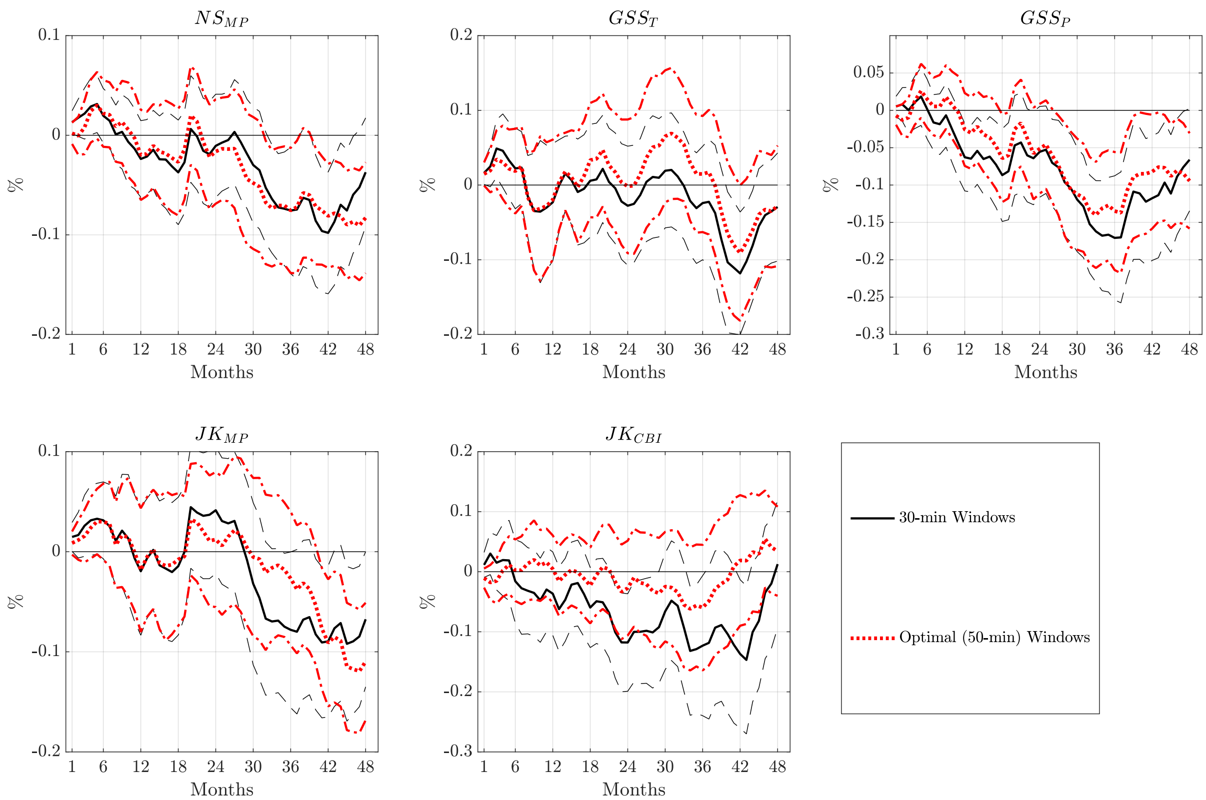


Figure 9: Effects of Event Window Choice on CPI Impulse Responses

Notes: Impulse responses are calculated using the lag-augmented local projection from Olea and Plagborg-Møller (2021). Confidence bands are at the 90% level. Standard errors are Eicker-Huber-White heteroscedasticity-robust standard errors. The responses are to a 100 basis point increase in the monetary policy shock series. The monetary policy shocks are constructed using the full set of monetary policy surprise instruments.

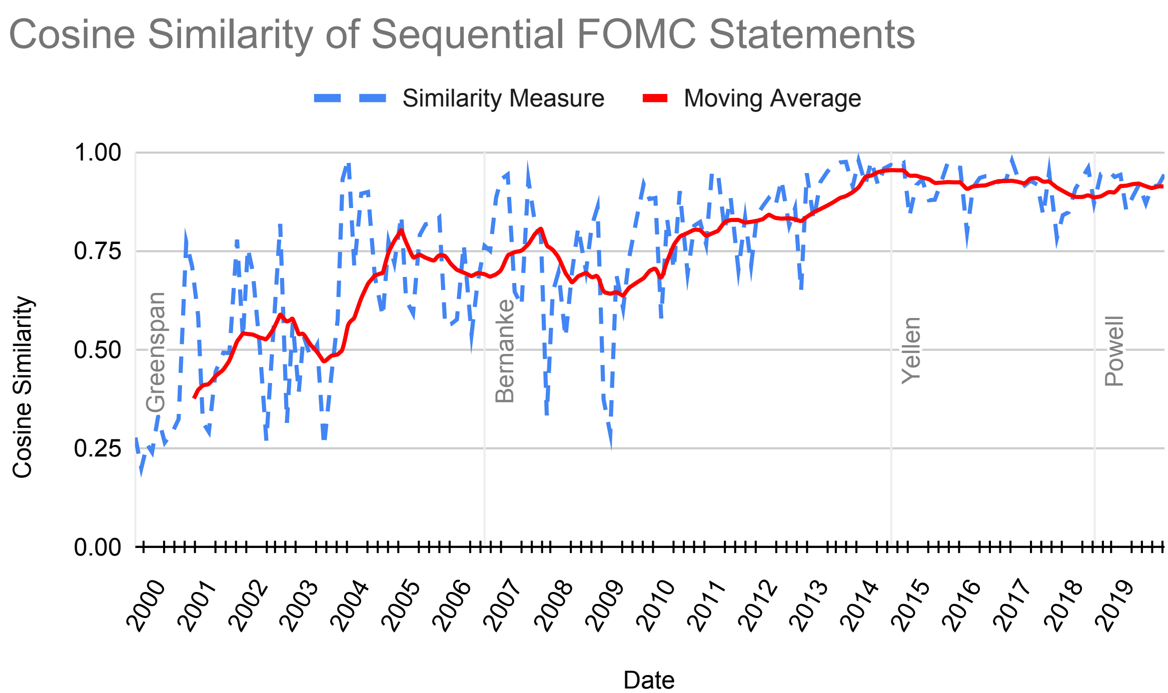


Figure 10: S^1 Cosine Similarity of Sequential FOMC Statements

Notes: S^1 is the cosine similarity between the TFIDF value of an FOMC statement and that of the immediately previous FOMC statement. Cosine similarity values closer to one (zero) mean the statements share more (less) common term usage. The moving average in solid red is calculated with a period of 10.

	30-minute Window				Optimal Window				**Difference**		
	ΔTY_1	ΔTY_2	ΔTY_5	ΔTY_{10}	ΔTY_1	ΔTY_2	ΔTY_5	ΔTY_{10}	ΔTY_2	ΔTY_5	ΔTY_{10}
GSS_T	1.00***	0.82***	0.15	-0.37	1.00***	0.78***	0.08	-0.42	**-0.04**	**-0.07**	**-0.05**
	(0.28)	(0.34)	(0.43)	(0.41)	(0.23)	(0.28)	(0.33)	(0.32)			
GSS_P	1.00***	1.46***	1.89***	1.64***	1.00***	1.51***	1.93***	1.66***	+**0.05**	+**0.04**	+**0.02**
	(0.09)	(0.09)	(0.25)	(0.36)	(0.11)	(0.10)	(0.21)	(0.30)			
NS_{MP}	1.00***	1.24***	1.29***	0.95***	1.00***	1.30***	1.39***	1.06***	+**0.06**	+**0.11**	+**0.11**
	(0.07)	(0.09)	(0.19)	(0.25)	(0.09)	(0.10)	(0.18)	(0.24)			
JK_{MP}	1.00***	1.30***	1.39***	0.99***	1.00***	1.35***	1.52***	1.16***	+**0.04**	+**0.13**	+**0.17**
	(0.11)	(0.15)	(0.28)	(0.36)	(0.12)	(0.16)	(0.32)	(0.43)			
JK_{CBI}	1.00***	1.04***	1.00***	0.82***	1.00***	1.20***	1.14***	0.85***	+**0.16**	+**0.14**	+**0.03**
	(0.25)	(0.30)	(0.31)	(0.29)	(0.23)	(0.25)	(0.27)	(0.26)			

Table 1: Differences in Responses of Nominal Interest Rates to Shocks from Event Window Choice

Notes: Each estimate comes from a separate OLS regression. For each regression, the dependent variable is the one-day change in end-of-day Treasury yields, represented by columns. The independent variable is the change in monetary policy shock series over a given event window around the time of scheduled FOMC announcements. Starting from the left, columns 2–5 represent the OLS regressions performed on shock series constructed within 30-minute windows. Columns 6–9 are equivalent, but for shocks constructed within the median optimal window length of 50 minutes. The sample period consists of all scheduled FOMC statement release dates from May 1999 through October 2019, resulting in 165 observations. Newey and West (1987) standard errors are reported in parentheses. Columns 10–12 represent the differences between the coefficients of the regressions using the optimal windows v. those using 30-minute windows. *** sig. at the 1% level.

	30-minute Window			Optimal Window			**Difference**		
	$\Delta TIPS_2$	$\Delta TIPS_5$	$\Delta TIPS_{10}$	$\Delta TIPS_2$	$\Delta TIPS_5$	$\Delta TIPS_{10}$	$\Delta TIPS_2$	$\Delta TIPS_5$	$\Delta TIPS_{10}$
GSS_T	-0.81	0.02	-0.19	-0.90	0.09	-0.16	**-0.09**	**+0.07**	**+0.03**
	(1.44)	(0.51)	(0.45)	(1.57)	(0.46)	(0.37)			
GSS_P	2.21***	1.96***	1.74***	2.20***	2.03***	1.75***	**-0.00**	**+0.06**	**+0.01**
	(0.49)	(0.46)	(0.44)	(0.37)	(0.38)	(0.37)			
NS_{MP}	1.17***	1.29***	1.08***	1.31***	1.47***	1.20***	**+0.14**	**+0.18**	**+0.13**
	(0.73)	(0.36)	(0.30)	(0.63)	(0.31)	(0.28)			
JK_{MP}	1.40***	1.40***	1.15***	1.66***	1.64***	1.38***	**+0.26**	**+0.24**	**+0.23**
	(0.83)	(0.47)	(0.41)	(0.63)	(0.49)	(0.46)			
JK_{CBI}	0.51	0.99***	0.85***	0.60	1.13***	0.84***	**+0.09**	**+0.14**	**-0.01**
	(0.85)	(0.33)	(0.25)	(0.92)	(0.33)	(0.25)			

Table 2: Differences in Responses of Real Interest Rates to Shocks from Event Window Choice

Notes: Each estimate comes from a separate OLS regression. For each regression, the dependent variable is the one-day change in end-of-day TIPS yields, represented by columns. The independent variable is the change in monetary policy shock series over a given event window around the time of scheduled FOMC announcements. Starting from the left, columns 2–4 represent the OLS regressions performed on shock series constructed within 30-minute windows. Columns 5–7 are equivalent, but for shocks constructed within the median optimal window length of 50 minutes. The sample period consists of all scheduled FOMC statement release dates from May 1999 to October 2019, resulting in 165 observations. Newey and West (1987) standard errors are reported in parentheses. Columns 8–10 represent the differences between the coefficients of the regressions using the optimal windows v. those using 30-minute windows. *** sig. at the 1% level.

	30-minute Window			Optimal Window			**Difference**		
	ΔBEI_2	ΔBEI_5	ΔBEI_{10}	ΔBEI_2	ΔBEI_5	ΔBEI_{10}	ΔBEI_2	ΔBEI_5	ΔBEI_{10}
GSS_T	1.63***	0.13	-0.18	1.67***	-0.01	−0.26**	**+0.05**	**-0.14**	**-0.08**
	(1.46)	(0.29)	(0.17)	(1.58)	(0.29)	(0.17)			
GSS_P	-0.75	-0.08	-0.10	−0.69*	-0.10	-0.09	**+0.06**	**-0.02**	**+0.01**
	(0.46)	(0.24)	(0.12)	(0.36)	(0.23)	(0.12)			
NS_{MP}	0.07	-0.01	−0.13*	-0.01	-0.07	−0.14**	**-0.08**	**-0.07**	**-0.01**
	(0.70)	(0.23)	(0.12)	(0.65)	(0.23)	(0.12)			
JK_{MP}	-0.09	-0.01	−0.17**	-0.31	-0.12	−0.22***	**-0.22**	**-0.11**	**-0.05**
	(0.81)	(0.29)	(0.13)	(0.61)	(0.25)	(0.11)			
JK_{CBI}	0.54	0.01	-0.02	0.60	0.01	0.02	**+0.07**	**+0.00**	**+0.04**
	(0.76)	(0.29)	(0.20)	(0.86)	(0.30)	(0.23)			

Table 3: Differences in Responses of Break-even Inflation to Shocks from Event Window Choice

Notes: Each estimate comes from a separate OLS regression. For each regression, the dependent variable is the one-day change in end-of-day break-even inflation, represented by columns. The independent variable is the change in monetary policy shock series over a given event window around the time of scheduled FOMC announcements. Starting from the left, columns 2–4 represent the OLS regressions performed on shock series constructed within 30-minute windows. Columns 5–7 are equivalent, but for shocks constructed within the median optimal window length of 50 minutes. The sample period consists of all scheduled FOMC statement release dates from May 1999 to October 2019, resulting in 165 observations. Newey and West (1987) standard errors are reported in parentheses. Columns 8–10 represent the differences between the coefficients of the regressions using the optimal windows v. those using 30-minute windows. * sig. at the 10% level. ** sig. at the 5% level. *** sig. at the 1% level.

	30-minute Window			Optimal Window			**Percentage Difference**		
	DP_{SPX}	DP_{ESc1}	DP_{ESc2}	DP_{SPX}	DP_{ESc1}	DP_{ESc2}	DP_{SPX}	DP_{ESc1}	DP_{ESc2}
GSS_T	-8.40^{***}	-8.83^{***}	-7.25^{***}	-7.39^{***}	-7.43^{***}	-7.34^{***}	**-11.99%**	**-15.92%**	**+1.25%**
	(2.71)	(2.68)	(2.78)	(3.10)	(3.15)	(3.12)			
GSS_P	-6.14^{***}	-6.27^{***}	-6.12^{***}	-6.85^{***}	-6.96^{***}	-7.63^{***}	**+11.51%**	**+11.00%**	**+24.61%**
	(1.81)	(1.83)	(1.76)	(2.88)	(2.91)	(2.81)			
NS_{MP}	-6.92^{***}	-7.15^{***}	-6.51^{***}	-7.00^{***}	-7.10^{***}	-7.55^{***}	**+1.23%**	**-1.00%**	**+16.00%**
	(1.32)	(1.37)	(1.31)	(1.85)	(1.89)	(1.84)			
JK_{MP}	-14.76^{***}	-15.08^{***}	-13.73^{***}	-17.46^{***}	-17.77^{***}	-17.30^{***}	**+18.25%**	**+17.88%**	**+26.00%**
	(0.81)	(0.91)	(0.94)	(1.04)	(1.08)	(1.06)			
JK_{CBI}	15.19^{***}	13.84^{***}	15.18^{***}	14.08^{***}	14.44^{***}	12.12^{***}	**-7.36%**	**-4.90%**	**-12.43%**
	(2.29)	(2.35)	(2.39)	(2.11)	(2.14)	(2.07)			

Table 4: Differences in Responses of Stock Prices to Shocks from Event Window Choice

Notes: Each estimate comes from a separate OLS regression. For each regression, the dependent variable is the price log-difference of the S&P 500 Index or E-mini futures, represented by columns. The independent variable is the change in monetary policy shock series over a given event window around the time of scheduled FOMC announcements. Starting from the left, columns 2–4 represent the OLS regressions performed on shock series constructed within 30-minute windows. Columns 5–7 are equivalent, but for shocks constructed within the median optimal window length of 50 minutes. The sample period consists of all scheduled FOMC statement release dates from May 1999 to October 2019, resulting in 165 observations. Newey and West (1987) standard errors are reported in parentheses. Columns 8–10 represent the percentage differences between the coefficients of the regressions using the optimal windows v. those using 30-minute windows, where positive (negative) values represent a stronger (weaker) effect in the same direction. *** sig. at the 1% level.

Metric	Simple	**Complicated**	**Different**	Similar	Unity	**Dissents**
Minimised MSE						
Average	1.25e-5	**1.06e-5**	**1.18e-5**	1.13e-5	9.57e-6	**1.43e-5**
Event Window Length (Minutes)						
Average	60	**71**	**62**	51	53	**83**

Table 5: Average MSEs and Event Window Lengths Using the "One Signal" Approach, Conditioned by FOMC Statement Complexity, Similarity, and Presence of Dissents
Notes: The complexity of FOMC statements is measured by the Flesch-Kincaid Grade Level, defined as: $0.39 \times$ average sentence length $+ 11.8 \times$ average number of syllables per word $-$ 15.59, and displayed in the first two columns. Changes in FOMC statements are measured using a pairwise-statement cosine similarity measure and displayed in the third and fourth columns from the left. The event window lengths are displayed in minutes. "Simple" statements have grade levels up to 16.5. "Complicated" statements have grade levels above 16.5. "Different" are sequential statements with a cosine similarity of less than to 0.885. "Similar" are sequential statements with a cosine similarity of more than 0.885. "Unity" statements are those without votes of dissent. "Dissents" are statements with recorded dissent votes. For all futures contracts, event window lengths considered as outliers under the "one signal" approach are set equal to the median of the sub-set window lengths in order to lessen their effects.

A Motivating Framework Simulations

For the motivating framework discussed in Section 2, I simulate the asset price process over time for multiple news announcements. The simulation illustrates the effect of cognitive noise and unrelated news on the optimal time horizon for price responses and demonstrates why a good signal is required for this time horizon estimation, both motivating the necessity of my methodology.

The simulation assumes that news is released at time $t = 0$. The fundamental price component, $P^f_{i,t}$, responds by jumping to and remaining at real value $P^f_i \in [-100, 100]$. Simultaneously, the cognitive noise component jumps to a random value within $[-100, 100]$ whilst unrelated news is equal to zero at $t = 0$. The noisy signal observed by the econometrician, s_i, is assumed to have its error component distributed normally with mean zero and finite variance. To analysis this on average, I simulate the price process and the noisy signal for $N = 10,000$ news announcements up to time $t = 100$, calculating both the true and estimated MSEs using P^f_i and signal s_i, respectively. Finally, the true and estimated times at which the market fully reacts on average, t^* and $\hat{t}$, can be determined according to Equation 4.

A.1 Simulation Results

Three scenarios are considered for the asset market price responding to the release of news: one that exhibits cognitive noise but little unrelated news, one with unrelated news but little cognitive noise, and one with both cognitive noise and unrelated news. The simulation parameters corresponding to these three scenarios can be found in the topmost eight rows of Appendix Table F1.

Reading the columns of the bottom two rows of Appendix Table F1 from left to right, one can see that an asset market possessing a lot of cognitive noise with little unrelated news will have a longer optimal time horizon. The middle column shows the opposite: when faced with a lot of unrelated news but little cognitive noise, t^* becomes very small. Both these results follow the dynamics predicted by Equation 3. Intuitively, when markets need more time to fully react to news, this means the optimal time horizon should be longer. In contrast, when markets are able to react quickly to the news and there is the presence of unrelated information coming out over time, using a short event window would be more appropriate. The rightmost column displays the most "realistic" scenario for an asset market: both cognitive noise and unrelated news exist. In such an event, the optimal time horizon lies somewhere between the two extreme scenarios.

