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

Paul L. Tran*

This version: 12 November, 2025

[Click here for the latest version]

Abstract

This paper shows that financial markets need more time than the standard 30 minutes to fully react to monetary policy announcements. I systematically estimate these optimal window lengths using a neural network text analysis method. I find that the required time increases with asset underlying maturity (reaching 50–60 minutes for maturities at least two quarters ahead) and FOMC statement characteristics: complexity, novelty, and dissents. This timing difference has economic consequences: optimal windows correct for the attenuated impact of monetary policy shocks about forward guidance on interest rates, break-even inflation, and equity prices, and yield more precise macroeconomic responses.

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, seminar participants at UT Austin, and participants at the 2025 Texas Macro Job Candidate Conference. I am grateful to Chris Boehm and Niklas Kroner for sharing data with me. I also appreciate the Empirical Macroeconomics Policy Center of Texas for their financial support to cover some of the computational costs behind my paper. 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)

Following the August 1999 Federal Open Market Committee (FOMC) announcement, the S&P 500 Index did not instantaneously settle at a new value. Instead, as shown in Figure 1, its price swung in both positive and negative directions for over an hour. This example highlights a key question: how much time do financial markets need to fully react to monetary policy announcements? Applied macroeconomists often measure these price changes within narrow windows around news events such as monetary policy announcements. Small windows avoid contamination from unrelated news, but assume that the market's first response is its final response. On the other hand, wide windows allow more time for the market to react, but risk mixing the impact of monetary policy announcements with other information. Despite these trade-offs, the monetary policy literature has often defaulted to an "ad-hoc" 30-minute window. This "one-size-fits-all" approach is problematic because it assumes all markets take the same amount of time to react to monetary policy news, even though markets for longer-maturity assets might need more time to understand its more "abstract" information. If the chosen window is wrong, measurements of monetary policy surprises and shocks can become attenuated and less relevant.

I propose a systematic method that pins down these "optimal" event windows. First, I train a neural network for text analysis to approximate the underlying relationship between the text of monetary policy announcements and the market price reaction. This process creates a "text-based signal", which represents the network's predictions of the market price response based only on the FOMC statements, word-for-word. Second, I have the network learn this relationship and generate this text-based signal for different event window lengths

(e.g., a 10-minute window, 20-minute window, and so on). The "optimal" window is defined as the length where the neural network can make the most accurate predictions. This "Goldilocks" window is the time where the market has fully reacted to the monetary policy announcement, but not so much time that the price becomes overwhelmed by noise.[1]

My method yields four key findings. First, financial markets require more time to fully react to monetary policy announcements than the 30 minutes typically assumed. Specifically, I find that regardless of contract maturity, financial markets for futures and equities fully react within an event window starting 10 minutes before the announcement and *ending at least 30 minutes* after.

Second, the optimal window length is not "one-size-fits-all". It varies systematically with underlying asset maturity. Figure 2 summarises this result, showing the estimated optimal window length increases with the underlying maturity of the asset, rising from 40 minutes for short-horizon assets to 50–60 minutes for assets with underlying maturities at least two quarters ahead. Relatedly, I find that statements with greater complexity, less similarity to previous texts, and the presence of dissents also require longer event windows on average.

Third, the correlation between monetary policy surprises measured with my optimal windows versus the standard 30-minute window decreases as asset maturity increases. While interest-rate surprises for current and next FOMC meetings are similar, the correlation for long-maturity assets (e.g., fifteen years out) can decline by as much as 10%. These differences primarily alter the forward guidance component of monetary policy shocks, leading to larger estimated peaks and troughs or "shifts in importance" in the shock's composition.

Fourth, using the optimal window increases the estimated impact of monetary policy shocks about forward guidance on interest rates, break-even inflation, and equity prices. Furthermore, the estimated responses of macroeconomic variables also become more precise. This suggests that previously documented effects of monetary policy may be attenuated and imprecise due to the use of suboptimal window lengths.

[1] 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.

This paper is not the first to investigate the appropriate event window length. For example, Hillmer and Yu (1979) find that windows should last several hours, while Chang and Chen (1989) find they should span several days. Das and King (2021) find that asymmetric, narrower one-day event windows exhibit greater information content for earnings announcements. More recently, Casini and McCloskey (2025) argues that the window should be selected based on when the monetary policy surprise "dominates" other noise.[2] Using a heteroscedasticity-based statistical method, Boehm and Kroner (2025) select a fourteen-hour window. My paper contributes to this literature by introducing a new source of information. While the aforementioned studies infer the optimal window using only observed price movements, my method combines these dynamics with a "text-based signal" from a neural network that reads the FOMC statement itself. This allows my method to systematically estimate the window length where the "text-driven" price reaction is fully captured and noise is minimised on average. I also contribute to the high-frequency identification literature by detailing the effects of window choice on measured monetary policy shocks and surprises, particularly regarding forward guidance.

This paper joins a growing literature that uses text analysis to study the effects of central bank communication on expectations and the macroeconomy. Traditional methods in economics often relied on word counts or sentiment dictionaries to classify text (e.g., Gentzkow, B Kelly, and Taddy, 2019; Husted, Rogers, and Sun, 2020; Acosta, 2023; and others). A key limitation of these methods is that they can miss the nuances of context. In contrast, modern neural networks are designed to understand the full context and interdependencies of words. Recent papers have used these advanced tools to read FOMC statements and speeches, creating new and improved measures of monetary policy shocks (e.g., Handlan, 2022b; Doh, Song, and SK Yang, 2023; Piller, Schranz, and Schwaller, 2025; and others).

I contribute to this new literature from a different angle. While the studies mentioned above use text-analysis neural networks to directly create policy shocks, I use these methods

[2]The authors call this condition relative exogeneity. This is a technical condition which means the policy news is so large that it effectively drowns out all other market noise within that window.

to first address the event window timing problem. I apply the neural network to determine the optimal window length where the network can best understand the relationship between monetary policy communication and market price reactions. By first identifying the proper window, my method allows for the construction of more precise and relevant monetary policy surprises and shocks.

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

This section presents a motivating framework to show why the shortest or longest event windows are not always optimal. The framework illustrates how conflicting noise components can distort asset price dynamics following a news release. For simplicity, this framework imposes strong linearity and orthogonality assumptions on the noise components. These assumptions are purely illustrative, serving only to show how noise affects asset prices and, in turn, the optimal window length. My systematic estimation method does not impose these assumptions because the true composition of market price responses is unknown. Therefore,

this framework and its related simulations serve to motivate the natural language processing approach used to extract a signal from the text of the 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, meaning it is assumed that $P_t^f = P^f \in \mathbb{R}$.

Cognitive noise represents the collective "market noise" from participants' limited capacity to process information. Immediately after a news release, factors like algorithmic trading, over-reactions, under-reactions, or liquidity constraints mean market participants may have incomplete information or differing beliefs about the news. As time progresses, processing costs fall and investors learn from each other's price responses, leading to a convergence towards the fundamental price component. In other words, cognitive noise and its error "decay" over time, eventually converging 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. I ensure the cognitive noise process exhibits the decaying behaviour over time by assuming 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

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

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 $P_{i,t}^f$ for all N news:

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

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, $(\varepsilon_t^c + \varepsilon_t^n)$, and the signal noise are *independent*, I

[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 while holding the other parameters constant.

am able to derive the following expression for the minimisation problem of the MSE:

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

This derivation implies that the econometrician can still solve for the optimal time horizon, t^*, using only the noisy signal s_i. This result is asymptotic: given infinite news events, the signal's precision has no effect on finding t^*. However, in finite samples like the monetary policy literature,[5] the signal's precision *does* determine the feasibility of estimating t^*. An equivalent and important interpretation is that t^* is the time where the average impact of noise is minimised. At this point, the fundamental component P_i^f (and its signal s_i) has the largest "share" in the observed asset price.

2.3 Framework Takeaways

I simulate the asset price process over time for multiple news announcements to illustrate the effects of cognitive and unrelated noise on the optimal time horizon, and to demonstrate why a good signal is needed for this MSE minimisation. The main implications are discussed in the main text, while Appendix A provides full details on the process.

I consider three scenarios for the asset price response: one with cognitive noise, one with unrelated news, and one with both. The simulations show that in all scenarios, the optimal time horizon (t^*) estimated from the noisy signal s_i is essentially equal to the true horizon

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

estimated from the fundamental price P_i^f. A "good" signal, therefore, makes it possible to estimate the time horizon that reflects a full market reaction. However, recall that this framework and its simulations impose strong linearity and orthogonality assumptions purely to demonstrate the effects of noise. In reality, the true compositions and processes of the fundamental and noise components are unknown, yet still influence the overall asset price.

Abstracting from this simple framework, I argue that the text-analysis neural network approach used in this paper does not need to know the true noise processes to approximate the underlying mapping between monetary policy communications and asset price changes. Therefore, the method is still able to extract a "good" signal.

3 Data

The sample period for my analysis runs from May 1999 to October 2019. The start date coincides with the FOMC's adoption of the practice of issuing a statement after every scheduled meeting, regardless of whether policy rates were changed. The end date provides a pre-COVID-19 benchmark for the analysis, thereby avoiding the structural breaks and unprecedented market dynamics associated with the subsequent global pandemic.[6] As the principal entity determining U.S. monetary policy, the Federal Open Market Committee (FOMC) is closely watched by markets. Of particular importance are the released FOMC statements, which are the initial and primary source of monetary policy news for market participants.[7] Because markets are known to dissect and react to these statements word-by-word,[8] understanding the time it takes for them to fully react is critical for monetary policy event studies. However, the mapping from FOMC statements to asset price reactions is complex, nonparametric, and difficult to capture with traditional text analysis methods. Approximating this

[6]Future versions of this paper could extend the sample period by adapting the systematic estimation to incorporate FOMC statements released after October 2019.

[7]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.

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

underlying relationship motivates the text-analysis neural network approach in this paper. Before detailing the systematic estimation, I provide descriptive information on the FOMC statements (i.e., the inputs) and the financial market asset prices (i.e., the outputs) used in this paper, as summarised in Appendix Table F2.

3.1 Inputs: FOMC Statements

FOMC decisions are announced in a press release after its eight scheduled annual meetings. The FOMC occasionally holds unscheduled meetings, but I drop these statements from my sample. This exclusion is to ensure that measured asset price changes are driven by the statement's content, not by the surprise of an unscheduled meeting. I source all scheduled statements from the Board of Governors of the Federal Reserve System website, resulting in a sample size of 165 statements.[9]

Text Pre-processing To prepare the FOMC statements for the neural network, I pre-process all 165 statements by standardising them to plain UTF-8 text with uniform spacing. I also remove URLs, release timestamps, and the list of regional bank request approvals. Most importantly, I remove the FOMC member voting record from the end of each statement. This removal is prioritised for two main reasons. First, it ensures the neural network, which has a hard constraint on input length, focuses entirely on the main economic discussion (see Subsection 4.3). Second, while voting records are known to affect stock markets (C Madeira and J Madeira, 2019), they have less impact on interest-rate assets. Given my paper's focus on interest-rate futures, this exclusion should not significantly affect the estimation of the optimal event windows.[10]

Content and Evolution of Statements A typical FOMC statement discusses current macroeconomic conditions, communicates the Committee's economic expectations, and concludes with the new Federal Funds and discount rates. Following the Great Recession of 2008–2009, these statements began to include discussions of unconventional monetary policy,

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

[10] Nonetheless, future versions of this paper will explore including the voting records as inputs.

such as quantitative easing. This post-Crisis period saw an increase in statement complexity and length, particularly as the Committee relied more on communication to influence market expectations when it could no longer use rate changes. As seen in Figure 3, statement length grew rapidly until 2014. This trend partially reverses after 2014, but statements remain longer on average than 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 outputs for the neural network are interest-rate and equity futures prices. I use intraday (tick) data for all assets, sourced from the Thomson Reuters Tick History database (LSEG). This high-frequency data is necessary to analyse the short event windows (e.g., 30 minutes) common in the high-frequency identification literature. Within these windows, the variation in interest-rate futures prices measures the change in expected interest rate paths in response to FOMC announcements.

I select a range of futures contracts commonly used to construct high-frequency monetary policy surprises (Kuttner, 2001; Gürkaynak, Sack, and Swanson, 2005; Gertler and Karadi, 2015; Nakamura and Steinsson, 2018; Jarociński and Karadi, 2020; Handlan, 2022b; and others). For short-term expectations, I use Federal Funds futures, which are liquid for measuring expectations up to three months out ($FF1$–$FF4$). To capture expectations from four months to one year ahead, I use Eurodollar futures ($EDcm2, EDcm3, EDcm4$), following

the methodology of Acosta, Brennan, and Jacobson (2024). For longer horizons, I use Treasury futures with two ($TUc1, TUc2$), five ($FVc1, FVc2$), ten ($TYc1, TYc2$), and thirty-year ($USc1, USc2$) contracts, which correspond to underlying maturities of approximately two, four, seven, and fifteen years, respectively (Gürkaynak, Kisacikoğlu, and Wright, 2020). Finally, I include the S&P 500 Index and its E-mini futures ($SPX, ESc1, ESc2$). I use equities for two reasons: they are a central asset in the monetary policy literature, and their futures trade outside regular hours, providing the high-quality data necessary to test a wide variety of event window lengths.

Dependent Variable Construction I collect futures contract price levels at 10-minute intervals, from 10 minutes before to 18 hours after an FOMC statement release. The output of interest for the neural network is the log-price difference for each contract, constructed for event windows starting 10 minutes before and ending n minutes after the announcement:

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

The neural network approach considers price log-differences for event windows up to 70 minutes in length.[11]

3.3 Data to Investigate Impacts of Monetary Policy Shocks

To investigate the effects of event window choice on monetary policy shocks, I use several additional datasets. Daily data for nominal Treasury yields are from Gürkaynak, Sack, and Swanson (2005).[12] Daily data for Treasury-inflation-protected-security (TIPS) yields and break-even inflation come from Gürkaynak, Sack, and Wright (2010).[13] Monthly data for industrial production (IP) and the consumer price index (CPI) are from Federal Reserve

[11] Longer event windows are not considered due to computational and financial constraints, but remain an interest for future versions of this paper.

[12] Treasuries: https://www.federalreserve.gov/data/nominal-yield-curve.htm

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

Economic Data. The monthly excess bond premium (EBP) measure is from Gilchrist and Zakrajšek (2012) and sourced from the Federal Reserve Board.[14]

4 Systematically Estimating Optimal Event Windows

Systematically estimating the optimal event window length hinges on approximating the underlying, nonparametric relationship between monetary policy communications and financial market price reactions. Obtaining the text-based signal needed for this estimation requires a precise approximation of this mapping. This precision, in turn, demands a method that can quantify the *full* information content of the FOMC statement, capturing all dimensions of communication that drive market reactions.

Traditional text analysis methods, such as "fitting predictive models on simple counts of text features" (Gentzkow, B Kelly, and Taddy, 2019), typically fail to capture the full information content of text. These methods often miss context and word interdependencies. For example, a "bag-of-words" approach would treat the phrases "*employment* went up, but *inflation* did not" and "*inflation* went up, but *employment* did not" as identical. Similarly, "n-gram" methods, which measure neighbouring words, miss information spread across a sentence. A trigram might capture "economic growth slowed" but miss the crucial reversal in the phrase "economic growth slowed, but is expected to pick up pace later this year". Therefore, these popular methods cannot realistically capture the "full" information content of FOMC statements.[15]

In contrast, text-analysis neural networks *can* capture complex relationships like context and interdependencies. For example, the words "slowed" and "pick up pace" could both be associated with "economic growth" for prediction, even if they are not adjacent. The power of these networks is mathematically justified by the Universal Approximation Theorem (Hornik, Stinchcombe, and White, 1989; and others), which states that any continuous

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

[15] Papers like Babur and Cleophas (2017) have shown that model performance does not increase monotonically with the size of "n-gram" methods.

function (such as the underlying mapping from FOMC statements to price reactions) can be approximated by a network with at least one hidden layer to any desired accuracy.[16] Given their success in other social sciences, these methods offer similar benefits for studying monetary policy communication in economics (Gentzkow, B Kelly, and Taddy, 2019).

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

I employ the XLNet-Base language model from Z Yang et al. (2019), which builds upon the transformer architecture of Vaswani et al. (2017).[17,18] This model abandons the common "masking" technique, which assumes word independence and can create training discrepancies.[19] Instead, XLNet-Base uses a permutation-based pre-training phase, allowing it to learn context bi-directionally without the weaknesses of masking. This makes XLNet-Base better suited for semantic tasks than unidirectional generative models (e.g., ChatGPT).[20] The model's versatility makes it ideal for fine-tuning on new regression tasks, such as approximating the underlying relationship between FOMC statements and price reactions. I start with the pre-trained XLNet-Base model and fine-tune it to "transfer" its general language knowledge to this specific task, allowing me to obtain the text-based signal.[21]

[16] The Universal Approximation Theorem is an existence theorem; it guarantees *existence* of a network with at least one hidden layer, but does not specify its structure. In practice, deep networks (multiple layers) are often used to reduce parameters and computational requirements for complex approximations.

[17] While XLNet-Base is chosen for its suitability to approximating the underlying relationship, the specific choice of model architecture could potentially influence the quantitative results. However, the core methodology of my paper—identifying the optimal window as the point of maximum predictive accuracy—is flexible, and could be implemented using alternative approaches capable of approximating the complex mapping from text to price changes. Future research could explore the sensitivity of my findings to different model architectures or approximation techniques.

[18] The open-source, pre-trained neural network has 12 layers, 12 attention heads, 117 million parameters, and a 32,000 token vocabulary. See Hugging Face for details.

[19] E.g., if "target" and "rate" are masked in "...increase its target rate", their implicit relation is ignored.

[20] Generative models are "left-to-right" and predict a word using only preceding words.

[21] Transfer learning, as shown by Z Yang et al. (2019) and others, reduces data requirements for new tasks while maintaining high accuracy.

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

Recall from the motivating framework that the optimal event window is where noise is minimised on average, allowing the fundamental price component to have the largest "share" of the price change. The text-based signal is the neural network's *predicted* price change for a given window, using only the FOMC statement text. It follows that the optimal window length is where the network's approximation simultaneously has the highest predictive performance, because only within this window is noise minimised on average, and produces the most precise signal. This simultaneity drives my systematic estimation procedure:

1. For a given event window length, XLNet-Base regresses futures price changes on the FOMC statement text to assess its predictive performance and signal precision.

2. I repeat this regression for all event windows up to 70 minutes in length.

3. The optimal event window for that asset is the length that yields the network's best predictive performance.

4.3 Stratified-sampling Cross Validation

To prepare the neural network for systematic estimation, I split the sample five times, with 20 per cent of FOMC statements in each testing subsample. This splitting is stratified by the FOMC's rate decision, the FOMC Chair, pre-/post-2007, and statement word count to ensure equal distributions. This process yields five distinct training-testing folds such that no statement is shared across testing subsamples.[22]

These splits consist of input-output pairs: the FOMC statements (inputs) and the futures price log-differences (outputs). For each event window length and futures contract, I train XLNet-Base on a training subsample to approximate the underlying relationship. I then test its performance 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

[22]This process is known as stratified 5-fold cross-validation in the machine learning literature.

all five test splits. This method ensures a robust measure of out-of-sample performance and accounts for prediction variance across the splits.[23]

I restrict the context window of XLNet-Base (i.e., the number of word tokens it considers per input) to 512 tokens for each FOMC statement.[24] Given the average statement length is roughly 327 words (peaking at 800), the 512-token window provides adequate headroom to understand most statements. A list of the main hyperparameters for fine-tuning is available in Appendix Table F3.

4.4 Accuracy Metrics

The primary accuracy metric for neural network fine-tuning is a *generalised* R^2 statistic from Hawinkel, Waegeman, and Maere (2024), denoted R^2_{OOS}. This metric is chosen for two main reasons. First, the conventional R^2 (i.e., the proportion of variance explained by a model) breaks down for non-linear methods like neural networks. For such models, the total variance no longer neatly decomposes into model and residual components, meaning the squared Pearson correlation coefficient does not equal R^2. Second, the R^2_{OOS} formula is specifically designed to assess out-of-sample performance, not in-sample fit.[25] Mathematical details of this statistic are in Appendix B.

As is standard in machine learning, model performance is compared to a baseline. This baseline assumes *no* relationship between the FOMC statement text and futures price changes, naively predicting with the in-sample average. Therefore, the R^2_{OOS} measures the reduction in predictive error achieved by the neural network compared to this naive baseline.

During fine-tuning, XLNet-Base also tracks other metrics: the out-of-sample Pearson correlation between predicted and actual values, the out-of-sample mean absolute error, and the in-sample mean squared error. This last metric is important for verifying that the net-

[23]Stratified 5-fold cross-validation minimises differences between the full-sample and subsample distributions, which is crucial for transfer learning on finite samples.

[24]Tokens are converted text inputs used to reduce the network's vocabulary size (e.g., "decreasing" and "increasing" might become {"de", "in", "creas", "ing"}).

[25]Using a conventional R^2 yields similar optimal window lengths and network quality, but R^2_{OOS} is a more appropriate metric.

work is genuinely learning the underlying relationship.

The decisive criterion for systematic estimation is the out-of-sample R^2 *averaged across sample splits* for each event window length and futures contract:

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

where $K = 5$ is the number of sample splits. This $\overline{R^2_{OOS}}$ statistic measures the neural network's average performance improvement within a given window, relative to the naive baseline. The optimal event window is therefore the length that yields the largest $\overline{R^2_{OOS}}$. This is the point where the network's predictive power and generalisability is highest, which only occurs when the impact of noise is minimised and the fundamental reaction is fully captured.

4.5 Fine-tuning the Neural Network for Approximation

Individual parameters in a neural network are not interpretable like coefficients in a parametric model (Athey and Imbens, 2019). Therefore, this paper does not describe the causal effect of specific words within FOMC communications on asset price reactions. Instead, the network's training process approximates this underlying, complex relationship. I then use its predictions to pin down the event window size that best reflects the market's full reaction to monetary policy communications.

Fine-tuning a 110-million-parameter model like XLNet-Base makes overfitting an inevitable issue, which appears as a gap between R^2_{OOS} and in-sample R^2. I mitigate this difference by tuning two key hyperparameters: the number of training iterations and the learning rate. Too many iterations cause the network to over-fit the training data, deteriorating out-of-sample accuracy. Too few iterations will cause it to under-fit, failing to learn the underlying relationship. The learning rate governs how quickly the network updates its parameters. A rate that is too high causes the neural network weights to update too dra-

matically, forgetting prior information and degrading out-of-sample predictions. Conversely, a rate that is too low will also result in underfitting, as the network never fully learns the mapping between FOMC communications and asset price changes. Because I use transfer learning, the network's initial parameters are already optimised for general English. My goal is therefore to limit parameter updates within and across training iterations in order to preserve this generalisability while adapting the network to the specific task.

Ultimately, training the neural network is a balancing act: XLNet-Base has to meaningfully approximate the underlying relationship without overfitting, which would make the predictions non-generalisable and create a bad signal. I perform this balancing act using a two-step process. First, I set the maximum training iterations to 2040 steps (120 epochs). Second, I perform a "hyperparameter sweep" on the learning rate using Bayesian optimisation.[26] During this sweep, accuracy metrics from Subsection 4.4 are tracked. Training stops via early stopping: When out-of-sample accuracy begins to degrade. After a 24-hour sweep for each split, I select the learning rate and training iteration that yield the highest R^2_{OOS}. Given the potential for volatile out-of-sample predictions on finite samples, I select the model from this iteration only if subsequent iterations show a permanent decline in out-of-sample accuracy. I use the Simple Transformers library from Rajapakse, Yates, and Rijke (2024), which simplifies fine-tuning Transformer models like XLNet-Base.

4.6 Addressing Look-ahead Bias

A valid concern for text-analysis neural networks is look-ahead bias. As explained in Sarkar and Vafa (2024), this bias occurs when a network is pre-trained on future data to predict past values. The risk stems from large, unknown pre-training corpora, which may include data from the future.[27] I argue that this bias is mitigated in my method for two reasons: the known pre-training corpora of XLNet-Base and the specific language composition of FOMC

[26] The set of considered learning rates is $[1e-5, 9e-5]$, assuming a uniform distribution.

[27] Look-ahead bias is especially prevalent for closed-source large language models (e.g., ChatGPT) whose pre-training data are not publicly accessible.

statements.

First, the authors of XLNet-Base mitigate this concern by restricting the pre-training data to *only* BookCorpus and the entire English Wikipedia. This restriction makes it highly unlikely that the network learned any specific information about the intersection of FOMC communications and futures market price reactions during its pre-training. Importantly, Z Yang et al. (2019) showed that this limited pre-training data does not negatively impact the model's generalisability compared to other models.

Second, look-ahead bias is mitigated because the information content in FOMC statements are temporally anonymised. Similar to Glasserman and Lin (2023), the pre-processed text contains no references to relative times (e.g., t and $t+1$). Therefore, even if XLNet-Base's pre-training data included future FOMC statements, the network has no chronological information to exploit. It cannot "figure out" the temporal order of the statements from the text alone, preventing it from learning from future information.

5 How Long Until Markets Fully React to MP News?

In short, the systematic estimation is searching for the event window length where the neural network's predictive power is highest. This is the optimal window: the length where, as argued in the motivating framework, noise has a minimal average impact on the asset price change. This window length allows the network to best approximate the underlying relationship, yielding the most precise and generalisable text-based signal.

Figure 4 presents the results for selected federal funds, Eurodollar, and Treasury futures, as well as the S&P 500 Index. Appendix Figures E2–E8 show the results for all futures contracts. In each sub-figure, the horizontal axis shows the end time of the event window lengths (e.g., $t+20$ implies a 30-minute window ending 20 minutes after release), and the vertical axis is the R^2_{OOS} metric. The cross points show the $\overline{R^2_{OOS}}$ statistic for each window, with the solid yellow point marking the maximum. Box-and-whisker plots show the distri-

bution of R^2_{OOS} across the five sample splits for each window length.[28]

The results show that regardless of futures maturity and asset type, the event window length yielding the largest $\overline{R^2_{OOS}}$ is always at least 40 minutes, ending 30 minutes after FOMC statement release. Examining other distributional metrics from the box-and-whisker plots, such as the 25^{th} or 75^{th} percentiles, leads to the same conclusion. Financial markets appear to fully react within a 40-minute event window *at minimum*, implying the popular 30-minute window is too short.[29]

The optimal event window length increases with the underlying maturity of the futures, a trend summarised in Figure 2. For Federal Funds futures (measuring policy expectations from the current FOMC meeting), the average window is 40 minutes for front-month contracts, rising to 45 minutes for two- and three-month-ahead contracts (measuring policy expectations for the upcoming FOMC meeting). For Eurodollar futures (two to four quarters ahead), the window is 50 minutes. This 50-minute length also applies to 2- and 5-year Treasury futures, but increases to 60 minutes for 10- and 30-year futures. One possible explanation is that traders of longer-horizon assets are more exposed to "soft" information (e.g., forward guidance), which is harder to interpret (Indriawan, Jiao, and Tse, 2021). This may increase belief and information uncertainty and result in higher risk premia (Piazzesi and Swanson, 2008). Relatedly, Okada (2025) document that longer-duration Treasury bonds are more affected by the FOMC announcement premium. Observed liquidity differences may be a symptom of these optimal event window lengths. For example, Fleming and Piazzesi (2005) find abnormal trading volumes persist longer in longer-maturity Treasury markets after FOMC announcements.

For equities, the S&P 500 Index and its front-month E-mini futures market ($ESc1$) fully react within 50 minutes, while the market for second-month E-mini futures ($ESc2$) takes

[28]Standard errors are not calculated due to the small number of sample splits ($K = 5$). While K could be increased (e.g., via leave-one-out cross-validation), the computational costs are prohibitive for this version of the paper.

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

60 minutes. The neural network estimates that the average optimal window length is 53.3 minutes. These longer-than-30-minute windows may be because equities are less directly impacted by policy rates, leading to more investor disagreement and misinterpretations about interest rate decisions (Zhang, Kappou, and Urquhart, 2025). The difference between $ESc1$ and $ESc2$, despite both sharing the S&P 500 Index's infinite underlying maturity, may stem from their differing contract expirations. $ESc2$ has a more distant expiration and lower liquidity than $ESc1$. This could cause $ESc2$ traders to weigh "soft" information more heavily, requiring more time to fully process the announcement.

Overall, the $\overline{R^2_{OOS}}$ broadly exhibits a "hump" shape. The neural network's predictive performance increases with event window length, peaks at 30–50 minutes after statement release, then declines. This pattern is consistent with the motivating framework. As the window initially lengthens, predictive performance improves as cognitive noise decays and the market's full reaction is captured. After the peak, performance declines as unrelated news begins to dominate the observed price reaction, reducing the network's predictive accuracy. The optimal event window length is therefore the point of maximum generalisability, which I interpret as the length where the average impact from noise is minimised.

Importantly, the neural network achieves a positive $\overline{R^2_{OOS}}$ for all event window lengths and assets, confirming a systematic relationship between FOMC statement text and asset price changes. Relative to the baseline of naively predicting with the in-sample average, the network's generalisability within the optimal window improves by an additional 2–17 percentage points (7.4 p.p. on average) compared to its performance in a conventional 30-minute window.[30] This improvement in predictive performance occurs at the optimal window length precisely because it is the point where the average impact of noise is minimised, allowing the network to best approximate the underlying relationship.

[30] Appendix Table F4 summarises these results for each futures contract.

5.1 Different Event Windows, Different Market Responses

I now compare the market price responses measured within the estimated optimal windows to those from the conventional 30-minute window. Figure 5 depicts this comparison for selected futures and the S&P 500 Index. Appendix Figures E9–E15 show the comparison for all interest-rate and equity futures. In each sub-figure, the horizontal axis plots the price log-difference within the optimal window, while the vertical axis plots the 30-minute window response. The blue dots represent market reactions on scheduled FOMC meeting dates, and the 45-degree line is shown in grey and dashed.

To characterise whether markets under- or over-react to FOMC statements, I regress the price log-differences from the optimal windows on those from a 30-minute window. If the slope coefficient is statistically less than one for at most $\alpha = 0.05$, I define the market as under-reacting to the FOMC statement. Otherwise, the reactions are not considered statistically different. In each sub-figure, the regression line is coloured red for under-reaction and grey otherwise.

Financial markets for all interest-rate futures appear to under-react to FOMC statements, ex post. Furthermore, the slope of the regression decreases as the underlying maturity increases. This is consistent with the post-FOMC announcement drift described in Indriawan, Jiao, and Tse (2021) and Brooks, Katz, and Lustig (2023), who both point to information processing limitations as a source of bond market under-reactions. Indriawan, Jiao, and Tse (2021) argue that interest-rate markets are more exposed to "soft" information (e.g., forward guidance), which is costlier to process. Brooks, Katz, and Lustig (2023) find that slow adjustments by mutual fund investors also play a role. The 16 December, 2008 FOMC meeting provides a notable example. In response to the Great Recession, the FOMC cut the federal funds rate to a 0–0.25% range. This decision was a surprise, as markets expected a smaller reduction to 0.25–0.5%. This surprise, combined with the statement's commitment to "employ all available tools to promote the resumption of sustainable economic growth and to preserve price stability", likely contributed to markets needing more time to fully react.

The top sub-figure of Figure 6 illustrates this: for all Eurodollar futures, the price change becomes increasingly positive as the event window expands beyond the initial 30-minute reaction.

The S&P 500 Index also appears to under-react, ex post, corroborating findings from Neuhierl and Weber (2024) and Golez, P Kelly, and Matthies (2025), who suggest market participants struggle to react to the full information in FOMC announcements. However, the S&P 500 Index also exhibited the most price reversals, often moving from negative to positive as the event window expanded. A key example is the 16 September, 2008 meeting, shown in the bottom sub-figure of Figure 6. The market initially responded negatively, but this trend reversed within a 40-minute window and continued to increase until market close. One explanation is that market beliefs reversed as participants processed the FOMC's decision to maintain interest rates. With more time, cognitive noise may have died out, allowing a negative "knee-jerk" reaction to reverse as the market fully processed the positive information. This aligns with Bordalo et al. (2024), who document that stock market overreactions are common and tend to reverse as beliefs correct with time.

Overall, these findings support the premise of my paper: the optimal event window length is not a parameter to be simply assumed. Instead, I document how the window length that best reflects the market's full reaction to FOMC statements varies systematically across asset types and underlying maturities. Conventional, shorter windows fail to capture the market's full processing of complex information, revealing systematic under-reactions in interest-rate markets and significant price reversals in equities that are otherwise missed. Furthermore, as explored in Section 7, specific characteristics of the FOMC statement text itself systematically correlate with these estimated optimal window lengths.

6 What Happens to Monetary Surprises and Shocks?

This paper has shown that the optimal event window is consistently longer than the standard 30 minutes. This section investigates how monetary policy surprises and shocks are affected by this change. To properly construct these shocks, I must use a common event window length for all assets. Standard macro-finance models that extract shocks from the yield curve rely on "invertibility", or "spanning" (Duffee, 2013), a condition that is well-defined for a temporal cross-section of yields but not for yields measured at different, asset-specific horizons. Therefore, I construct and test monetary policy surprises separately for each of the optimal window lengths found in Section 5: 40-, 50-, and 60-minute windows. The data and FOMC meeting dates are the same as in Section 3.

6.1 Effects of Window Choice on Monetary Policy Surprises

I construct interest-rate surprises for each FOMC meeting date using the standard 30-minute window and the 40-, 50-, and 60-minute optimal windows. Following Kuttner (2001), Gürkaynak, Sack, and Swanson (2005), Nakamura and Steinsson (2018), and others, I use federal funds futures for expectations up to three months out, Eurodollar futures for four months to one year ahead, and Treasury futures for expectations out to fifteen years. Appendix C details the conversion from futures price changes to interest-rate surprises.

Figure 7 plots the correlation between monetary policy surprises from the optimal windows versus the standard 30-minute window. The horizontal axis shows the underlying maturity of the surprises (e.g., "0–1M" for the FOMC meeting happening in the current or next month). The individual lines show the correlations for the 40-, 50-, and 60-minute optimal windows, while the blue-solid line shows their average.

The correlations consistently decline as the event window lengthens. This decline is more pronounced at longer underlying maturities. For example, while surprises for near-term meetings are highly correlated (above 95%) regardless of window choice, the correlation for

15-year-ahead surprises (measured using 50-minute windows) drops by 10%. This demonstrates that as expectations about monetary policy move further into the future, the resulting interest-rate surprises become more sensitive to the choice of event window.

6.2 Effects of Window Choice on Monetary Policy Shock Impacts

I now use the interest-rate surprises as financial instruments to identify exogenous variation in monetary policy, following several popular methods. I include Treasury futures in my instrument set for two reasons. First, recent studies (e.g., Brennan et al., 2024; An, Stedman, and Lusompa, 2025) show longer-maturity instruments help prevent understating policy stimulus during the effective lower bound period.[31] Second, my Figure 7 results show that Treasury futures are the most sensitive to window choice.[32]

For conciseness, I use the *median* of the optimal window lengths found by the neural network: 50 minutes (10 minutes before and 40 minutes after statement release).

Shock Construction I construct monetary policy shocks following the methods of Gürkaynak, Sack, and Swanson (2005) (GSS), Nakamura and Steinsson (2018) (NS), and Jarociński and Karadi (2020) (JK).[33] All three methods use principal component analysis (PCA) to reduce the high-dimensional set of interest-rate surprises into one or two factors that capture their common variation.[34,35]

GSS identify two factors. They use the first two principal components and apply a rotation such that the first factor, "target" (GSS_T), drives changes in the current federal funds rate, while the second factor, "path" (GSS_P), does not. GSS_T is re-scaled one-for-one with the current-meeting surprise, while GSS_P is scaled to have an equal impact on the four-quarter Eurodollar surprise.

[31] For my sample, the effective lower bound is 16 December, 2008, through 16 December, 2015.

[32] All exercises are robust to using the authors' original instrument sets, as shown in Appendix D.

[33] See Appendix Table F5 for a summary of the shocks.

[34] PCA projects the data onto new dimensions based on its variance. The first principal component captures the largest common variation, the second captures the next largest, and so on.

[35] 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

NS use only the first principal component, which Bauer and Swanson (2023) explain is essentially a weighted average of the GSS target and path factors. NS re-scale this component to be one-for-one with the daily change in the one-year nominal Treasury yield, which I denote NS_{MP}.

JK also use the first principal component, re-scaling it to match the standard deviation of the four-quarter Eurodollar surprise. To disentangle policy from central bank information effects, they impose sign restrictions based on the shock's co-movement with stock market changes.[36] A shock with a negative (positive) stock price co-movement is identified as a monetary policy (JK_{MP}) (central bank information, JK_{CBI}) shock.

For easy interpretation and comparison, I re-scale all shock series to be one-for-one with the daily change in the zero-coupon, nominal one-year Treasury yield.[37] Appendix Table F6 provides summary statistics for each shock series. Because these methods (GSS, NS, and JK) jointly account for the key characteristics of monetary policy—such as its multi-dimensional nature, its impact beyond the immediate horizon, and information effects—any observed differences in the resulting shocks can be attributed directly to the choice of event window.

Visual Shock Differences I begin by plotting the shock series from both event windows in Figure 8. The horizontal axis shows the FOMC meeting date, and the vertical axis is in percentage points (re-scaled to the one-year Treasury yield). The black-solid line represents the 30-minute window shocks, and the red-dotted line represents the optimal window shocks.

Regardless of construction method, the optimal window shocks (red-dotted line) frequently show larger peaks and troughs in magnitude than the 30-minute window shocks (black-solid line).[38] For example, at the critical December 2008 and March 2009 FOMC meetings, the optimal window shocks are 3–6 basis points larger in magnitude, suggesting that the 30-minute window fails to capture the full impact of these statements.

[36] Stock market changes are measured within the same event window as the surprises.

[37] Robustness checks in Appendix D show this scaling choice does not affect the main findings.

[38] Plots for GSS_T are omitted as the shocks are nearly identical, given that markets react almost instantly to the current target rate change.

Another notable episode is the August 2011 FOMC meeting, where the statement first included explicit, date-based forward guidance (Crump, Eusepi, and Moench, 2013). On this date, the optimal window is associated with a more negative shock for both GSS_P and NS_{MP}. When applying the JK sign restrictions, the negative central bank information shock (JK_{CBI}) increased in magnitude by 3.6 basis points, while the policy shock (JK_{MP}) was unchanged. Because the federal funds rate was already at the zero lower bound, this new language on the future path of rates could be interpreted as an information shock about weaker-than-anticipated economic conditions. The optimal window may be capturing the full extent of this negative information shock, which took longer than 30 minutes for the market to fully process.

The event window choice can also cause "shifts in importance" between shock components. A clear example is the September 2008 FOMC meeting. Using a 30-minute window, the JK_{MP} (policy shock) is the main driver, while the JK_{CBI} (information shock) is near zero. When using a 50-minute window, however, this interpretation flips: the information shock becomes the dominant factor.

This shift is consistent with the statement's content. The FOMC held rates steady but used forward guidance to signal moderate expected growth and efforts to promote liquidity. The market, likely needing more time to process this positive signal, updated its economic prospects upwards, consistent with a positive JK_{CBI} shock and the observed rise in the S&P 500.

Overall, when comparing the JK shocks, the central bank information shock (JK_{CBI}) more frequently increases in magnitude with the optimal window than the policy shock (JK_{MP}). While JK_{MP} also sees increases, the differences due to window choice are relatively smaller. A likely explanation for the larger JK_{CBI} differences is mechanical, stemming from the sign restriction method combined with the larger S&P 500 price changes observed in the optimal window.

Responses of Interest Rates I now examine how the event window choice affects the

impact of monetary policy shocks on nominal and real interest rates. I use the daily change in Treasury yields and TIPS yields as the dependent variables, calculated using end-of-day yields for the day of and the day before the FOMC announcement. Because the shocks (constructed with 30-minute or optimal windows) are measured *within* this daily window, the timing ensures that the yield changes do not 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 $y^{i,j}$ is the daily change in Treasury or TIPS yields (i) at a specific horizon (j). The regressor, $(Shock)^{k,l}$, is the monetary policy shock (k) constructed using either the optimal or 30-minute window (l).

Regression results for the responses of nominal and real interest rates are presented in Table 1 and Table 2, respectively. Each estimate comes from a separate OLS regression specified in Equation 8. The tables compare the coefficients (β_1) from shocks constructed using the 30-minute window against those constructed using the median optimal window (50 minutes). The bolded columns (10–12 in Table 1, 8–10 in Table 2) show the difference between these coefficients, isolating the impact of the event window choice.

When opening up the event window from 30 minutes to the optimal length, monetary policy shocks that capture forward guidance (GSS_P, NS_{MP}, JK_{MP}) have larger effects on both nominal and real interest rates. For example, the impact of JK_{MP} on 2-, 5-, and 10-year real yields is 23–26 basis points larger, and the estimates retain their 1% statistical significance.

Conversely, the "target" factor (GSS_T) has a *smaller* impact on nominal rates at all horizons when using the optimal window, though its impact on real yields is unclear. Only the 2-year nominal yield estimates are statistically significant for GSS_T under both windows. A likely explanation for this pattern is that the sample includes the effective lower bound period, where monetary policy stimulus was driven by forward guidance, not by changes to the target rate.

The main takeaway from these results is that using too short of an event window prevents markets from fully processing the "soft" information in FOMC announcements, particularly regarding the future path of interest rates. As a result, the measured impact of forward guidance shocks on financial market variables is attenuated.

These findings support the interpretation suggested by Gürkaynak, Sack, and Swanson (2005): while markets can process simple target rate changes almost instantly, the more complex information about the economic outlook and future policy path requires more time to be fully priced in by market participants.

Responses of Break-even Inflation I next investigate the impact of monetary policy shocks on break-even inflation, defined as the difference between nominal Treasury and TIPS yields. The regressions follow the same specification as Equation 8, with the daily change in break-even inflation at various maturities as the dependent variable. Table 3 presents the results, comparing the impact of shocks constructed within the 30-minute window (columns 2–4) to those from the optimal window (columns 5–7). The final columns (8–10) show the difference between these corresponding coefficients, isolating the effect of the event window choice.

Similar to Nakamura and Steinsson (2018), most estimates are not statistically significant. However, the results provide two key insights. First, using the 50-minute optimal window increases the statistical significance of some estimates. For example, the decline in break-even inflation at the furthest horizon from NS_{MP} becomes statistically significant at the 5% level, as do some effects from GSS_T. Second, and more importantly, the results mirror those for interest rates: expanding the event window to the optimal length consistently makes the estimated impact of forward guidance shocks on break-even inflation *stronger* in magnitude and more *precise*.

Responses of Equities I finish my analysis by examining the impact of shock timing on equity prices. Table 4 presents analogous results for the S&P 500 Index and its E-mini futures, following the same regression specification as Equation 8, but with 100 times the

price log-difference of equities as the dependent variable. The table compares the coefficients from shocks constructed using the 30-minute window (columns 2–4) against those from the optimal window (columns 5–7). The final columns (8–10) show the percentage change between these corresponding estimates.

Expanding the event window from 30 minutes to the optimal length causes monetary policy shocks containing forward guidance to impact stock prices more negatively. Using the optimal window, the negative impact of JK_{MP} on the S&P 500 Index is 18.25% larger, while the impacts of GSS_P and NS_{MP} are 11.51% and 1.23% larger, respectively. These larger negative impacts are even more pronounced for the second-month S&P 500 E-mini futures contract. Notably, the magnitude of these differences for equities varies more widely across shock types than it did for interest rates.

Consistent with the findings for interest rates, the equity response to the GSS "target" factor (GSS_T) also becomes weaker (by 11.99%) in the optimal window. While the estimates remain statistically significant at the 1% level, this weakening suggests the longer window causes a "shift in importance" toward the "path" factor. This is intuitive: the optimal window captures the market's initial, fast reaction to the target rate, but then allows more time for the market to fully process the more complex forward guidance, which dominates the overall price change.

The effects on the central bank information shock (JK_{CBI}) are more complex. One explanation for the differences relates to the high idiosyncratic volatility of stock prices (relative to yields) combined with the JK sign restriction method. As shown in Subsection 5.1, equities exhibit both under-reaction and frequent, sharp price reversals. Expanding the window to 50 minutes captures more of these reversals. This volatility can lead to different shock decompositions under the sign restriction, amplifying the JK_{CBI} series in both directions and potentially weakening its net, directional impact on prices.

The large equity impact of the Jarociński and Karadi (2020) shocks (JK_{MP} and JK_{CBI}) should be interpreted with caution. The sign restriction method used to construct them is

based on their co-movement with equity prices, which mechanically ensures a large impact on the S&P 500. However, the finding is not purely mechanical. Other forward guidance shocks, such as GSS_P, do not use this restriction but still show similarly large negative impacts on equity prices, corroborating the overall result.

Responses of the Macroeconomy via Local Projections To study the macroeconomic impact of the shocks, I estimate impulse responses using local projections. Following Gertler and Karadi (2015) and Bauer and Swanson (2023), the model includes the log of CPI, the log of IP, the nominal two-year Treasury yield, and the excess bond premium (EBP). The EBP, measured by Gilchrist and Zakrajšek (2012), represents the financial sector's risk-bearing capacity and serves as the key transmission channel from financial markets to the real economy.

All variables are converted to a monthly frequency. Shock series are set to zero in non-FOMC months. Appendix Tables F7 and F8 provide summary statistics for the monthly data.

I estimate the impulse responses using the lag-augmented local projection method of Olea and Plagborg-Møller (2021). This approach is chosen because adding an additional lag to the specification as a control allows for correct inference using only Eicker-Huber-White heteroscedasticity-robust standard errors, avoiding the need to correct for serial correlation.

I run a separate regression for each outcome variable on each shock for each event window length (l) at each horizon (h):

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

where $y_{t+h}^{i,l}$ is the outcome variable (i) at horizon h, $(Shock)_t^{k,l}$ is the monetary policy shock (k) from window (l), and *controls* include the requisite lags. The impulse responses are the θ coefficients, plotted over horizon h.

For interpretation, all impulse responses are scaled to show the effect of a 100 basis point

contractionary policy shock, based on the common re-scaling to the one-year nominal Treasury yield.

Figure 9 compares the impulse responses of CPI to the monetary policy shocks constructed using both event windows.[39] At first glance, the point estimates of the impulse responses are qualitatively similar regardless of window choice. While some differences exist (e.g., the decline in CPI from NS_{MP} is more severe at farther horizons in the optimal window), the window choice does not dramatically alter the response path.

The most noticeable difference, however, is in the confidence intervals. The impulse responses derived from the optimal window shocks are consistently more precise than those from the 30-minute window. To formally measure this, I calculate the width of the confidence intervals for each time horizon, variable, and shock, for both the optimal and 30-minute windows.

I then compute the ratio of the optimal window's confidence interval width to the 30-minute window's width. A ratio less than one implies a more precise estimate from the optimal window. The average and median of this ratio across all impulse responses are 0.9173 and 0.9410, respectively.[40] This confirms that constructing monetary policy shocks within the optimal window results in *more precise* macroeconomic impulse responses.

Overall, the varied responses of interest rates, equities, and the macroeconomy support the findings of Brennan et al. (2024) that long-term rates and shock construction methods are important contributors. This paper provides strong evidence that the event window length is another critical dimension that must be considered.

My findings show that the standard 30-minute window is insufficient. It attenuates the measured impact of forward guidance shocks on financial variables because markets have not fully reacted to the policy announcements. In contrast, constructing shocks within the optimal window not only reveals these larger impacts but also results in more precise estimates of the macroeconomic responses. This optimal window is the length of maximum generalis-

[39] Results for the other macroeconomic variables are robust and presented in Appendix Figures E16–E18.
[40] These confidence interval width ratios are robust to the choice of monetary policy surprise instruments.

ability, which allows for these benefits because the average impact of noise is minimised.

7 Statement Features and Their Effects on Windows

Recall that the systematic estimation is an exercise where the neural network regresses asset price changes on FOMC statement text across various windows. This process can be viewed as "jointly" estimating the optimal event window length and the text-based signal that yield the greatest precision. Unfortunately, the financial and hardware constraints of this joint estimation currently limit the analysis to windows ending 60 minutes after the statement release.