Importantly, the simulations also present minimal differences between t^* and $\hat{t}$—the true and estimated time horizons, respectively—when using only the signal s_i. Therefore, a "good" signal allows for the possibility of estimating the time horizon that reflects when markets have fully reacted to news on average. However, recall that this the motivating framework and related simulations have strong linearity and orthogonality assumptions imposed to simply demonstrate the effects of noise on the estimation for the optimal time horizon. In reality, the compositions and processes of the fundamental price and noise components are unknown for equity markets, yet still influence the movement of the overall asset price.

Abstracting away from the simple framework, I argue in the rest of this paper that my text-analysis neural network approach is still able to extract a "good" signal from its approximation of the underlying mapping between monetary policy communications and asset price changes, regardless of what are the true noise processes at any point of time.

B The Primary Accuracy Metric for Judging XLNet

As mentioned in Subsection 4.4, the primary metric considered by the neural network to judge its accuracy during fine-tuning for each sample split is the generalised R^2 defined in Hawinkel et al. (2024):

$$R^2_{OOS} = 1 - \frac{\widehat{MSE}}{\widehat{MST}} = 1 - \frac{T^{-1}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2}{\frac{T+1}{T(T-1)}\sum_{i=1}^{T}(y_i - \overline{y_{IS}})^2}, \tag{B1}$$

where $\widehat{MSE}$ is the mean squared error of the neural network over the out-of-sample observations; $\widehat{MST}$ is the mean squared error calculated using the in-sample mean $\overline{y_{IS}}$ as prediction over all out-of-sample observations; and T is the size of the testing sample. Although the explicit objective function of the neural network during fine-tuning is to minimise $\widehat{MSE}$, this simultaneously maximises R^2_{OOS}. Other related metrics that are tracked by XLNet-Base during its fine-tuning process are the out-of-sample Pearson correlation coefficient ρ between the predicted and actual observations, the out-of-sample mean absolute error, and the in-sample mean squared error. The last criterion is important for determining if the network is learning the relationship between the FOMC statement text and the price log-differences or not.

The interpretation of R^2_{OOS} can be better understood through a brief discussion about the underlying relationship that the neural network is approximating and using to predict, which is the nonparametric function mapping the text of FOMC statements to asset price changes. As is standard in the machine learning literature, the predictive quality of models is compared to the common baseline that states there is *no* relationship between the FOMC statement text and asset price changes and naively predicts with the in-sample average. When using the neural network to instead approximate this relationship and predict asset price changes for statements in the testing subsample, the R^2_{OOS} therefore states how much predictive error from naively using the in-sample average has been reduced. This interpretation implies that R^2_{OOS} does *not* follow the conventional definition of R^2, which is the proportion of variance in the data that is explained by a given model.[46] First, the conventional definition breaks down when applied to non-linear methods like neural networks because the variance of the observations is no longer comprised of only the model and residual variances. Second, applying the conventional R^2 metric onto out-of-sample observations breaks the spirit of prediction because it would be using the average outcome of the testing sample, which is supposed to be inaccessible to models during fitting.

[46]Note that the conventional definition is equivalent to the definition comparing a given model to the "naive model" when applied to the in-sample observations.

C Construction of High-Frequency Monetary Surprises

C.1 Financial Data Overview and Price Selection

The intraday data on interest-rate and equity futures is sourced from the Thomson Reuters Tick History database and provided by LSEG. The sample period for my analysis is from May 1999 to October 2019. Table F2 provides an overview of the financial outputs data. For each futures contract, I have a minutely series which includes the price of the first and last trades for a given minute. For the rest of this appendix section, I detail the construction of the interest-rate surprises from the contracts. Following previous papers in the literature, I use federal funds futures contracts to measure interest rate expectations within three months, Eurodollar futures to capture expectations between two and four quarters out, and Treasury futures for expectations out to fifteen years.

As discussed in Subsection 3.2, I collect price levels for all futures contracts at 10-minute intervals, starting from 10 minutes before to 18 hours after an FOMC statement release. Specifically, the last recorded traded price 10 minutes before whilst the first recorded traded price at 10-minute intervals after a statement release are collected. I make the core assumption that the futures price does not change in-between times of recorded trades. For example, if the recorded traded prices for $FF1$ were at 15 and 9 minutes before an FOMC statement release, I assume that the first recorded trading price 15 minutes before the release is the same at 10 minutes beforehand. Therefore, I collect the price 15 minutes beforehand. The equivalent logic applies to collecting futures prices for the 10-minute intervals after a monetary policy announcement.

C.2 Surprises from Federal Funds Futures

On the last day of a given expiry month, a federal funds rate futures (FFF) contract pays out 100 minutes the average (effective) federal funds rate over the expiry month. In other words, let $p_{\tau,t}^{FFj}$ be the price of the $(j-1)$ month-ahead FFF at time t for FOMC meeting τ. Then, the expected average federal funds rate of the $(j-1)$ month ahead FFF at time t for FOMC meeting τ is calculated as $100 - p_{\tau,t}^{FFj}$.

C.2.1 Surprises from Federal Funds Futures: Current Meeting

I calculate the federal funds rate surprise for the current meeting as follows:

1. For each time t, calculate the implied federal funds rate, $100 - p_{\tau,t}^{FFj}$.

2. Calculate $mp1_{\tau,t+n} = \frac{m}{m-d}\left(FF1_{\tau,t+n} - FF1_{\tau,t-10}\right)$, the federal funds rate surprise for the current meeting for each FOMC announcement τ, where d is the day of the announcement and m is the number of days in the month of the announcement.

3. If $m - d + 1 \leq 7$, i.e., the FOMC announcement occurs in the last seven days of the month, use the change in the price of the next month's FFF. In other words, the current-meeting surprise would be calculated as $FF2_{\tau,t+n} - FF2_{\tau,t-10}$. The reason behind this criterion is to avoid multiplying the difference in implied rates with large $\frac{m}{m-d}$.

C.2.2 Surprises from Federal Funds Futures: Next Meeting

I calculate the surprise in the expected federal funds rate, at current FOMC meeting τ, for the next meeting as follows:

1. For a given FOMC announcement τ, find the number of months out $(j-1)$ that has the next scheduled meeting. This determines the maturity of the FFF to use, i.e., j. It was found that $j \in \{2, 3, 4\}$.

2. Calculate $mp2_{\tau,t+n} = \frac{m_2}{m_2-d_2}\left\{[FFj_{\tau,t+n} - FFj_{\tau,t-10}] - \frac{d2}{m2}mp1_{\tau,t+n}\right\}$, where d_2 is the day of the next scheduled FOMC meeting, m_2 is the number of days in the month containing the next scheduled meeting, and $mp1_{\tau,t+n}$ is the current-meeting surprise calculated within an event window beginning 10 minutes before and ending $t+n$ minutes after announcement τ.

3. If $m_2 - d_2 + 1 \leq 7$, meaning the next scheduled FOMC announcement occurs in the last seven days of the month, use the change in the price of the $j+2$-month-ahead FFF contract. In other words, $mp2_{\tau,t+n} = FF(j+1)_{\tau,t+n} - FF(j+1)_{\tau,t-10}$.

C.3 Surprises from Eurodollar Futures

Eurodollar contracts are quarterly contracts that pay out 100 minus the US Dollar BBA LIBOR interest rate at the time of expiration. The last trading day of the contracts is the second London bank business day (usually a Monday) before the third Wednesday of the last month of the expiry quarter. In other words, let $p^{edj}_{\tau,t}$ be the price of the j^{th} closest quarterly Eurodollar futures contract (i.e., March, June, September, December) at time t on FOMC meeting date τ. Note that the expiration date of this contract is between $j-1$ and j quarters into the future at any point of time. We can write the implied interest rate to be $edj_{\tau,t} = 100 - p^{edj}_{\tau,t}$ at time t in FOMC meeting date τ. For a given FOMC announcement τ, I calculate the surprise in the implied rate of the contract maturity j as follows:

1. For each time t, calculate the implied rate, $edj_{\tau,t} = 100 - p^{edj}_{\tau,t}$.

2. Calculate the surprise in the implied rate of contract maturity $j, edj_{\tau,t+n} = edj_{\tau,t+n} - edj_{\tau,t-10}$.

C.4 Surprises from Treasury Futures

At the time of expiration, Treasury futures contracts oblige the seller to deliver Treasury bonds within a range of maturities to the buyer. Let $p^{tkj}_{\tau,t+n}$ be the price at time t on FOMC announcement τ of the j^{th} nearest quarterly k-year Treasury futures contract (i.e., March, June, September, December). To calculate the implied yield surprise around the FOMC meeting $\tau, tk_{\tau,t+n}$, I divide the price change by the approximate maturity and flip the sign:

$$tk_{\tau,t+n} = -\left(p^{tkj}_{\tau,t+n} - p^{tkj}_{\tau,t-10}\right)/l, \tag{C.1}$$

where approximate maturities $l \in \{2, 4, 7, 15\}$ come from Gürkaynak, Kisacikoğlu, et al. (2020). If the FOMC meeting τ falls within the expiry month of the quarter (i.e., March, June, September, December), I use the $(j+1)^{th}$ nearest quarterly Treasury futures contract due to higher liquidity (Gorodnichenko and Ray, 2017).

D Robustness Checks

D.1 Original Surprise Instruments for MP Surprises & Shocks

I perform the same exercises in Subsection 6.2, but for monetary policy shocks constructed with the original instrument set of monetary policy surprises used by the authors.[47]

Visual Shock Differences I plot the shock series constructed in both window lengths using the original instrument set, displayed in Figure E28. For each sub-figure, the horizontal axis depicts the FOMC statement release dates. The vertical axis depicts the principal components re-scaled according to the original specifications of the authors. Interestingly, the differences in all four monetary policy shocks containing forward guidance have extremely similar visual characteristics as those shown in the main results of the paper, which use the full instrument set. The similar differences from event window choice also apply to JK_{CBI}, which still observe more "shifts of importance" when using the optimal event window length due to the sign restriction method of Jarociński and Karadi (2020).

Responses of Interest Rates Regression results following the specification of Equation 8 for the responses of nominal and real interest rates to monetary policy shocks within different event windows are presented in Appendix Table F11 and Appendix Table F12, respectively. Overall, we can see that when changing only the event window length, monetary policy shocks containing forward guidance have stronger point estimates in units of basis points. Interestingly, using the median optimal event window length of 50 minutes also increases the statistical significance of several coefficient estimates, such as regressing the daily change in the two-year TIPS yield on GSS_P or various nominal Treasury yields on the central bank information shock of Jarociński and Karadi (2020).

Responses of Break-even Inflation Regression results for break-even inflation using the original surprise instruments can be found in Appendix Table F13. Similar to the main results of the paper, monetary policy shocks constructed within the optimal event window length results in more negative impacts on break-even inflation for all maturities. Furthermore, moving away from the 30-minute window results in increased statistical significance for several estimates. These results show that the main takeaways of the paper are robust to the choice of surprise instruments.

Responses of Equities Using the original instrument set to construct all considered monetary policy shocks, I also investigate the effect of event window size on the impact of monetary policy shocks on equity prices. Appendix Table F14 presents analogous results for the S&P 500 Index and its E-mini futures following the regression form of Equation 8. The sole

[47] Technically, Jarociński and Karadi (2020) use $FF4$ instead of the next-meeting surprise, $mp2_\tau$. Because both instruments broadly represent the same underlying maturity and the other authors use this instrument over $FF4$, I construct monetary policy shocks with the method of Jarociński and Karadi (2020) using $mp2_\tau$ for convenience.

difference is that the dependent variable is now 100 times the price log-difference of equities. Compared to the main results in Subsection 6.2, the effects of event window choice on the impact of monetary policy shocks are more mixed. For example, whilst GSS_P results in a decline in the S&P 500 Index by an additional 2.52% when using the optimal window length, other shocks like NS_{MP} see a weaker effect by 3.32%. Whilst all monetary policy shocks still have statistically significant effects on equity prices which go in the expected direction, the effects of window size choice are overall smaller. In contrast, the percentage changes to the responses of the second-month S&P 500 E-mini futures contract are similar to those seen in the main results.

The overall results from constructing monetary policy shocks with the original instrument set suggests that the shorter underlying maturity of the surprises could be understating the shock effects. Because my sample period includes the effective lower bound, the effects of forward guidance and other non-conventional monetary policy tools might not be fully captured within the implied surprises of interest-rate surprises whose expected path only goes out one year into the future. Indeed, recent studies (e.g., Brennan et al., 2024; An et al., 2025) have shown that incorporating instruments with longer underlying maturities can help prevent this understatement in constructed monetary policy shocks.

Responses of the Macroeconomy via Local Projections I use the lag-augmented local projection method from Olea and Plagborg-Møller (2021) to estimate the responses of log CPI, log IP, the nominal two-year Treasury yield, and the EBP to the considered monetary policy shock series. The separate regressions of each outcome variable on each shock follow the form specified in Equation 9. Figures E29–E32 visually display the impulse responses for all outcome variables to all shock series. Similar to the main results of the paper, the effects of event window length don't really change most point estimates of the responses, although there are some exceptions to this trend (e.g., the EBP increases positively with statistical significance to the NS_{MP} shock starting at 36 months after initial impact). Also similar to the main results is the fact that the confidence intervals of the impulse responses using the optimal window length appear smaller than those using 30-minute windows. When calculating the average and median confidence interval width ratios discussed in Subsection 6.2, using the original instrument set yields ratios of 0.9033 and 0.9506, respectively. In other words, using the optimal event window length results in estimated responses of macroeconomic variables to be more precise.

D.2 Optimal Event Window Lengths Beyond 70 Minutes?

An additional advantage with the "one signal" approach is that I am able to crudely check if event windows beyond the optimal window length from the "joint" estimation have a larger $\overline{R^2_{OOS}}$. Appendix Figures E21–E27 visually display the results of this exercise for maturities of federal funds futures, Eurodollar futures, Treasury futures of two, five, ten, and thirty years, and both the S&P 500 Index and its E-mini futures. The horizontal axis depicts the event window lengths, starting from the systematically estimated event window length and ending at an event window starting 10 minutes before and ending 18 hours after FOMC statement release. The cross points represent the event window length associated with the highest $\overline{R^2_{OOS}}$. One can see that for most maturities and assets, the event window length with the largest $\overline{R^2_{OOS}}$ is the length systematically estimated by the "joint" approach. In other words, the results from this exercise can be viewed as a robustness check in support of XLNet-Base providing a "good" signal overall for systematic event window length estimation. However, there are few futures contracts that show a difference between the "joint" and "one signal" estimated event window lengths, such as how the "one signal" approach estimates the optimal window length for the 1-month-ahead federal fund futures ($FF2$), 3-month-ahead federal fund futures $FF4$), front-month 2-year Treasury futures ($TUc1$), second-month 10-year Treasury futures ($TYc2$), and front-month 30-year Treasury futures ($USc1$) are longer than those systematically estimated. To confirm the validity of these estimated window lengths, I calculate $\overline{R^2_{OOS}}$ under the "joint" approach for these futures. The results are depicted as the rightmost box-and-whisker plot in the respective sub-figures of systematic estimation for these futures contracts. Two insights can be taken away from this robustness exercise. First, the optimal event window for these futures contracts is still within 60 minutes in length. However, "joint" estimation for all these markets showed that the considered window lengths beyond 60 minutes after statement released yielded $\overline{R^2_{OOS}}$ greater than that obtained within a 30-minute window. This leads to the second insight, which is the interesting possibility that the global optimal event window length might possibly be between 60 minutes and the amount of time predicted by the "one signal" approach after FOMC statement release. Regardless, a 30-minute window is never predicted to be the optimal event window length.

Appendix Figures

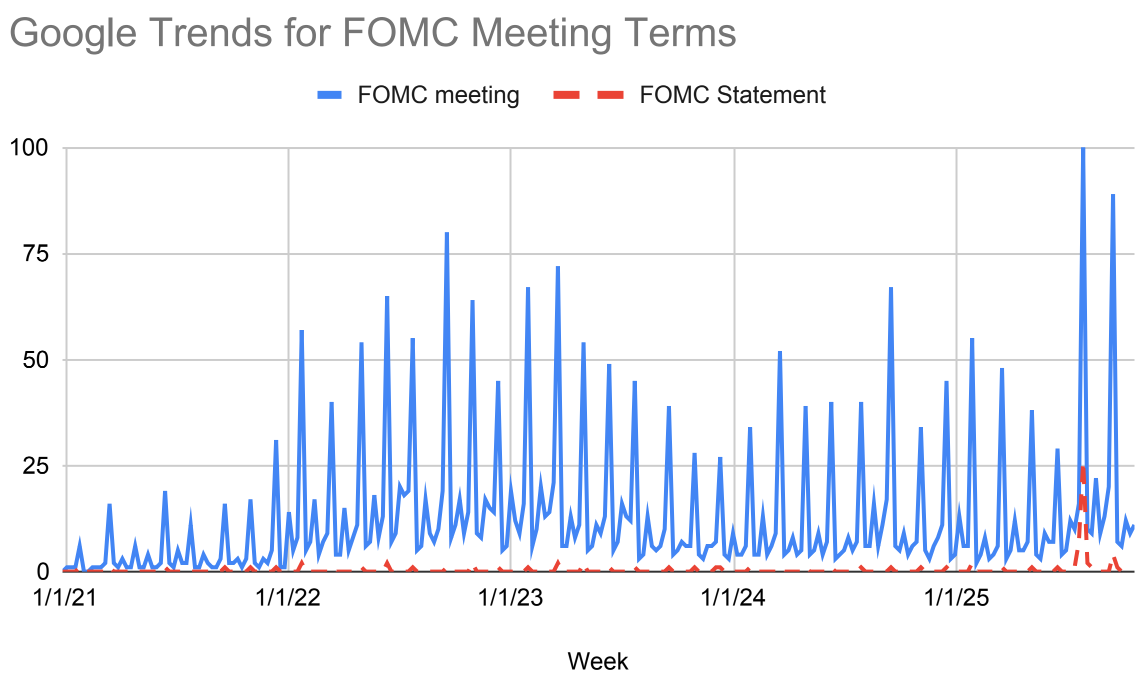
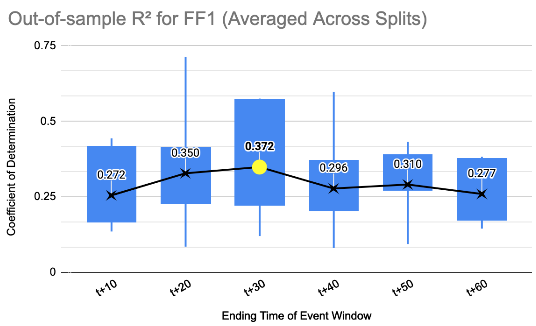
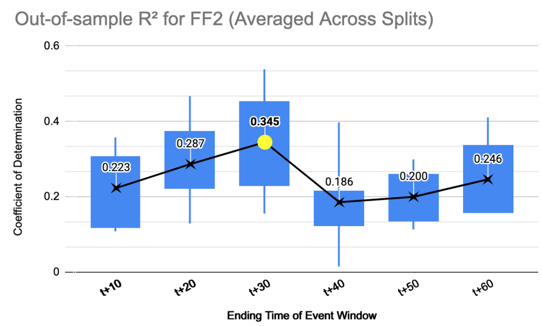


Figure E1: Google Trends Interest About FOMC Statements over Time, January 2021–October 2025

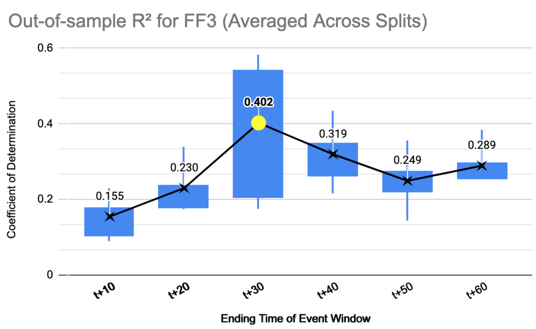
Notes: The blue and solid (red and dashed) line represents the search interest for the phrase "FOMC meeting" ("FOMC statement") in the U.S. The interest gauge is substantially greater than