An alternative method that potentially circumvents these constraints and provides initial insights is the "one signal" approach. This approach follows the motivating framework more closely by first assuming that the predicted price change from the optimal window (i.e., the most precise signal) is *constant* for all time. This allows for the estimation of event windows longer than one hour after the statement release.

However, the core assumption of this "one signal" approach is strong. Because the neural network's weights are sensitive to both input and output variables, it is possible that the "jointly" estimated optimal window differs from the one found using this "one signal" approach.[41]

The "one signal" approach uses the text-based signal from the "jointly" estimated optimal window to calculate $\overline{R^2_{OOS}}$ for event windows beyond the 60-minute constraint. This shortcut provides two advantages. First, it allows me to observe how different dimensions of monetary policy communication potentially affect the optimal event window length. Second, it allows me to check if a longer event window yields a larger $\overline{R^2_{OOS}}$, a possibility I discuss in Appendix D.

[41] A rough analogy is how iterative generalised method of moments (GMM) does not always converge.

7.1 How FOMC Statement Characteristics Affect Window Lengths

I next investigate whether specific features of an FOMC statement affect its optimal event window length. Using the "one signal" approach, I examine how statement complexity, similarity to previous statements, and the presence of dissents correlate with the window length. This heterogeneity analysis relies on two strong assumptions: first, that the text-based signal from the "jointly" estimated optimal window is constant; and second, that this signal remains relevant when conditioning the statements into subsamples.[42]

Text Complexity First, I examine how the optimal event window length for various futures maturities changes when conditioned on statement complexity. I assess this *semantic* complexity using the Flesch-Kincaid Grade Level index. This standard readability formula determines the U.S. grade level required to comprehend a text based on sentence length, word difficulty, and syllable count.[43]

I first split the FOMC statements into two subsamples based on their Flesch-Kincaid Grade Level relative to the median (16.5). I then calculate the MSEs under the "one signal" approach for each subsample across all futures maturities.

Table 5 (columns 1-2) summarises this exercise, showing the minimised MSEs and associated event window lengths, averaged across all assets. The results show a clear difference: statements with a "complex" readability level (≥ 16.5) have a longer average optimal event window of 71 minutes, compared to 60 minutes for "simpler" statements (< 16.5).

Interestingly, both subsamples have similar average MSEs. This finding suggests that the price response to "complex" statements is relatively more affected by cognitive noise. This is consistent with the idea that complex information generates more disagreement and market volatility (Smales and Apergis, 2017), ultimately requiring more time for financial markets

[42] Ideally, one would systematically re-estimate the optimal window for large subsamples of each characteristic, but this is computationally infeasible for the current paper.

[43] Appendix Figure E19 displays the evolution of the Flesch-Kincaid Grade Level for my sample. Descriptive statistics are in Appendix Table F9.

to fully process the announcement.[44]

Text Similarity FOMC statements, while relatively concise, have language and semantics that vary over time (Acosta and Meade, 2015; Handlan, 2022a). Using the "one signal" approach, I therefore analyse how the optimal event window length changes with the *similarity* between consecutive FOMC statements. To measure this, I use a standard bag-of-words model, which represents each statement as a vector of weighted word frequencies. This method assumes that documents with similar word frequencies are discussing similar topics, allowing me to calculate a similarity score between any two statements.

The specific weighted frequency I use is Term Frequency-Inverse Document Frequency (TFIDF), implemented via the Python `sklearn` library. TFIDF computes a weight for each word by balancing its frequency within a single document (Term Frequency) against its frequency across all documents (Inverse Document Frequency).

This weighting serves a crucial purpose: common words that provide little unique meaning (e.g., "a", "the") appear in all documents, so their high document-wide frequency significantly reduces their final weight.

Similarly, standard policy words (e.g., "unemployment", "inflation") also have their weights reduced, though less so, as their relative frequency can still differentiate statements. Conversely, unique or rare words (e.g., "persists", "tight") that appear in few documents receive a higher weight, as they are more informative. These differential weights are what allow the vector analysis to effectively measure the degree of similarity between any two 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. Let a single document and term be indexed by $d \in D$ and

[44]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.

$t \in T$, respectively. The term frequency is:

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

where $tc_{t,d}$ is the count of term t in document d, nt_d is the total number of words in d, and the log and addition terms smooth the frequency measure. The inverse document frequency is:

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

where nd is the total number of documents and $df_{t,d}$ is the number of documents in which term t appears. Multiplying the two terms gives the final weighted frequency:

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

Calculating this value for all terms and documents yields a $D \times T$ matrix, where each element $TFIDF_{d,t}$ is the final weighted frequency. Essentially, the higher the TFIDF value for a given term, the more informative that term is for distinguishing the content of a specific FOMC statement from all others.

The TFIDF matrix from the *current* pre-processed text would be inaccurate. This simple model is not sophisticated enough to handle context, meaning it would incorrectly treat "Federal Funds Rate" and "federal funds rate" as different terms. I mitigate this with further pre-processing. First, I convert all text to lowercase. Second, I remove common "stop words" (articles, pronouns, conjunctions) that convey little semantic meaning.[45] Third, I convert all words to their "base" or "stem" form (e.g., "increases", "increasing", and "increase" all become "increas").

Calculating the TFIDF matrix on this fully cleaned sample (May 1999–October 2019)

[45] I retain numbers, as they are crucial for conveying semantic content like interest rates or forward guidance.

yields a final matrix of 165 rows (statements) and 966 columns (terms).[46]

With the TFIDF matrix, I can now measure the similarity between any two FOMC statements. I create a document similarity matrix by multiplying the TFIDF matrix by its transpose:

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

Because the TFIDF vectors are normalised, this multiplication is equivalent to calculating the dot product between every pair of document row vectors. This allows me to define the similarity between any two statements (A and B) as their *cosine similarity*:

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

where $\mathbf{A}$ and $\mathbf{B}$ are the TFIDF vectors for statements A and B. A cosine similarity of one means the statements are identical, while a value of zero means they have no common informative terms. The more informative base terms two statements share, the closer their cosine similarity will be to one.

The Document Similarity Matrix allows me to compare any two statements. I focus on the similarity between sequential statements, $(d, d-1)$, which I denote S^1 and plot in Figure 10.[47] Figure 10 shows that the cosine similarity between sequential FOMC statements increases over time. This trend of increased standardisation in FOMC language begins around 2009–2010, coinciding with the federal funds rate approaching the zero lower bound. As discussed in Handlan (2022a), one explanation is that the FOMC sought to reduce market surprises by standardising its statement vocabulary during and after the financial crisis. This trend of high similarity continued through 2019, even after the zero lower bound ended, reflecting the FOMC's broader efforts to increase policy transparency under Chairs Yellen and Powell.

I split the FOMC statements into two subsamples based on their S^1 value relative to the

[46] Appendix Table F10 lists the 30 base terms with the highest TFIDF scores, which are the most informative for differentiating statements.

[47] Appendix Figure E20 presents the full Document Similarity Matrix as a heat map. Descriptive statistics for S^1 are in Appendix Table F9.

median (0.826): "different" (below median) and "similar" (above median). I then calculate the MSEs for each subsample using the "one signal" approach. Table 5 (columns 3-4) summarises the results, averaged across all assets. The results show that "different" statements have a longer average optimal event window of 62 minutes, while "similar" statements have a shorter average window of 51 minutes. As with text complexity, both subsamples have similar minimised MSEs.

This finding is intuitive: if an FOMC statement is "different" from its predecessor, it contains more new information that requires more time for markets to discern and process. Conversely, "similar" statements contain less new information, allowing markets to react more quickly as most of the content has already been priced in.

Presence of Dissents The final characteristic I consider is the presence of dissents, which are recorded in roughly 40% of the statements in my sample. While not rare, dissents are significant, as they traditionally signal that a Board member finds the majority's opinion unacceptable (C Madeira and J Madeira, 2019). The very existence of a dissent thus provides additional, complex information for markets to process.

To investigate this, I split the sample into statements with unanimous votes and those with at least one dissent.[48] Table 5 (columns 5-6) presents the results of the "one signal" approach for these subsamples, averaged across all assets.

The results show a large difference: statements with dissents have an average optimal window of 83 minutes, far longer than the 53-minute window for unanimous statements. This suggests the additional information and internal disagreement signalled by a dissent requires significantly more time for markets to fully process. While prior work has shown dissents have a nuanced impact (e.g., C Madeira and J Madeira, 2019; Tsang and Z Yang, 2024 find stronger effects on equities than on interest rates), my finding shows that, when averaged across all asset types, the presence of a dissent is associated with a longer market reaction time.

[48]This analysis does not distinguish between the *reasons* for the dissent (e.g., for tighter or easier policy).

Overall, this section's findings show that monetary policy communication affects markets along multiple dimensions, not just the "hard" information of the federal funds rate decision. The "soft" information characteristics of the statement—such as its semantic complexity, its similarity to prior announcements, and the presence of dissents—systematically influence the time financial markets need to fully process the communication, particularly its forward guidance.

Indeed, future research into how these textual dimensions affect market processing times could provide deeper insight into the transmission of monetary policy.

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 paper challenges this convention by asking a simple question: how does one systematically choose the appropriate window? I answer this by developing a new methodology that combines observed price dynamics with a text-based signal derived from a neural network. This method allows me to systematically estimate the event window length that best reflects the market's full reaction to the information content of FOMC statements.

My systematic estimation yields two key findings. First, the common 30-minute window is insufficient. My results show that, on average, markets fully react within an event window *ending at least 30 minutes after release* (a 40-minute total window), and a 30-minute window is never found to be optimal for any asset. Second, the optimal window length is not "one-size-fits-all". Instead, the optimal duration increases systematically with asset underlying maturity, rising from 40 minutes for short-term assets to 50–60 minutes for assets with maturities of two quarters or more. Relatedly, I find that statements with greater complexity, less similarity, and the presence of dissents are all associated with longer event windows on average.

The implications of these findings are not merely methodological. The choice of event window has a tangible impact on monetary policy surprises and shocks. I find that the correlation between surprises measured with optimal versus 30-minute windows decreases with asset underlying maturity. These discrepancies are economically meaningful, altering the forward guidance component of policy shocks. Consequently, the measured impact of these forward guidance shocks on interest rates, break-even inflation, and equity prices becomes larger, suggesting that conventional, shorter windows attenuate the true effects of policy.

The benefits of the optimal window extend to the broader macroeconomy. I document that when monetary policy shocks are constructed using the optimal event window, the estimated impulse responses of macroeconomic variables 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. This is achieved because the optimal window is the length of maximum generalisability and minimised average noise impact, leading to a more precise understanding of monetary policy's effects.

As the literature expands to study broader forms of central bank communication (e.g., Neuhierl and Weber, 2019; Swanson and Jayawickrema, 2023; Bauer and Swanson, 2023; and others), this paper's methodology can be applied to measure the full reaction time to any text-based news event. Doing so offers a path toward constructing 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*.

Babur, Önder and Loek Cleophas (2017). "Using n-grams for the Automated Clustering of Structural Models". In: *SOFSEM 2017: Theory and Practice of Computer Science*. Ed. by Bernhard Steffen, Christel Baier, Mark van den Brand, Johann Eder, Mike Hinchey, and Tiziana Margaria. Cham: Springer International Publishing, pp. 510–524.

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.

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

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.

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.

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.

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.

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 Announcenments: 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 (2025). "Post-FOMC Drift in the Equity Options Market". SSRN Working Paper No 5464630.

Figures and Tables

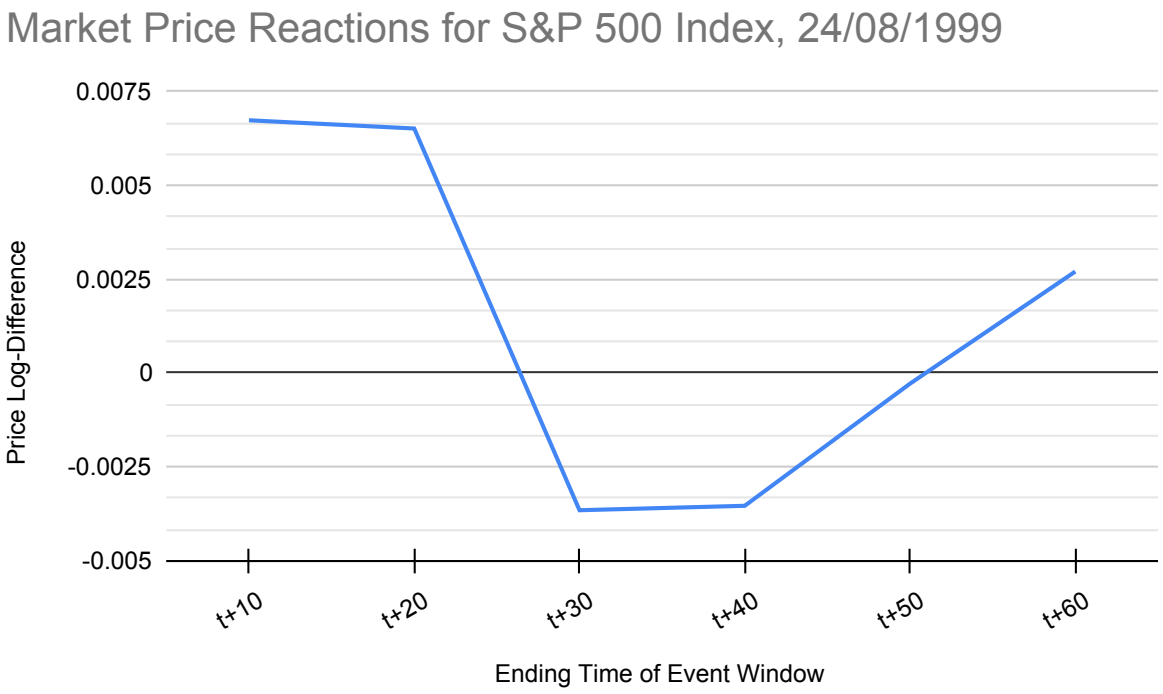


Figure 1: Responses of S&P 500 Index to FOMC Statements

Notes: The price log-differences of the S&P 500 Index for different event window lengths after the release of the 24 August, 1999 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. Every event window starts at 10 minutes before the FOMC statement release.

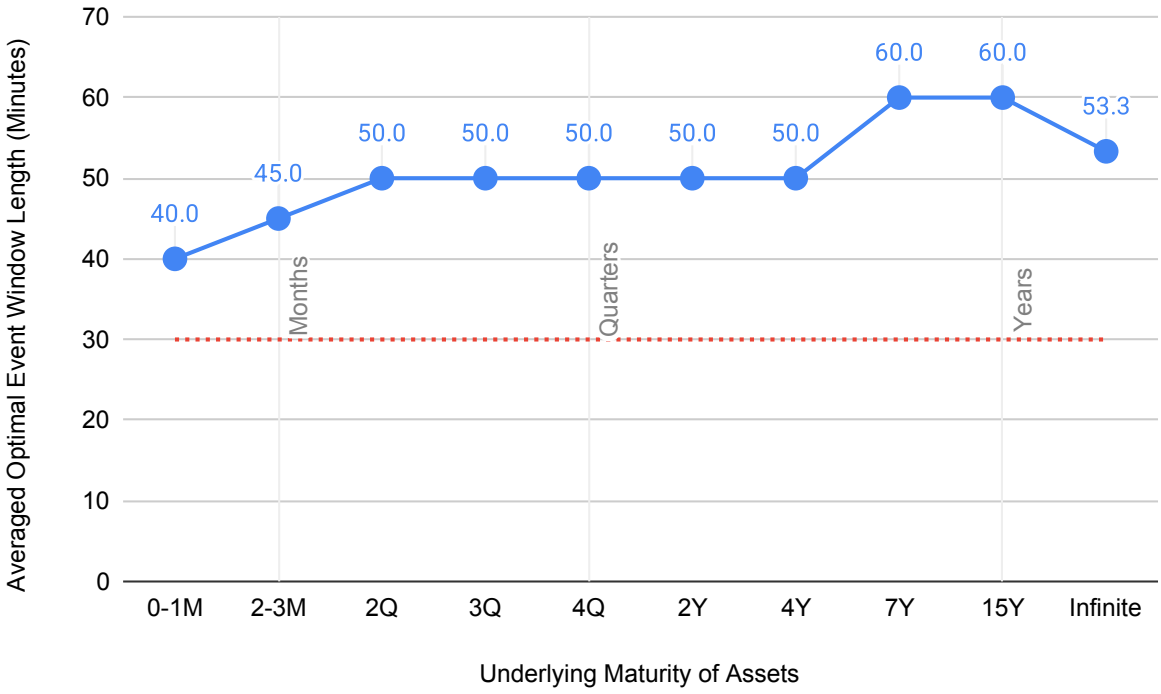


Figure 2: Averaged Optimal Event Window Lengths for Various Asset Underlying Maturities

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, and Wright (2020).

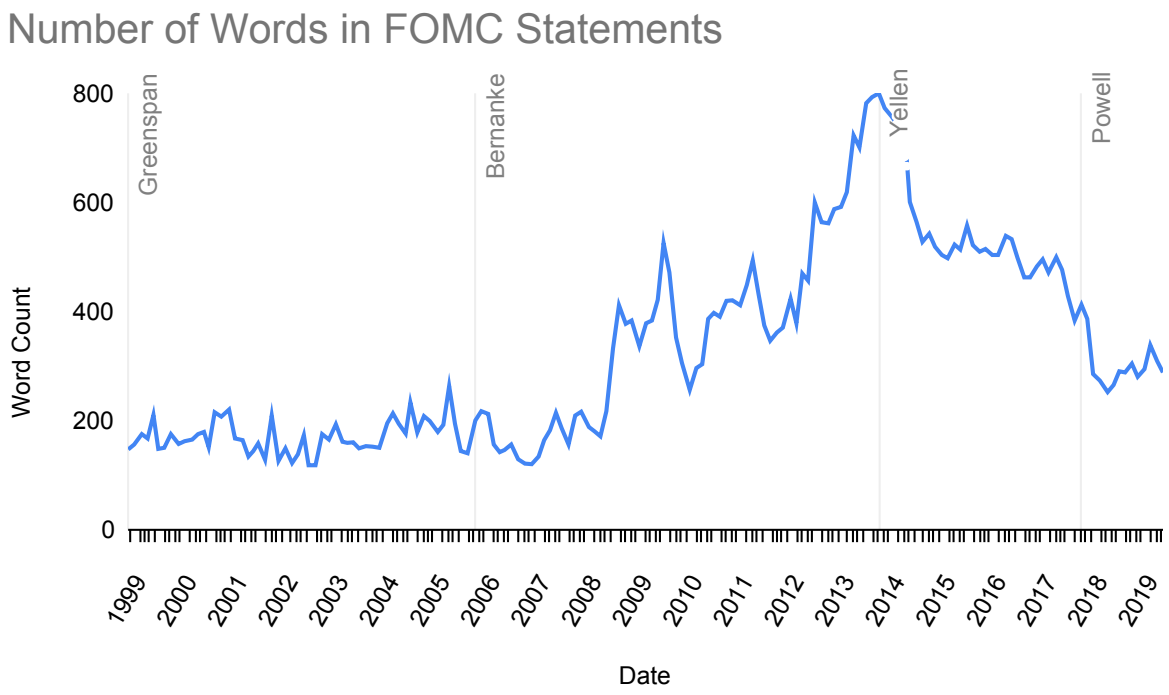
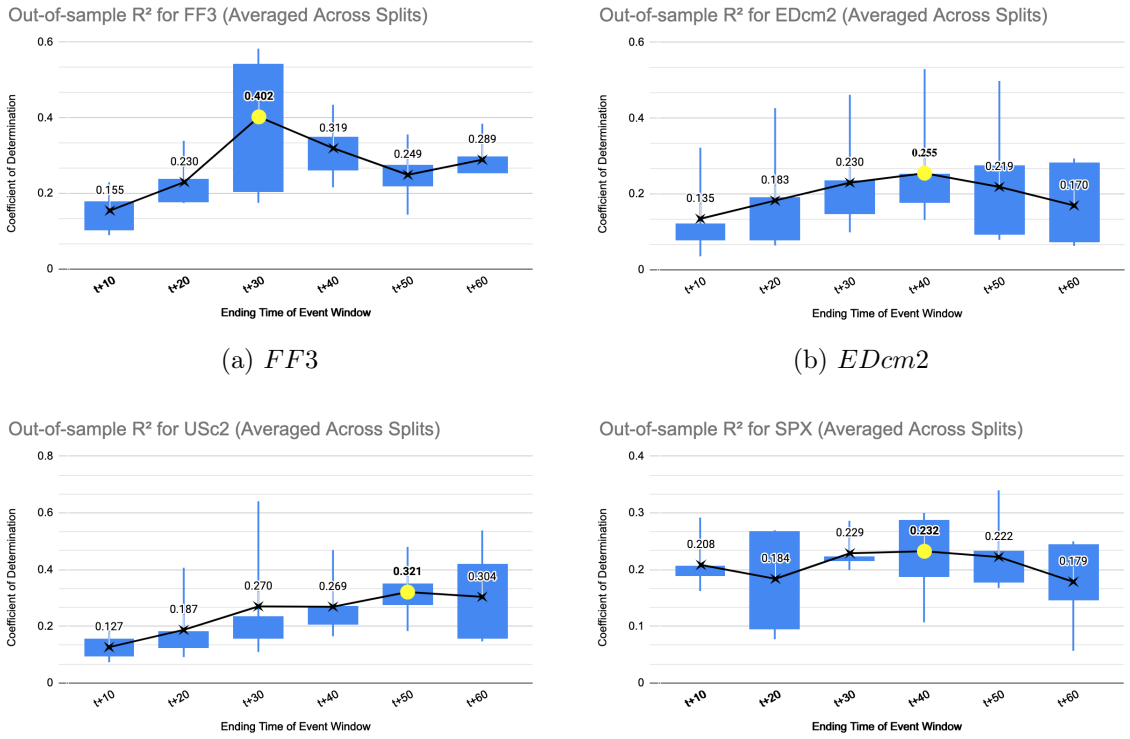


Figure 3: Number of Words in FOMC Statements, May 1999–October 2019

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.



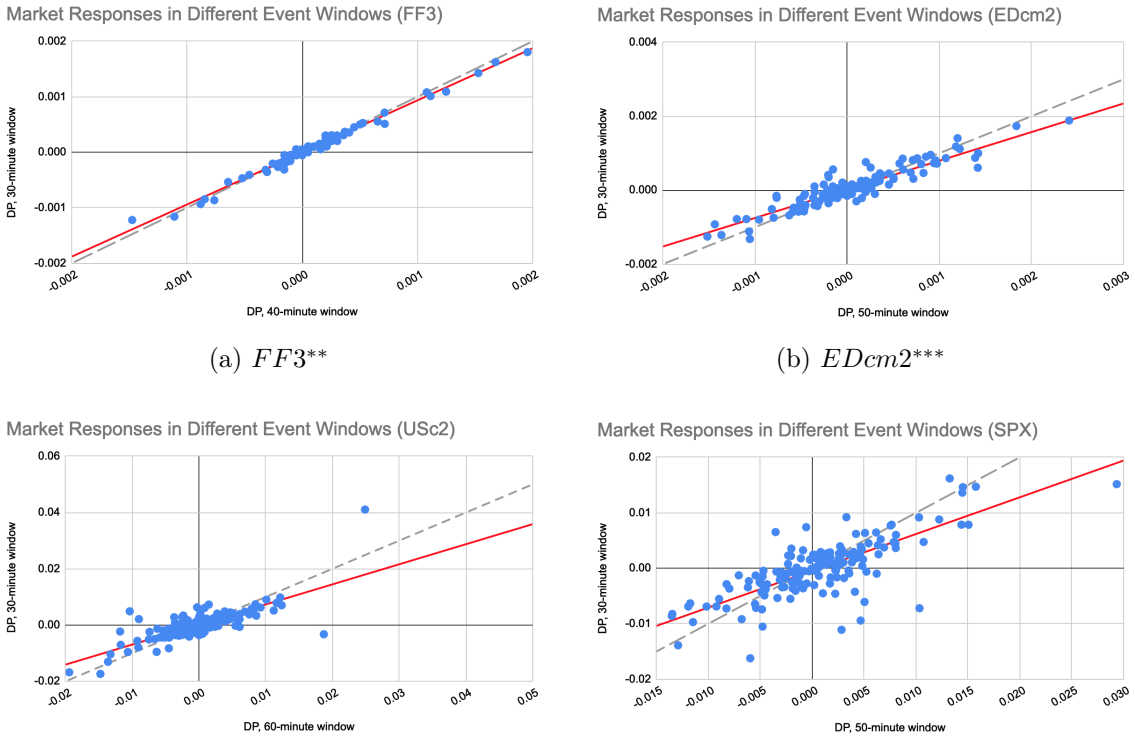
(a) $FF3$

(b) $EDcm2$

(c) $USc2$

(d) SPX

Figure 4: Optimal Event Window Lengths 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 coefficient of determination, R^2_{OOS}. 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. $\overline{R^2_{OOS}}$ measures the average reduction in predictive error achieved by the neural network compared to the naive baseline.