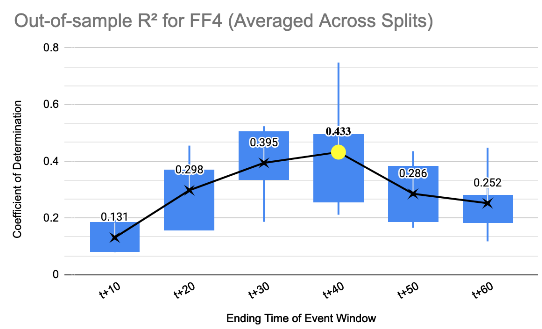
(a) $FF1$



(b) $FF2$

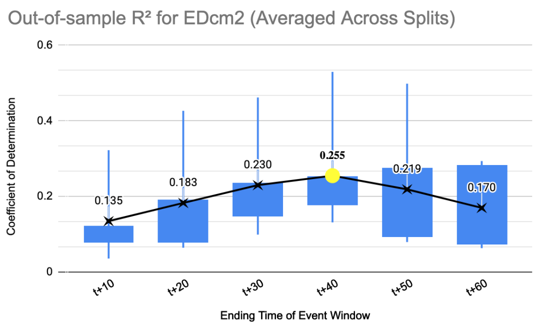


(c) $FF3$

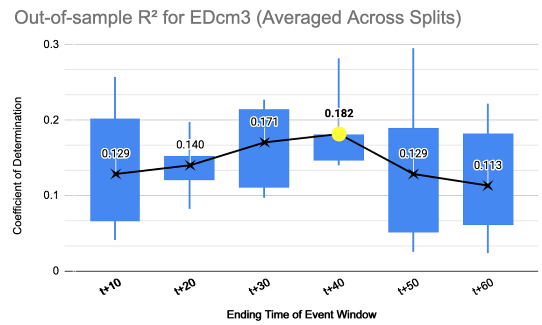


(d) $FF4$

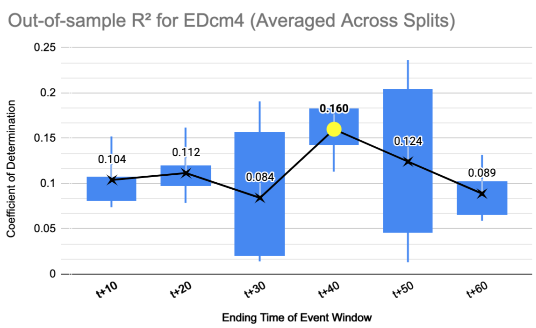
Figure E2: Systematic Estimation of Optimal Window Sizes for Federal Funds Futures
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the percentage points of price changes by the market directly and only due to the FOMC statement text. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $0 \leq \overline{R^2_{OOS}} \leq 1$ represents the proportion of the mean squared error from predicting with the in-sample average explained by the superior out-of-sample performance of XLNet-Base.



(a) $EDcm2$



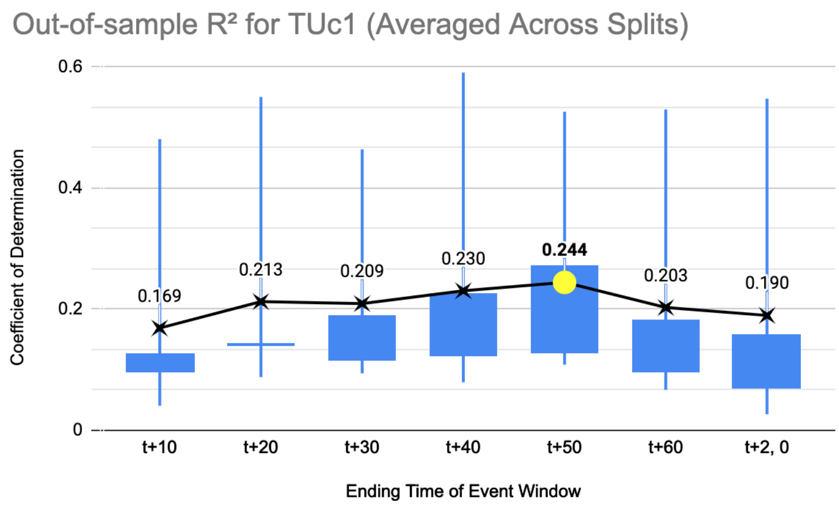
(b) $EDcm3$



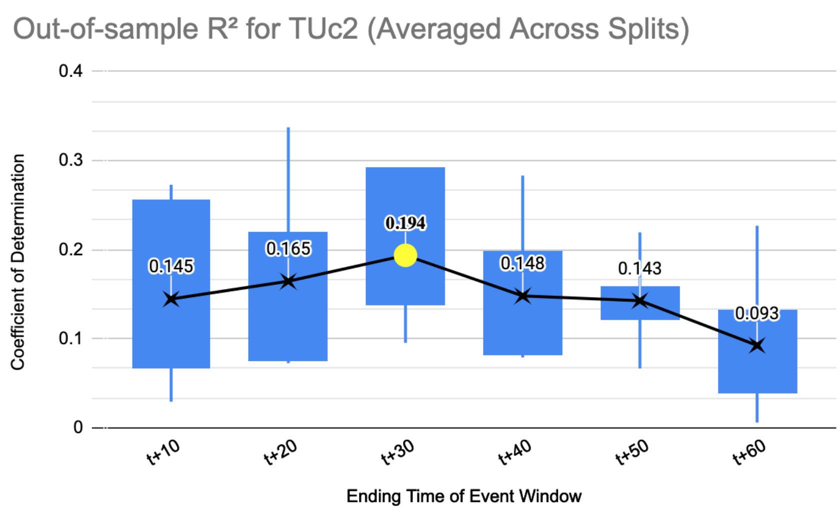
(c) $EDcm4$

Figure E3: Systematic Estimation of Optimal Window Sizes for Eurodollar Futures

Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the percentage points of price changes by the market directly and only due to the FOMC statement text. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $0 \leq \overline{R^2_{OOS}} \leq 1$ represents the proportion of the mean squared error from predicting with the in-sample average explained by the superior out-of-sample performance of XLNet-Base.

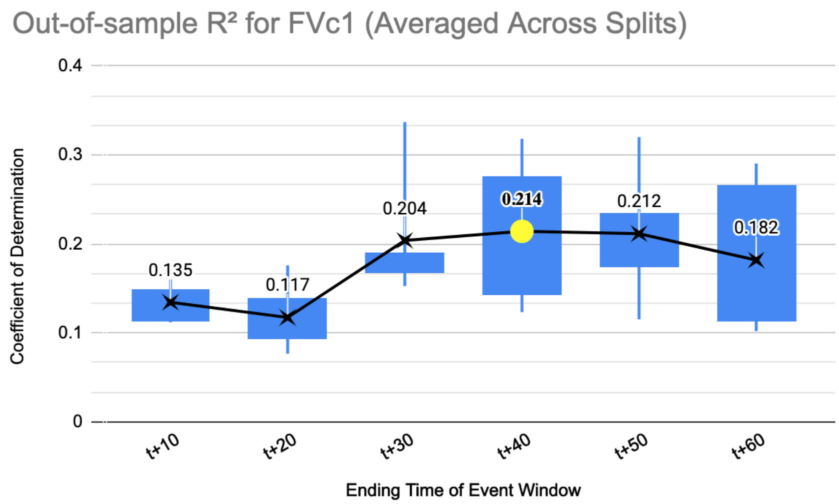


(a) $TUc1$

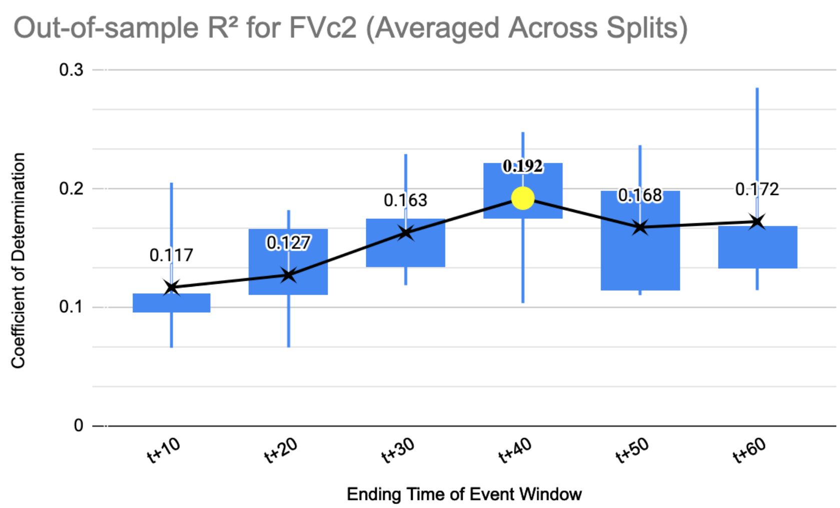


(b) $TUc2$

Figure E4: Systematic Estimation of Optimal Window Sizes for 2-Year Treasury Futures
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the percentage points of price changes by the market directly and only due to the FOMC statement text. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $0 \leq \overline{R^2_{OOS}} \leq 1$ represents the proportion of the mean squared error from predicting with the in-sample average explained by the superior out-of-sample performance of XLNet-Base.

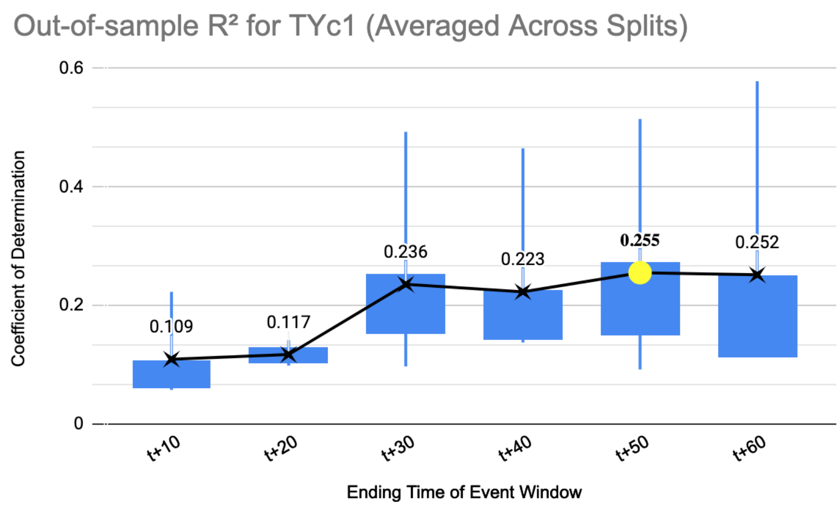


(a) $FVc1$

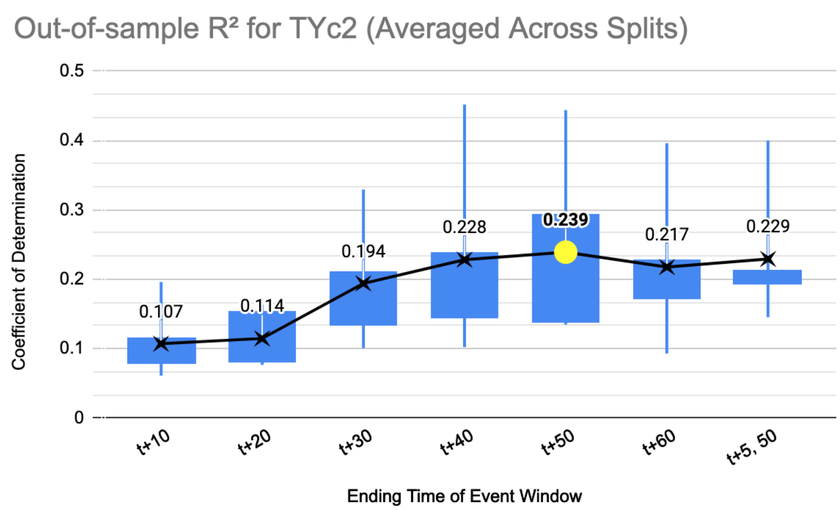


(b) $FVc2$

Figure E5: Systematic Estimation of Optimal Window Sizes for 5-Year Treasury Futures
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the percentage points of price changes by the market directly and only due to the FOMC statement text. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $0 \leq \overline{R^2_{OOS}} \leq 1$ represents the proportion of the mean squared error from predicting with the in-sample average explained by the superior out-of-sample performance of XLNet-Base.

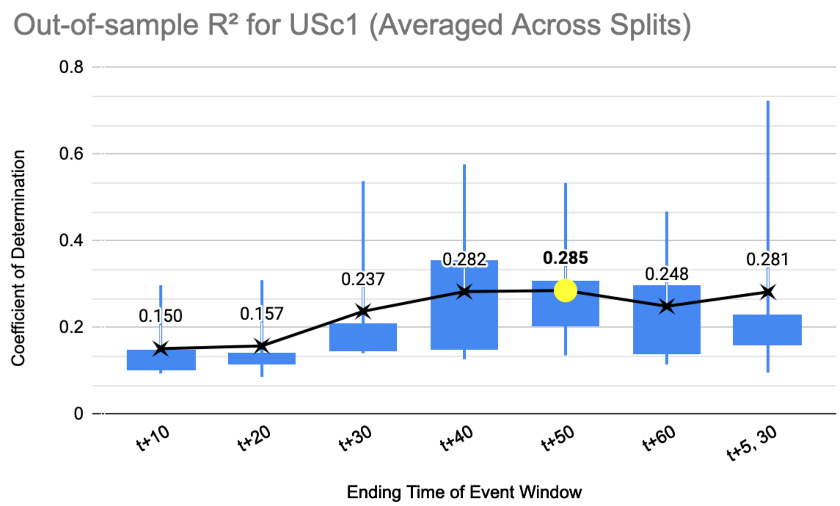


(a) $TYc1$

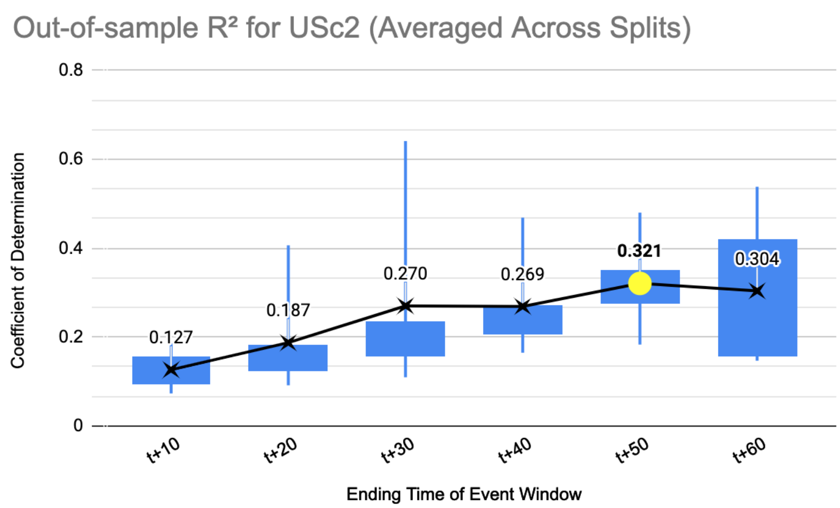


(b) $TYc2$

Figure E6: Systematic Estimation of Optimal Window Sizes for 10-Year Treasury Futures
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the percentage points of price changes by the market directly and only due to the FOMC statement text. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $0 \leq \overline{R^2_{OOS}} \leq 1$ represents the proportion of the mean squared error from predicting with the in-sample average explained by the superior out-of-sample performance of XLNet-Base.

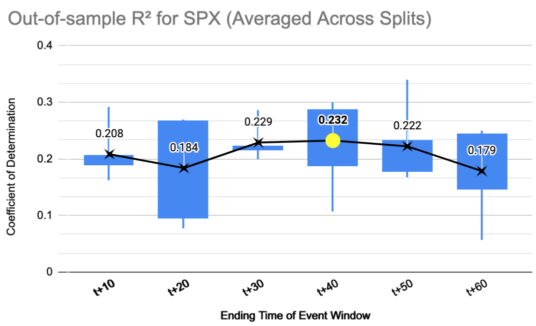


(a) $USc1$

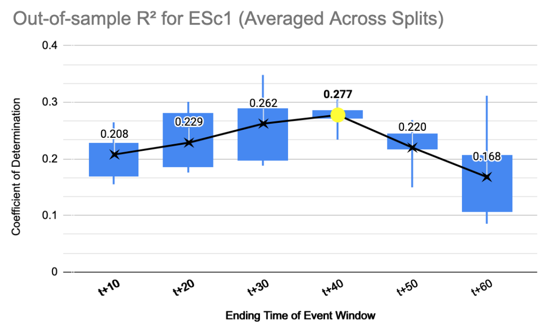


(b) $USc2$

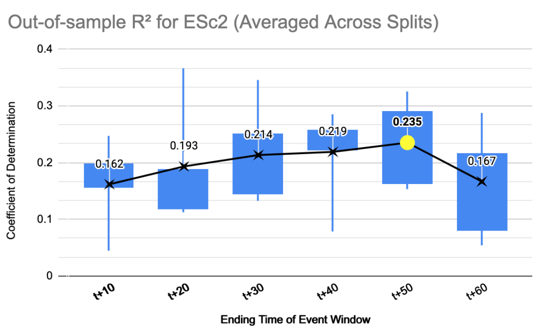
Figure E7: Systematic Estimation of Optimal Window Sizes for 30-Year Treasury Futures
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the percentage points of price changes by the market directly and only due to the FOMC statement text. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $0 \leq \overline{R^2_{OOS}} \leq 1$ represents the proportion of the mean squared error from predicting with the in-sample average explained by the superior out-of-sample performance of XLNet-Base.



(a) SPX

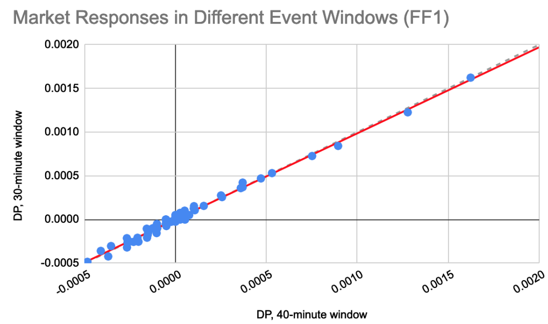


(b) $ESc1$

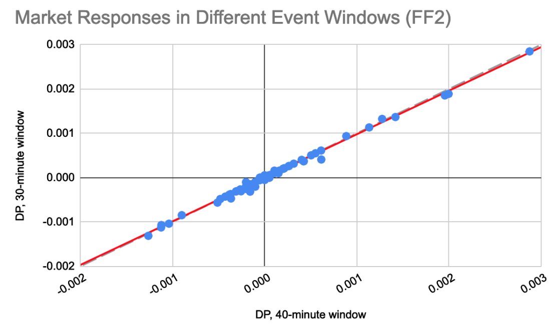


(c) $ESc2$

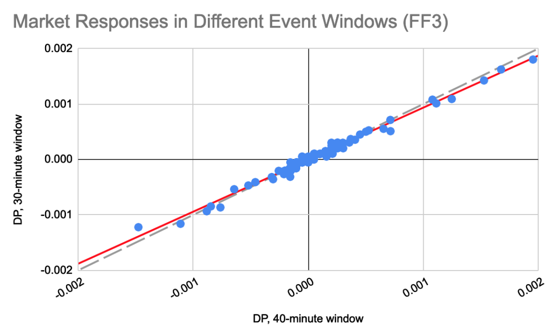
Figure E8: Systematic Estimation of Optimal Window Sizes for S&P 500 and E-mini Futures
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the percentage points of price changes by the market directly and only due to the FOMC statement text. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $0 \leq \overline{R^2_{OOS}} \leq 1$ represents the proportion of the mean squared error from predicting with the in-sample average explained by the superior out-of-sample performance of XLNet-Base.



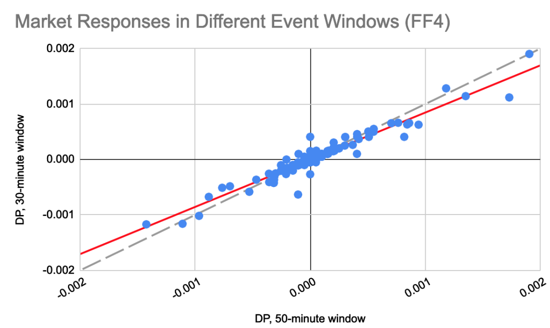
(a) $FF1^{**}$



(b) $FF2^{***}$



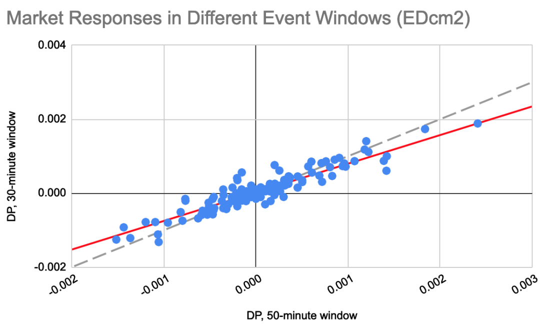
(c) $FF3^{***}$



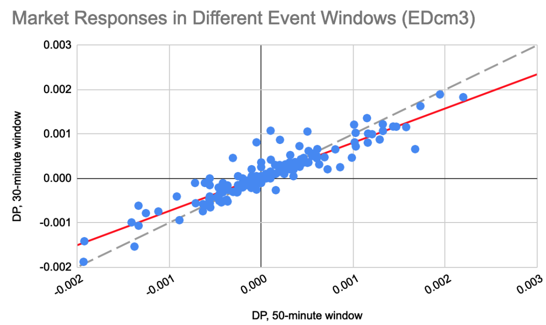
(d) $FF4^{***}$

Figure E9: Comparing Market Responses in Different Event Windows for Federal Funds Futures

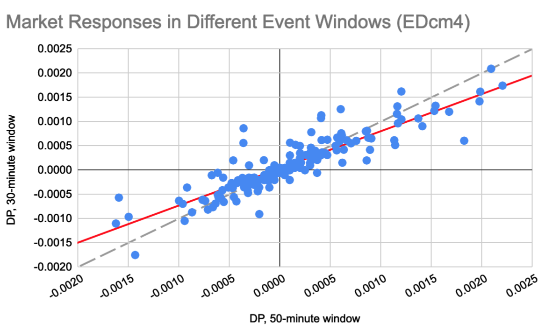
Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient is greater (less) than one and its difference with one is statistically significant for at most $\alpha = 0.10$, then the financial market under-reacts to the information content of FOMC statements on release, ex post, and is depicted in red. ** sig. at the 5% level, *** sig. at the 1% level.



(a) $EDcm2^{***}$



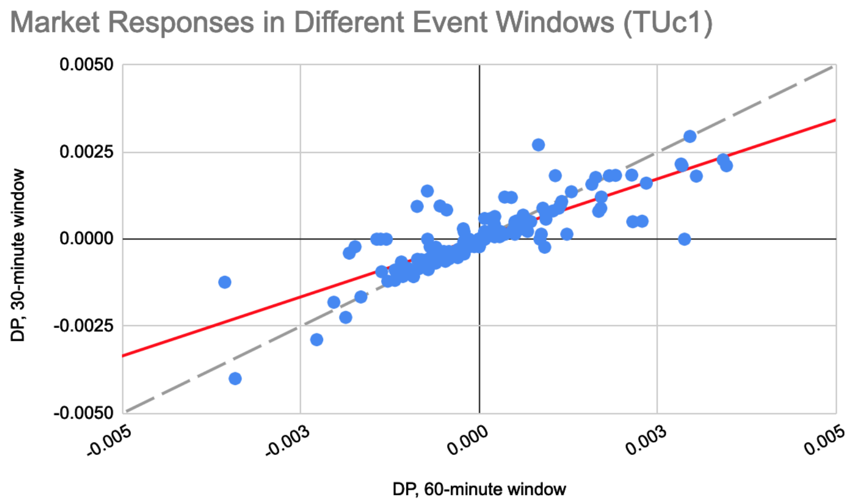
(b) $EDcm3^{***}$



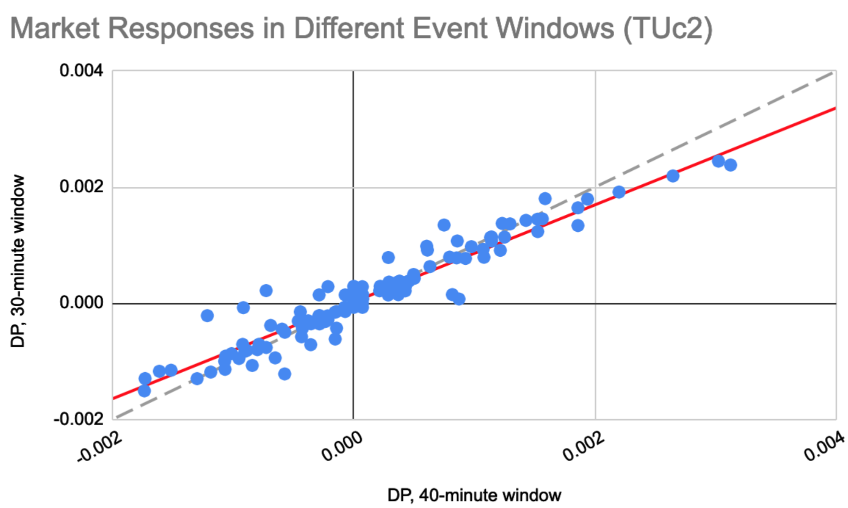
(c) $EDcm4^{***}$

Figure E10: Comparing Market Responses in Different Event Windows for Eurodollar Futures

Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient is greater (less) than one and its difference with one is statistically significant for at most $\alpha = 0.10$, then the financial market under-reacts to the information content of FOMC statements on release, ex post, and is depicted in red. ** sig. at the 5% level, *** sig. at the 1% level.

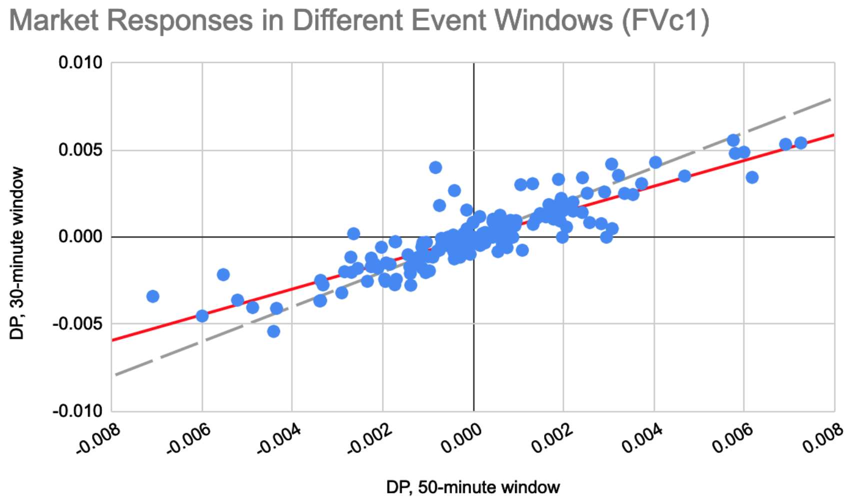


(a) $TUc1^{***}$

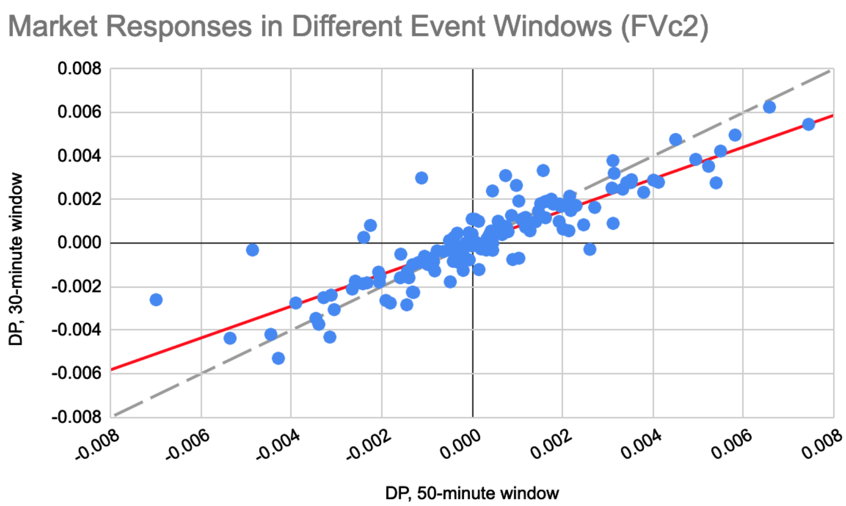


(b) $TUc2^{***}$

Figure E11: Comparing Market Responses in Different Event Windows for 2-Year Treasury Futures
Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient is greater (less) than one and its difference with one is statistically significant for at most $\alpha = 0.10$, then the financial market under-reacts to the information content of FOMC statements on release, ex post, and is depicted in red. ** sig. at the 5% level, *** sig. at the 1% level.



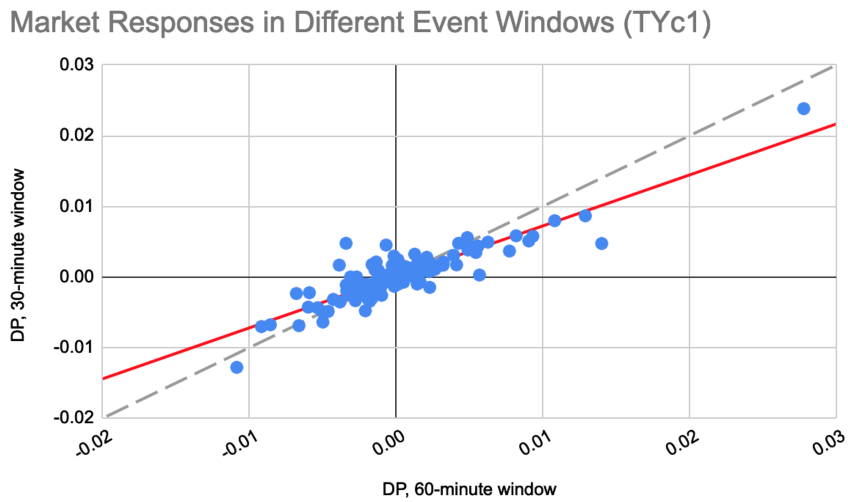
(a) $FVc1^{***}$



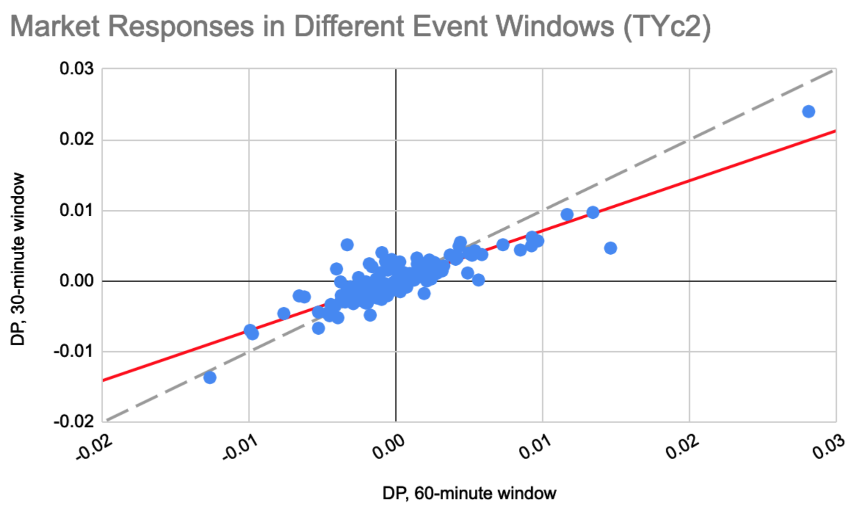
(b) $FVc2^{***}$

Figure E12: Comparing Market Responses in Different Event Windows for 5-Year Treasury Futures

Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient is greater (less) than one and its difference with one is statistically significant for at most $\alpha = 0.10$, then the financial market under-reacts to the information content of FOMC statements on release, ex post, and is depicted in red. ** sig. at the 5% level, *** sig. at the 1% level.

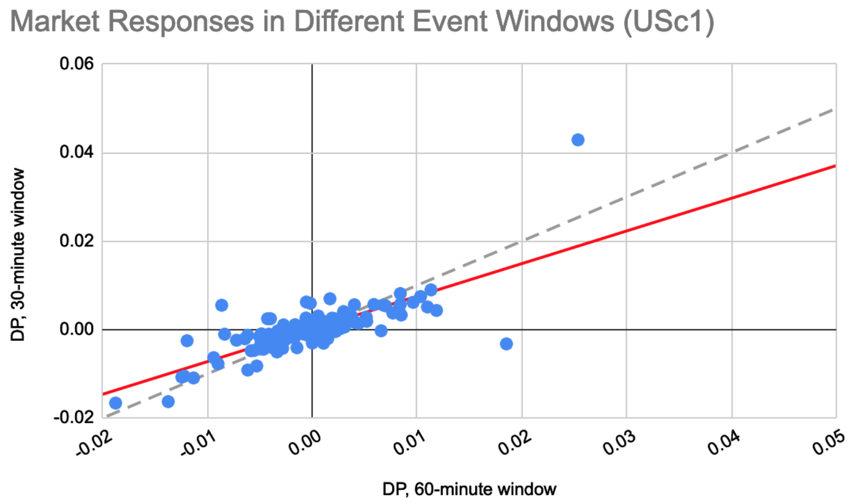


(a) $TYc1^{***}$

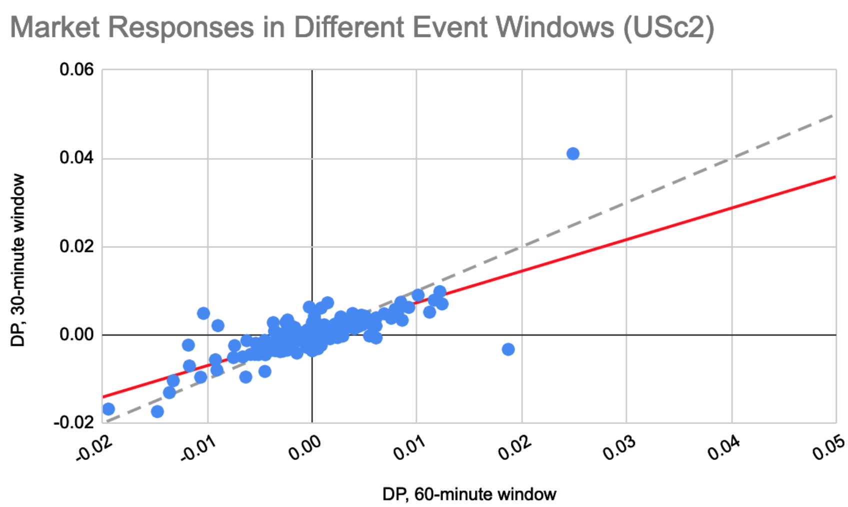


(b) $TYc2^{***}$

Figure E13: Comparing Market Responses in Different Event Windows for 10-Year Treasury Futures
Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient is greater (less) than one and its difference with one is statistically significant for at most $\alpha = 0.10$, then the financial market under-reacts to the information content of FOMC statements on release, ex post, and is depicted in red. ** sig. at the 5% level, *** sig. at the 1% level.

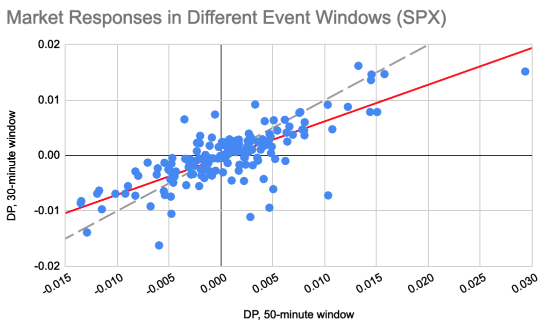


(a) $USc1^{***}$

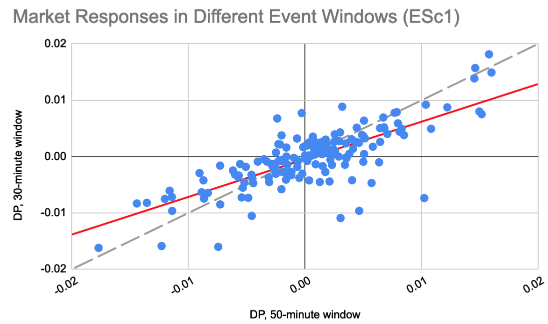


(b) $USc2^{***}$

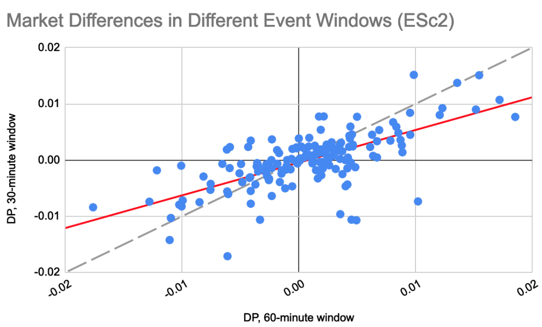
Figure E14: Comparing Market Responses in Different Event Windows for 30-Year Treasury Futures
Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient is greater (less) than one and its difference with one is statistically significant for at most $\alpha = 0.10$, then the financial market under-reacts to the information content of FOMC statements on release, ex post, and is depicted in red. ** sig. at the 5% level, *** sig. at the 1% level.



(a) SPX^{***}



(b) $ESc1^{***}$



(c) $ESc2^{***}$

Figure E15: Comparing Market Responses in Different Event Windows for S&P 500 and E-mini Futures

Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient is greater (less) than one and its difference with one is statistically significant for at most $\alpha = 0.10$, then the financial market under-reacts to the information content of FOMC statements on release, ex post, and is depicted in red. ** sig. at the 5% level, *** sig. at the 1% level.