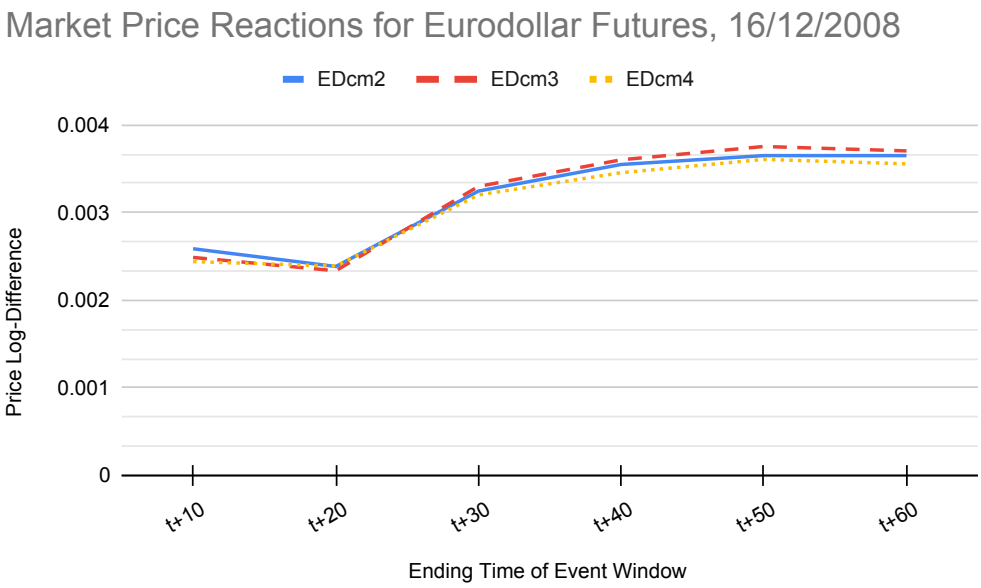


(a) $FF3^{**}$ (b) $EDcm2^{***}$

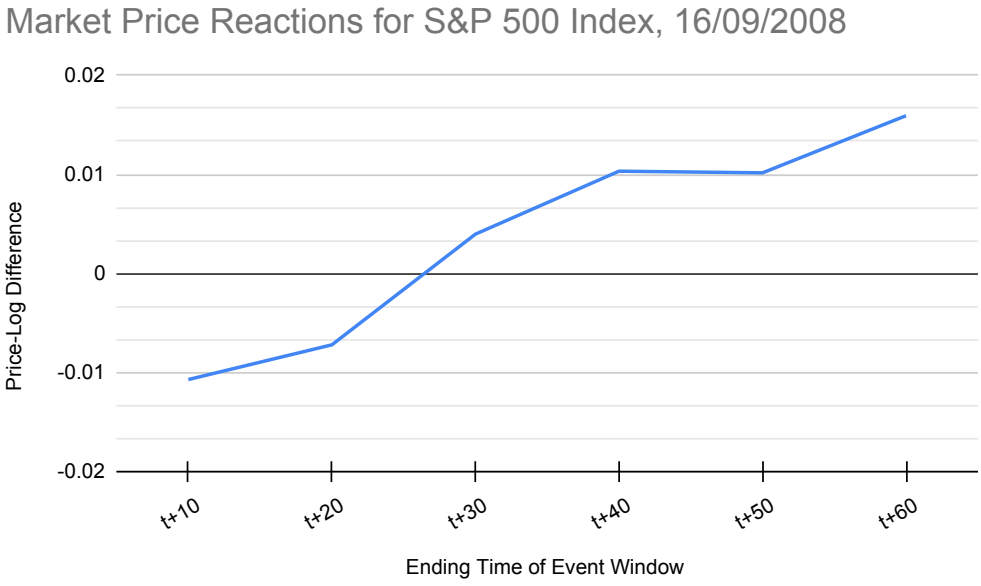
(c) $USc2^{***}$ (d) SPX^{***}

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 $\alpha = 0.05$, 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 Index to FOMC Statements
Notes: The top (bottom) sub-figure depicts the price log-differences of Eurodollar futures markets (the S&P 500 Index) 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.

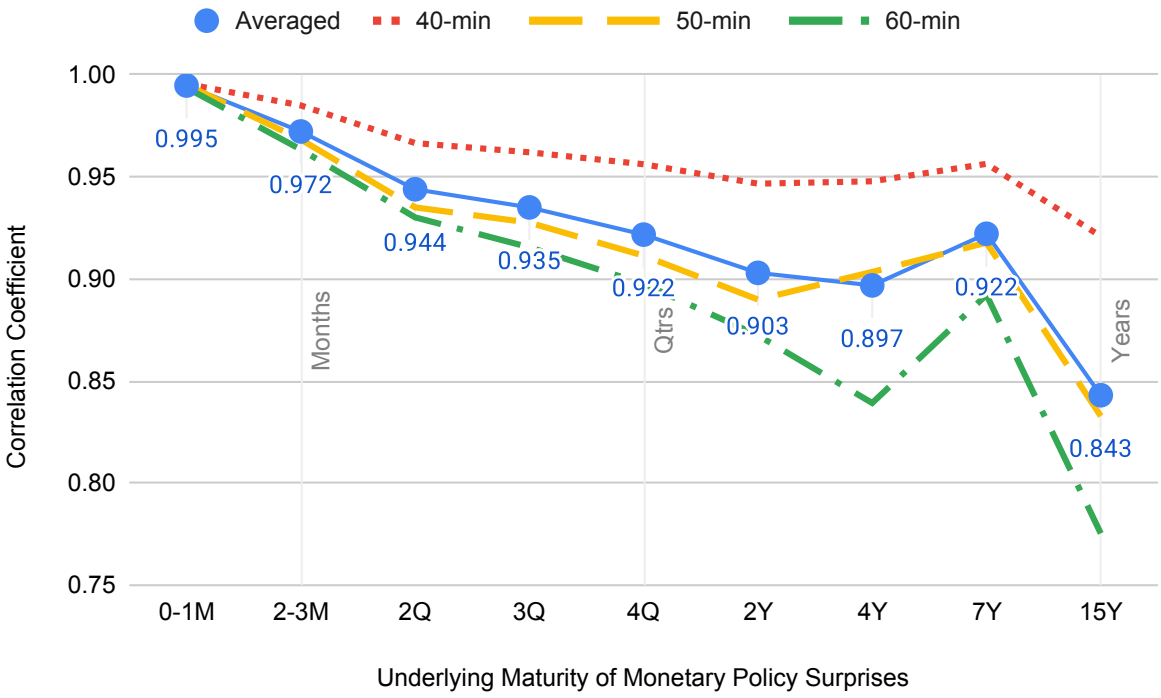
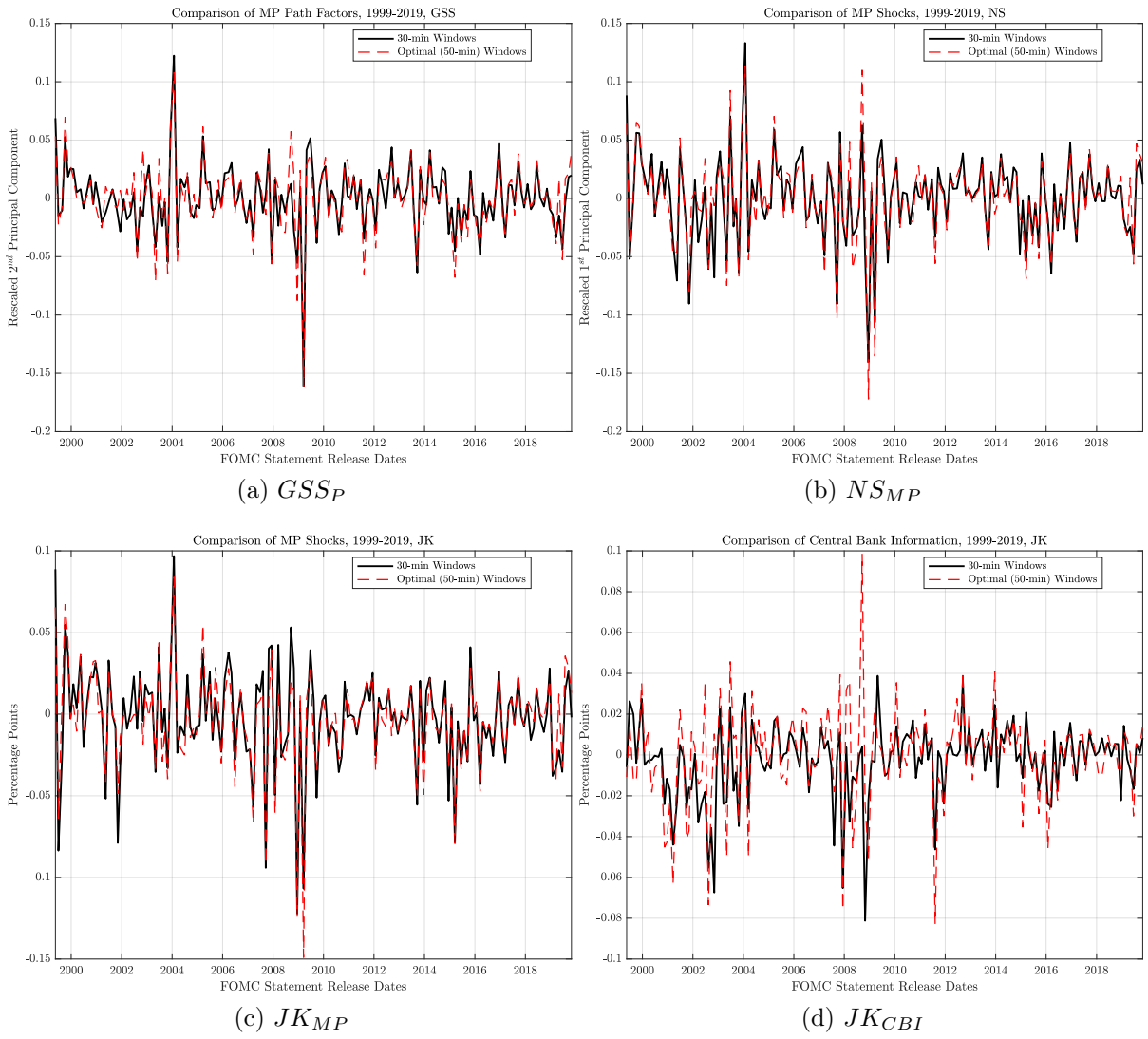


Figure 7: Correlation Between Monetary Policy Surprises Calculated in Optimal v. 30-minute Window Lengths

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, and Wright (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.

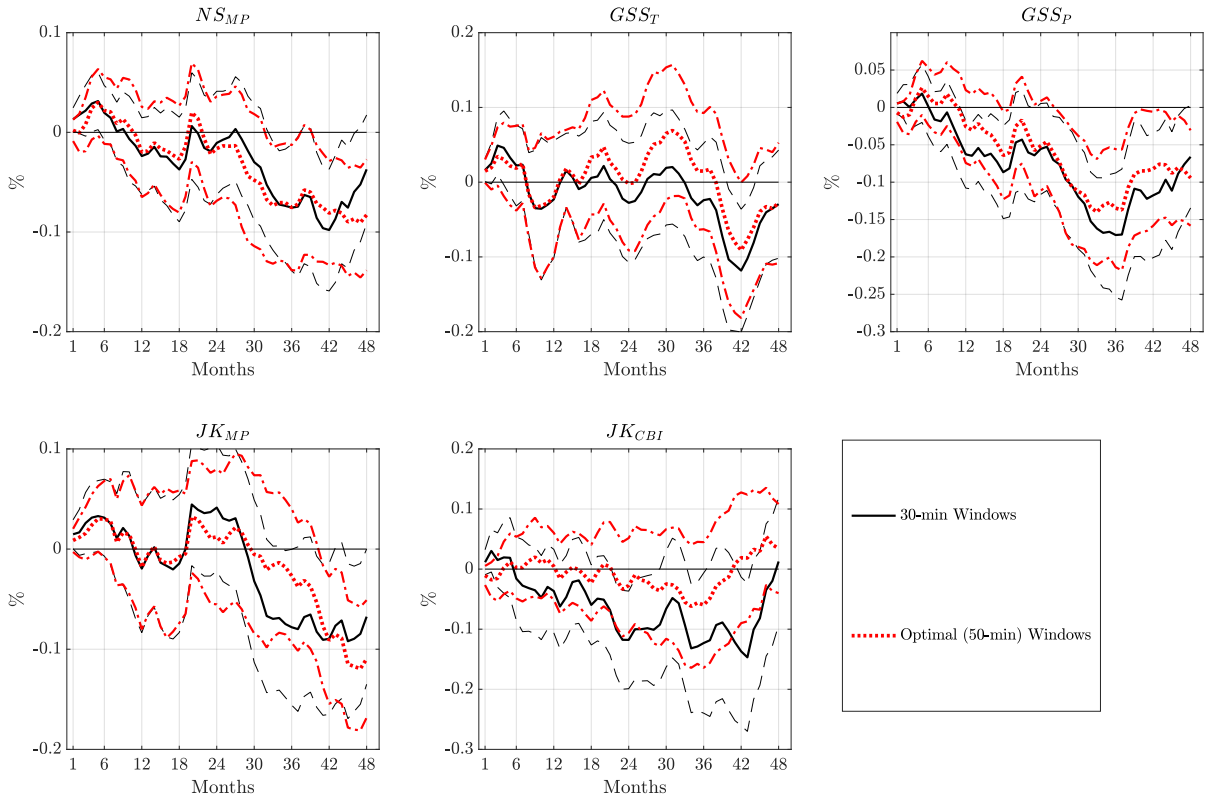


Figure 9: Effects of Event Window Choice on CPI Impulse Responses

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.

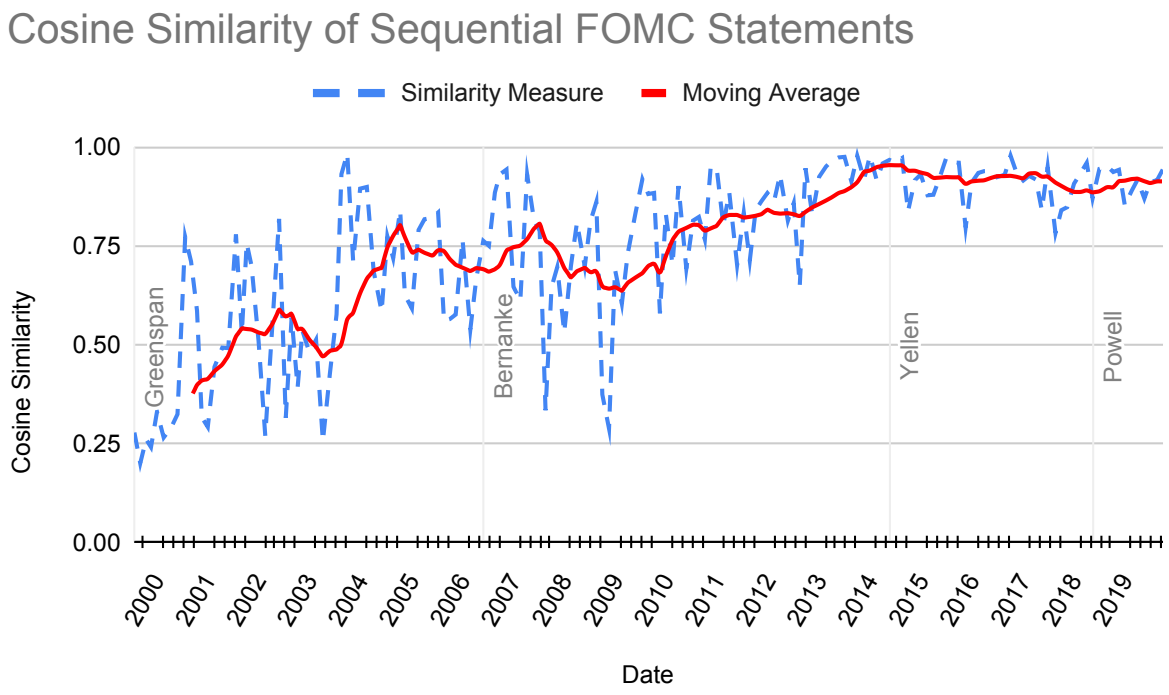


Figure 10: S^1 Cosine Similarity of Sequential FOMC Statements

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 0.826. "Similar" are sequential statements with a cosine similarity of more than 0.826. "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

This appendix details the simulation for the motivating framework in Section 2. I simulate the asset price process for multiple news announcements to illustrate how cognitive noise and unrelated news affect the optimal time horizon. The simulation also demonstrates why a "good" signal is necessary for this estimation, providing a rationale for the paper's methodology.

The simulation assumes news is released at time $t = 0$. The fundamental price, $P^f_{i,t}$, jumps to a random value $P^f_i \in [-100, 100]$ and remains there. Simultaneously, the cognitive noise component jumps to a random value in $[-100, 100]$, while the unrelated news component is zero. The econometrician observes a noisy signal, s_i, with a normally distributed error (mean zero, finite variance). I simulate this process for $N = 10,000$ news announcements up to $t = 100$. For each simulation, I calculate both the true MSE (using P^f_i) and the estimated MSE (using s_i). The true optimal time (t^*) and the estimated optimal time ($\hat{t}$) are the points that minimise these respective MSEs, per Equation 4.

A.1 Simulation Results

I consider three scenarios for the asset price response: (1) high cognitive noise and low unrelated news, (2) low cognitive noise and high unrelated news, and (3) high levels of both noise components. The simulation parameters for these three scenarios are listed in the topmost rows of Appendix Table F1.

The bottom rows of Appendix Table F1 show the results. A market with high cognitive noise and low unrelated news (left column) has a longer optimal time horizon (t^*). Conversely, a market with high unrelated news and low cognitive noise (middle column) has a very small t^*. These results are consistent with Equation 3. Intuitively, a longer reaction process requires a longer optimal time horizon. In contrast, when markets react quickly but unrelated news accumulates, a short window is optimal. The rightmost column shows the most realistic scenario, which includes both cognitive and unrelated noise. Here, the optimal time horizon lies between the two extremes.

Importantly, the simulations show minimal differences between the true (t^*) and estimated ($\hat{t}$) time horizons, even when using only the noisy signal s_i. This demonstrates that a "good" signal allows for a precise estimation of the market's full reaction time.

This motivating framework, however, relies on strong linearity and orthogonality assumptions to isolate the effects of noise. In reality, the true compositions and processes of the fundamental and noise components are unknown, yet they still drive the overall asset price.

Abstracting from this simple framework, I argue the paper's neural network approach can extract a "good" signal by approximating the mapping between communications and price changes, regardless of the true, unobserved noise processes.

B The Primary Accuracy Metric for Judging XLNet

As mentioned in Subsection 4.4, the primary metric for neural network fine-tuning is a generalised R^2 statistic from Hawinkel, Waegeman, and Maere (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 neural network's out-of-sample mean squared error; $\widehat{MST}$ is the out-of-sample MSE from a naive model that predicts using only the in-sample mean ($\overline{y_{IS}}$); and T is the testing sample size. While the network's explicit objective function is to minimise $\widehat{MSE}$, this action simultaneously maximises R^2_{OOS}.

This metric is chosen for two main reasons. First, the conventional R^2 (i.e., the proportion of variance explained by a model) breaks down for non-linear methods like neural networks. For such models, the total variance no longer neatly decomposes into model and residual components, meaning the squared Pearson correlation coefficient does not equal R^2. Second, the R^2_{OOS} formula is specifically designed to assess out-of-sample performance, not in-sample fit.[49]

As is standard in machine learning, model performance is compared to a baseline. This baseline assumes *no* relationship between the FOMC statement text and futures price changes, naively predicting with the in-sample average. Therefore, the R^2_{OOS} measures the reduction in predictive error achieved by the neural network compared to this naive baseline.

During fine-tuning, XLNet-Base also tracks other metrics: the out-of-sample Pearson correlation between predicted and actual values, the out-of-sample mean absolute error, and the in-sample mean squared error. This last metric is important for verifying that the network is genuinely learning the underlying relationship.

The decisive criterion for systematic estimation is the out-of-sample R^2 *averaged across sample splits* for each event window length and futures contract:

$$\overline{R^2_{OOS}} = \frac{\sum_{i=1}^{K} R^2_{OOS}}{K}, \tag{B2}$$

where $K = 5$ is the number of sample splits. This $\overline{R^2_{OOS}}$ statistic measures the neural network's average performance improvement within a given window, relative to the naive baseline. The optimal event window is therefore the length that yields the largest $\overline{R^2_{OOS}}$. This is the point where the network's predictive power and generalisability is highest, which only occurs when the impact of noise is minimised and the fundamental reaction is fully captured.

[49] Using a conventional R^2 yields similar optimal window lengths and network quality, but R^2_{OOS} is a more appropriate metric.

C Construction of High-Frequency Monetary Surprises

C.1 Financial Data Overview and Price Selection

The intraday data for interest-rate and equity futures (May 1999 to October 2019) is from the Thomson Reuters Tick History database via LSEG. Table F2 provides an overview of this data. For each futures contract, I use a minutely price series. The remainder of this appendix details the construction of interest-rate surprises from these contracts. Following the literature, I use federal funds futures for expectations up to three months, Eurodollar futures for two to four quarters, and Treasury futures for expectations out to fifteen years.

As discussed in Subsection 3.2, I collect prices at 10-minute intervals, from 10 minutes before to 18 hours after an FOMC statement release. Because trades do not occur at every exact interval, I apply two rules. For the pre-announcement price ($t-10$), I use the last recorded trade *at or before* that timestamp. For all post-announcement intervals (e.g., $t+10, t+20$), I use the first recorded trade *at or after* that interval's timestamp. This method assumes the price is constant between recorded trades.

C.2 Surprises from Federal Funds Futures

A federal funds rate futures (FFF) contract's settlement price is based on the average effective federal funds rate over its expiry month. The contract pays out 100 minus this average rate. Therefore, the market's expected average federal funds rate is implied by its price. Let $p_{\tau,t}^{FFj}$ be the price of the $(j-1)$ month-ahead FFF at time t for FOMC meeting τ. The expected average federal funds rate is then calculated as $100 - p_{\tau,t}^{FFj}$.

C.2.1 Surprises from Federal Funds Futures: Current Meeting

I calculate the "current meeting" monetary policy surprise ($mp1$) for each FOMC announcement τ at time $t+n$ as:

$$mp1_{\tau,t+n} = \frac{m}{m-d}\left(FF1_{\tau,t+n} - FF1_{\tau,t-10}\right), \tag{B3}$$

where $FF1$ is the implied federal funds rate from the current-month contract ($100 - p_{\tau,t}^{FF1}$), d is the day of the announcement, and m is the number of days in the month.

A special case applies if the FOMC announcement occurs in the last seven days of the month ($m-d+1 \leq 7$). To avoid distortion from the large $\frac{m}{m-d}$ scaling factor, the surprise is calculated as the simple change in the next month's futures contract: $mp1_{\tau,t+n} = FF2_{\tau,t+n} - FF2_{\tau,t-10}$.

C.2.2 Surprises from Federal Funds Futures: Next Meeting

The surprise for the next scheduled FOMC meeting ($mp2$) is more complex, as it must isolate the new information for that specific meeting. First, the relevant futures contract (j) is identified, which corresponds to the month of the next scheduled meeting ($j \in \{2,3,4\}$).