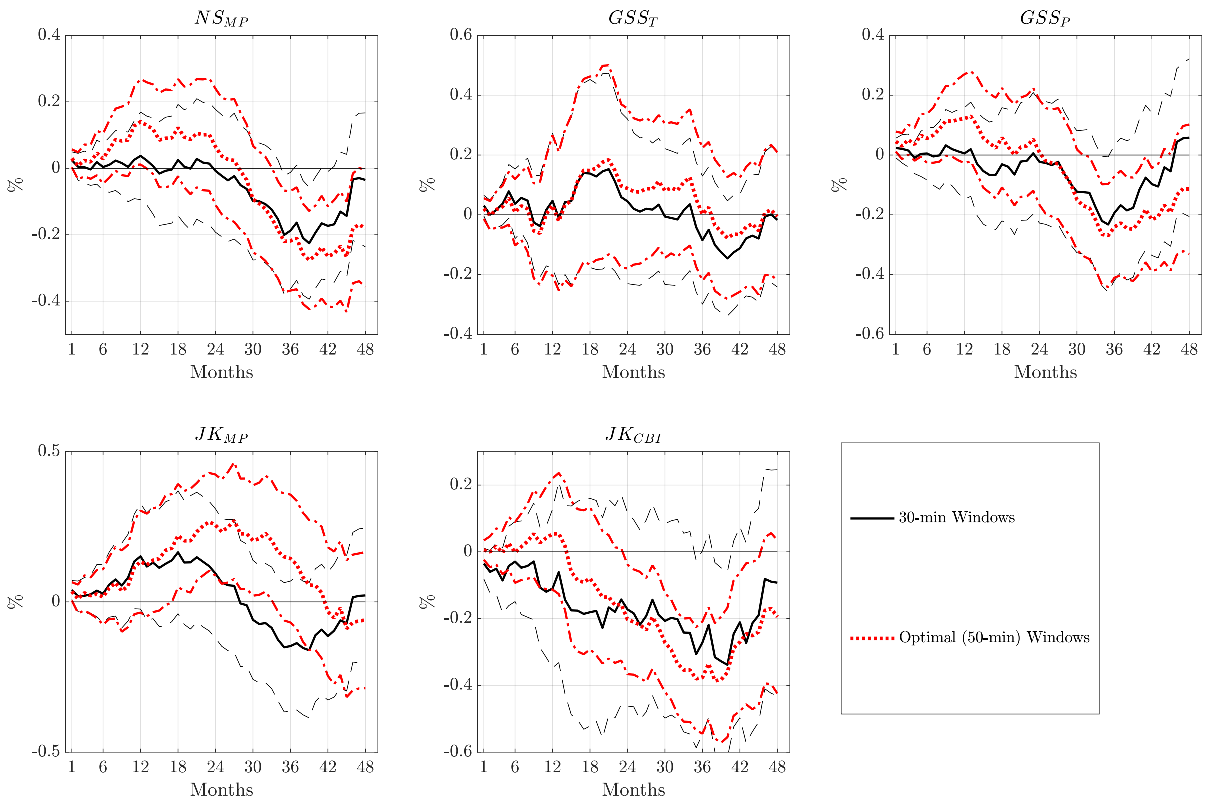


Figure E16: Effects of Event Window Choice on IP Impulse Responses

Notes: Impulse responses are calculated using the lag-augmented local projection from Olea and Plagborg-Møller (2021). Confidence bands are at the 90% level. Standard errors are Eicker-Huber-White heteroscedasticity-robust standard errors. The responses are to a 100 basis point increase in the monetary policy shock series. The monetary policy shocks are constructed using the full set of monetary policy surprise instruments.

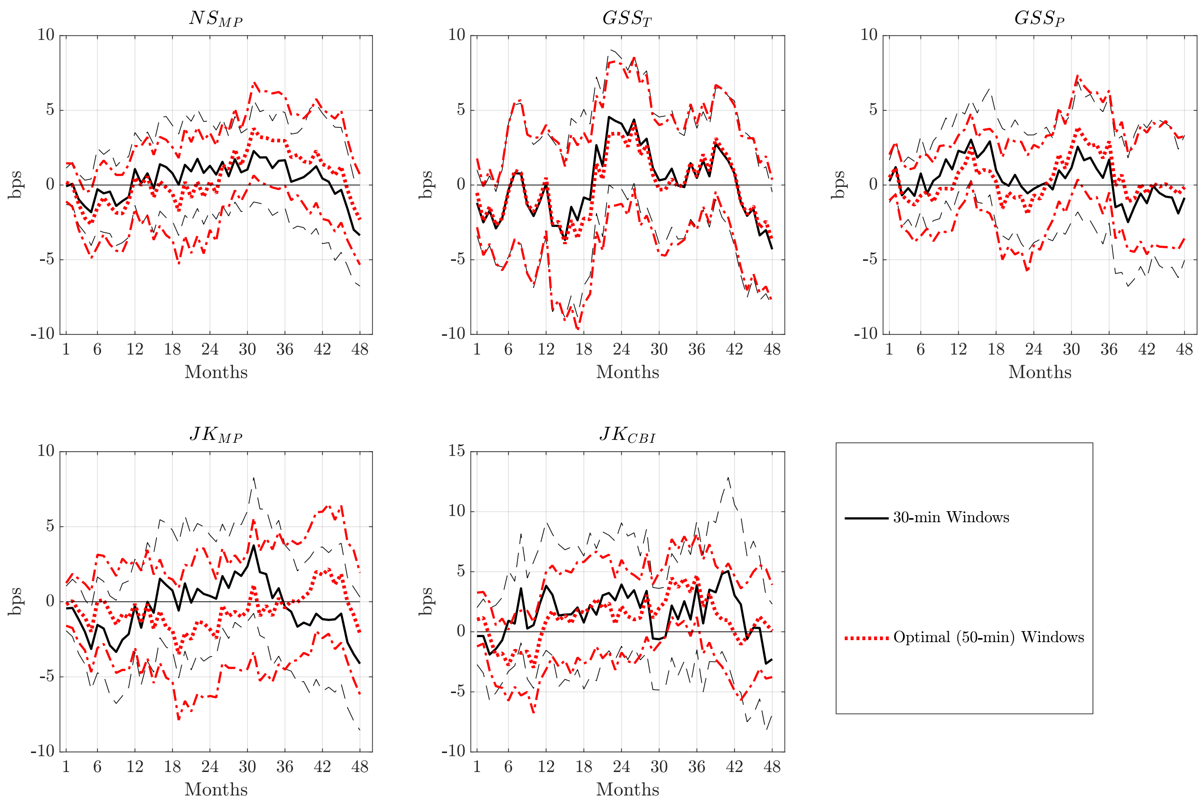


Figure E17: Effects of Event Window Choice on Excess Bond Premium Impulse Responses

Notes: Impulse responses are calculated using the lag-augmented local projection from Olea and Plagborg-Møller (2021). Confidence bands are at the 90% level. Standard errors are Eicker-Huber-White heteroscedasticity-robust standard errors. The responses are to a 100 basis point increase in the monetary policy shock series. The monetary policy shocks are constructed using the full set of monetary policy surprise instruments.

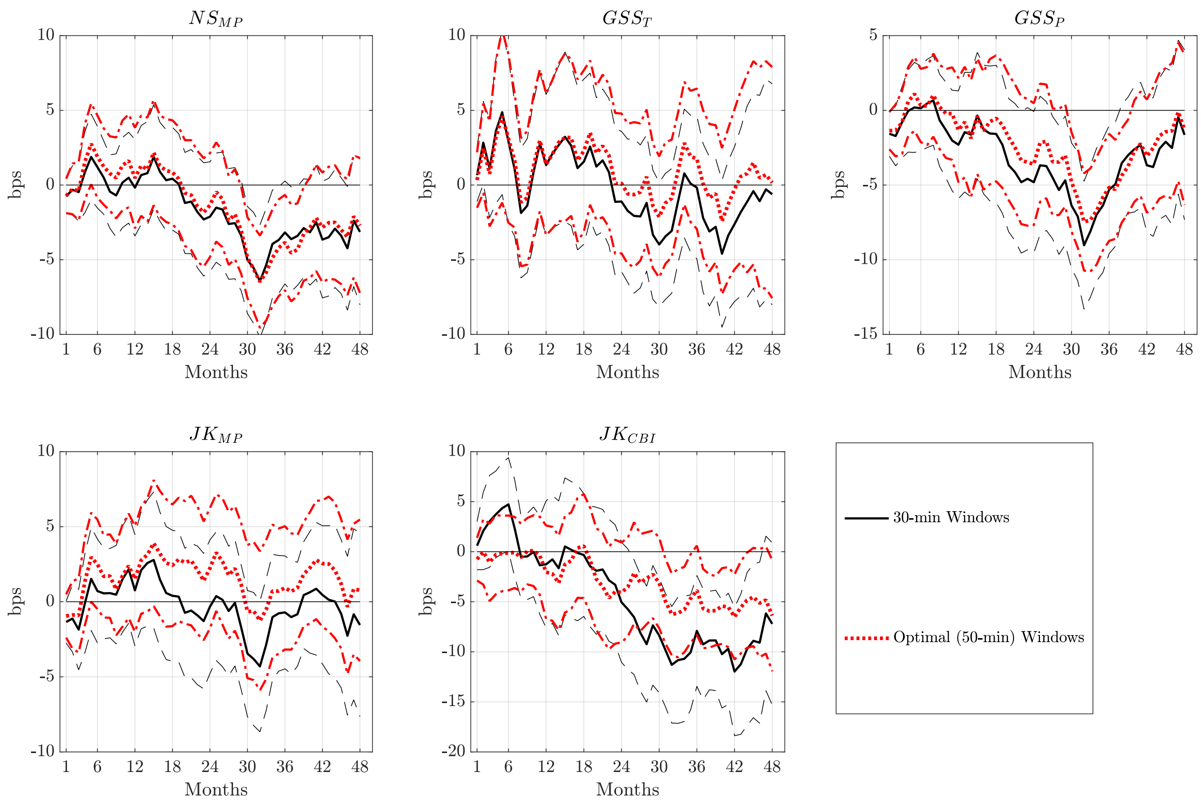


Figure E18: Effects of Event Window Choice on Two-year Treasury Yield Impulse Responses

Notes: Impulse responses are calculated using the lag-augmented local projection from Olea and Plagborg-Møller (2021). Confidence bands are at the 90% level. Standard errors are Eicker-Huber-White heteroscedasticity-robust standard errors. The responses are to a 100 basis point increase in the monetary policy shock series. The monetary policy shocks are constructed using the full set of monetary policy surprise instruments.

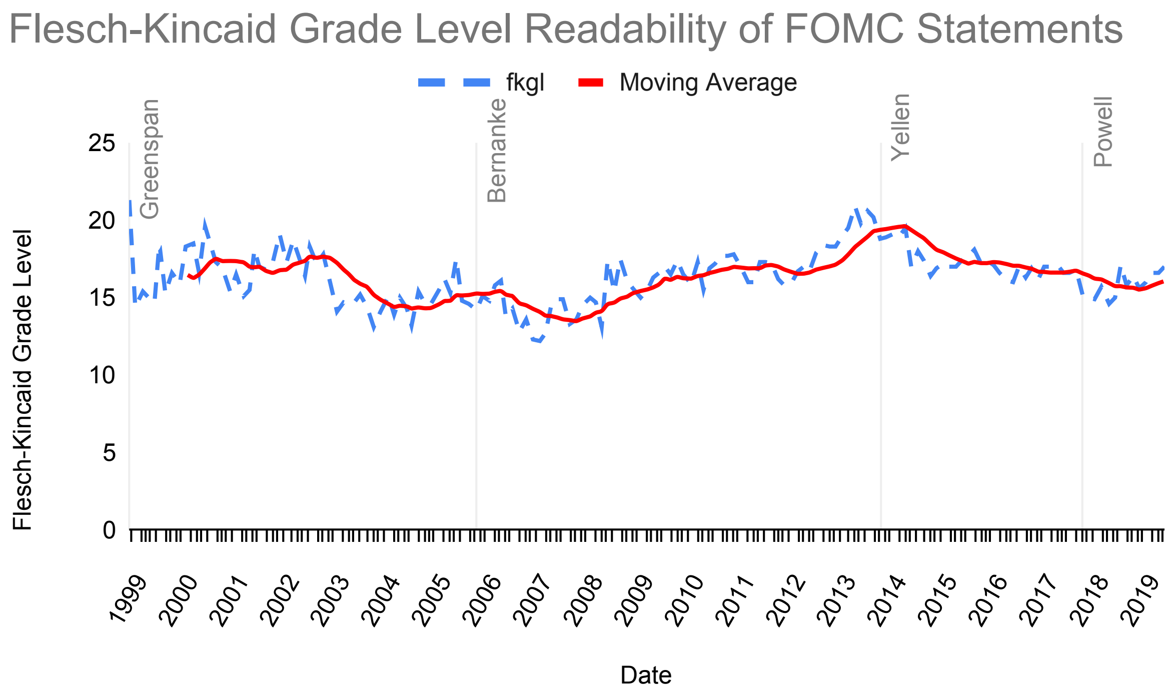


Figure E19: Flesch-Kincaid Grade Level Readability of FOMC Statements

Notes: The complexity of FOMC statements is measured by the Flesch-Kincaid Grade Level, defined as: $0.39 \times$ average sentence length $+ 11.8 \times$ average number of syllables per word $- 15.59$. From left to right, the vertical grey lines indicate the first FOMC meeting with Greenspan, Bernanke, Yellen, and Powell as Fed Chair. The moving average in solid red is calculated with a period of 10. For a description of U.S. education grade levels, see https://en.wikipedia.org/wiki/Education_in_the_United_States.

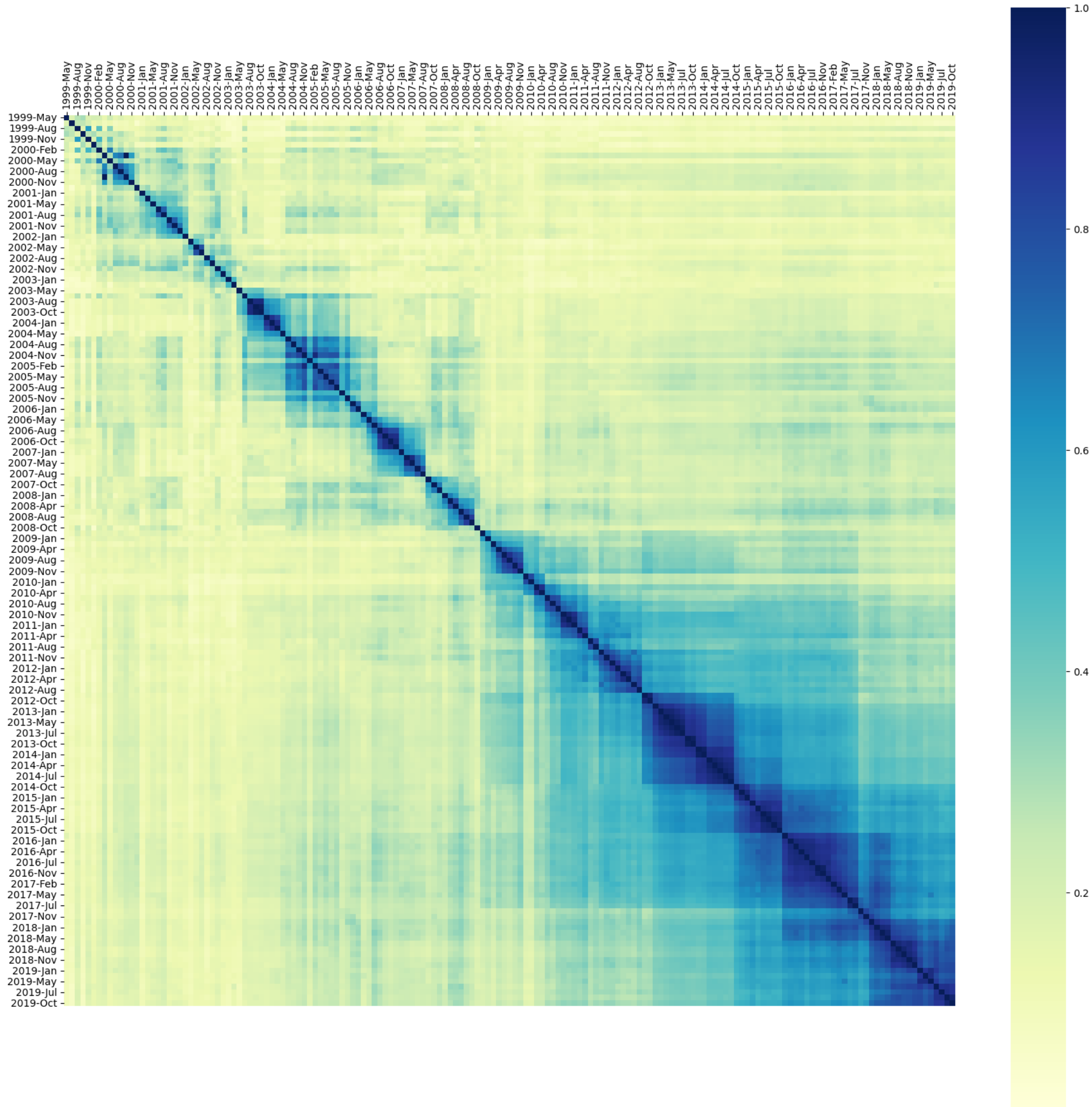


Figure E20: Document Similarity Matrix for FOMC Statements

Notes: Each element is the cosine similarity between two FOMC statements of that row and column. The cosine similarity value measures how similar the terms of both statements are. The darker shade of blue represents pairs of documents that have higher similarity measures closer to one, whilst the light shade of green represents pairs of documents that do not have terms in common and similarity values closer to zero. The main diagonal has all ones because the cosine similarity value is calculated between an FOMC statement with itself. All FOMC statements have been preprocessed such that all words are lowercase, all words are reduced to base form, and all stop-words are removed.

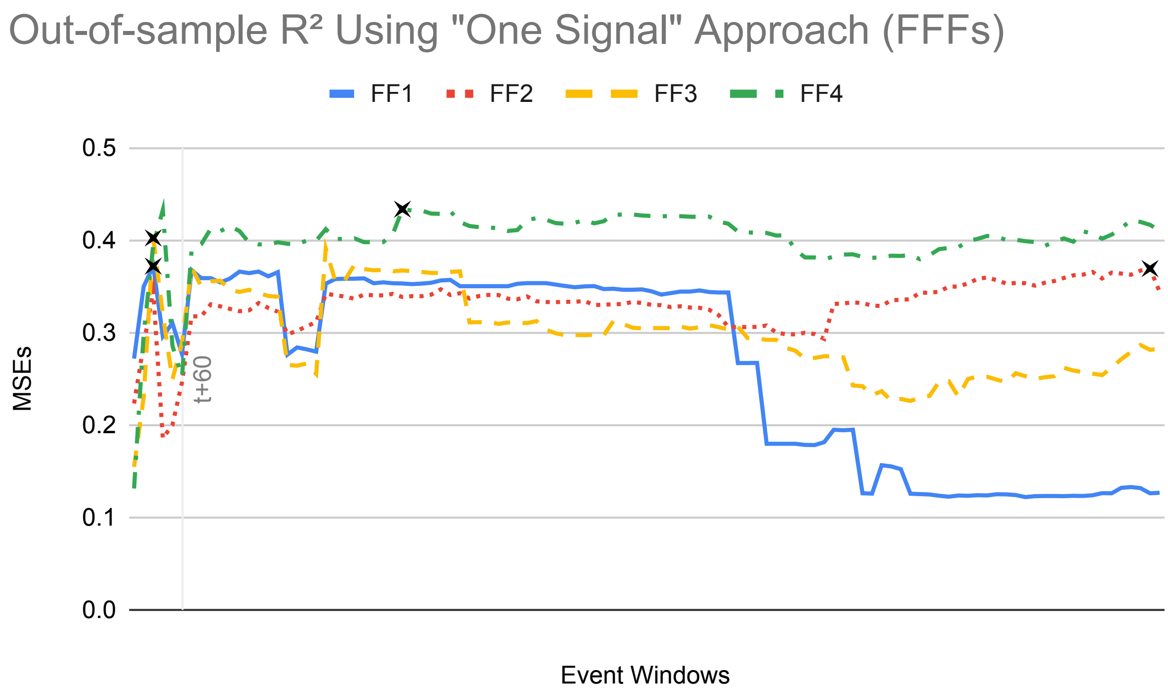


Figure E21: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for Federal Funds Futures

Notes: The horizontal axis depicts the event window lengths, starting from a 20-minute event window and ending at an event window starting 10 minutes before and ending 18 hours after FOMC statement release. The cross points represent the event window length associated with the largest $\overline{R^2_{OOS}}$. Estimates to the left of the vertical grey line labelled "$t + 60$" are calculated using the "joint' approach. The 1-month and 3-month-ahead federal funds futures contract are the two exceptions that sees the largest $\overline{R^2_{OOS}}$ within a window length than the systematically estimated length. As a robustness check, $\overline{R^2_{OOS}}$ was calculated with systematic estimation for both futures and considered event window lengths, which can be found in the rightmost box-and-whisker plot of Appendix Sub-figures E2b and E2d, respectively.

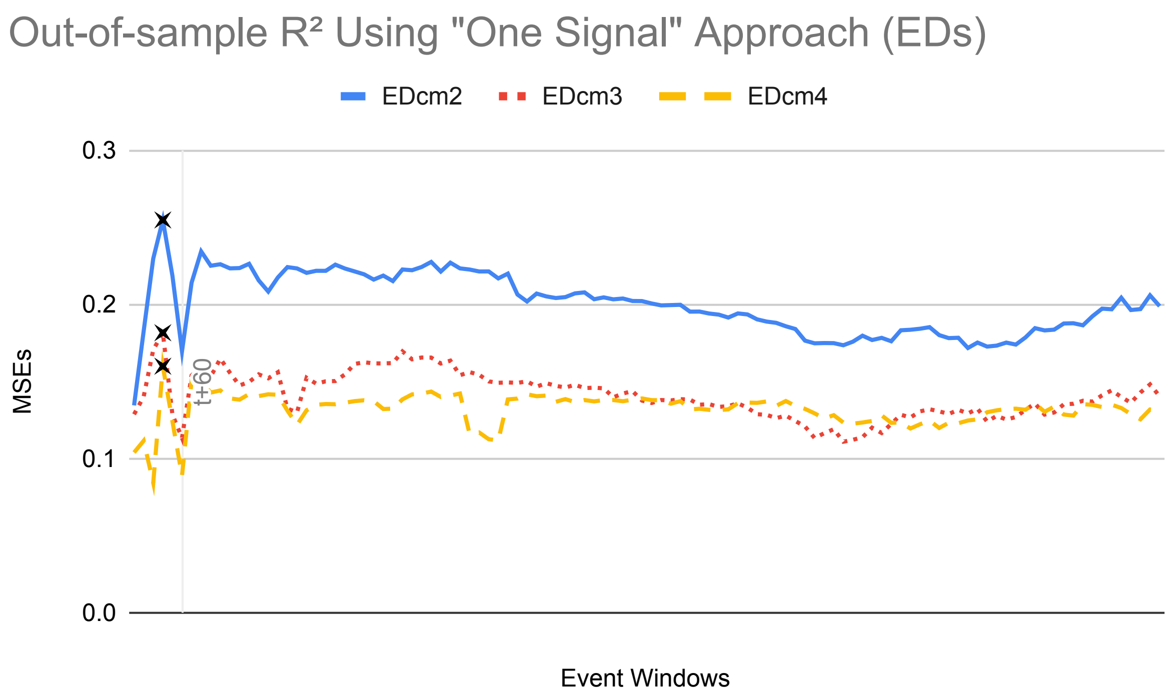


Figure E22: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for Eurodollar Futures

Notes: The horizontal axis depicts the event window lengths, starting from a 20-minute event window and ending at an event window starting 10 minutes before and ending 18 hours after FOMC statement release. The cross points represent the event window length associated with the largest $\overline{R^2_{OOS}}$. Estimates to the left of the vertical grey line labelled "$t + 60$" are calculated using the "joint" approach.

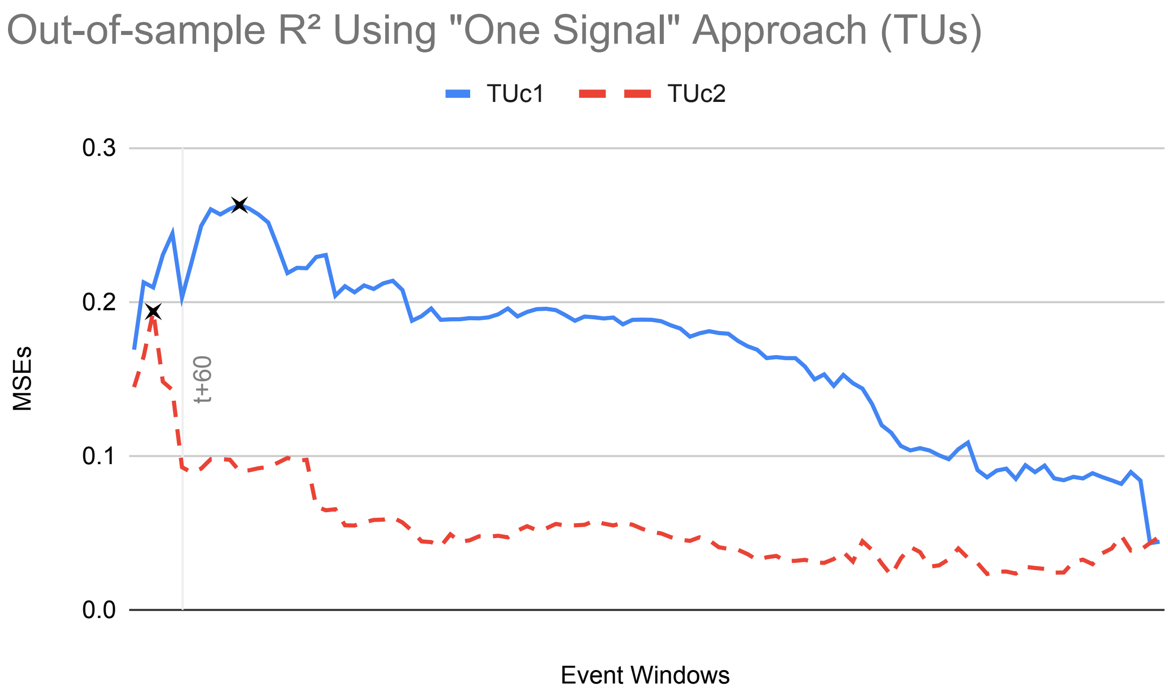


Figure E23: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for 2-year Treasury Futures

Notes: The horizontal axis depicts the event window lengths, starting from the systematically estimated event window length and ending at an event window starting 10 minutes before and ending 18 hours after FOMC statement release. The cross points represent the event window length associated with the largest $\overline{R^2_{OOS}}$. Estimates to the left of the vertical grey line labelled "$t + 60$" are calculated using the "joint" approach. The front-month 2-year Treasury futures contract is the one exception that sees the largest $\overline{R^2_{OOS}}$ within a window length than the systematically estimated length. As a robustness check, $\overline{R^2_{OOS}}$ was calculated with systematic estimation for the futures contract and considered event window length, which can be found in the rightmost box-and-whisker plot of Appendix Sub-figure E4a.

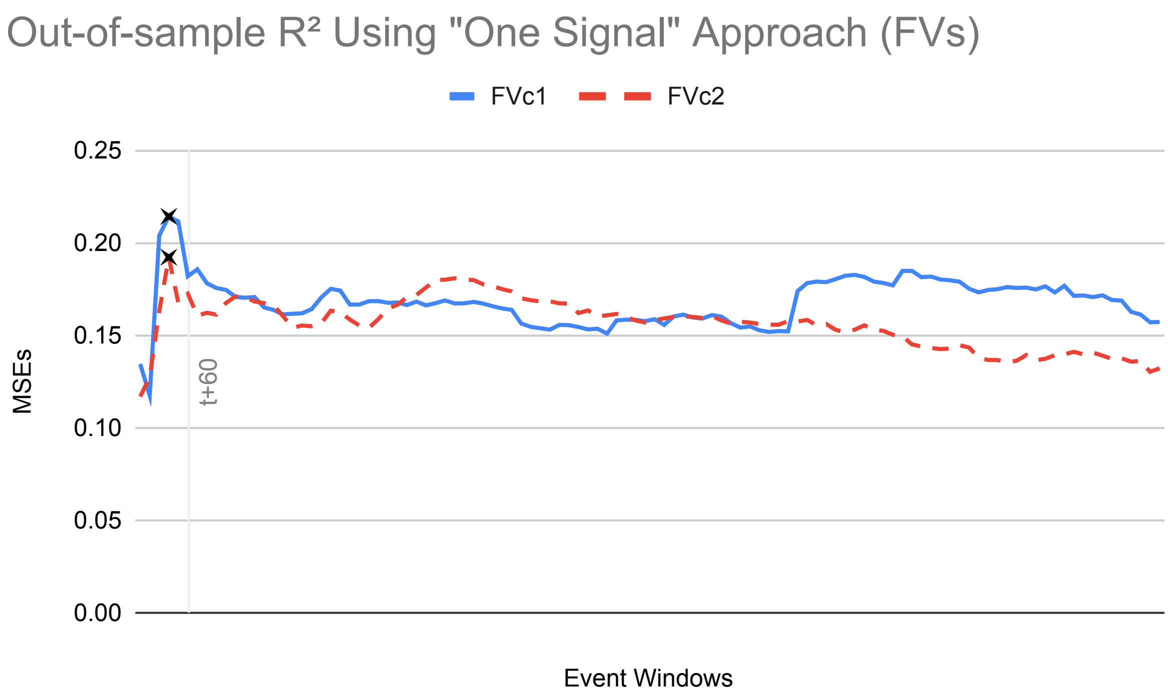


Figure E24: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for 5-year Treasury Futures

Notes: The horizontal axis depicts the event window lengths, starting from the systematically estimated event window length and ending at an event window starting 10 minutes before and ending 18 hours after FOMC statement release. The cross points represent the event window length associated with the largest $\overline{R^2_{OOS}}$. Estimates to the left of the vertical grey line labelled "$t + 60$" are calculated using the "joint" approach.

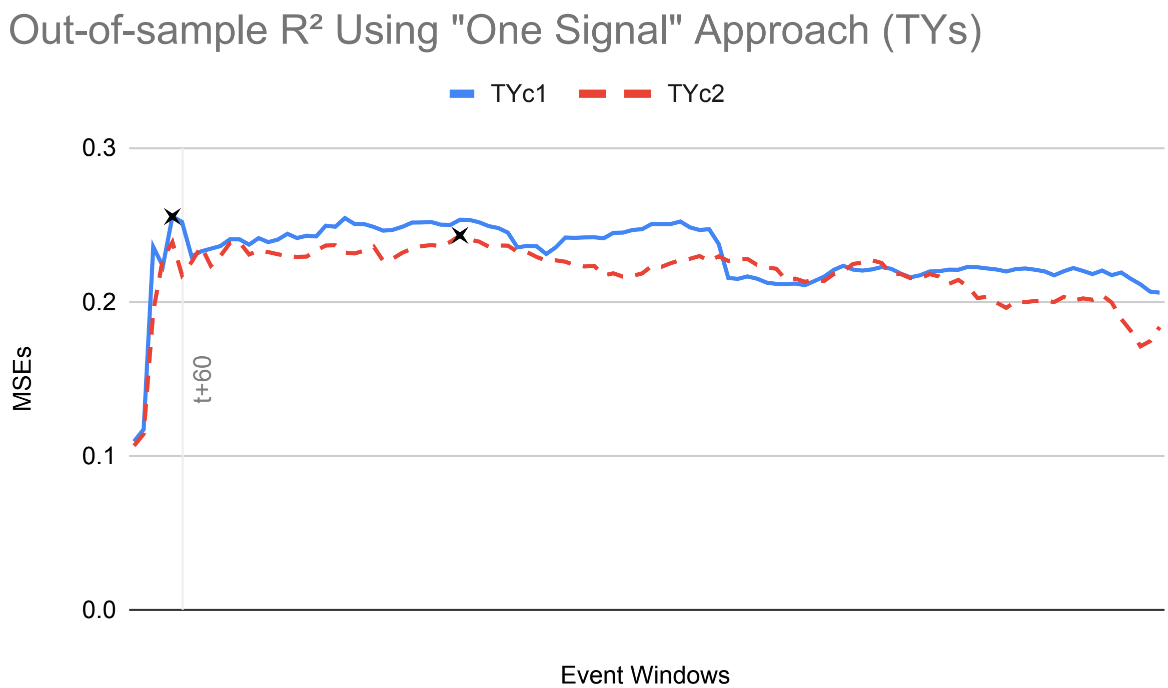


Figure E25: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for 10-year Treasury Futures

Notes: The horizontal axis depicts the event window lengths, starting from the systematically estimated event window length and ending at an event window starting 10 minutes before and ending 18 hours after FOMC statement release. The cross points represent the event window length associated with the largest $\overline{R^2_{OOS}}$. Estimates to the left of the vertical grey line labelled "$t + 60$" are calculated using the "joint" approach. The second-month 10-year Treasury futures contract is the one exception that sees the largest $\overline{R^2_{OOS}}$ within a window length than the systematically estimated length. As a robustness check, $\overline{R^2_{OOS}}$ was calculated with systematic estimation for the futures contract and considered event window length, which can be found in the rightmost box-and-whisker plot of Appendix Sub-figure E6b.

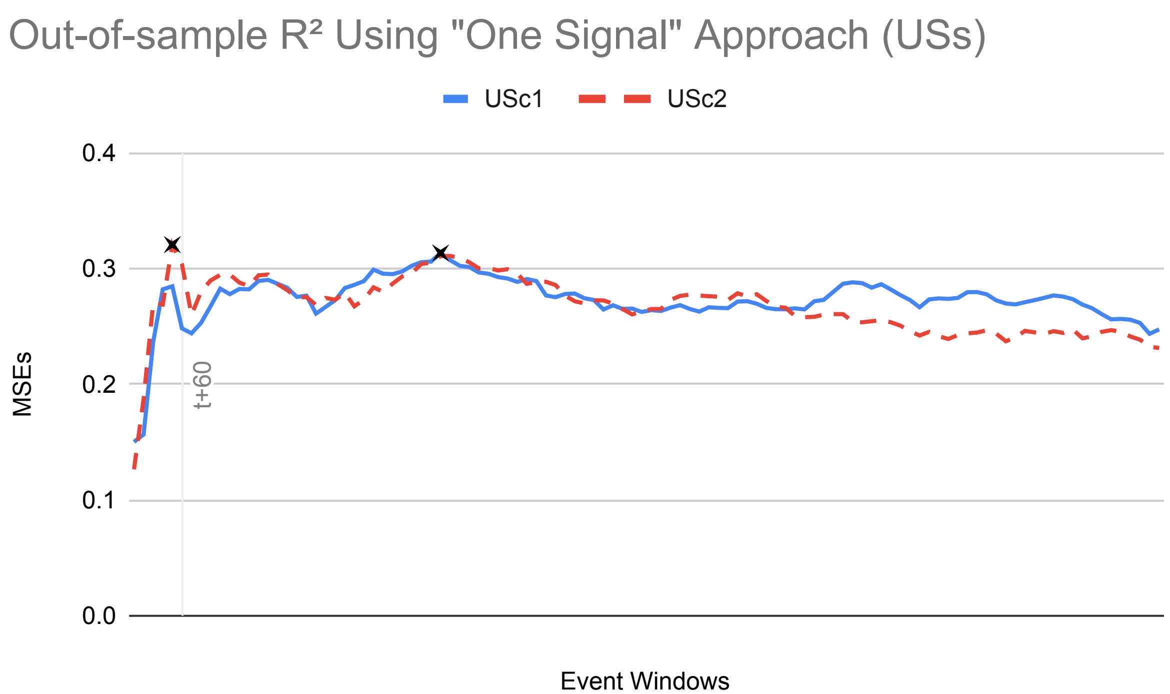


Figure E26: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for 30-year Treasury Futures

Notes: The horizontal axis depicts the event window lengths, starting from the systematically estimated event window length and ending at an event window starting 10 minutes before and ending 18 hours after FOMC statement release. The cross points represent the event window length associated with the largest $\overline{R^2_{OOS}}$. Estimates to the left of the vertical grey line labelled "$t + 60$" are calculated using the "joint" approach. The front-month 30-year Treasury futures contract is the one exception that sees the largest $\overline{R^2_{OOS}}$ within a window length than the systematically estimated length. As a robustness check, $\overline{R^2_{OOS}}$ was calculated with systematic estimation for the futures contract and considered event window length, which can be found in the rightmost box-and-whisker plot of Appendix Sub-figure E7a.

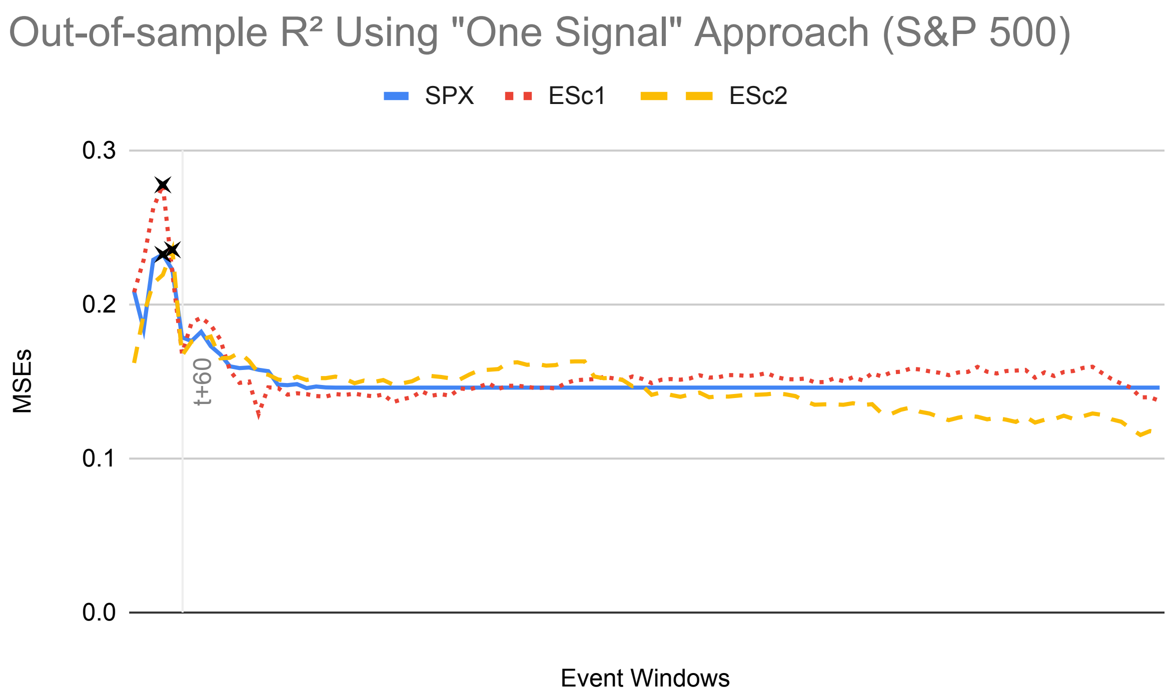
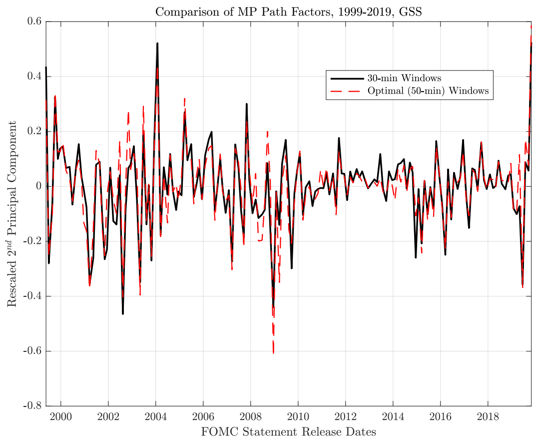
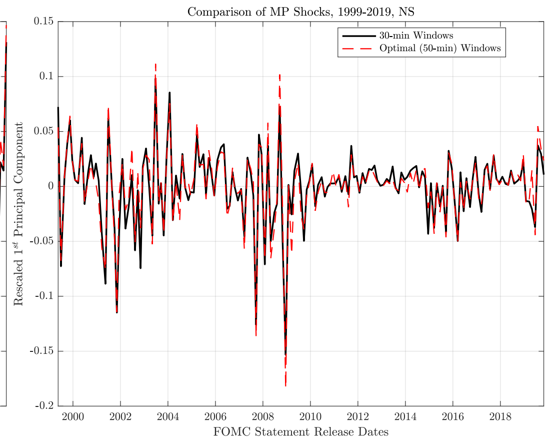


Figure E27: $\overline{R^2_{OOS}}$ Calculated Using the “One Signal” Approach for S&P 500 and E-mini Futures

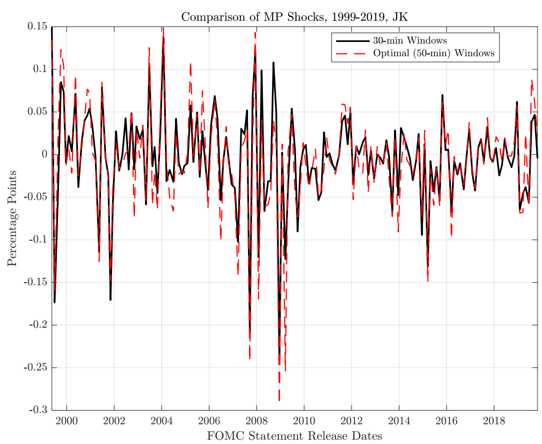
Notes: The horizontal axis depicts the event window lengths, starting from the systematically estimated event window length and ending at an event window starting 10 minutes before and ending 18 hours after FOMC statement release. The cross points represent the event window length associated with the largest $\overline{R^2_{OOS}}$. Estimates to the left of the vertical grey line labelled “$t + 60$” are calculated using the “joint” approach.