The surprise is then calculated as:

$$mp2_{\tau,t+n} = \frac{m_2}{m_2 - d_2}\left\{[FFj_{\tau,t+n} - FFj_{\tau,t-10}] - \frac{d2}{m2} mp1_{\tau,t+n}\right\}, \tag{B4}$$

where FFj is the implied rate from the relevant contract, d_2 and m_2 are the day of and total days in the month of the next meeting, and $mp1_{\tau,t+n}$ is the current-meeting surprise.

As with the current-meeting surprise, a special case applies if the next meeting falls in the last seven days of its month ($m_2 - d_2 + 1 \leq 7$). To avoid distortion, the surprise is calculated as the simple change in the following month's contract: $mp2_{\tau,t+n} = FF(j+1)_{\tau,t+n} - FF(j+1)_{\tau,t-10}$.

C.3 Surprises from Eurodollar Futures

Eurodollar contracts are quarterly contracts (expiring in March, June, September, or December) that pay out 100 minus the US Dollar BBA LIBOR interest rate at expiration. Let $p^{edj}_{\tau,t}$ be the price of the j^{th} closest quarterly futures contract at time t on FOMC meeting date τ.

The implied interest rate is $edj_{\tau,t} = 100 - p^{edj}_{\tau,t}$. The monetary policy surprise for this contract, mp_{edj}, is calculated as the simple change in this implied rate within the event window:

$$mp_{edj,\tau,t+n} = edj_{\tau,t+n} - edj_{\tau,t-10}. \tag{B5}$$

C.4 Surprises from Treasury Futures

Treasury futures contracts oblige the seller to deliver Treasury bonds within a specific maturity range at expiration. Let $p^{tk^j}_{\tau,t}$ be the price of the j^{th} nearest quarterly k-year Treasury futures contract at time t on FOMC meeting date τ.

The implied yield surprise ($tk_{\tau,t+n}$) for an event window ending at $t+n$ is calculated by dividing the price change by its approximate maturity (l) and flipping the sign:

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

where the approximate maturities, $l \in \{2, 4, 7, 15\}$, are from Gürkaynak, Kisacikoğlu, and Wright (2020). Following Gorodnichenko and Ray (2017), if the FOMC meeting occurs during an expiry month (March, June, etc.), I use the next-nearest contract ($j+1$) due to its higher liquidity.

D Robustness Checks

D.1 Original Surprise Instruments for MP Surprises & Shocks

This appendix repeats the exercises from Subsection 6.2, but constructs the monetary policy shocks using the *original* instrument sets from the cited authors.[50]

Visual Shock Differences I plot the shock series constructed from the original instrument sets in Figure E28. The axes and scaling in each sub-figure match the original authors' specifications.

The results are robust. The visual differences in the forward guidance shocks caused by the event window choice are similar to those in the main text (Figure 8), which used the full instrument set. This robustness also applies to the JK_{CBI} shock, which continues to show "shifts of importance" in the optimal window, a result of the Jarociński and Karadi (2020) sign restriction method.

Responses of Interest Rates Appendix Tables F11 and F12 present the regression results for nominal and real interest rates, respectively, following the specification in Equation 8. The results are robust. Opening the event window to the optimal 50 minutes yields stronger point estimates (in basis points) for all monetary policy shocks containing forward guidance. Furthermore, using the optimal window also increases the statistical significance of several coefficient estimates, such as the 2-year TIPS yield response to GSS_P or the nominal Treasury yield responses to the JK_{CBI} shock.

Responses of Break-even Inflation Regression results for break-even inflation using the original surprise instruments are in Appendix Table F13. The results are robust to the instrument choice. Similar to the main text, monetary policy shocks constructed within the optimal window have more negative impacts on break-even inflation at all maturities. Furthermore, moving from the 30-minute to the optimal window again increases the statistical significance for several estimates.

Responses of Equities I next examine the impact of monetary policy shocks on equity prices, again using the original instrument sets. Appendix Table F14 presents analogous results for the S&P 500 Index and its E-mini futures, following the specification in Equation 8, with the dependent variable as 100 times the log-price difference.

Compared to the main results (Subsection 6.2), the effects of the window choice are more mixed for equities. For example, while the optimal window strengthens the impact of GSS_P on the S&P 500 Index by 2.52%, it weakens the impact of NS_{MP} by 3.32%. Although all shocks remain statistically significant and have the expected sign, the overall effects of the window choice are smaller than in the main text. In contrast, the results for the second-month S&P 500 E-mini futures contract remain similar to those in the main text.

The overall results suggest that the original instrument sets, which rely on shorter-maturity surprises, may understate the shock effects. This limitation is particularly relevant as my sample period includes the effective lower bound, where forward guidance is not fully captured by surprises with maturities of only one year. This finding aligns with recent studies (e.g., Brennan et al., 2024; An, Stedman, and Lusompa, 2025; and others) which

[50] For consistency with the other shock constructions, I use the next-meeting surprise ($mp2_\tau$) to build the Jarociński and Karadi (2020) shocks, as it represents a similar underlying maturity to their original $FF4$ instrument.

show that incorporating longer-maturity instruments is necessary to prevent understating monetary policy shocks in this era.

Responses of the Macroeconomy via Local Projections I estimate the macroeconomic impulse responses using the lag-augmented local projection method of Olea and Plagborg-Møller (2021). The model estimates the responses of log consumer price index, log industrial production, the nominal two-year Treasury yield, and the excess bond premium (EBP) to each monetary policy shock, following the specification in Equation 9.

Figures E29–E32 plot the impulse responses. Similar to the main results, the event window choice does not qualitatively change most point estimates, though some exceptions exist (e.g., the EBP response to the NS_{MP} shock becomes statistically significant and positive after 36 months).

Also consistent with the main findings, the confidence intervals of the impulse responses from the optimal window are visibly smaller. To formalise this, I calculate the confidence interval width ratios as discussed in Subsection 6.2. Using the original instrument set, the average and median ratios are 0.9033 and 0.9506, respectively. This confirms that using the optimal event window length yields more precise estimates of the macroeconomic responses, even with the original instrument set.

D.2 Optimal Event Window Lengths Beyond 70 Minutes?

The "one signal" approach also allows for a crude check of whether a larger $\overline{R^2_{OOS}}$ exists beyond the 70-minute window used in the "joint" estimation. Appendix Figures E21–E27 display this exercise, extending the event window out to 18 hours after the FOMC statement release. For most assets, the event window with the largest $\overline{R^2_{OOS}}$ is the same one found via the "joint" approach, supporting the overall validity of the 70-minute estimation limit.

However, for a few contracts (e.g., $FF2$, $FF4$, $TUc1$, $TYc2$, and $USc1$), the "one signal" approach does suggest an optimal window longer than 70 minutes. To confirm these, I calculate the $\overline{R^2_{OOS}}$ using the full "joint" approach for these specific longer windows. This check yields two insights. First, the true optimal window for these contracts is still within 60 minutes.

Second, even though the "one signal" optimum was incorrect, the "joint" estimation for these longer windows (beyond 60 minutes) still yielded a higher $\overline{R^2_{OOS}}$ than the conventional 30-minute window. This suggests the global optimum for some assets may lie somewhere between 60 minutes and the time predicted by the "one signal" approach. Regardless, this exercise confirms that a 30-minute window is never optimal.

Appendix Figures

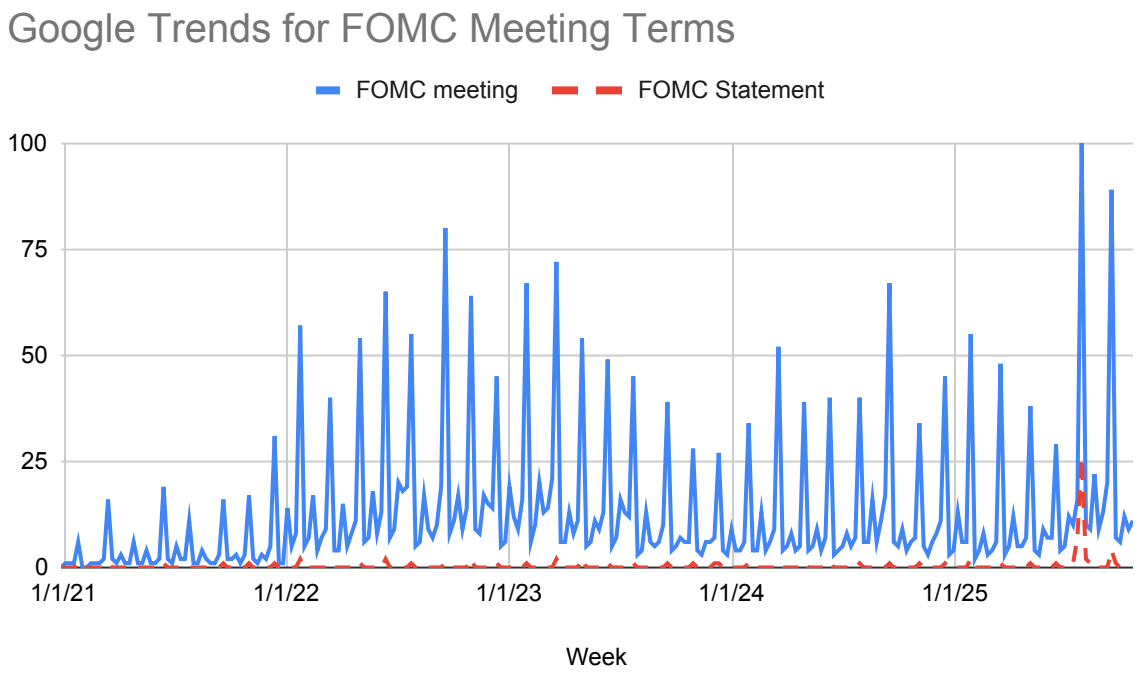
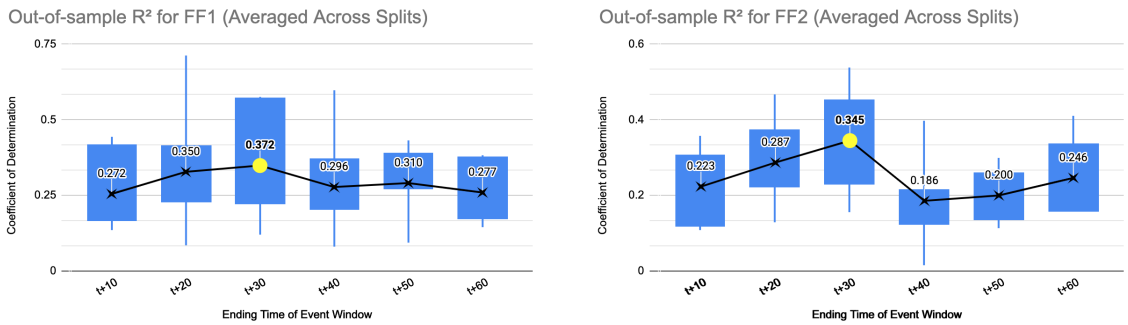


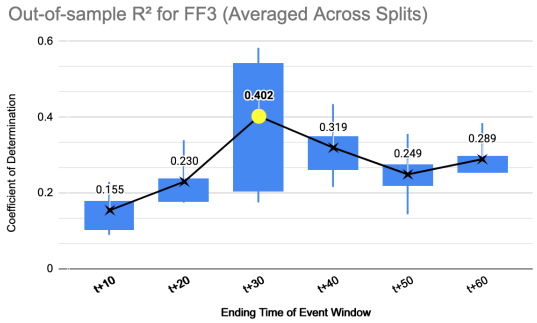
Figure E1: Google Trends Interest About FOMC Statements over Time, January 2021–October 2025

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

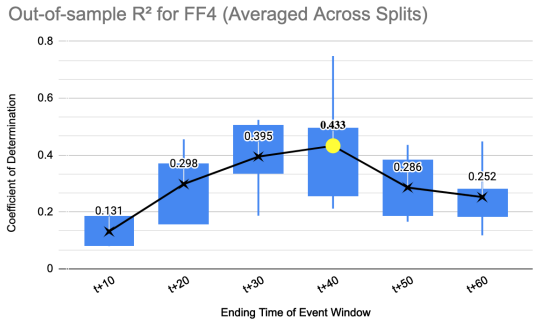


(a) $FF1$

(b) $FF2$



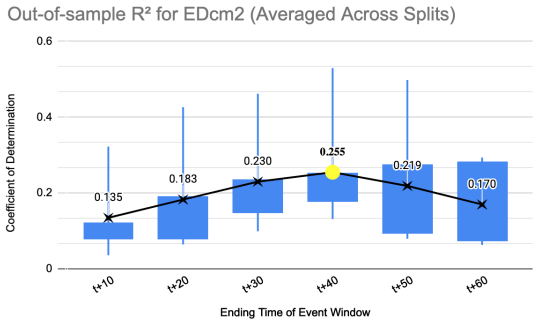
(c) $FF3$



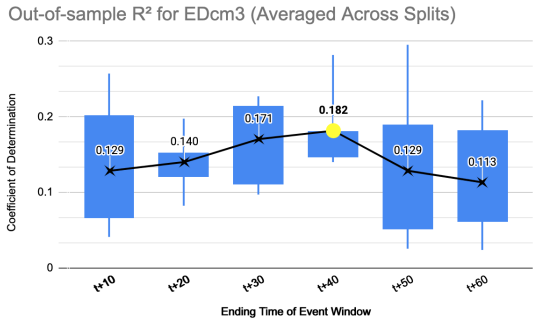
(d) $FF4$

Figure E2: Optimal Event Window Lengths for Federal Funds Futures

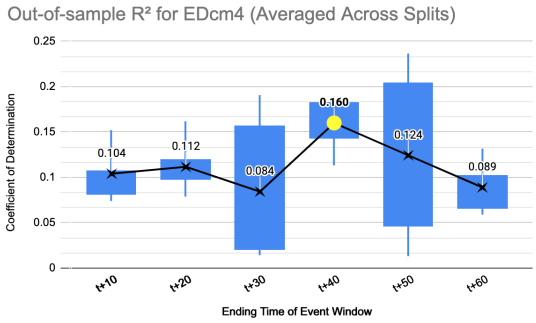
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the coefficient of determination, R^2_{OOS}. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $\overline{R^2_{OOS}}$ measures the average reduction in predictive error achieved by the neural network compared to the naive baseline.



(a) $EDcm2$



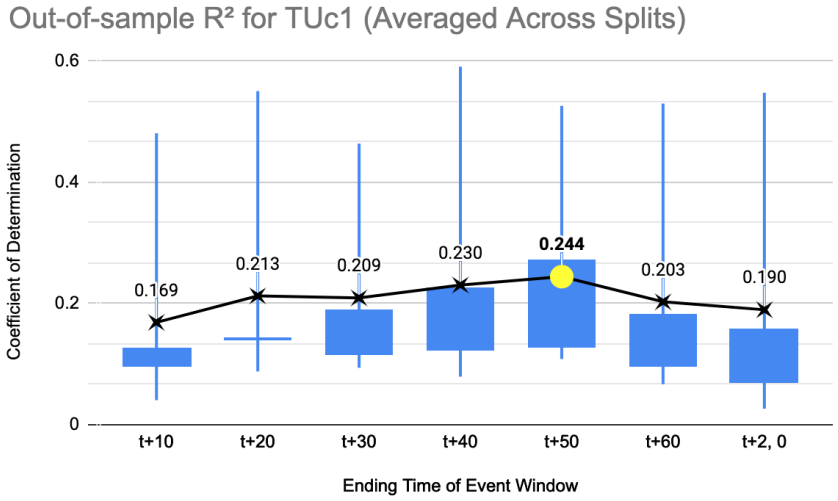
(b) $EDcm3$



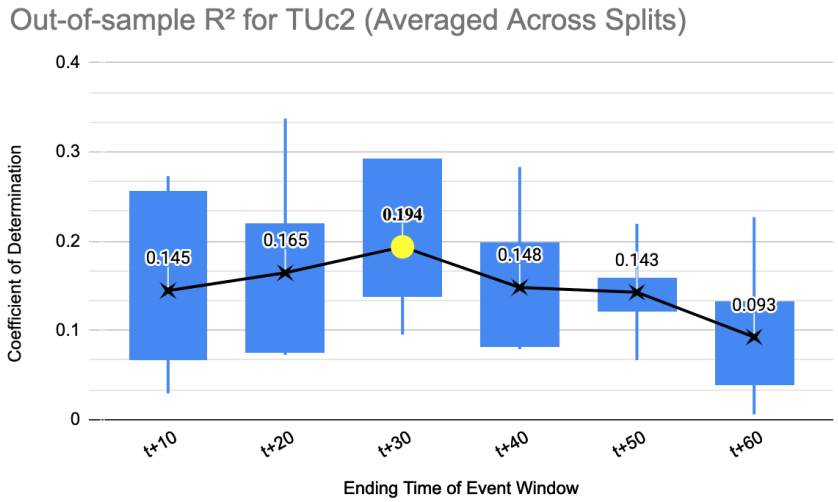
(c) $EDcm4$

Figure E3: Optimal Event Window Lengths 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 coefficient of determination, R^2_{OOS}. 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. $\overline{R^2_{OOS}}$ measures the average reduction in predictive error achieved by the neural network compared to the naive baseline.



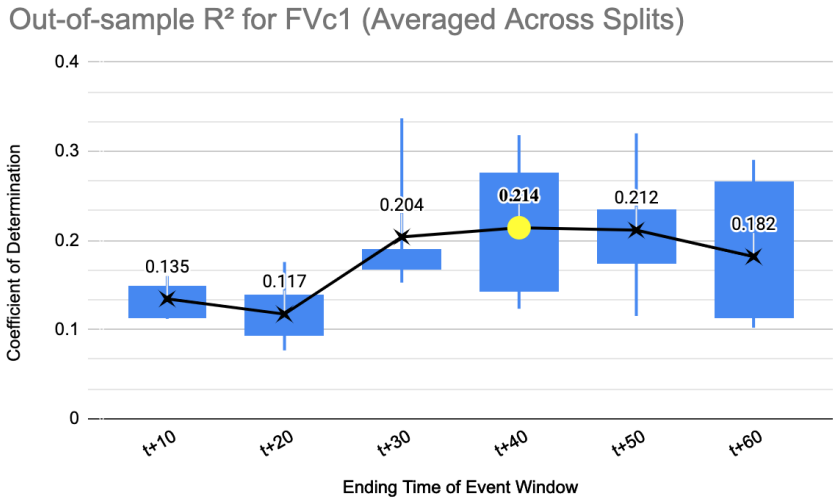
(a) $TUc1$



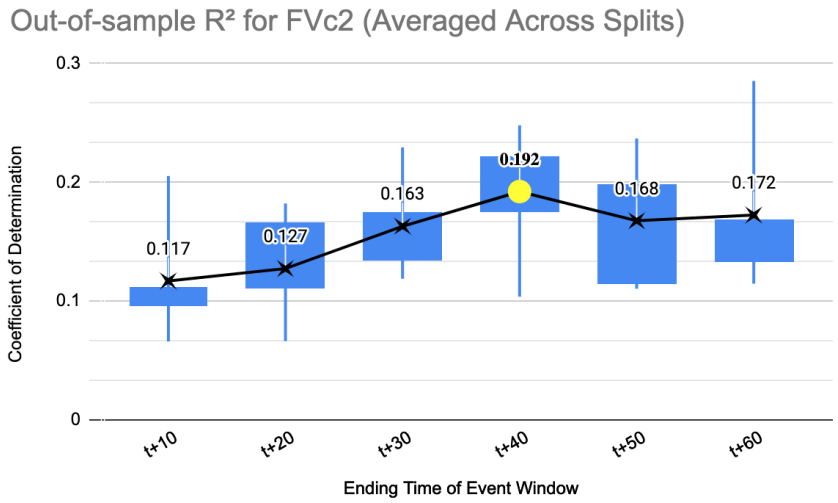
(b) $TUc2$

Figure E4: Optimal Event Window Lengths 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 coefficient of determination, R^2_{OOS}. 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. $\overline{R^2_{OOS}}$ measures the average reduction in predictive error achieved by the neural network compared to the naive baseline.



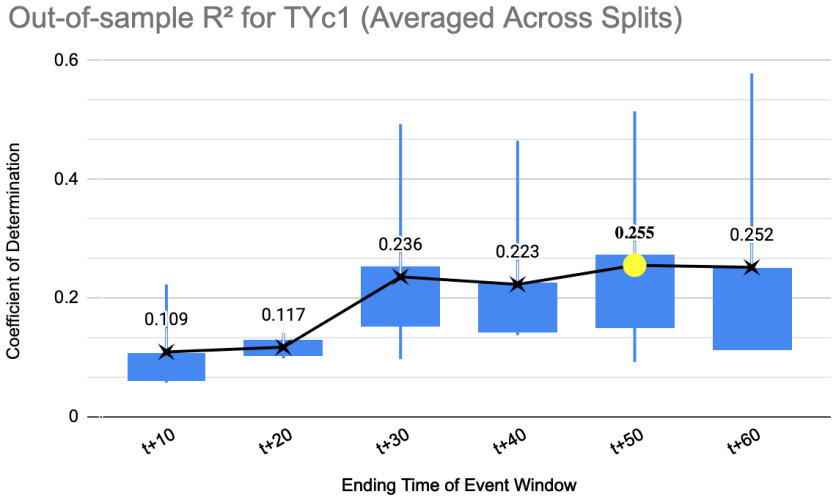
(a) $FVc1$



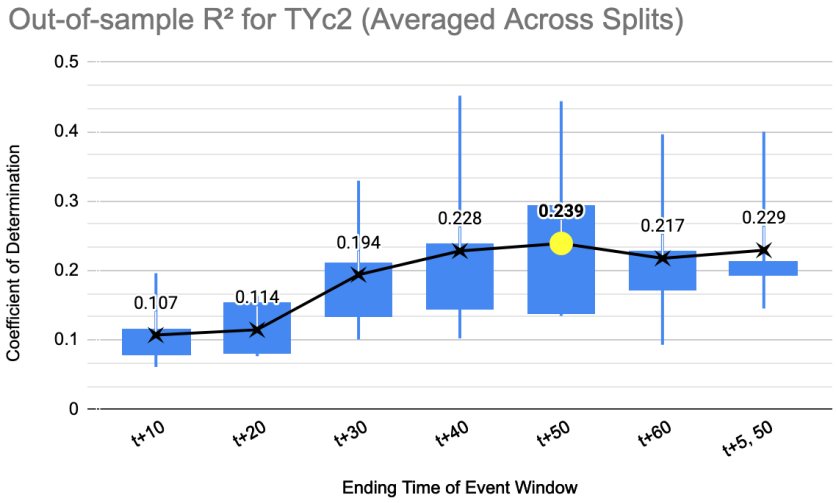
(b) $FVc2$

Figure E5: Optimal Event Window Lengths 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 coefficient of determination, R^2_{OOS}. 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. $\overline{R^2_{OOS}}$ measures the average reduction in predictive error achieved by the neural network compared to the naive baseline.



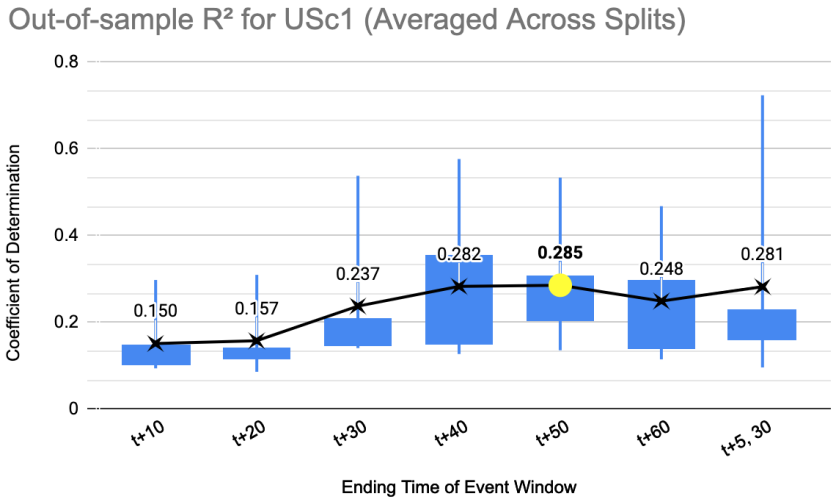
(a) $TYc1$



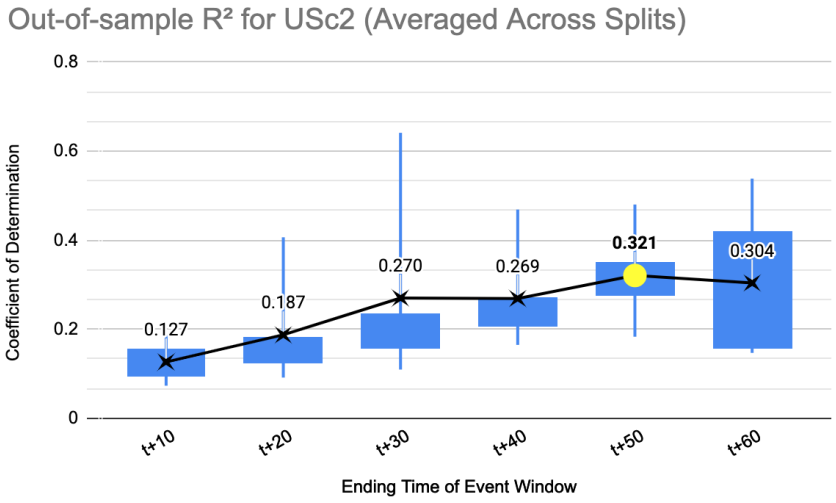
(b) $TYc2$

Figure E6: Optimal Event Window Lengths 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 coefficient of determination, R^2_{OOS}. 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. $\overline{R^2_{OOS}}$ measures the average reduction in predictive error achieved by the neural network compared to the naive baseline.



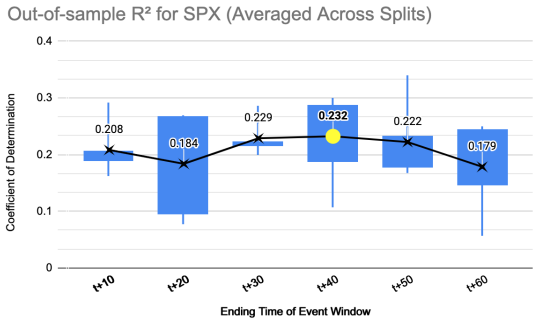
(a) $USc1$



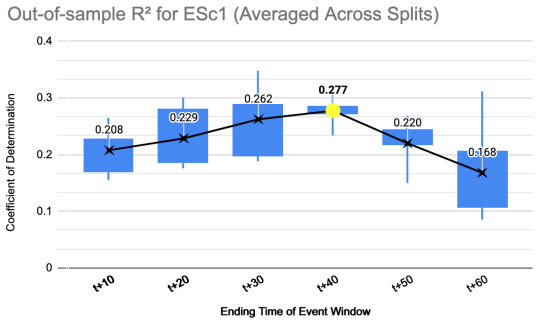
(b) $USc2$

Figure E7: Optimal Event Window Lengths for 30-Year Treasury Futures

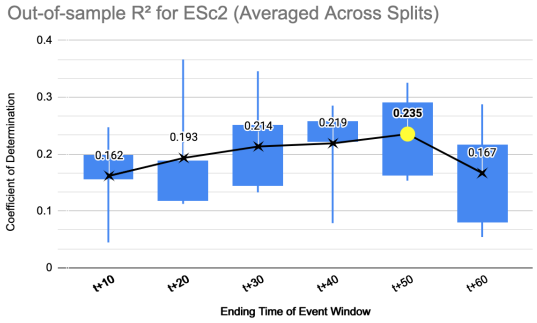
Notes: The horizontal axis of each figure depicts the end time of the considered event window lengths and the vertical axis represents the coefficient of determination, R^2_{OOS}. The cross points represent $\overline{R^2_{OOS}}$ for each event window size, where the solid yellow point represents the event window with the largest $\overline{R^2_{OOS}}$. For each event window, box-and-whisker plots are shown surrounding the corresponding averages. $\overline{R^2_{OOS}}$ measures the average reduction in predictive error achieved by the neural network compared to the naive baseline.



(a) SPX



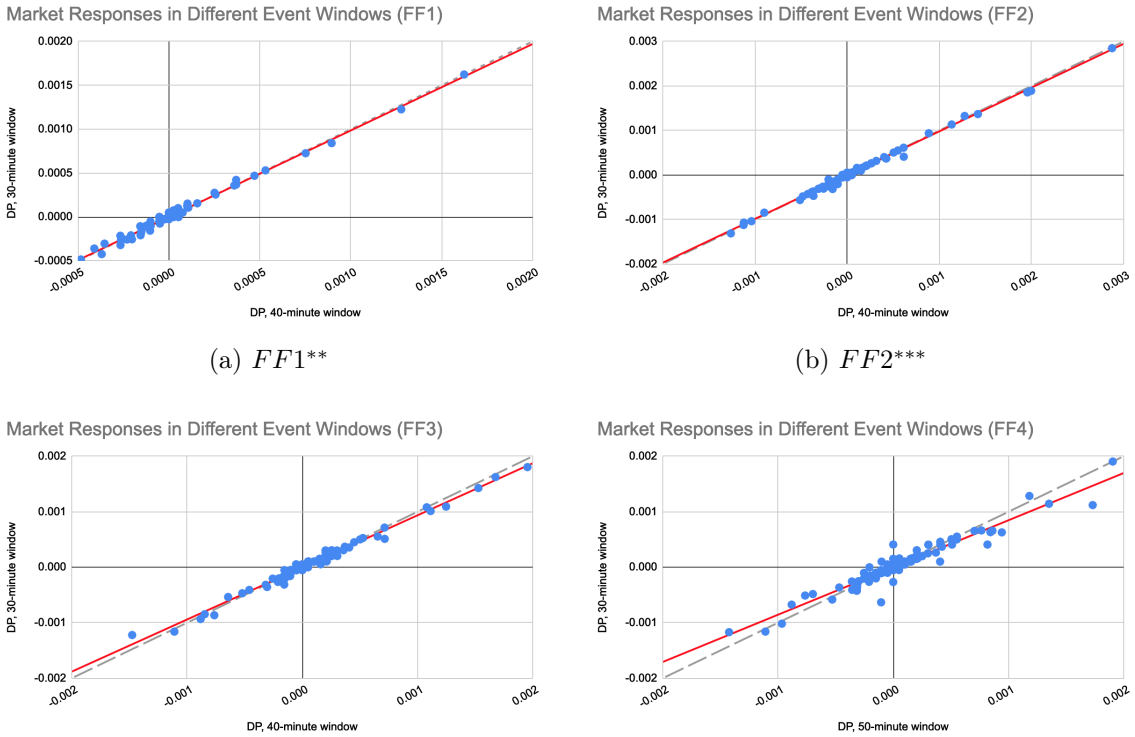
(b) $ESc1$



(c) $ESc2$

Figure E8: Optimal Event Window Lengths 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 coefficient of determination, R^2_{OOS}. 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. $\overline{R^2_{OOS}}$ measures the average reduction in predictive error achieved by the neural network compared to the naive baseline.

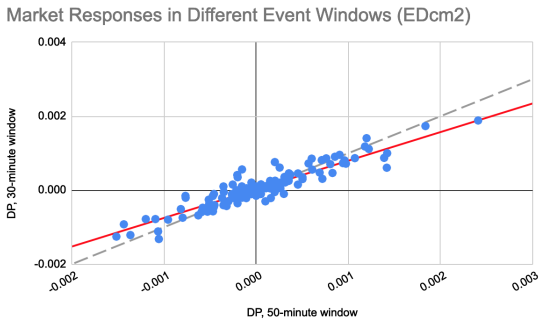


(a) $FF1^{**}$ (b) $FF2^{***}$

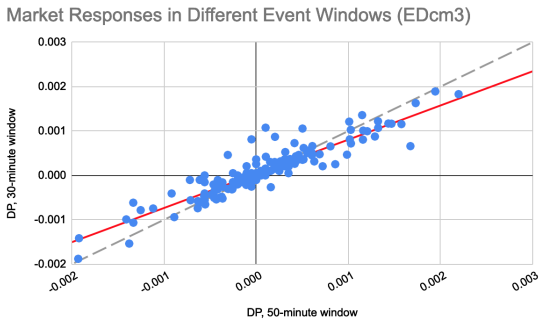
(c) $FF3^{***}$ (d) $FF4^{***}$

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

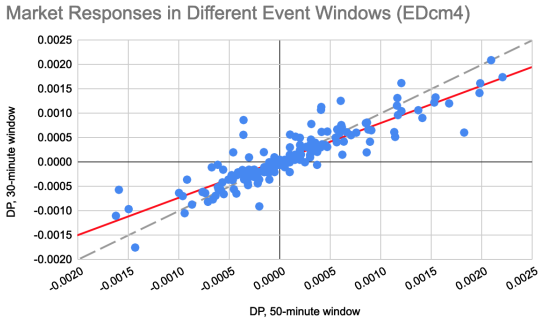
Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient of the regression line is statistically less than 1, it confirms under-reaction 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) $EDcm2^{***}$



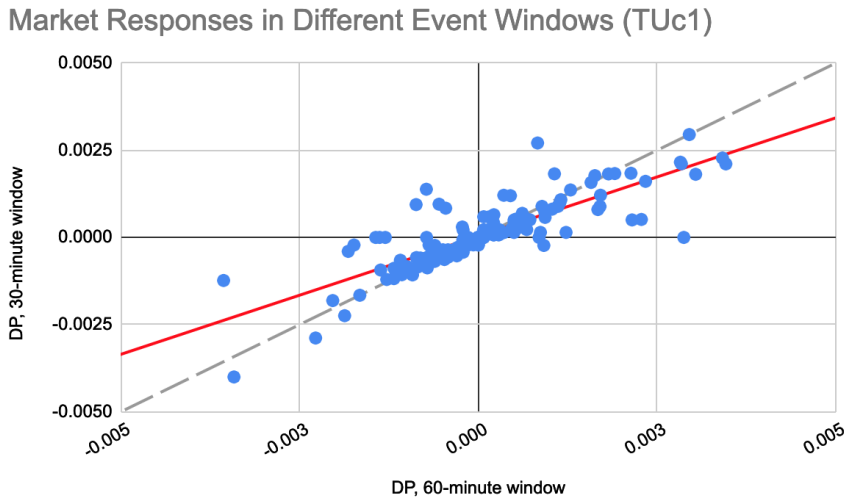
(b) $EDcm3^{***}$



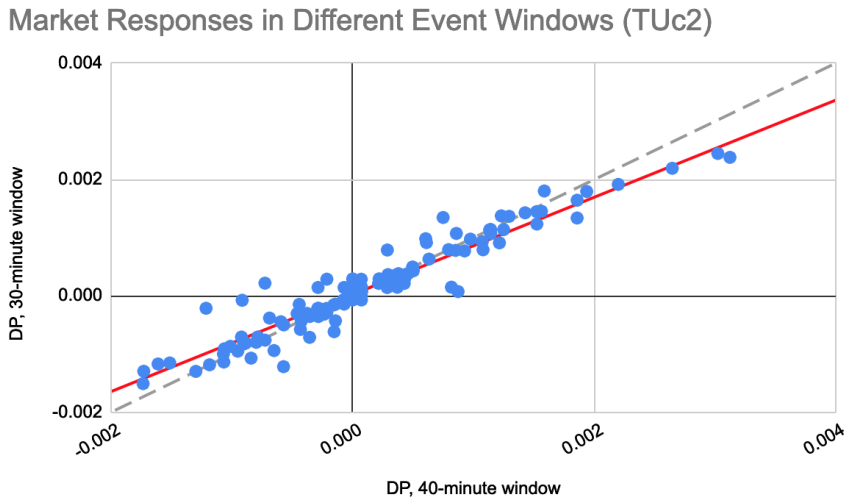
(c) $EDcm4^{***}$

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 of the regression line is statistically less than 1, it confirms under-reaction 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.



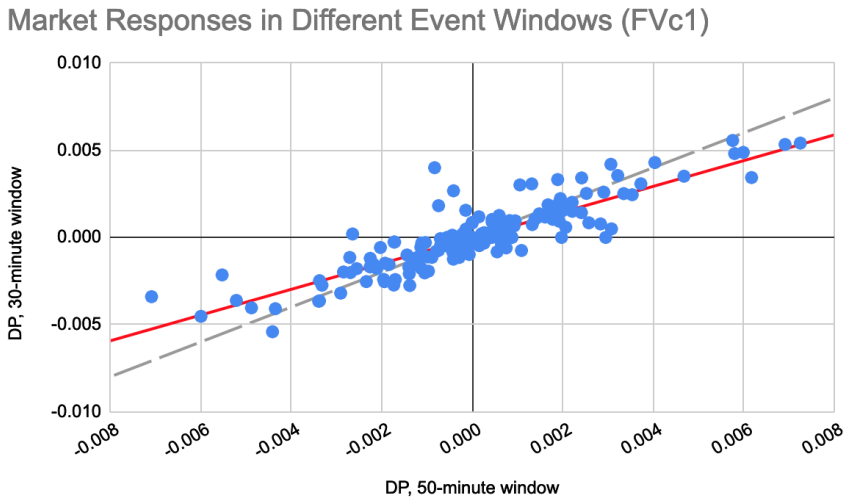
(a) $TUc1^{***}$



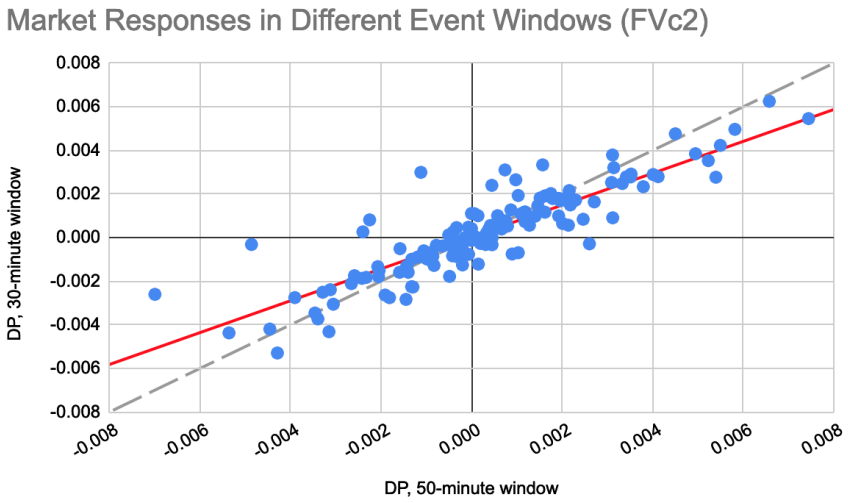
(b) $TUc2^{***}$

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 of the regression line is statistically less than 1, it confirms under-reaction 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.



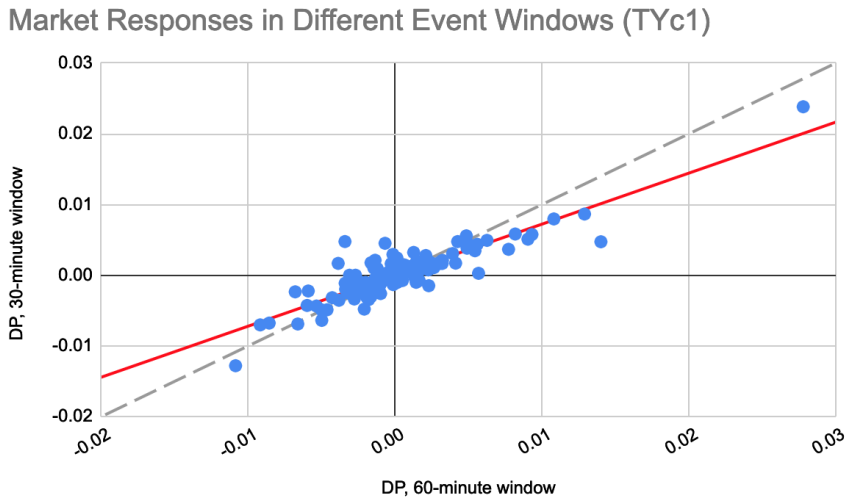
(a) $FVc1^{***}$



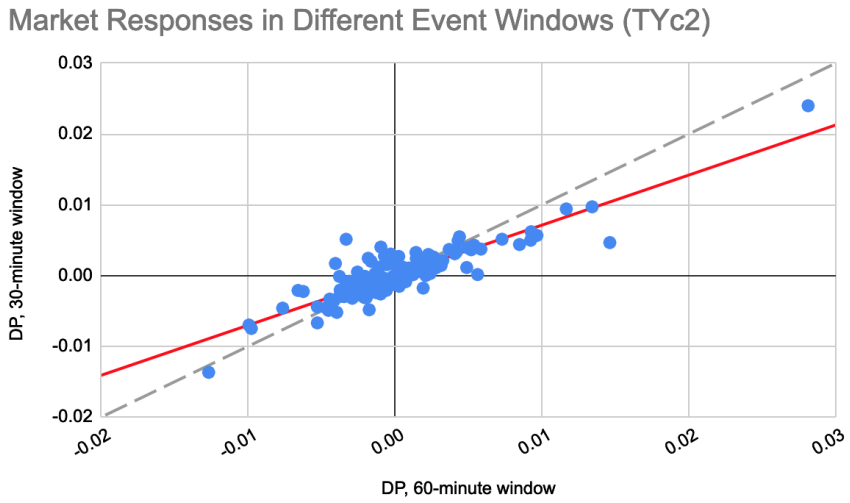
(b) $FVc2^{***}$

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 of the regression line is statistically less than 1, it confirms under-reaction 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.



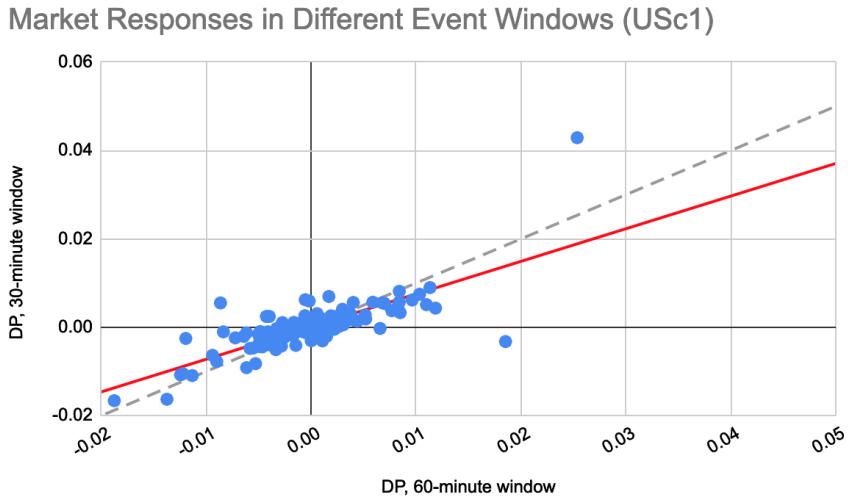
(a) $TYc1^{***}$



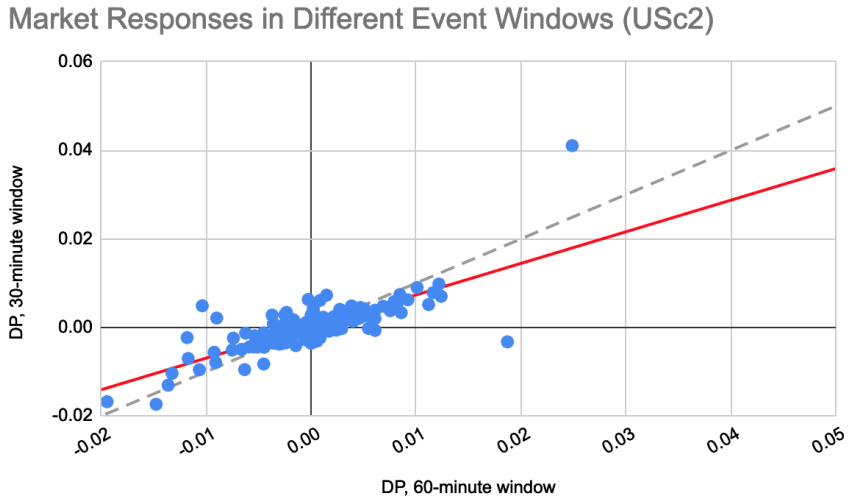
(b) $TYc2^{***}$

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 of the regression line is statistically less than 1, it confirms under-reaction 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.



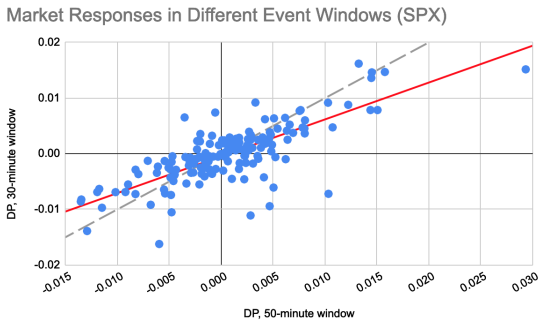
(a) $USc1^{***}$



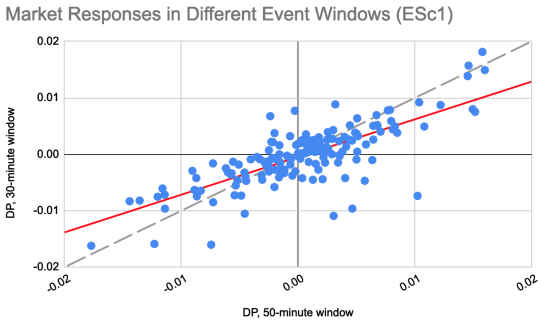
(b) $USc2^{***}$

Figure E14: Comparing Market Responses in Different Event Windows for 30-Year Treasury Futures

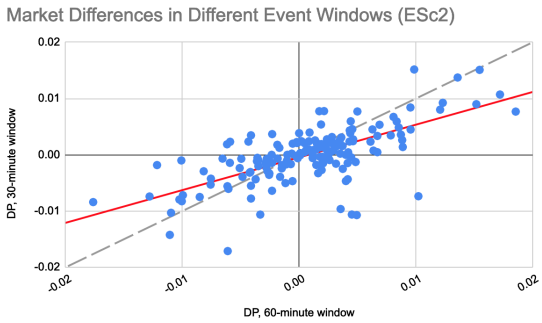
Notes: The horizontal axis depicts the price log-difference within the systematically estimated event window length. The vertical axis represents DP_{t+20}. The 45-degree line is depicted as grey and dashed. The blue dots are market reactions on scheduled FOMC meetings. The price-log differences calculated within the optimal window lengths are regressed on DP_{t+20} through OLS. If the slope coefficient of the regression line is statistically less than 1, it confirms under-reaction 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 of the regression line is statistically less than 1, it confirms under-reaction 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.

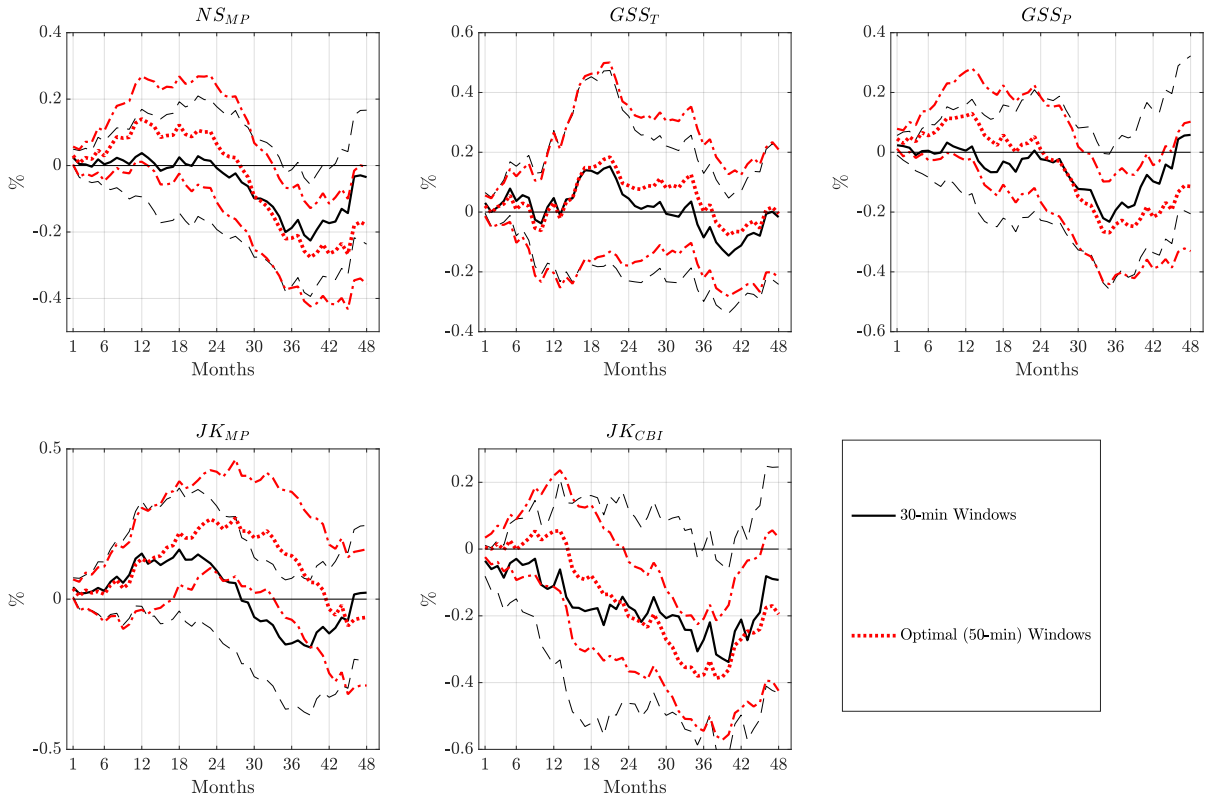


Figure E16: Effects of Event Window Choice on IP Impulse Responses

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.