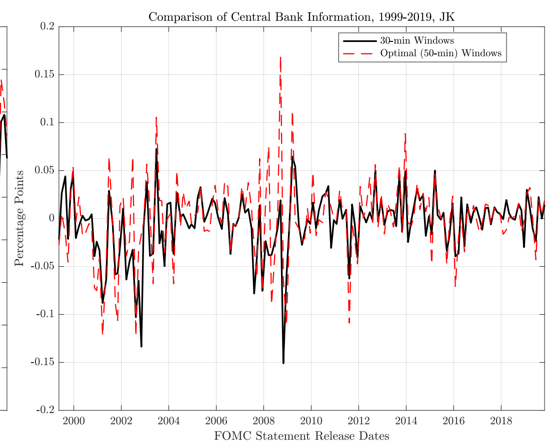
(a) GSS_P



(b) NS_{MP}



(c) JK_{MP}



(d) JK_{CBI}

Figure E28: Comparing Monetary Policy Shock Series Derived from Optimal Window Length v. 30 Minutes

Notes: For each sub-figure, the horizontal axis represents FOMC statement release dates. The vertical axis depicts the principal components scaled according to the original specification of the authors. For all construction methods, the black-solid and red-dotted lines represent the shocks derived from surprises measured within 30 minutes and the median optimal event window length of 50 minutes, respectively.

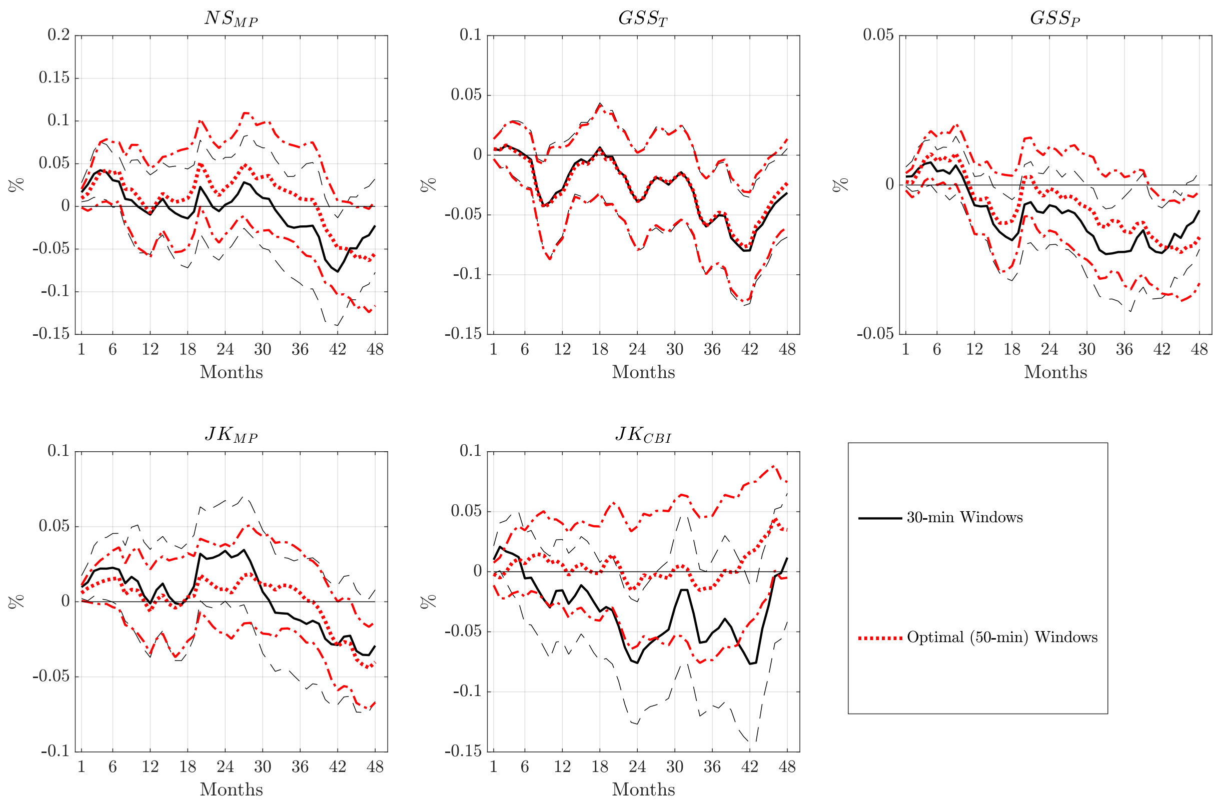


Figure E29: Effects of Event Window Choice on CPI Impulse Responses

Notes: Impulse responses are calculated using the lag-augmented local projection from Olea and Plagborg-Møller (2021). Confidence bands are at the 90% level. Standard errors are Eicker-Huber-White heteroscedasticity-robust standard errors. The responses are to a 100 basis point increase in the monetary policy shock series. The monetary policy shocks are constructed using the full set of monetary policy surprise instruments.

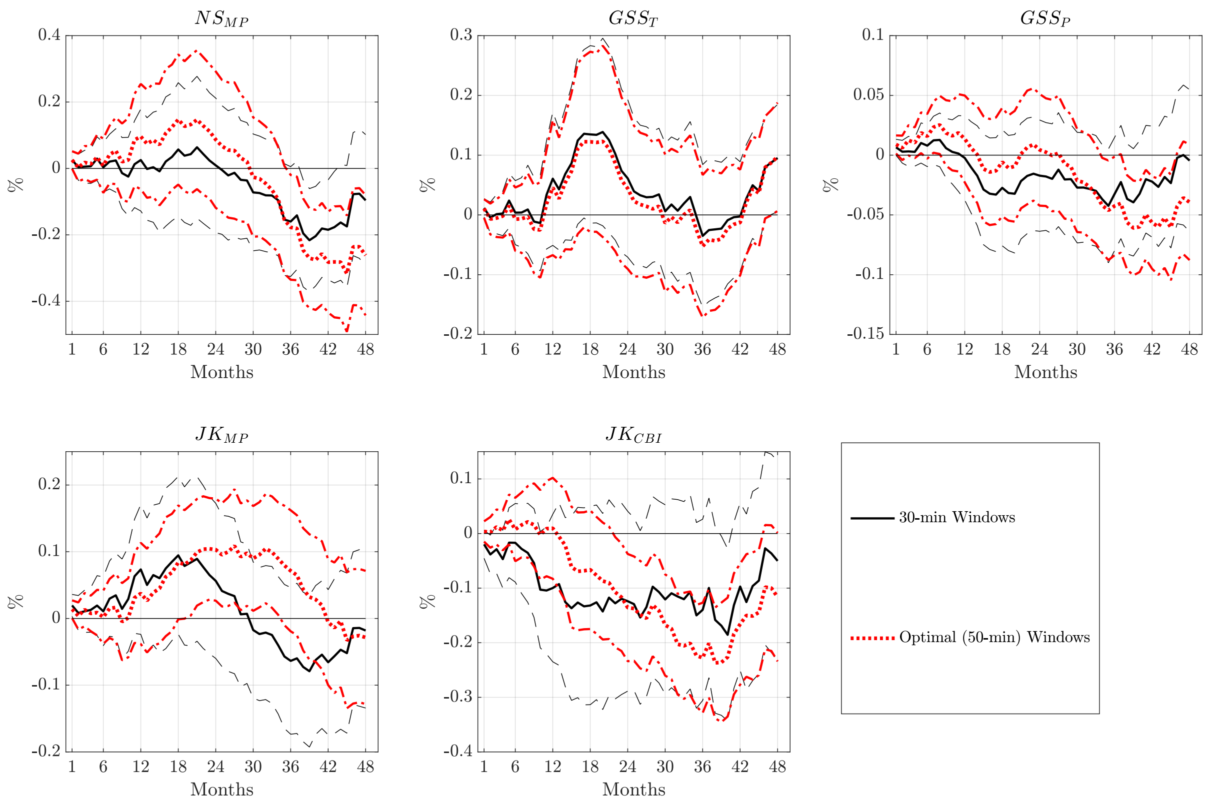


Figure E30: Effects of Event Window Choice on IP Impulse Responses

Notes: Impulse responses are calculated using the lag-augmented local projection from Olea and Plagborg-Møller (2021). Confidence bands are at the 90% level. Standard errors are Eicker-Huber-White heteroscedasticity-robust standard errors. The responses are to a 100 basis point increase in the monetary policy shock series. The monetary policy shocks are constructed using the full set of monetary policy surprise instruments.

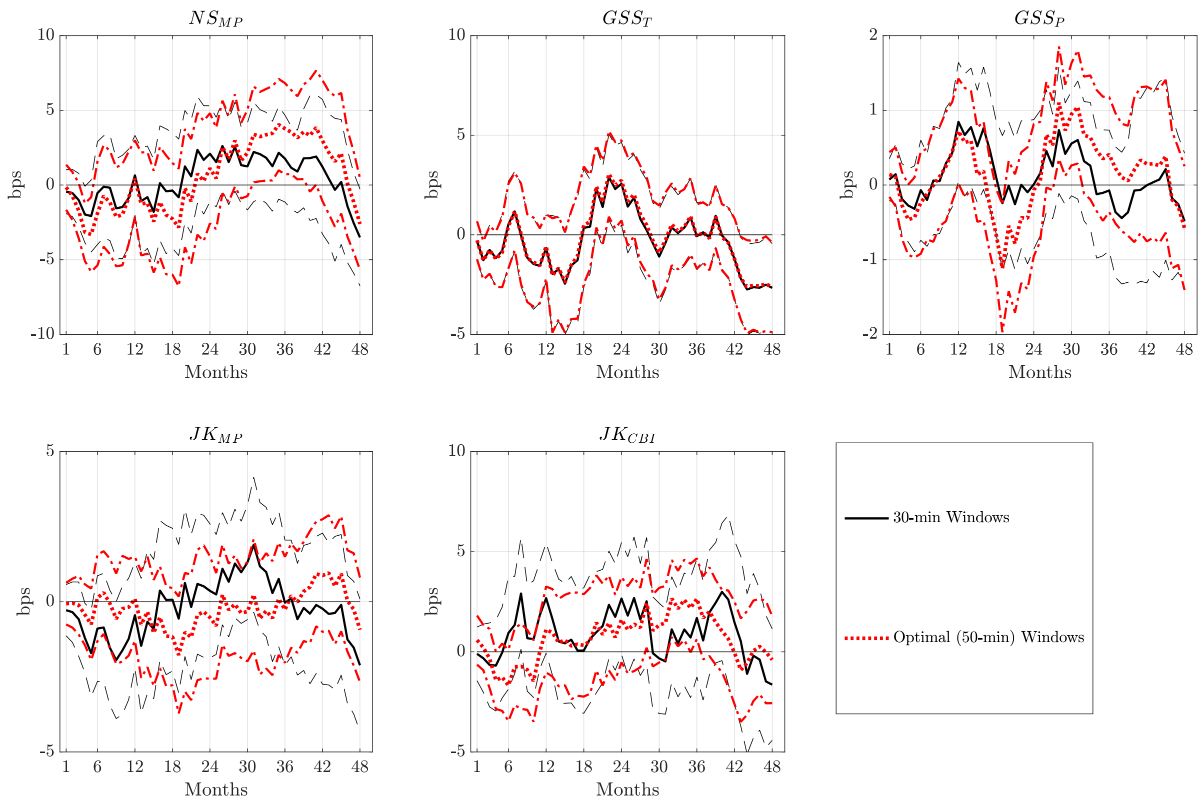


Figure E31: Effects of Event Window Choice on Excess Bond Premium Impulse Responses

Notes: Impulse responses are calculated using the lag-augmented local projection from Olea and Plagborg-Møller (2021). Confidence bands are at the 90% level. Standard errors are Eicker-Huber-White heteroscedasticity-robust standard errors. The responses are to a 100 basis point increase in the monetary policy shock series. The monetary policy shocks are constructed using the full set of monetary policy surprise instruments.

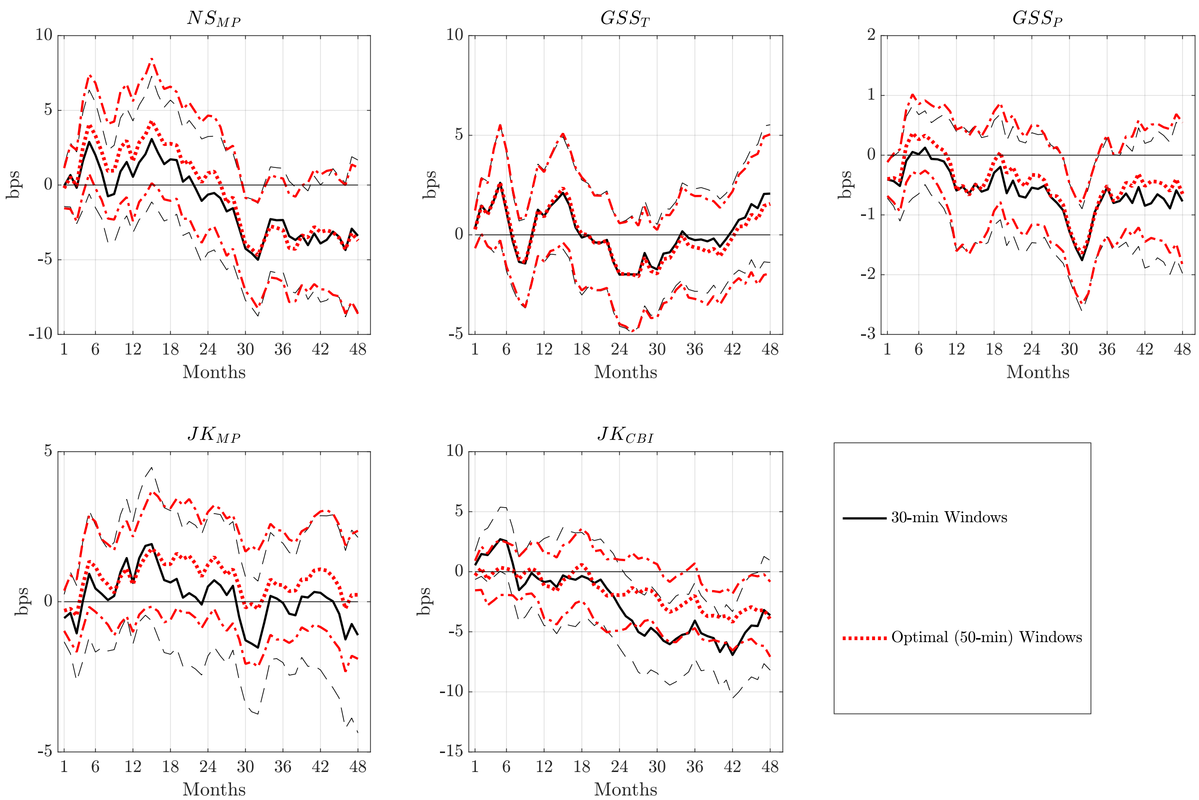


Figure E32: Effects of Event Window Choice on Two-year Treasury Yield Impulse Responses

Notes: Impulse responses are calculated using the lag-augmented local projection from Olea and Plagborg-Møller (2021). Confidence bands are at the 90% level. Standard errors are Eicker-Huber-White heteroscedasticity-robust standard errors. The responses are to a 100 basis point increase in the monetary policy shock series. The monetary policy shocks are constructed using the full set of monetary policy surprise instruments.

Appendix Tables

	Scenario 1	Scenario 2	Scenario 3
Framework Simulation Parameters			
P_i^f	$\in [-100, 100]$	$\in [-100, 100]$	$\in [-100, 100]$
$\varepsilon_{i,0}^c$	$\in [-100, 100]$	$\in [-100, 100]$	$\in [-100, 100]$
$\varepsilon_{i,0}^n$	0	0	0
σ_c	100	0.1	50
$\mathcal{D}$	0.5	1	0.75
σ_n	0.1	10	1
ρ_c	0.47	0.47	0.47
σ_s	$\in \mathbb{R}$	$\in \mathbb{R}$	$\in \mathbb{R}$
Simulation Results			
t^*	16	2	10
$\hat{t}$	15	2	10

Table F1: Framework Simulation Parameters and Results for Different Asset Market Scenarios
Notes: The results are from 10,000 simulations of the asset price framework for the three asset market scenarios considered. t^* and $\hat{t}$ are defined as the time at which the asset market fully reacts to news on average, where the former is calculated using the fundamental price component and the latter with the observed signal according to the motivating framework.

Name	Maturity	Ticker	Sample	Observations
Inputs				
FOMC Statements	N/A	N/A	1999–2019	165
Outputs				
Federal Funds Rate Futures	Front-month	$FFc1$	1999–2019	165
Federal Funds Rate Futures	1-month-ahead	$FFc2$	1999–2019	165
Federal Funds Rate Futures	2-month-ahead	$FFc3$	1999–2019	165
Federal Funds Rate Futures	3-month-ahead	$FFc4$	1999–2019	165
Eurodollar Futures	2-quarter	$EDcm2$	1999–2019	165
Eurodollar Futures	3-quarter	$EDcm3$	1999–2019	165
Eurodollar Futures	4-quarter	$EDcm4$	1999–2019	165
2-year Treasury Futures	Front-month	$TUc1$	1999–2019	165
2-year Treasury Futures	Second-month	$TUc2$	1999–2019	165
5-year Treasury Futures	Front-month	$FVc1$	1999–2019	165
5-year Treasury Futures	Second-month	$FVc2$	1999–2019	165
10-year Treasury Futures	Front-month	$TYc1$	1999–2019	165
10-year Treasury Futures	Second-month	$TYc2$	1999–2019	165
30-year Treasury Futures	Front-month	$USc1$	1999–2019	165
30-year Treasury Futures	Second-month	$USc2$	1999–2019	165
S&P 500 Index	N/A	SPX	1999–2019	165
S&P 500 E-mini Futures	Front-month	$ESc1$	1999–2019	165
S&P 500 E-mini Futures	Second-month	$ESc2$	1999–2019	165

Table F2: Independent and Dependent Variables for Systematic Estimation of Appropriate Event Window Lengths by Neural Networks

Notes: The table shows the FOMC statements and financial market asset prices considered as independent and dependent variables, respectively, in my analysis. The statements are collected from the Board of Governors of the Federal Reserve System website and the asset prices are from the Thomson Reuters Tick History database. All series begin in May 1999 and end in October 2019. *Ticker* refers to the Reuters Instrument Code, which uniquely identifies the financial instrument. The letters *c* and *cm* standard for continuous futures contracts.