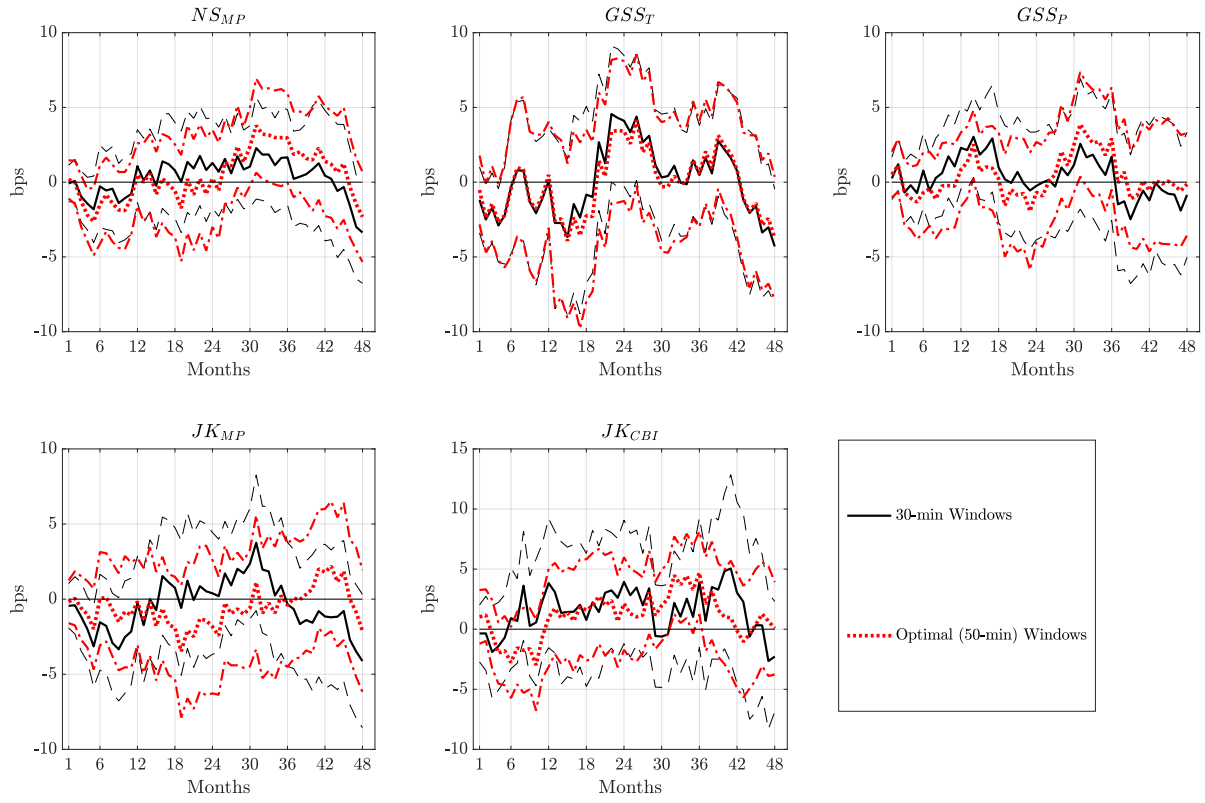


Figure E17: Effects of Event Window Choice on Excess Bond Premium Impulse Responses

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.

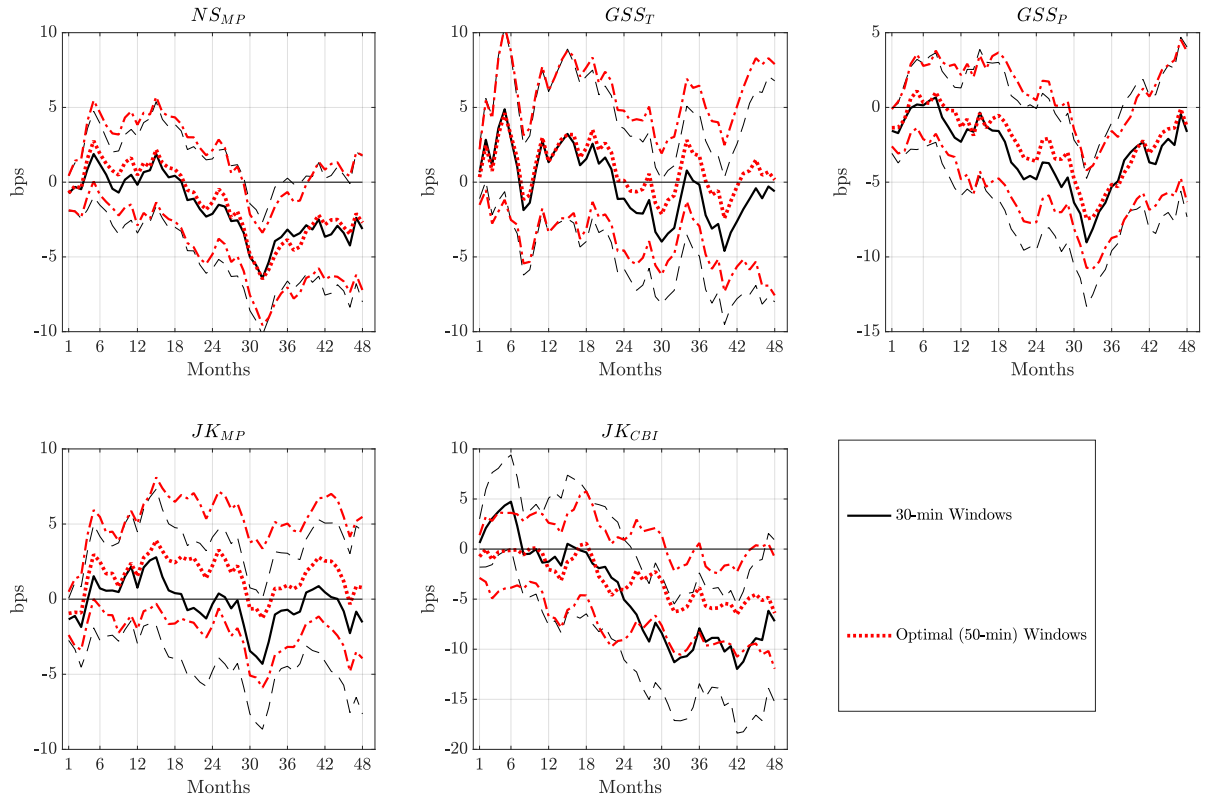


Figure E18: Effects of Event Window Choice on Two-year Treasury Yield Impulse Responses

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.

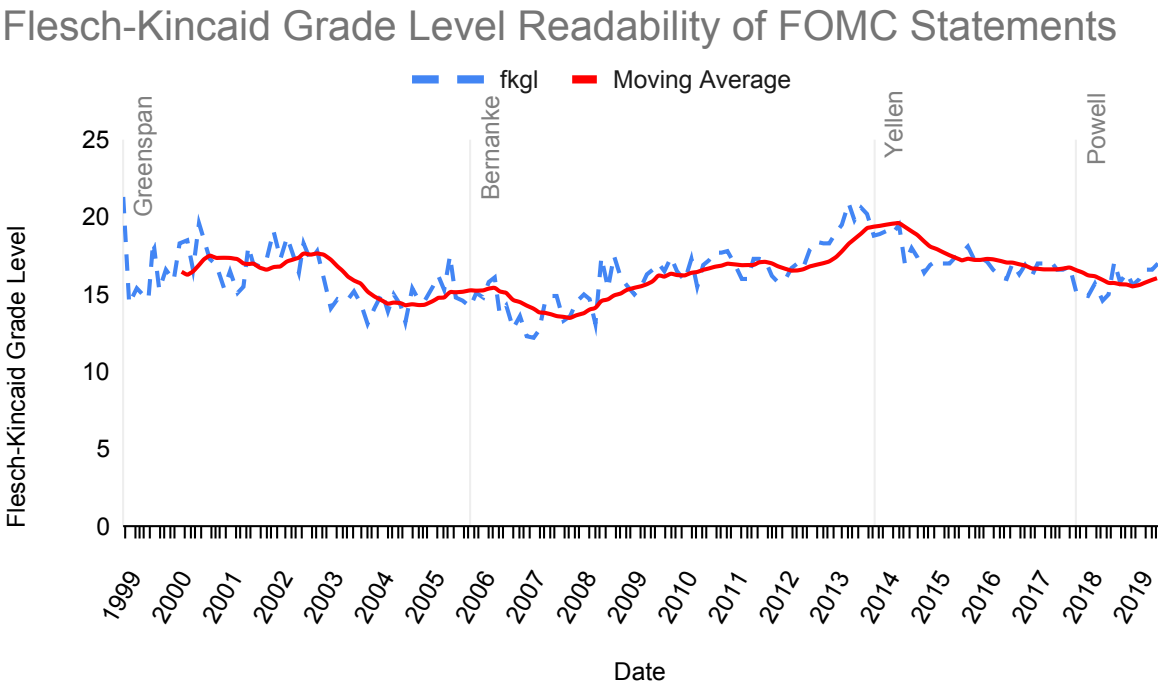


Figure E19: Flesch-Kincaid Grade Level Readability of FOMC Statements

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.

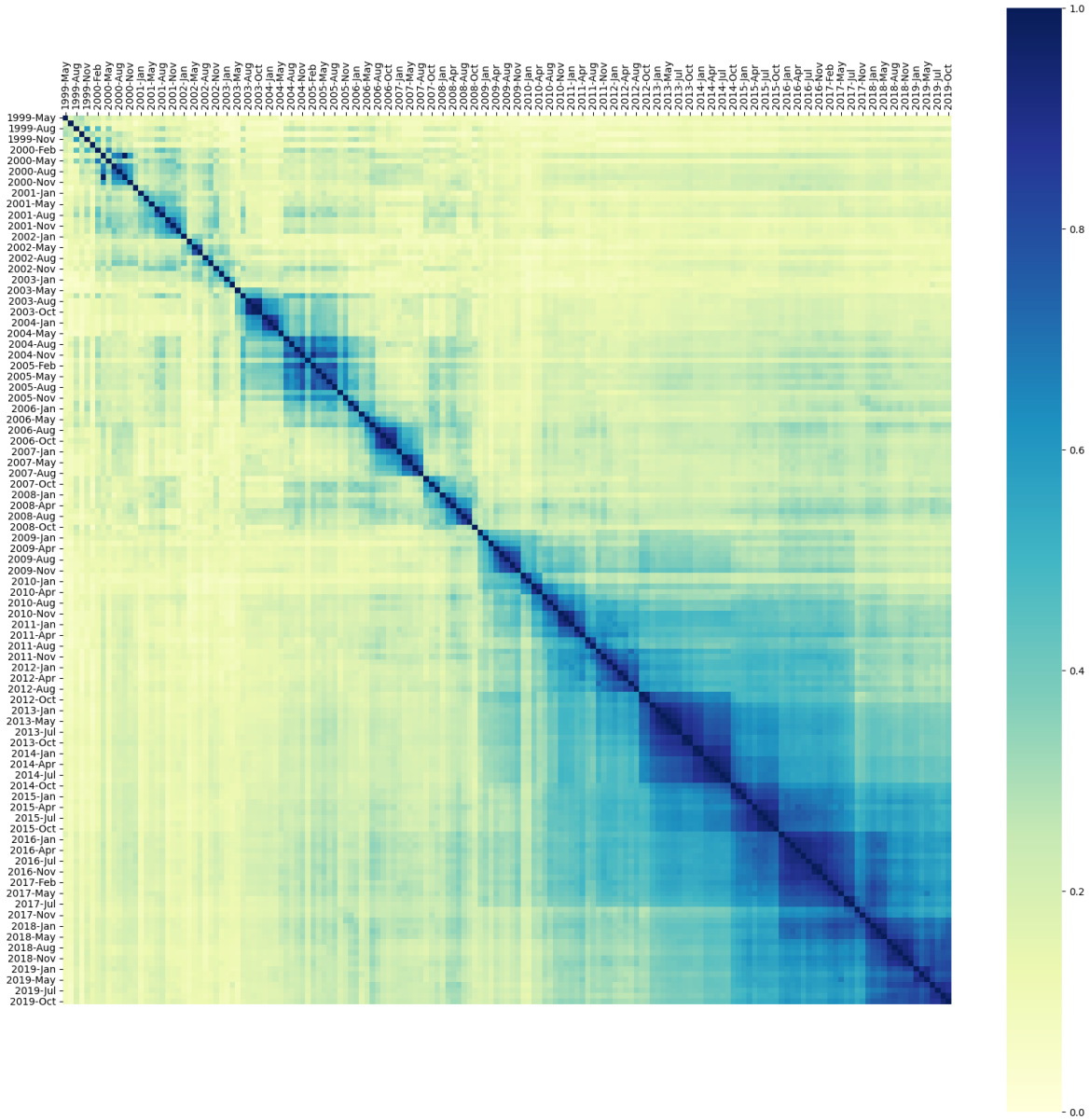


Figure E20: Document Similarity Matrix for FOMC Statements

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, while 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.

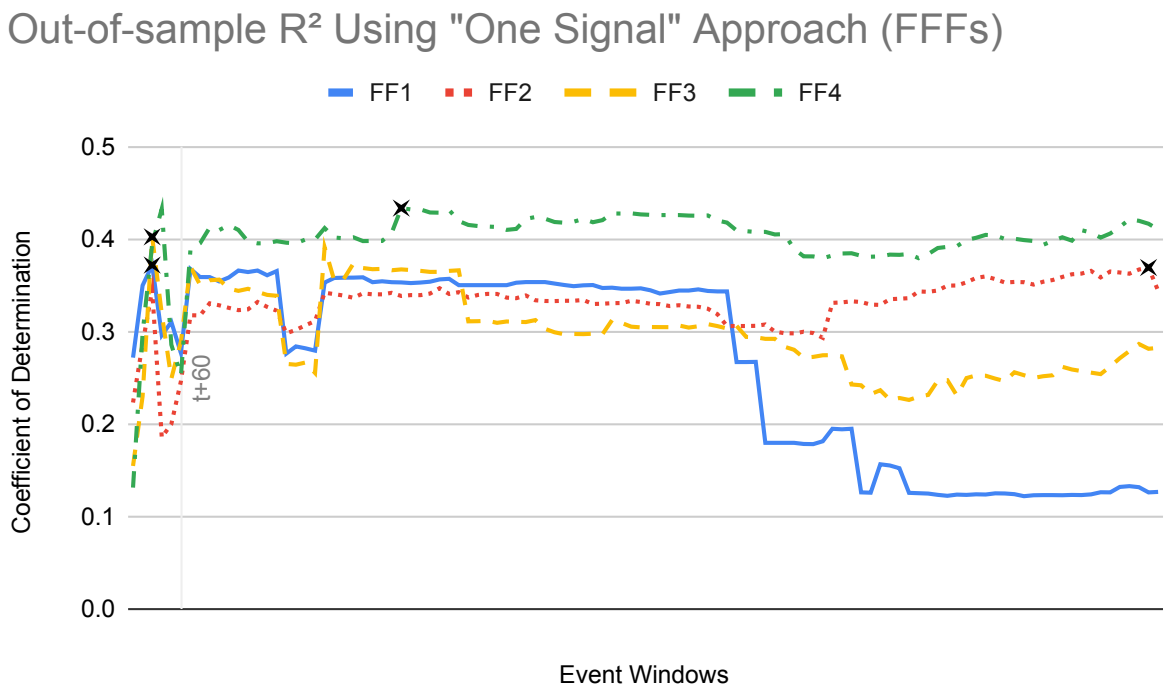


Figure E21: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for Federal Funds Futures

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 using 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.

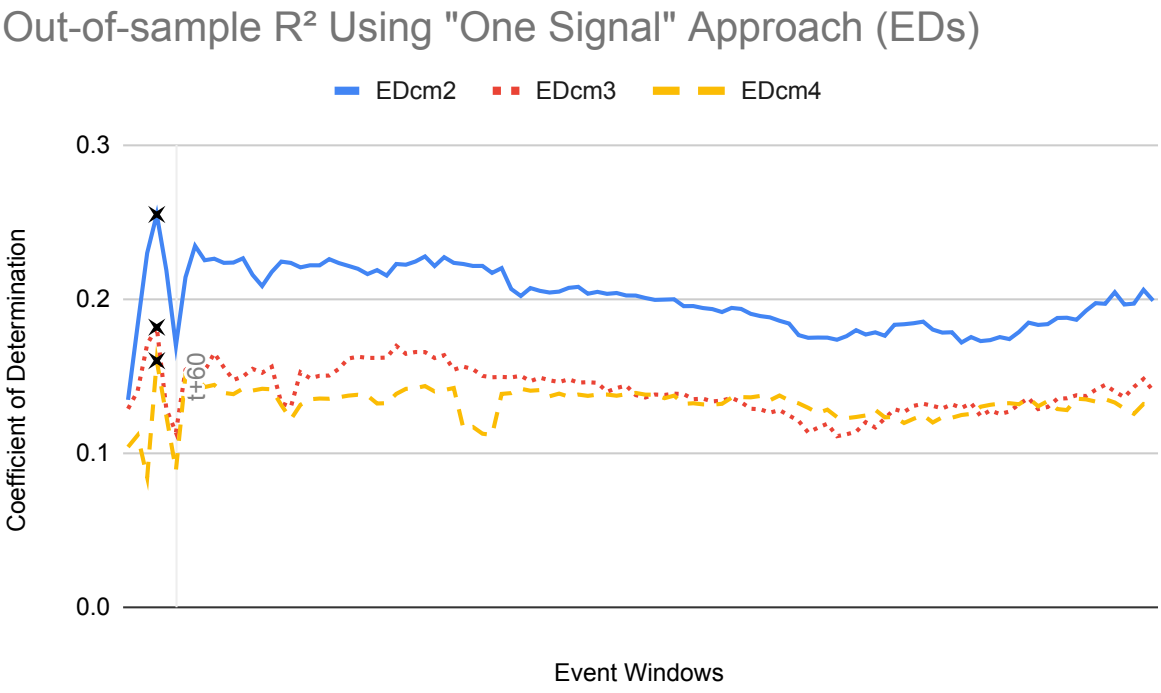


Figure E22: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for Eurodollar Futures

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.

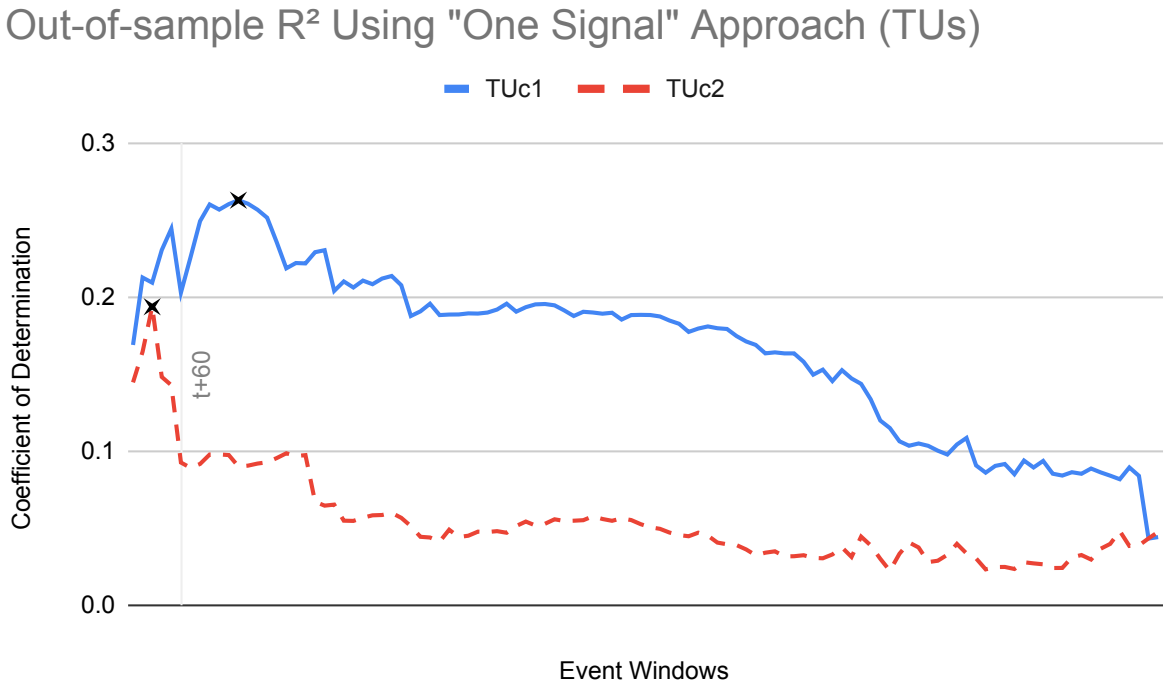


Figure E23: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for 2-year Treasury Futures

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.

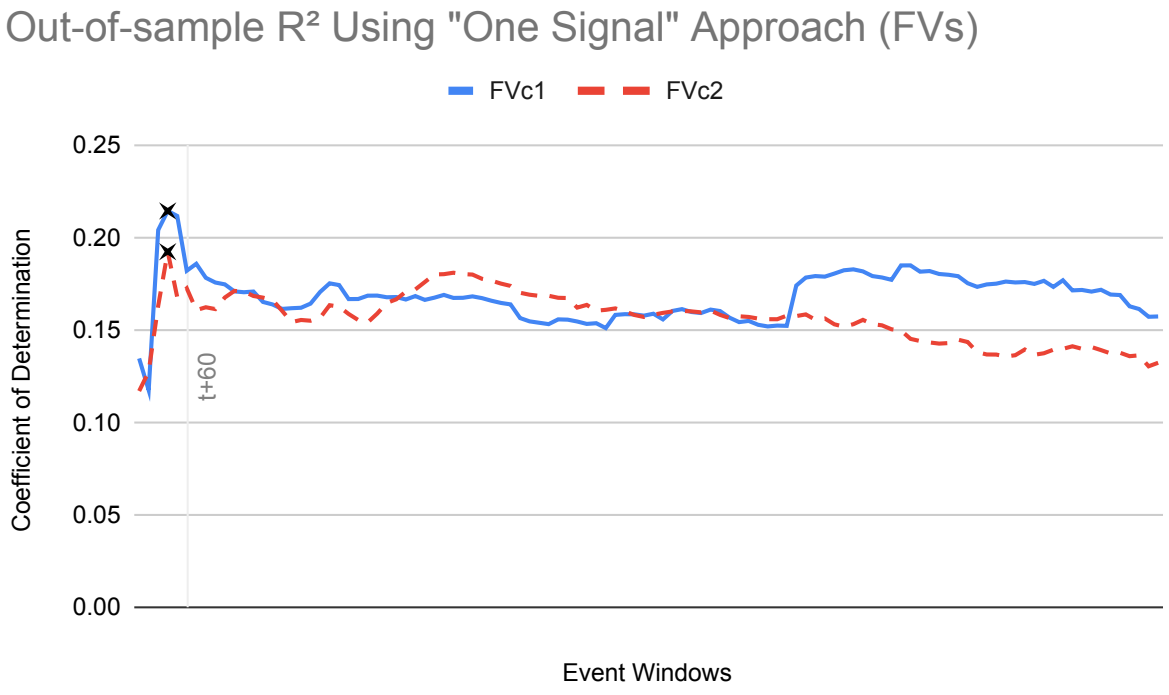


Figure E24: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for 5-year Treasury Futures

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.

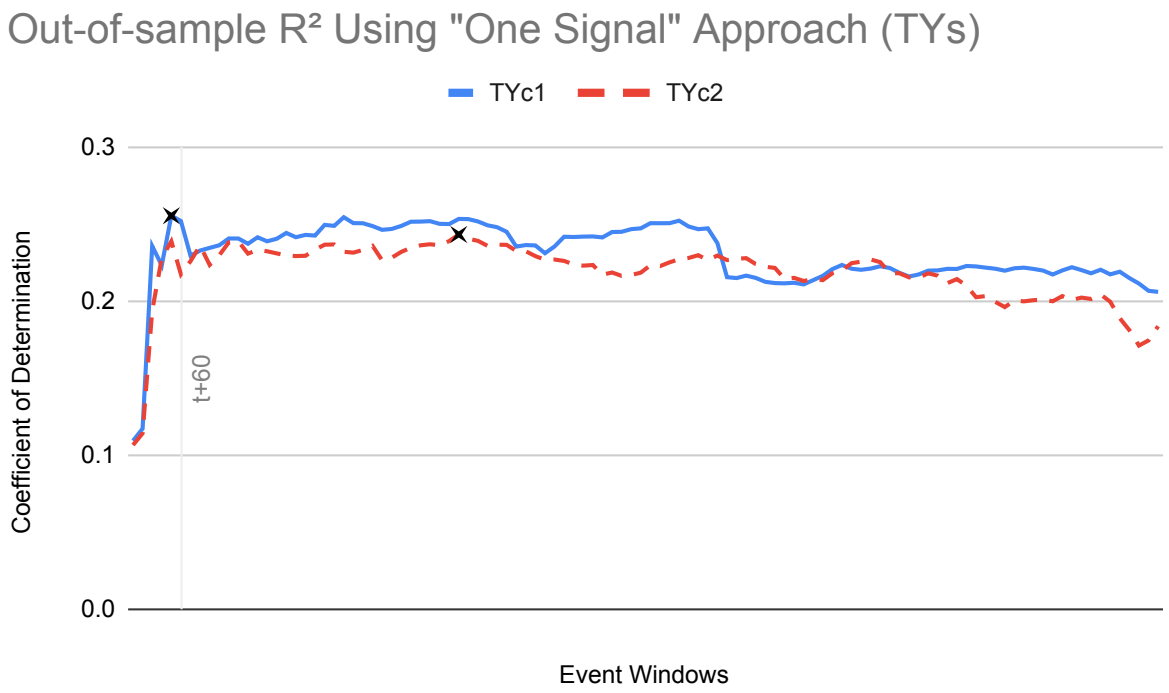


Figure E25: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for 10-year Treasury Futures

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.

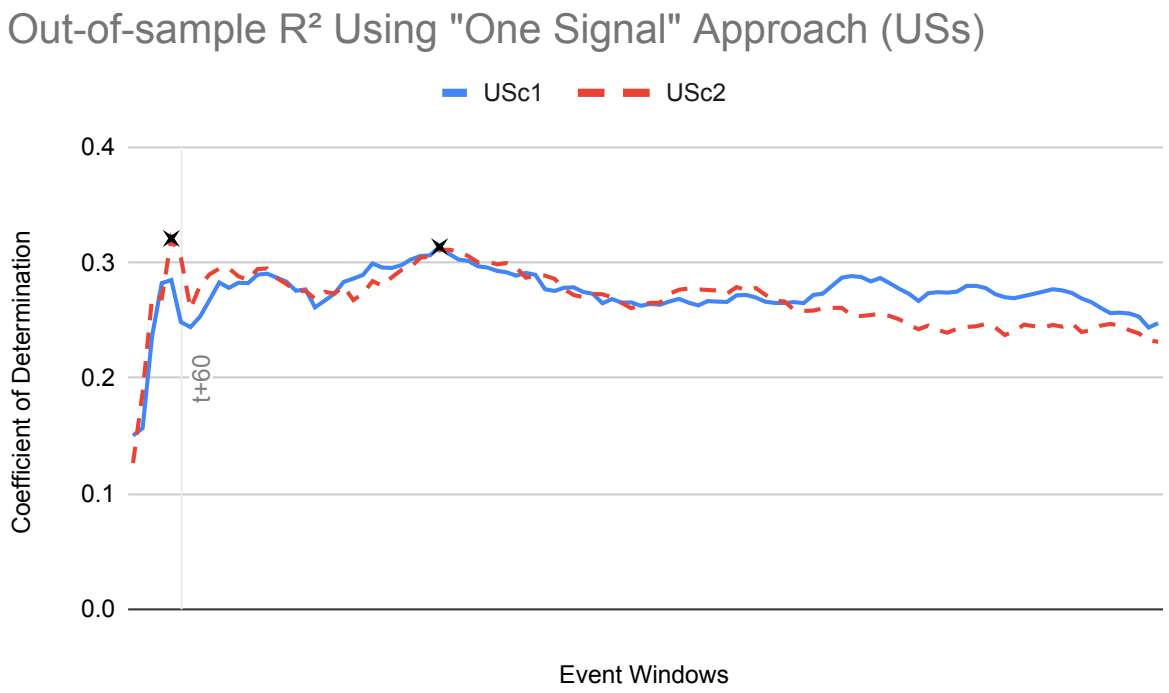


Figure E26: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for 30-year Treasury Futures

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.

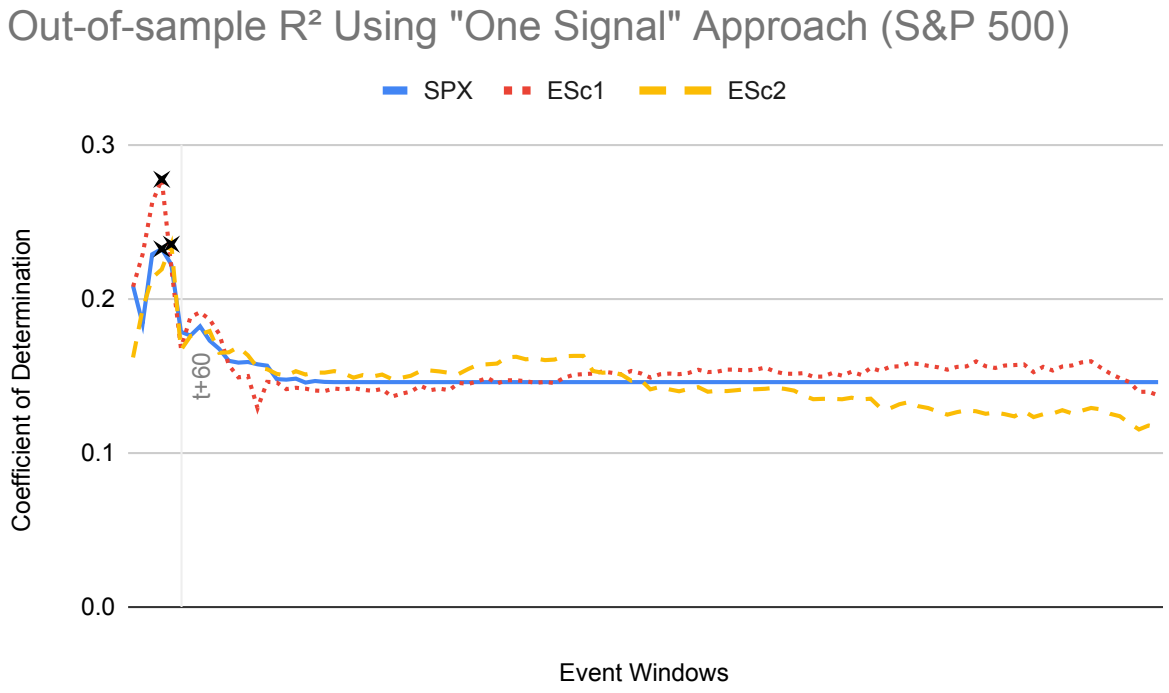
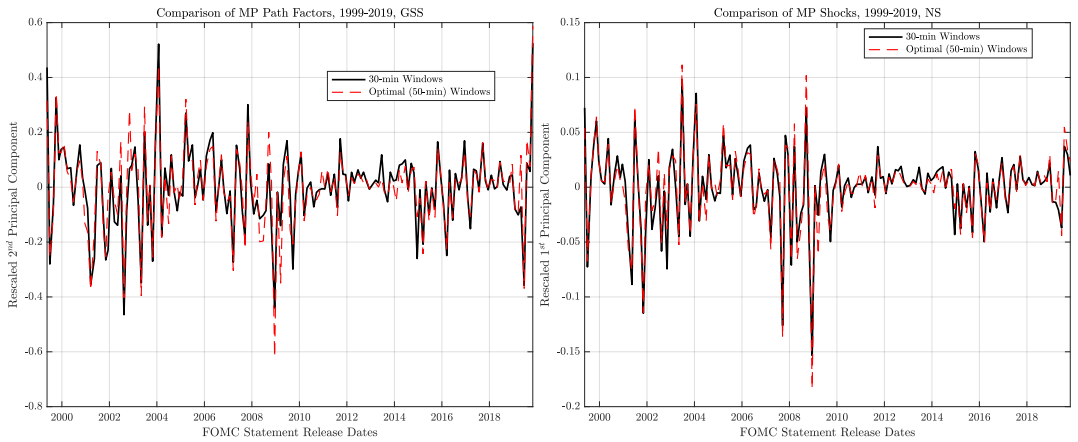


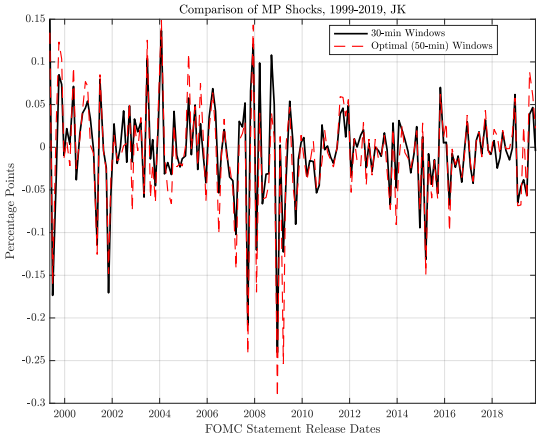
Figure E27: $\overline{R^2_{OOS}}$ Calculated Using the "One Signal" Approach for S&P 500 and E-mini Futures

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.

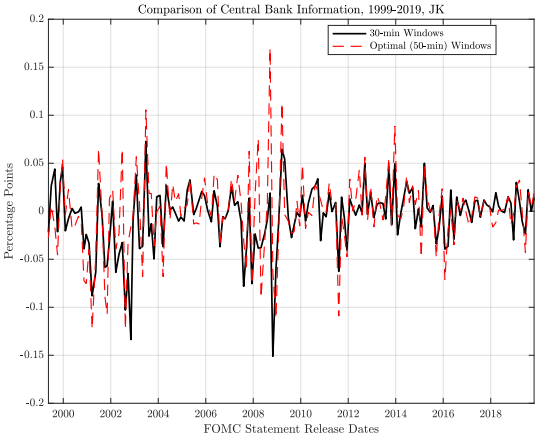


(a) GSS_P

(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.

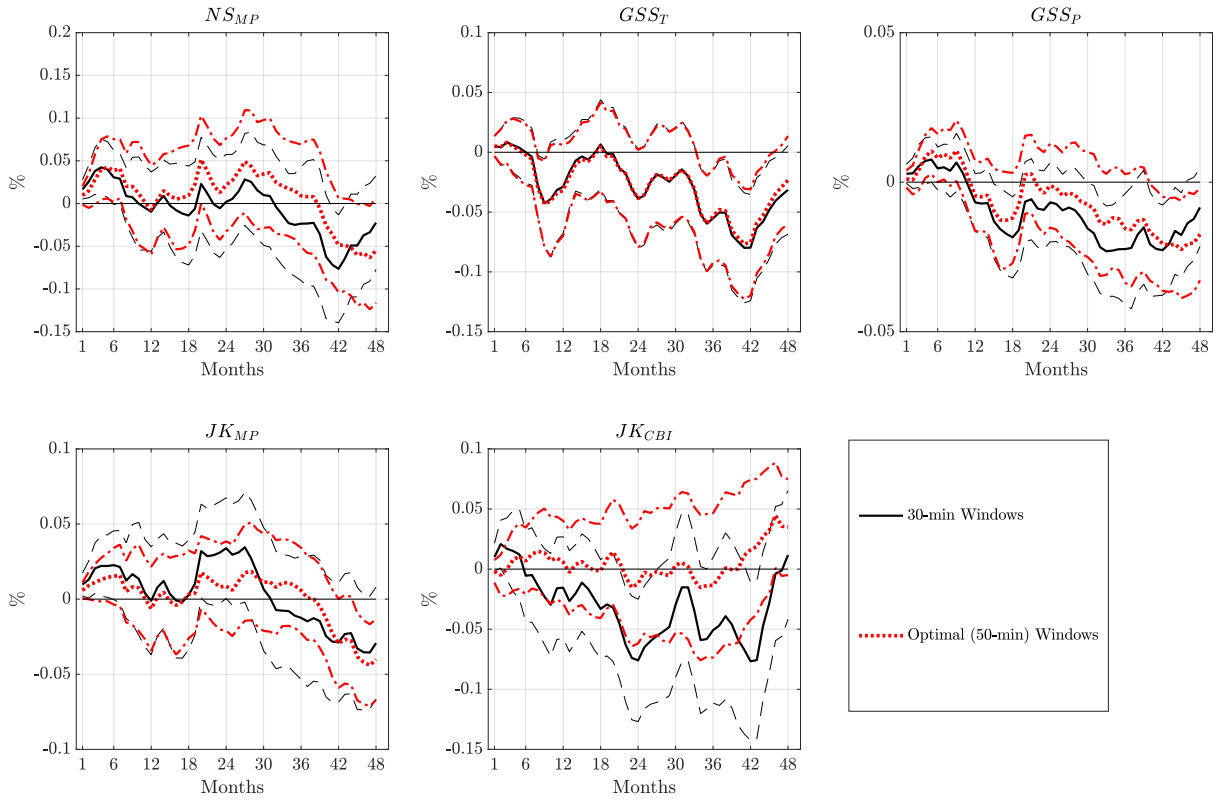


Figure E29: Effects of Event Window Choice on CPI Impulse Responses

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.

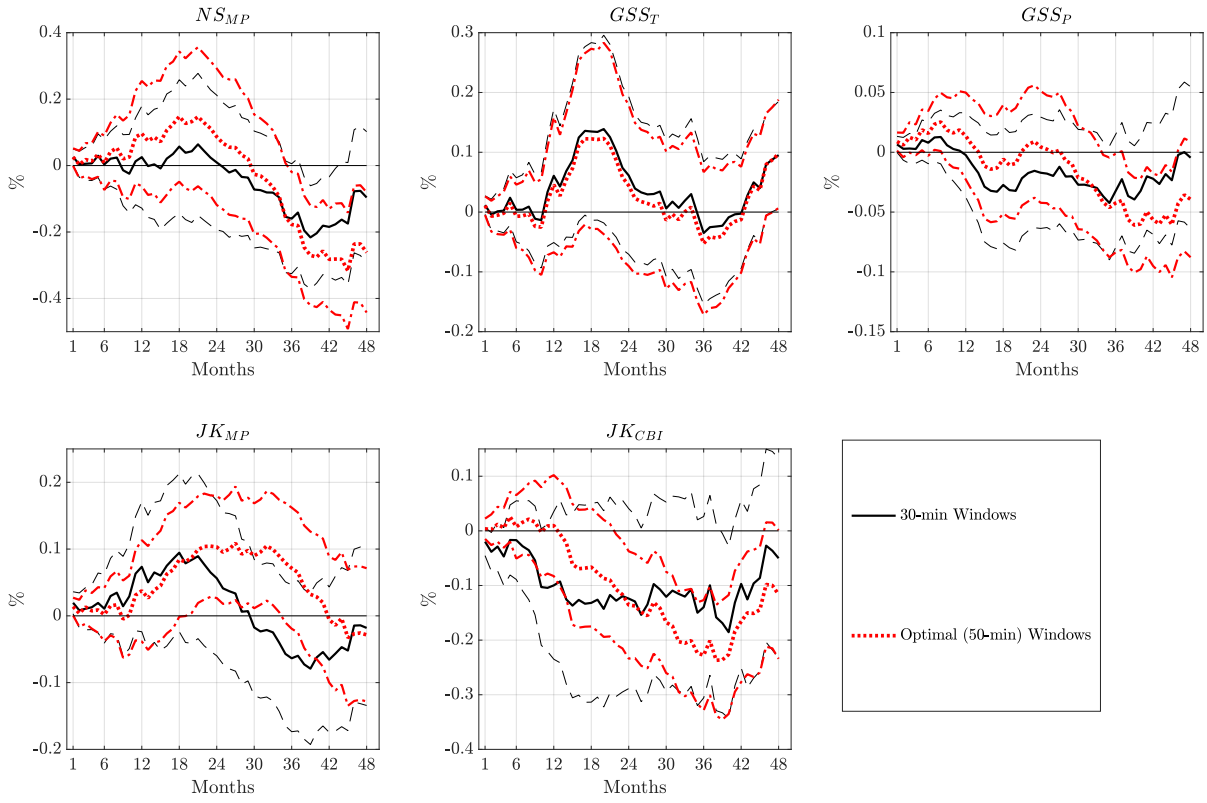


Figure E30: Effects of Event Window Choice on IP Impulse Responses

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.

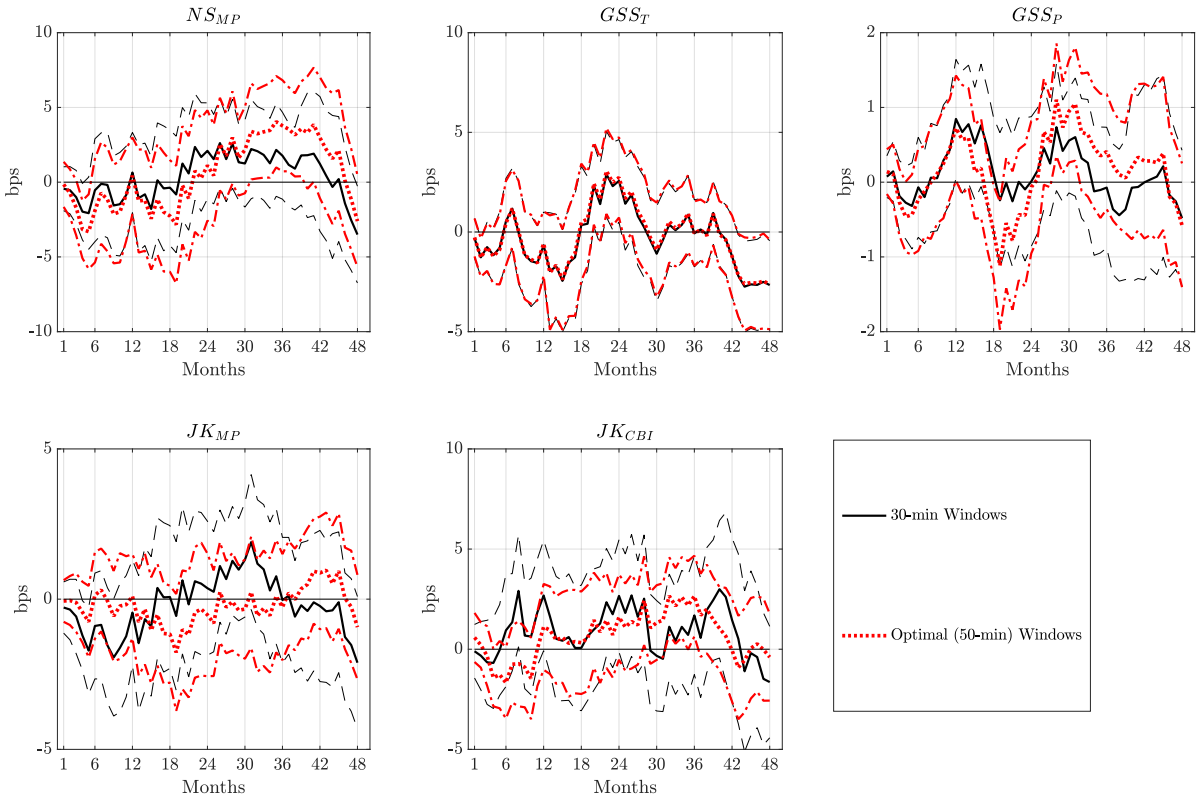


Figure E31: Effects of Event Window Choice on Excess Bond Premium Impulse Responses

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.

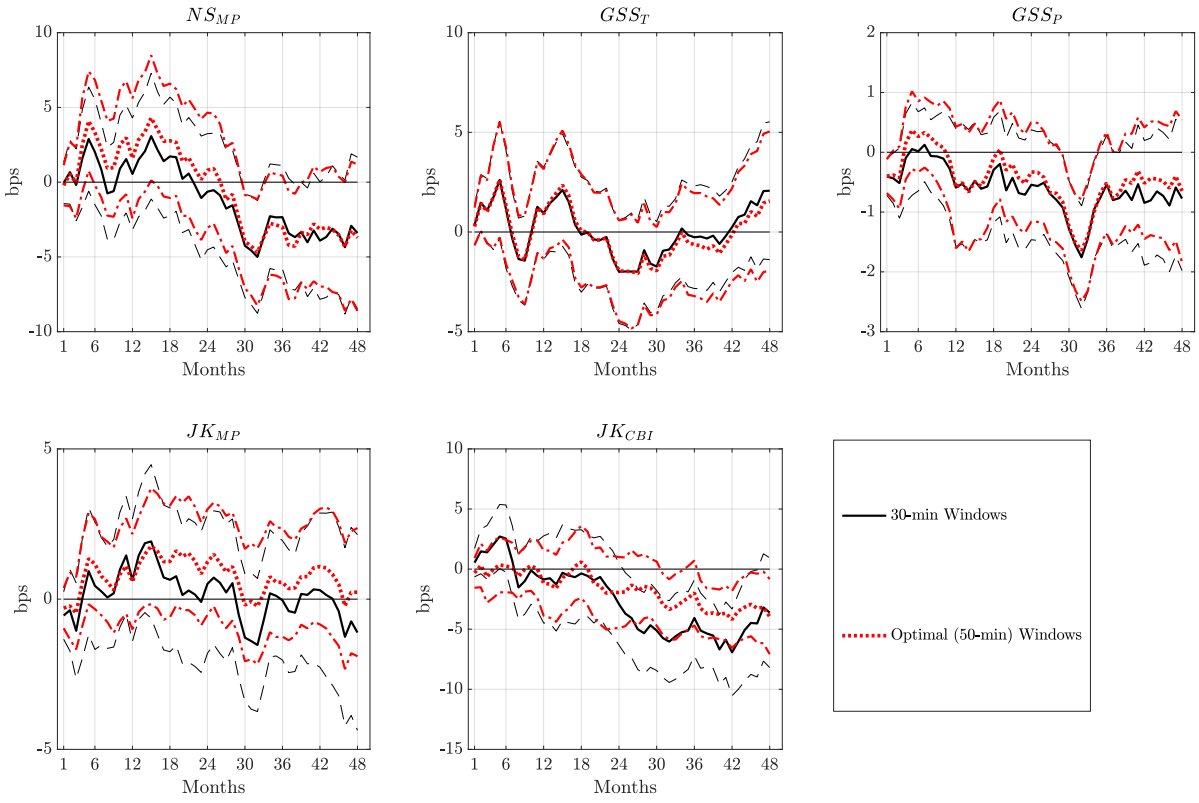


Figure E32: Effects of Event Window Choice on Two-year Treasury Yield Impulse Responses

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: $0.39 \times$ average sentence length $+ 11.8 \times$ 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.