Hyperparameter	Value
Number of Layers	12
Hidden Size	768
Number of Attention Heads	12
Attention Head Size	64
Hidden Dropout	0.1
Attention Dropout	0.1
Max Sequence Length	512
Vocab Size	32,000
Training Batch Size	8
Evaluation Batch Size	8
Max Num of Steps	2040
Max Num of Epochs	120
Manual Random Seed	47
Learning Rate	*
Warmup Ratio	0.06
Adam Epsilon	1e-8
Learning Rate Decay	Linear

Table F3: XLNet-Base Hyperparameters for Fine-tuning

Notes: * denotes the hyperparameter that undergoes tuning during the fine-tuning process. The random seed is set as a constant to ensure replicability by always initialising random components (e.g., weights and biases) to the same initial values.

Asset	$\overline{R^2_{OOS}}$, 30-min	$\overline{R^2_{OOS}}$, Optimal	**Difference**
$FF1$	35.0%	37.2%	+**2.2 p.p.**
$FF2$	28.7%	34.5%	+**5.8 p.p.**
$FF3$	23.0%	40.2%	+**17.2 p.p.**
$FF4$	29.8%	43.3%	+**13.5 p.p.**
$EDcm2$	18.3%	23.3%	+**5 p.p.**
$EDcm3$	14.0%	18.2%	+**4.2 p.p.**
$EDcm4$	11.2%	16.0%	+**4.8 p.p.**
$TUc1$	21.3%	24.4%	+**3.1 p.p.**
$TUc2$	16.5%	19.4%	+**2.9 p.p.**
$FVc1$	11.7%	21.4%	+**9.7 p.p.**
$FVc2$	12.7%	19.2%	+**6.5 p.p.**
$TYc1$	11.7%	25.5%	+**13.8 p.p.**
$TYc2$	11.4%	23.9%	+**12.5 p.p.**
$USc1$	15.7%	28.5%	+**12.8 p.p.**
$USc2$	18.7%	32.1%	+**13.4 p.p.**
SPX	18.4%	23.2%	+**4.8 p.p.**
$ESc1$	22.9%	27.7%	+**4.8 p.p.**
$ESc2$	19.3%	23.5%	+**4.2 p.p.**

Table F4: Differences of $\overline{R^2_{OOS}}$ between 30-minute and Optimal Event Windows

Series Name	Notation	Description
Gürkaynak, Sack, and Swanson (2005) Target Shocks	GSS_T	First principal component of monetary policy surprises that is rotated such that it drives unexpected changes in the surprise in the current federal funds rate. The authors re-scale the shock originally to be one-for-one with the surprise in the current federal funds rate.
Gürkaynak, Sack, and Swanson (2005) Path Shocks	GSS_P	Second principal component of monetary policy surprises that is rotated such that on average, it has no effect on the surprise in the current federal funds rate. The authors re-scale the shock originally to have equal effect with GSS_T on the surprise in the four-quarter Eurodollar futures.
Nakamura and Steinsson (2018) Shocks	NS_{MP}	First principal component of monetary policy surprises. The authors re-scale the shock originally to be one-for-one with the daily change in the zero-coupon, nominal one-year Treasury yield.
Jarociński and Karadi (2020) Shocks	JK_{MP}	First principal component of monetary policy surprises that have negative co-movement with stock market changes. The authors make the first principal component to have equal standard deviation with that of the surprise in the four-quarter Eurodollar futures.
Jarociński and Karadi (2020) Central Bank Information	JK_{CBI}	First principal component of monetary policy surprises that have positive co-movement with stock market changes. The authors make the first principal component to have equal standard deviation with that of the surprise in the four-quarter Eurodollar futures.

Table F5: Monetary Policy Shock Series

Notes: All monetary policy shocks are re-scaled to be one-for-one with the daily change in the zero-coupon, nominal one-year Treasury yield for easy interpretation and comparison. Before applying sign restrictions to obtain JK_{MP} and JK_{CBI}, I follow Jarociński and Karadi (2020) by first re-scaling the first principal component of monetary policy surprises in terms of the standard deviation of the zero-coupon, nominal one-year Treasury yield.

Metric	GSS_T	GSS_P	NS_{MP}	JK_{MP}	JK_{CBI}
Count	165	165	165	165	165
	(165)	(165)	(165)	(165)	(165)
Mean	0	0	0	-0.0023	-0.0028
	(0)	(0)	(0)	(-0.0035)	(-0.0004)
SD	0.0219	0.0282	0.0356	0.0305	0.0180
	(0.0216)	(0.0309)	(0.0375)	(0.0297)	(0.0233)
Max	0.0756	0.1223	0.1333	0.0968	0.0388
	(0.0652)	(0.1079)	(0.1135)	(0.0840)	(0.0984)
75^{th}	0.0076	0.0134	0.0193	0.0175	0.0069
	(0.0084)	(0.0167)	(0.0203)	(0.0124)	(0.0112)
Median	0.0011	-0.0010	0.0033	-0.0017	-0.0003
	(0.0023)	(0.0019)	(0.0003)	(-0.0009)	(0.0022)
25^{th}	-0.0069	-0.0124	-0.0187	-0.0153	-0.0086
	(-0.0077)	(-0.0127)	(-0.0164)	(-0.0138)	(-0.0103)
Min	-0.1038	-0.1611	-0.1403	-0.1222	-0.0812
	(-0.1025)	(-0.1618)	(-0.1722)	(-0.1491)	(-0.0828)

Table F6: Descriptive Statistics for Monetary Policy Shock Series, FOMC Statement Frequency for 1999–2019
Notes: Numbers in parentheses are summary statistics for each shock series derived from monetary policy surprises calculated within the median optimal event window length of 50 minutes. All shocks series have been re-scaled to be one-for-one with the daily change in the zero-coupon, nominal one-year Treasury yield.

Metric	$ln(CPI)$	$ln(IP)$	EBP	TY_2
Count	246	246	246	246
Mean	5.356	4.597	0.110	2.210
SD	0.126	0.057	0.715	1.794
Max	5.551	4.706	3.283	6.650
75^{th}	5.466	4.642	0.341	3.569
Median	5.381	4.605	-0.108	1.673
25^{th}	5.241	4.548	-0.334	0.662
Min	5.112	4.467	-1.140	0.188

Table F7: Descriptive Statistics for Impulse Response Variables, Monthly for 1999–2019
Notes: All logs are natural logarithms. Consumer Price Index (CPI) and Industrial Production (IP) are sourced from the Federal Reserve Economic Data. The Excess Bond Premium (EBP) is from Gilchrist and Zakrajšek (2012) and is expressed in percentage points. The two-year Treasury yield (TY_2) is from Gürkaynak, Sack, and Swanson (2005).

Metric	GSS_T	GSS_P	NS_{MP}	JK_{MP}	JK_{CBI}
Count	165	165	165	165	165
	(165)	(165)	(165)	(165)	(165)
Mean	0	0	0	-0.0015	-0.0019
	(0)	(0)	(0)	(-0.0024)	(-0.0002)
SD	0.0180	0.0230	0.0291	0.0250	0.0148
	(0.0177)	(0.0253)	(0.0307)	(0.0243)	(0.0190)
Max	0.0756	0.1223	0.1333	0.0968	0.0388
	(0.0652)	(0.1079)	(0.1135)	(0.0840)	(0.0984)
75^{th}	0.0042	0.0078	0.0114	0.0055	0.0031
	(0.0042)	(0.0089)	(0.0110)	(0.0058)	(0.0075)
Median	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)
25^{th}	-0.0021	-0.0064	-0.0036	-0.0073	-0.0037
	(-0.0011)	(-0.0052)	(-0.0039)	(-0.0070)	(-0.0035)
Min	-0.1038	-0.1611	-0.1403	-0.1222	-0.0812
	(-0.1025)	(-0.1618)	(-0.1722)	(-0.1491)	(-0.0828)

Table F8: Descriptive Statistics for Monetary Policy Shock Series, Monthly for 1999–2019
Notes: Numbers in parentheses are summary statistics for each shock series derived from monetary policy surprises calculated within the median optimal event window length of 50 minutes. All shocks series have been re-scaled to be one-for-one with the daily change in the zero-coupon, nominal one-year Treasury yield. All shocks are zero for any month that does not have an FOMC meeting.

Metric	FKGL	S^1
Count	165	164
Mean	16.361	0.751
SD	1.715	0.212
Max	21.3	0.984
75^{th}	17.3	0.920
Median	16.5	0.826
25^{th}	15.1	0.622
Min	12.2	0.200

Table F9: Descriptive Statistics for Heterogeneity Analyses
Notes: The complexity of FOMC statements is measured by the Flesch-Kincaid Grade Level (FKGL), defined as: $0.39 \times \text{average sentence length} + 11.8 \times \text{average number of syllables per word} - 15.59$. S^1 is the cosine similarity measure between sequential FOMC statements.

domestic	alreadytight	recov	tilt	buildup
alert	alter	foreign	imbal	undermin
direct	excess	quit	favor	perform
strength	fall	trend	eas	concern
background	firm	gener	demand	potenti
core	subdu	cost	longer	gain

Table F10: FOMC Statement Base Terms with Top 30 TFIDF Scores

Notes: TFIDF is a weighted frequency combining word and document counts such that greater weight is given to words that are more informative about the information content of the FOMC statements relative to other statements where the words are not found. Specifically, the weighted term frequency gives higher weight to terms that occur more frequently in a given document. The term frequency is then divided by the number of documents that has this term appear. The more documents that have the word, the less importance and weight will be given to the word as it is less informative for distinguishing documents from one another.

	30-minute Window				Optimal Window				**Difference**		
	ΔTY_1	ΔTY_2	ΔTY_5	ΔTY_{10}	ΔTY_1	ΔTY_2	ΔTY_5	ΔTY_{10}	ΔTY_2	ΔTY_5	ΔTY_{10}
GSS_T	0.36***	0.28**	0.15	0.05	0.35***	0.27**	0.13	0.03	**-0.01**	**-0.01**	**-0.02**
	(0.11)	(0.13)	(0.12)	(0.11)	(0.11)	(0.13)	(0.11)	(0.10)			
GSS_P	0.21***	0.25***	0.24***	0.15***	0.21***	0.28***	0.27***	0.19***	**+0.01**	**+0.02**	**+0.03**
	(0.02)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.06)			
NS_{MP}	1.00***	1.11***	0.95***	0.57***	1.00***	1.16***	1.04***	0.67***	**+0.05**	**+0.09**	**+0.10**
	(0.09)	(0.14)	(0.18)	(0.19)	(0.11)	(0.15)	(0.20)	(0.22)			
JK_{MP}	0.53***	0.64***	0.61***	0.39***	0.44***	0.55***	0.56***	0.39***	**-0.09**	**-0.09**	**-0.06**
	(0.08)	(0.11)	(0.16)	(0.18)	(0.08)	(0.11)	(0.18)	(0.21)			
JK_{CBI}	0.51***	0.39**	0.15	0.03	0.52***	0.50***	0.29**	0.12	**+0.01**	**+0.11**	**+0.15**
	(0.16)	(0.19)	(0.23)	(0.24)	(0.15)	(0.19)	(0.25)	(0.24)			

Table F11: Differences in Responses of Nominal Interest Rates to Shocks from Event Window Choice

Notes: Each estimate comes from a separate OLS regression. For each regression, the dependent variable is the one-day change in end-of-day Treasury yields, represented by columns. The independent variable is the change in monetary policy shock series over a given event window around the time of scheduled FOMC announcements. Starting from the left, columns 2–5 represent the OLS regressions performed on shock series constructed within 30-minute windows. Columns 6–9 are equivalent, but for shocks constructed within the median optimal window length of 50 minutes. The sample period consists of all scheduled FOMC statement release dates from May 1999 through October 2019, resulting in 165 observations. Newey and West (1987) standard errors are reported in parentheses. Columns 10–12 represent the differences between the coefficients of the regressions using the optimal windows v. those using 30-minute windows. ** sig. at the 5% level. *** sig. at the 1% level.

	30-minute Window			Optimal Window			**Difference**		
	$\Delta TIPS_2$	$\Delta TIPS_5$	$\Delta TIPS_{10}$	$\Delta TIPS_2$	$\Delta TIPS_5$	$\Delta TIPS_{10}$	$\Delta TIPS_2$	$\Delta TIPS_5$	$\Delta TIPS_{10}$
GSS_T	-0.43	0.21	0.15	-0.40	0.21	0.14	**+0.03**	**-0.00**	**-0.01**
	(0.75)	(0.21)	(0.15)	(0.71)	(0.20)	(0.14)			
GSS_P	0.25**	0.21***	0.16***	0.28***	0.26***	0.21***	**+0.02**	**+0.06**	**+0.04**
	(0.12)	(0.06)	(0.05)	(0.13)	(0.08)	(0.07)			
NS_{MP}	0.56***	0.89***	0.69***	0.68***	1.07***	0.83***	**+0.12**	**+0.18**	**+0.13**
	(0.93)	(0.30)	(0.23)	(0.93)	(0.32)	(0.27)			
JK_{MP}	0.53**	0.60***	0.48***	0.54**	0.60***	0.49***	**+0.01**	**-0.00**	**+0.01**
	(0.52)	(0.24)	(0.20)	(0.38)	(0.25)	(0.22)			
JK_{CBI}	-0.39	0.07	0.03	-0.23	0.25	0.11	**+0.16**	**+0.17**	**+0.08**
	(0.56)	(0.25)	(0.23)	(0.71)	(0.29)	(0.25)			

Table F12: Differences in Responses of Real Interest Rates to Shocks from Event Window Choice

Notes: Each estimate comes from a separate OLS regression. For each regression, the dependent variable is the one-day change in end-of-day TIPS yields, represented by columns. The independent variable is the change in monetary policy shock series over a given event window around the time of scheduled FOMC announcements. Starting from the left, columns 2–4 represent the OLS regressions performed on shock series constructed within 30-minute windows. Columns 5–7 are equivalent, but for shocks constructed within the median optimal window length of 50 minutes. The sample period consists of all scheduled FOMC statement release dates from May 1999 to October 2019, resulting in 165 observations. Newey and West (1987) standard errors are reported in parentheses. Columns 8–10 represent the differences between the coefficients of the regressions using the optimal windows v. those using 30-minute windows. ** sig. at the 5% level. *** sig. at the 1% level.

	30-minute Window			Optimal Window			**Difference**		
	ΔBEI_2	ΔBEI_5	ΔBEI_{10}	ΔBEI_2	ΔBEI_5	ΔBEI_{10}	ΔBEI_2	ΔBEI_5	ΔBEI_{10}
GSS_T	0.71**	-0.06	−0.10*	0.67**	-0.07	−0.11**	**-0.04**	**-0.02**	**-0.01**
	(0.75)	(0.13)	(0.09)	(0.71)	(0.13)	(0.09)			
GSS_P	0.00	0.03	-0.01	-0.00	0.01	-0.02	**-0.00**	**-0.02**	**-0.01**
	(0.11)	(0.05)	(0.03)	(0.12)	(0.05)	(0.03)			
NS_{MP}	0.55	0.06	−0.13*	0.48	-0.03	−0.15**	**-0.07**	**-0.09**	**-0.03**
	(0.90)	(0.24)	(0.13)	(0.92)	(0.24)	(0.12)			
JK_{MP}	0.11	0.02	−0.09**	0.01	-0.04	−0.10***	**-0.10**	**-0.05**	**-0.01**
	(0.50)	(0.16)	(0.07)	(0.36)	(0.12)	(0.13)			
JK_{CBI}	0.78**	0.07	-0.00	0.73**	0.04	0.01	**-0.05**	**-0.03**	**+0.01**
	(0.55)	(0.17)	(0.11)	(0.67)	(0.18)	(0.13)			

Table F13: Differences in Responses of Break-even Inflation to Shocks from Event Window Choice

Notes: Each estimate comes from a separate OLS regression. For each regression, the dependent variable is the one-day change in end-of-day break-even inflation, represented by columns. The independent variable is the change in monetary policy shock series over a given event window around the time of scheduled FOMC announcements. Starting from the left, columns 2–4 represent the OLS regressions performed on shock series constructed within 30-minute windows. Columns 5–7 are equivalent, but for shocks constructed within the median optimal window length of 50 minutes. The sample period consists of all scheduled FOMC statement release dates from May 1999 to October 2019, resulting in 165 observations. Newey and West (1987) standard errors are reported in parentheses. Columns 8–10 represent the differences between the coefficients of the regressions using the optimal windows v. those using 30-minute windows. * sig. at the 10% level. ** sig. at the 5% level. *** sig. at the 1% level.

	30-minute Window			Optimal Window			**Percentage Difference**		
	DP_{SPX}	DP_{ESc1}	DP_{ESc2}	DP_{SPX}	DP_{ESc1}	DP_{ESc2}	DP_{SPX}	DP_{ESc1}	DP_{ESc2}
GSS_T	-3.46^{***}	-3.54^{***}	-2.68^{***}	-3.24^{***}	-3.25^{***}	-2.67^{**}	**-6.52%**	**-8.15%**	**-0.57%**
	(1.36)	(1.39)	(1.40)	(1.23)	(1.25)	(1.20)			
GSS_P	-1.39^{***}	-1.47^{***}	-1.39^{***}	-1.43^{***}	-1.45^{***}	-1.69^{***}	**+2.52%**	**-1.12%**	**+21.24%**
	(0.39)	(0.40)	(0.36)	(0.51)	(0.53)	(0.48)			
NS_{MP}	-7.57^{***}	-7.88^{***}	-6.96^{***}	-7.33^{***}	-7.41^{***}	-7.82^{***}	**-3.32%**	**-6.00%**	**+12.30%**
	(1.52)	(1.58)	(1.57)	(1.81)	(1.87)	(1.82)			
JK_{MP}	-8.16^{***}	-8.35^{***}	-7.53^{***}	-8.01^{***}	-8.15^{***}	-7.91^{***}	**-1.75%**	**-2.35%**	**+5.07%**
	(0.43)	(0.47)	(0.57)	(0.58)	(0.60)	(0.62)			
JK_{CBI}	8.27^{***}	8.20^{***}	7.68^{***}	8.04^{***}	8.25^{***}	7.00^{***}	**-2.84%**	**+0.66%**	**-8.97%**
	(1.25)	(1.29)	(1.24)	(1.44)	(1.46)	(1.47)			

Table F14: Differences in Responses of Stock Prices to Shocks from Event Window Choice

Notes: Each estimate comes from a separate OLS regression. For each regression, the dependent variable is the price log-difference of the S&P 500 Index or E-mini futures, represented by columns. The independent variable is the change in monetary policy shock series over a given event window around the time of scheduled FOMC announcements. Starting from the left, columns 2–4 represent the OLS regressions performed on shock series constructed within 30-minute windows. Columns 5–7 are equivalent, but for shocks constructed within the median optimal window length of 50 minutes. The sample period consists of all scheduled FOMC statement release dates from May 1999 to October 2019, resulting in 165 observations. Newey and West (1987) standard errors are reported in parentheses. Columns 8–10 represent the percentage differences between the coefficients of the regressions using the optimal windows v. those using 30-minute windows, where positive (negative) values represent a stronger (weaker) effect in the same direction. *** sig. at the 1% level.